LIMIT THEOREMS FOR FAST-SLOW PARTIALLY HYPERBOLIC SYSTEMS

JACOPO DE SIMOI AND CARLANGELO LIVERANI

Abstract. We prove several limit theorems for a simple class of partially hyperbolic fast-slow systems. We start with some well know results on averaging, then we give a substantial refinement of known large (and moderate) deviation results and conclude with a completely new result (a local limit theorem) on the distribution of the process determined by the fluctuations around the average. The method of proof is based on a mixture of standard pairs and transfer operators that we expect to be applicable in a much wider generality.

1. Introduction

In this paper we analyze various limit theorems for a class of partially hyperbolic systems of the fast-slow type. Such systems are very similar to the ones studied by Dolgopyat in [20]: in such paper the fast variables are driven by an hyperbolic diffeomorphism or flow (see also [4, 3, 38, 45, 27] for related results), here we consider the case in which they are driven by an expanding map. Notwithstanding the fact that we are not aware of an explicit treatment of the latter case, the difference is not so relevant to justify, by itself, a paper devoted to it. In fact, we chose to deal with one dimensional expanding maps only to simplify the exposition. The point here is that, on the one hand, we propose a different approach and, more importantly, on the other hand, we show that by such an approach it is possible to obtain much sharper results: a *Moderate and Large Deviation* Theorem and a *Local Limit* Theorem. To the best of our knowledge, this is the first time a rate function is computed with such a precision to yield moderate deviations of the paths and a local limit type theorem is obtained for a deterministic evolution converging to a diffusion process with non constant diffusion. Admittedly, the present is not the most general case one would like to deal with, it is just a primer. However, it shows that local limit results are attainable with an appropriate combination/refinement of present days techniques (see the discussion below on how general our approach really is).

The importance of local limit theorems hardly needs to be emphasized but, for the skeptical reader, it is nicely illustrated in [12, 14]. Indeed, in such papers the present large and moderate deviations and local limit results are used in a fundamental way to obtain a precise understanding of the statistical properties (e.g. existence and properties of the SRB measure, decay of correlations, metastability etc...) for the same class of systems for a small, but fixed, rate between the speeds of the slow and fast motions. This provides a class of partially hyperbolic

2000 *Mathematics Subject Classification.* 37A25, 37C30, 37D30, 37A50, 60F17.

Key words and phrases. Averaging theory, Large deviations, Limit Theorems, partially hyperbolicity.

We thank the Fields Institute, Toronto were this project started a very long time ago. Both authors have been partially supported by the ERC Grant MALADY (ERC AdG 246953). JDS acknowledges partial NSERC support. Most of all we would like to thank Dmitry Dolgopyat for his many comments and suggestion; without his advice this paper would certainly not exist. Finally, we are deeply indebted to the anonymous referees that we thank for their amazing job.

systems for which very precise *quantitative* statistical properties can be established. In addition, contrary to other cases, our results apply to an open set of systems (in the $\mathcal{C}^4$ topology).

For partially hyperbolic fast-slow systems several results concerning limit laws have already been obtained. In [4, 37] it is proven that the motion converges in probability to the motion determined by the averaged equation (morally a *law of large numbers*). In [20] there are important results on the fluctuations around the average (at a given time). In particular, both *large deviations* and converges in law to a diffusion for the fluctuation field (morally a *central limit theorem*) are obtained. In [4, 3] one can find very sharp results on normal fluctuation and moderate deviation at a given time. In particular, in [3] Bakhtin provides Cramer asymptotics for the distribution of the slow variable at a fixed time for a system with fast motion given by a mixing hyperbolic attractor. Such Cramer asymptotics gives estimates for moderate deviation, at a fixed time, sharper than the one obtained here, but they do not provide directly a rate function in path space, they hold only under the assumption that the dynamics is $\mathcal{C}^r$ for a very large r (contrary to our $\mathcal{C}^4$ assumption) and they are not sufficient to establish a local central limit theorem. In [38] more general large deviation results (in path space) are obtained. In particular, a variational formula for the rate function is established. Yet, Kifer's results are not precise enough to treat moderate deviations. To obtain a rate function for moderate deviations it is necessary to compute the exponential momenta with a precision considerably higher than the $o(1)$ achieved in [38]. Here we present independent proofs of the above facts (or, better, of the aforementioned substantial refinements of the above facts) and, most importantly, we make a further step forward by addressing the issue of the local central limit theorem, a result out of the reach of all previous approaches.

The lesson learned from [20] is that the *standard pair* technique is the best suited to investigate these type of partially hyperbolic systems.[1] Nevertheless, in the uniformly hyperbolic case, techniques based on the study of the spectrum of the *transfer operator* are usually much more efficient. It is then tempting to try to mix the two points of view as much as possible. This was partially done already in [3] and is also one of the goals of our work. To simplify matters, we carry it out it in the simplest possible setting (one dimensional expanding maps). Nevertheless, we like to remark that extending many of the present results to hyperbolic maps or flows is just a technical, not a conceptual problem. Indeed, till the recent past the use of transfer operators was limited to the expanding case (or could be applied only after coding the system via Markov partitions, greatly reducing the effectiveness of the method). Yet, recently, starting with [9] and reaching maturity with [29, 6, 30, 43, 51, 24, 16, 17, 22, 23, 26], it has been clarified how to fully exploit the power of transfer operators in the hyperbolic, partially hyperbolic and piecewise smooth setting. Accordingly, it is now totally reasonable to expect that any proof developed in the expanding case can be extended to the hyperbolic one, whereby making the following arguments of a much more general interest.[2]

The structure of the paper is as follows: we first describe the class of systems we are interest in, and state precisely the main results. Then we discuss in detail the standard pair technology. This must be done with care as we will need *higher*

---

[1] In particular, as far as we know, it represents the most efficient way to "condition" with respect to the past in a field (deterministic systems) where conditioning poses obvious conceptual problems.

[2] The only exception being the "Dolgopyat estimate" necessary to compute the error term in the local limit theorem which still poses a conceptual challenge in the general hyperbolic case, but see [50] for recent progresses.

smoothness as well as *complex* standard pairs, which have not been previously considered. In the following section we use the tools so far introduced to establish an averaging theorem. As already explained this result is not new, but it serves the purpose of illustrating the generals strategy to the reader and the proof contains several facts needed in the following arguments. Section 5 is devoted to the precise computation of the logarithmic moment generation function. This allows, in section 6, to establish the large and moderate deviations of our dynamics from the average. We compute with unprecedented precision the rate function of the large deviation principle. We stop short of providing a full large and moderate deviations theory only to keep the exposition simple and since it is not needed for our later purposes. Nevertheless, we improve considerably on known results. Finally in Section 8 we build on the previous work and prove a local limit result for our dynamics. The proof is a bit lengthy but it follows the usual approach: compute the Fourier transform of the distribution. This computation is very similar to the one in section 5 only now we want to compute the expectation of a complex exponential rather then a real one, also we aim at a better precision. Yet, the strategy is essentially the same: we divide the time interval in shorter blocks (this is done in Section 9), then estimate carefully the contribution of each block (this is done in Sections 10 and 11) and we conclude by combining together the contributions of the single blocks (done in section 13). Some fundamental technical tools needed to perform such computations are detailed in the appendices. Appendix A contains a manifold of results on transfer operators and their perturbation theory. In fact, not only it collects, for the reader convenience, many results scattered in the literature, but also provides some new results. In addition, it contains a discussion of the genericity of various conditions used in the paper including the, to us, unexpected results that for smooth maps aperiodicity and not being cohomologous to a constant are equivalent. Appendix B provides a detailed discussion of transfer operators associated to semiflows that, although essentially present in the literature, was not in the form needed for our needs (in particular we need uniform results for a one parameter family of systems). Finally, Appendix C contains some simple and uneventful, but a bit lengthy, computations needed in the text.

Notation. Through the paper we will use $C_\#$ and $c_\#$ to designate an arbitrary positive constant, depending only on our dynamical system, whose value can change form an occurrence to the next even in the same line. We will use $C_{a,b,\cdots}$ to designate arbitrary constants that depend on the quantities $a, b, \cdots$ while constants with other decorations (e.g. numbers as subscript) stand for a fixed specific value.

Also we write $\mathcal{O}(X)$ to denote a number which is bounded by $C_\# X$ for any $\varepsilon < \varepsilon_\#$, where $\varepsilon_\#$ depends only on the dynamics (note that X might not depend on ε, so that the second requirement becomes empty). While we will use $\mathcal{O}_{\mathcal{B}}(X)$, where $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is a Banach space, to denote an element of $\mathcal{B}$ with norm bounded by $C_\# \|X\|_{\mathcal{B}}$, again for all $\varepsilon \leq \varepsilon_\#$. We will always assume ε to be so small that this condition is met for every instance of the expression $\mathcal{O}(\cdot)$.

Finally, for $a \in \mathbb{R}$ we will use $\lfloor a \rfloor$ to designate its integer part, that is the largest integer that is smaller or equal to a.

Contents

2. The system and the results

For $\varepsilon > 0$ let us consider the map $F_\varepsilon \in \mathcal{C}^4(\mathbb{T}^2, \mathbb{T}^2)$ defined by

$$F_\varepsilon(x, \theta) = (f(x, \theta), \theta + \varepsilon\omega(x, \theta)), \tag{2.1}$$

where $\|\omega\|_{\mathcal{C}^4} = 1$. We assume that the $f_\theta = f(\cdot, \theta)$ are uniformly expanding, i.e.:

$$\inf_{(x,\theta)\in\mathbb{T}^2} \partial_x f(x, \theta) \geq \lambda, \tag{2.2}$$

for some $\lambda > 1$; indeed by considering a suitable iterate of F_ε, we will assume without loss of generality that $\lambda > 2$.

This fact is well known to imply that each $f(\cdot, \theta)$ has a unique invariant probability measure that is absolutely continuous with respect to the Lebesgue measure. We denote this measure (often called the SRB measure) by μ_θ. Also, we assume that, for every $\theta \in \mathbb{T}$, ω is not f_θ-cohomologous to a constant function, i.e.

(A1) For any $\theta \in \mathbb{T}$ there exist no measurable[3] function $g_\theta : \mathbb{R} \to \mathbb{R}$ and constant $a_\theta \in \mathbb{R}$ so that $\omega(x, \theta) = g_\theta(f(x, \theta)) - g_\theta(x) + a_\theta$.

Note that the latter equation can hold only if for any invariant probability measure μ of f_θ, $\mu(\omega(\cdot, \theta)) = \mu(a_\theta) = a_\theta$. In particular, if $\omega(\cdot, \theta)$ has different averages along two different periodic orbits of f_θ, then (A1) is satisfied. It is then fairly easy to check such a condition. In particular note that the assumption above holds generically (see Section A.5 for a more complete discussion of these issues).

Given $(x, \theta) \in \mathbb{T}^2$, let us define the trajectory $(x_n, \theta_n) = F_\varepsilon^n(x, \theta)$ for any $n \in \mathbb{N}$.

Here we describe a sequence of increasingly sharper results on the behavior of the dynamics for times of order ε^{-1}.[4] We start with well known facts, but we provide complete proofs both for the reader's convenience and because they are a necessary preliminary to tackle our main results.

2.1. **The skew product case.**

For the reader convenience, we first give a brief, impressionistic, discussion of the case in which $\partial_\theta f = \partial_\theta \omega = 0$. In this case the map F_ε is a *skew product*: $F_\varepsilon(x, \theta) = (f(x), \theta + \varepsilon\omega(x))$. This case is fairly well understood as it can be reduced to the study of the statistical properties of the map f, let us recall why.

We are interested in the evolution of the slow variable θ_n. Clearly we must wait for a time at least ε^{-1} in order for a change of order one to be possible. It is then natural to rescale the time and introduce the *macroscopic time* $t = \varepsilon n$. The idea is then to fix some arbitrary $T > 0$ and then define, for all $t \in [0, T]$,

$$\theta_\varepsilon(t) = \theta_{\lfloor t\varepsilon^{-1}\rfloor} + (t\varepsilon^{-1} - \lfloor t\varepsilon^{-1}\rfloor)[\theta_{\lfloor t\varepsilon^{-1}\rfloor+1} - \theta_{\lfloor t\varepsilon^{-1}\rfloor}] \mod 1.$$

Note that $\theta_\varepsilon \in \mathcal{C}^0([0, T], \mathbb{T})$. The point here is twofold: on the one hand it is clear that we cannot expect, at first, to control the behavior of θ_n for arbitrarily large n. Hence we fix a time horizon $T\varepsilon^{-1}$, T being arbitrary but independent on ε. On the other hand, it is natural to introduce a continuous interpolation of the evolution of θ_n since $|\theta_{\lfloor t\varepsilon^{-1}\rfloor} - \theta_{\lfloor s\varepsilon^{-1}\rfloor}| \leq |t - s|\|\omega\|_\infty$ hence, once rescaled, the trajectory is Lipschitz on $\varepsilon\mathbb{Z}$ and it is then naturally interpolated by a Lipschitz function on $\mathbb{R}$. Since

$$F_\varepsilon^n(x, \theta) = \left(f^n(x), \theta + \varepsilon \sum_{k=0}^{n-1} \omega(f^k(x))\right)$$

[3] It is well known by the Livšic Theorems that if g_θ is measurable, it is actually as smooth as the map f_θ (see also the proof of Lemma A.16).

[4] In some cases it is also possible to obtain information for times of the order ε^{-2} (see [20]). Yet, as far as we currently see, not of the quantitative type we are interested in.

it follows that

$$\left|\theta_\varepsilon(t) - \frac{t}{\lfloor \varepsilon^{-1}t\rfloor}\sum_{k=0}^{\lfloor \varepsilon^{-1}t\rfloor -1}\omega(f^k(x))\right| \leq C_\# \varepsilon.$$

By the Birkhoff Ergodic Theorem, the sum converges almost surely with respect to each invariant measure. This raises the issue of which measures to consider. In general this is an issue open to debate, however here we take the point of view that the fast variable x is originally distributed according to a probability measure absolutely continuous with respect to Lebesgue and with a smooth density. This means that we are interested in the so called *physical measures*. It is then well known that the distribution of x will tend exponential fast to the unique absolutely continuous invariant measure of f, call it μ, hence, x almost surely,[5]

$$\bar\theta(t) = \lim_{\varepsilon\to 0}\theta_\varepsilon(t) = t\mu(\omega) =: t\bar\omega =: \bar\theta(t).$$

That is, the limit satisfies the autonomous differential equation $\dot{\bar\theta} = \bar\omega$.

Next, one is interested in the deviations from such a limit. This leads us to the study of large deviations for an ergodic average. Such a problem has been intensively studies starting with [52] and the situation can be summarized as follows: consider the initial condition $\theta = \theta_0$ and x distribute as above, then $\gamma_\varepsilon(\cdot) = \theta_\varepsilon(\cdot) - \theta_0$ can be considered a random variable in $\mathcal{C}^0_*([0,T],\mathbb{T}) = \{\gamma \in \mathcal{C}^0 \;:\; \gamma(0) = 0\}$. Let $\mathbb{P}_\varepsilon$ be its law, then for a sufficiently regular set $Q \in \mathcal{C}^0_*([0,T],\mathbb{T})$

$$\mathbb{P}_\varepsilon(Q) \sim e^{-\varepsilon^{-1}\inf_{\gamma\in Q}\mathscr{I}(\gamma)}$$

where the *rate function* $\mathscr{I}$ is defined as

$$\mathscr{I}(\gamma) = \begin{cases} -\infty & \text{if } \gamma \text{ is not Lipschitz}\\ \int_0^T \mathcal{Z}(\gamma'(t))dt & \text{otherwise}\end{cases}$$
$$\mathcal{Z}(b) = -\sup_{\nu\in\mathcal{M}_\theta(b)}\{h_{\mathrm{KS}}(\nu) - \nu(\log f'_\theta)\},$$

$\mathcal{M}(b) = \{\nu \in \mathcal{M} \;:\; \nu(\omega) = b\}$, $\mathcal{M}$ denotes the set of (ergodic) f-invariant probability measures, and $h_{\mathrm{KS}}(\nu)$ is the Kolmogorov-Sinai entropy of the measure ν. The above formula is very suggestive: if the statistics of a point x is described by an invariant measure ν, then it will give rise to a trajectory θ_ε with velocity $\nu(\omega)$; moreover points that start in a $e^{-c_\# T\varepsilon^{-1}}$ neighborhood will have essentially the same trajectory for a time $T\varepsilon^{-1}$, hence the probability of order $e^{-c_\# T\varepsilon^{-1}}$ for such a trajectory. Unfortunately, the formula for $\mathcal{Z}$ is not very handy to compute. However the connection between the pressure and the maximal eigenvalue of the Ruelle transfer operator [5] allows to compute the rate function for smooth γ in a small neighborhood of $t\bar\omega$ yielding

$$\mathscr{I}(\gamma) \sim \frac{1}{2}\int_0^T \sigma^{-2}\left[\gamma'(s) - \bar\omega\right]^2 ds,$$

where, setting $\hat\omega = \omega - \bar\omega$, $\sigma^2 = \mu\left(\hat\omega^2\right) + 2\sum_{m=1}^\infty \mu_\theta\left(\hat\omega \circ f^m\hat\omega\right)$ is the variance of $\hat\omega$.

The above formula suggests that typical deviations are of order $\sqrt{\varepsilon}$. It is then natural to study the fluctuations $\zeta_\varepsilon(t) = \varepsilon^{-\frac{1}{2}}[\theta_\varepsilon(t) - \bar\theta(t)]$. This corresponds to the Central Limit theorem and its refinements (Local CLT, Berry-Essen estimates etc.). The CLT in this context states that

$$\mathbb{E}(\varphi(\zeta_\varepsilon(t)) \sim \int \varphi(x)\frac{e^{-x^2/2\sigma^2}}{\sigma^2\sqrt{2\pi}}.$$

[5] Note that, by the uniform Lipschitz property of the θ_ε it suffices to control the limit on countably many t to control it for all t.

Of course, for the applications it is essential to know *quantitatively* what the $\sim$ in the previous equations really means, that is we need an explicit estimate of the error. This is the task of the present paper as is explained shortly in the general case in which f, ω depend on the slow variable.

The basic idea used to extend results like the above to the general case is that the fast variable goes to its equilibrium (i.e. the physical measure) at an exponential speed. Hence in a times interval of size ε^α for some $\alpha \in (1, 0)$, the slow variable is almost constant and so is the dynamics. Since the invariant measure changes smoothly with the dynamics (*linear response*), then the statistical properties of the fast variable are more or less the same in the considered interval and large deviation results and LCLT hold. One can then use the Markov properties of the dynamics (here expressed in the formalism of *standard families*) to extend the result to longer times.

Note however that, while carrying out the above program, one must keep track of the mistakes in the various approximations and this is rather taxing. Especially if one needs to obtain very precise results like the ones achieved here. In fact, to understand if such error terms could be efficiently controlled was one of the motivations of the present paper. Finally, we should remark that most of our results are new even in the trivial skew product case discussed here.

2.2. **The Law of Large Numbers.** If we take the formal average with respect to the SRB measure of (2.1) we obtain the following first order differential equation

$$\frac{d\bar\theta}{dt} = \bar\omega(\bar\theta) \qquad \bar\theta(0) = \theta_0, \tag{2.3}$$

where $\bar\omega(\theta) = \mu_\theta(\omega(\cdot, \theta))$. For future use, let us also define the function $\hat\omega(x, \theta) = \omega(x, \theta) - \bar\omega(\theta)$.

Remark. *Note that, since $F_\varepsilon \in \mathcal{C}^4$, we can apply* [29, Theorem 8.1] *with the Banach spaces $\{\mathcal{C}^i\}_{i=0}^s$, $s = 3$ and obtain that $\bar\omega \in \mathcal{C}^{3-\alpha}$, for any $\alpha > 0$.*

Accordingly, the above equation has a unique solution, which we denote by $\bar\theta(t, \theta_0)$. This can be generalized: let $d \in \mathbb{N}$, $B = (B_1, \cdots, B_{d-1}) \in \mathcal{C}^2(\mathbb{T}^2, \mathbb{R}^{d-1})$, and fix $\zeta_0 = 0$; for any $k \in \mathbb{N}$ let us define

$$\zeta_{k+1} = \zeta_k + \varepsilon B(x_k, \theta_k). \tag{2.4}$$

This equation describes the evolution of a passive quantity and it is relevant in many situations (see e.g. [12, 14]). Then ζ_k should be *close* (in a sense that will be detailed shortly) to $\bar\zeta(\varepsilon k, \theta_0)$, the unique solution of

$$\frac{d\bar\zeta(t, \theta_0)}{dt} = \bar B(\bar\theta(t, \theta_0)) \qquad \bar\zeta(0, \theta_0) = 0, \tag{2.5}$$

where we introduced the averaged function $\bar B(\theta) = \mu_\theta(B(\cdot, \theta))$. It is then convenient to introduce the variables $z = (\theta, \zeta) \in \mathbb{R}^d$ (for convenience we have lifted θ to its universal cover) and $A = (A_1, \cdots, A_d) \in \mathcal{C}^3(\mathbb{T}^2, \mathbb{R}^d)$, with $A_1(x, \theta) = \omega(x, \theta)$ and $A_{i+1} = B_i$ for $i \in \{1, \cdots, d-1\}$. Then the evolution of the variables (x, z) is described by the map

$$\mathbb{F}_\varepsilon(x, z) = (f(x, \theta), z + \varepsilon A(x, \theta)); \tag{2.6}$$

again we set $(x_k, z_k) = \mathbb{F}_\varepsilon^k(x, z)$, for $k \in \mathbb{N}$. A first relevant fact is that the above averaging approximation can be justified rigorously. These type of results are well known and go back, at least, to Anosov [1]. Fix $T > 0$ and, for $t \in [0, T]$, let

$$z_\varepsilon(t) = (\theta_\varepsilon(t), \zeta_\varepsilon(t)) = z_{\lfloor t\varepsilon^{-1}\rfloor} + (t\varepsilon^{-1} - \lfloor t\varepsilon^{-1}\rfloor)[z_{\lfloor t\varepsilon^{-1}\rfloor+1} - z_{\lfloor t\varepsilon^{-1}\rfloor}]. \tag{2.7}$$

Observe that in the above definition we scale t in such a way that the slow variable moves of $O(1)$ for times t of order one. This in turns corresponds to study the F_ε^n

for $n \sim t\varepsilon^{-1}$. In fact, given $T > 0$, we will study the evolution of F_ε up to iterates of order $T\varepsilon^{-1}$. Then $z_\varepsilon \in C^0([0,T],\mathbb{R}^d)$, and we can consider it as a random element of $C^0([0,T],\mathbb{R}^d)$, the randomness being determined by the distribution of the initial condition.

Theorem 2.1 (Averaging). *Let $\theta_0 \in \mathbb{T}^1$ and x_0 be distributed according to a smooth distribution μ; then for all $T > 0$:*

$$\lim_{\varepsilon\to 0} z_\varepsilon = \bar{z}(\cdot,\theta_0)$$

where $\bar{z}(\cdot,\theta_0) = (\bar{\theta}(\cdot,\theta_0),\bar{\zeta}(\cdot,\theta_0))$ and the limit is in probability with respect to the measure μ and the uniform topology in $C^0([0,T],\mathbb{R}^d)$.

The proof is more or less standard. We provide it in Section 4 for reader's convenience. Indeed, our proof contains, in an elementary form, some of the ideas that will be instrumental in the following. The reader not very familiar with the transfer operator or standard pairs technology is advised to read Sections 3 and 4 first.

2.3. **Large and Moderate Deviations.** We will consider d and $A \in \mathcal{C}^3(\mathbb{T}^2,\mathbb{R}^d)$, $A_1(x,\theta) = \omega(x,\theta)$, to be fixed throughout the paper and to be data associated to the dynamical systems; although many quantities will depend on A, we do not add subscripts emphasize this dependence. In particular constants indicated with $C_\#$ or $c_\#$ may indeed depend on A.

We find convenient to define γ_ε to be the random element of $C^0([0,T],\mathbb{R}^d)$ obtained by subtracting to z_ε its (random) initial condition $z_\varepsilon(0)$, i.e. we let

$$\gamma_\varepsilon(t) = z_\varepsilon(t) - z_\varepsilon(0); \tag{2.8}$$

similarly, we define

$$\bar{\gamma}(t,\theta) = \bar{z}(t,\theta) - \bar{z}(0,\theta). \tag{2.9}$$

The next natural question concerns the behavior of deviations from the average. To this end it is more convenient to consider the fundamental probability space to be the classical Wiener space $C^0([0,T],\mathbb{R}^d)$ endowed with the Borel σ-algebra and the probability measure given by $\mathbb{P}_{\mu,\varepsilon} = (\gamma_\varepsilon)_*\mu$, where μ is the distribution of initial conditions on $\mathbb{T}^2$; in other words $\mathbb{P}_{\mu,\varepsilon}$ is the law of γ_ε under μ.

Note that the paths γ_ε are all Lipschitz with Lipschitz constant bounded by $\|A\|_{\mathcal{C}^0}$. To obtain stronger results we need some extra hypotheses:

(A1') for any $\theta \in \mathbb{T}$ and $\sigma \in \mathbb{R}^d$, the function $\langle\sigma, A(\cdot,\theta)\rangle$ is not f_θ-cohomologous to a constant (in particular, this implies (A1)).

Note that such condition is implied by the existence of $d+1$ periodic orbits for which the differences of the averages of A span $\mathbb{R}^d$. In particular, condition (A1') is generic.

Given this assumption, we prove upper and lower bounds for the probability of large and moderate deviations. The result we are after is much sharper than the one contained in [38]. It is of a more quantitative nature (in the spirit of [20] where the rate function is only estimated near zero and in a much rougher manner). In particular, we provide bounds on the rate function that allow to treat both large and moderate deviations for all ε small enough (not just asymptotically for $\varepsilon \to 0$). We refrain from developing a more complete theory[6] because on the one hand it would not change substantially the result, on the other hand it would increase the length of an already long paper and, finally, since the results presented here already

[6] For example, we do not strive for optimal results (such as the equivalence of the lower and upper bounds for all possible events in all the regimes under discussion, or the best possible estimate of the error terms).

more than suffices for our purposes (i.e. both for our later use and to pedagogically illustrates some ideas used in the following). In fact, the theorem that we state next does not contain even the full force of what we prove in Section 6, nevertheless its statement requires already quite a bit of preliminary notations. We advise the reader that wants a quick, but sub-optimal, idea of the type of results that can be obtained to jump directly to the Corollaries 2.6 and 2.7.

The first objects we need, as in any respectable large deviation theory, are *rate functions*. Their precise properties will be specified in detail in Section 6.1; here we summarize some basic facts. For any $\theta \in \mathbb{T}$ we define the set[7]

$$\mathbb{D}(\theta) = \{\mu(A(\cdot,\theta)) \ : \ \mu \text{ is a } f_\theta - \text{invariant probability}\}.$$

In other words $\mathbb{D}(\theta)$ is the set of all possible averages of A with respect to f_θ-invariant measures. Observe that $\mathbb{D}(\theta)$ can be determined with arbitrary precision by studying the periodic orbits of the dynamics (see Lemma 6.8 for details). The set $\mathbb{D}(\theta)$ is a compact convex subset of $\mathbb{R}^d$; it is also non-empty, since for any $\theta \in \mathbb{T}$, $\bar{A}(\theta) \in \mathbb{D}(\theta)$, where $\bar{A}(\theta) = \mu_\theta(A(\cdot,\theta))$ (observe that $\bar{A}(\theta)$ is deterministic, i.e. it is a non-random vector). Additionally, condition (A1') implies (see Lemma 6.2 for details) that $\bar{A}(\theta) \in \operatorname{int}\mathbb{D}(\theta)$ for any $\theta \in \mathbb{T}$. Let us now define the $d \times d$ matrix

$$\begin{aligned}\Sigma^2(\theta) =&\mu_\theta\left(\hat{A}(\cdot,\theta)\otimes\hat{A}(\cdot,\theta)\right) + \sum_{m=1}^{\infty}\mu_\theta\left(\hat{A}(f_\theta^m(\cdot),\theta)\otimes\hat{A}(\cdot,\theta)\right)\\ &+\sum_{m=1}^{\infty}\mu_\theta\left(\hat{A}(\cdot,\theta)\otimes\hat{A}(f_\theta^m(\cdot),\theta)\right),\end{aligned}$$

where $\hat{A} = A - \bar{A}$. Then $\Sigma \in \mathcal{C}^1(\mathbb{T}, M_d)$,[8] where M_d is the space of $d \times d$ symmetric non negative matrices. If (A1') holds, then Σ is invertible (see Lemma A.16).

In the following statement (and in the rest of the paper) we adopt the convention that $\inf\emptyset = +\infty$ (resp. $\sup\emptyset = -\infty$).

Proposition 2.2 (*Asymptotic* Large Deviation Principle)**.** *Let $\theta \in \mathbb{T}$, μ be a measure with smooth density on $\mathbb{T}$ and $\mathbb{P}_{\mu,\varepsilon} = (\gamma_\varepsilon)_*(\mu\times\delta_\theta)$. There exists a lower semicontinuous function $\mathscr{I}_\theta : C^0([0,T],\mathbb{R}^d) \to \mathbb{R}_{\geq 0}\cup\{+\infty\}$ (see (6.17) for an explicit definition) so that $\mathbb{P}_{\mu,\varepsilon}$ satisfies the Large Deviation Principle with rate function $\mathscr{I}_\theta$, that is: given any event $Q \subset C^0([0,T],\mathbb{R}^d)$ we have*

$$\begin{aligned}-\inf_{\gamma\in\operatorname{int}Q}\mathscr{I}_\theta(\gamma) &\leq \liminf_{\varepsilon\to 0}\varepsilon\log\mathbb{P}_{\mu,\varepsilon}(Q)\\ &\leq \limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}_{\mu,\varepsilon}(Q) \leq -\inf_{\gamma\in\overline{Q}}\mathscr{I}_\theta(\gamma).\end{aligned}\tag{2.10}$$

Note that $\mathscr{I}_\theta$ is not necessarily convex, yet it satisfies the following properties:

(a) *the* effective domain *$\mathfrak{D}(\mathscr{I}_\theta) := \{\gamma \in C^0([0,T],\mathbb{R}^d) \ : \ \mathscr{I}_\theta(\gamma) < \infty\}$ consists of Lipschitz paths such that $\gamma(0) = 0$ and, for almost all $t \in [0,T]$, the vector*[9] *$\gamma'(t) \in \mathbb{D}(\theta^\gamma(t))$, where $\theta^\gamma(t,\theta) = \theta + (\gamma(t))_1$.*[10] *In particular, this implies $\|\gamma'\|_{L^\infty} \leq \|A\|_{L^\infty}$ for any $\gamma \in \mathfrak{D}(\mathscr{I}_\theta)$.*

---

[7] For any $A \in \mathcal{C}^0(\mathbb{T},\mathbb{R}^d)$ and measure μ on $\mathbb{T}$ we define $\mu(A)$ to be the vector $(\mu(A_i)) \in \mathbb{R}^d$, where $\mu(A_i) = \int_{\mathbb{T}} A_i(x)\mu(dx)$.

[8] It follows from the fact that Σ can be seen as the second derivative of the eigenvalue of an appropriate transfer operator (A.12b), which, in turn is differentiable by Lemma A.9.

[9] Recall that by Rademacher's Theorem, any Lipschitz function is a.e. differentiable.

[10] Here $(\gamma(s))_1$ is the first component of the vector $\gamma(s)$: the one that corresponds to the θ motion. Also remark that, to ease notation, we will often suppress the θ dependency if no confusion arises.

(b) *for any* $\gamma \in \mathfrak{D}(\mathscr{I}_\theta)$*, the rate function* $\mathscr{I}_\theta$ *satisfies the following expansion:*

$$\left|\mathscr{I}_\theta(\gamma) - \frac{1}{2}\int_0^T \langle\gamma'(s) - \bar{A}(\theta^\gamma(s)), \left[\Sigma^2(\theta^\gamma(s))\right]^{-1}\left[\gamma'(s) - \bar{A}(\theta^\gamma(s))\right]\rangle ds\right| \le C_\#\|\gamma' - \bar{A}\circ\theta^\gamma)\|_{L^3}^3. \tag{2.11}$$

The above is the usual *asymptotic* large deviation principle, similar to what can be found in [38], although in a different setting. We are, however, interested in stating estimates valid for all, sufficiently small, ε not just in the limit $\varepsilon \to 0$.

In order to properly state results in the needed generality, we define, for each $\theta_0 \in \mathbb{T}$, a set $\mathcal{P}_\varepsilon(\theta_0)$ of *good* probability measures that are supported in a ε-neighborhood of θ_0. We refer to Section 3, in particular (3.23), for the precise definition, but, as an example, $\mu \times \delta_{\theta_0} \in \mathcal{P}_\varepsilon(\theta_0)$ where μ is a measure on $\mathbb{T}$ with a smooth distribution ρ, and the derivative of $\log\rho$ is bounded by some fixed constant. Here is a useful, but minimal, example of the kind of results we are after.

Proposition 2.3. *There exists* $T_{\max}, \varepsilon_0 \in (0,1)$*,* $\bar{C}, c_\star > 0$ *such that, for all* $\varepsilon \le \varepsilon_0$*,* $T \in [\varepsilon_0^{-4}\varepsilon, T_{\max}]$*,* $R \ge \bar{C}\sqrt{\varepsilon T}$ *and* $\theta_0^* \in \mathbb{T}$*, if we set*

$$Q_R = \{\gamma \in C^0([0,T],\mathbb{R}^d) \ : \ \|\gamma(\cdot) - \bar{\gamma}(\cdot,\theta_0^*)\|_{\mathcal{C}^0} \ge R\},$$

then, for any $\mu \in \mathcal{P}_\varepsilon(\theta_0^*)$ *and recalling* $\mathbb{P}_{\mu,\varepsilon} = (\gamma_\varepsilon)_*\mu$*, we have*

$$\mathbb{P}_{\mu,\varepsilon}(Q_R) \le \exp\left[-c_\star\varepsilon^{-1}T^{-1}R^2\right].$$

Proposition 2.3 is similar to [20, Theorem 6(b)], although in a different setting: our goal is to obtain stronger results encompassing the above ones. In particular, the previous results will be mere byproducts (see Section 7.5).

In order to properly present such result we introduce a slightly modified rate functions and we will state the result by saying that the probability of an event is controlled from above by the inf of the rate function on a slightly larger set and from below by the inf on a slightly smaller set. Also, if an event describes a small deviation from the average, then we can obtain effective bounds only if it is not too wild on a small given scale. Unfortunately, it is a bit tricky to make quantitatively precise these notions, so we ask for the reader patience.

First, for any $\Delta_* > 0$, we introduce functionals $\mathscr{I}^{\pm}_{\theta_0,\Delta_*}$ so that $\mathscr{I}^{-}_{\theta_0,\Delta_*} \le \mathscr{I}_{\theta_0} \le \mathscr{I}^{+}_{\theta_0,\Delta_*}$ but agree with $\mathscr{I}_{\theta_0}$ outside a Δ_* neighborhood of $\partial\mathfrak{D}(\mathscr{I}_{\theta_0})$.[11] Remark that we consider $\mathfrak{D}(\mathscr{I}_{\theta_0})$ as a subset of the Lipschitz functions with the associated topology, see Remark 6.14 for more details. In Lemma 6.15 we prove

$$\lim_{\Delta_*\to 0}\mathscr{I}^{-}_{\theta_0,\Delta_*} = \mathscr{I}_{\theta_0} \le \mathscr{I}^{+}_{\theta_0} = \lim_{\Delta_*\to 0}\mathscr{I}^{+}_{\theta_0,\Delta_*},$$

where $\mathscr{I}^{+}_{\theta_0}$ agrees with $\mathscr{I}_{\theta_0}$ everywhere apart from $\partial\mathfrak{D}(\mathscr{I}_{\theta_0})$ where it has value $+\infty$. Second, let $\theta_0 \in \mathbb{T}$, $\hat{\gamma}(t) = \gamma(t) - \bar{\gamma}(t,\theta_0)$ and define $R^{\pm} : C^0([0,T],\mathbb{R}^d) \to \mathbb{R}_{\ge 0}$ by

$$R^-(\gamma) = \varepsilon^{\frac{1}{2}} \qquad\qquad R^+(\gamma) = C_{\Delta_*,T}\left\{\varepsilon^{1/7}\|\hat{\gamma}\|_{L^\infty}^{5/7} + \sqrt{\varepsilon}\right\}, \tag{2.12}$$

for some appropriate constant $C_{\Delta_*,T}$. Then, for each $Q \subset \mathcal{C}^0([0,T],\mathbb{R}^d)$ let

$$Q^- = \{\gamma \in Q \ : \ B(\gamma, R^-(\gamma)) \subset Q\}\,; \qquad\qquad Q^+ = \bigcup_{\gamma\in Q} B(\gamma, R^+(\gamma)) \tag{2.13}$$

where $B(\gamma,r)$ is the standard C^0-ball in $C^0([0,T],\mathbb{R}^d)$. Obviously $Q^- \subset \operatorname{int} Q \subset \overline{Q} \subset Q^+$. Finally, we want to make precise what do we mean by event that are not

[11] Essentially $\mathscr{I}^{+}_{\theta_0,\Delta_*} = \infty$ in a Δ_*-neighborhood of $\partial\mathfrak{D}(\mathscr{I}_{\theta_0})$ while $\mathscr{I}^{-}_{\theta_0,\Delta_*} < \infty$ in the same neighborhood, see (6.16), (6.17), Section 6.1 and Lemma 6.6 for precise definitions.

too wild on a given scale. Let

$$\begin{aligned}\varrho(\theta_0, Q) &= \inf_{\gamma\in Q} \|\hat\gamma\|_\infty,\\ C_{\rm Lip}(\gamma) &= T^{-11/7}\varepsilon^{-2/7}\|\hat\gamma\|_{L^\infty}^{11/7}\\ \varsigma(\gamma) &= \sqrt{\varepsilon}\left(\frac{T^2\varepsilon}{\|\hat\gamma\|^2_{L^\infty}}\right)^{1/14}.\end{aligned} \tag{2.14}$$

Given a measure $\mathbb{P}$ on $\mathcal{C}^0$, we call an event a $Q\subset\mathcal{C}^0$ $\mathbb{P}$-*regular* if for $\mathbb{P}$-almost all $\gamma\in Q$ we have

$$|s-s'|\leq \frac{\varsigma(\gamma)}{2C_{\rm Lip}(\gamma)} \quad\Longrightarrow\quad \|\gamma(s)-\gamma(s')\|\leq\frac{\varsigma(\gamma)}{4}. \tag{2.15}$$

In other words, for each $\beta\in(0,1]$, points at a distance $\frac{\beta\varsigma(\gamma)}{2C_{\rm Lip}(\gamma)}$ yield a Lipschitz constant bounded by $C_{\rm Lip}(\gamma)/(2\beta)$.

We are now ready to state our first main result. Essentially, it is a quantitative version Proposition 2.2 which allows to, rather precisely, estimate the probability of events when ε is small, but non zero. In particular, it provides bounds for the speed at which the limits in Proposition 2.2 take place.

Theorem 2.4 (Large and Moderate deviations). *Let $T>0$; for all $\Delta_*>0$, ε small enough (depending on T and Δ_*), $\theta_0\in\mathbb{T}$, $\mu\in\mathcal{P}_\varepsilon(\theta_0)$, and for any $\mathbb{P}_{\mu,\varepsilon}$-regular event Q_ε (possibly depending on ε), we have*

$$\begin{aligned}\mathbb{P}_{\mu,\varepsilon}(Q_\varepsilon)&\leq e^{-\varepsilon^{-1}\left[\left(1-C_{\Delta_*,T}\varepsilon^{1/7}\varrho(\theta_0,Q_\varepsilon)^{-2/7}\right)\inf_{\gamma\in Q_\varepsilon^+}\mathscr{I}^-_{\theta_0,\Delta_*}(\gamma)\right]}\\ \mathbb{P}_{\mu,\varepsilon}(Q_\varepsilon)&\geq e^{-\varepsilon^{-1}\left[(1+C_{\Delta_*,T}\varepsilon^{1/2})\inf_{\gamma\in Q_\varepsilon^-}\mathscr{I}^+_{\theta_0,\Delta_*}(\gamma)+C_{\Delta_*,T}\varepsilon^{1/8}\right]}.\end{aligned} \tag{2.16}$$

The proof can be found in Section 7.4.

Remark 2.5. *Note that $\mathbb{P}_{\mu,\varepsilon}$ almost surely the paths have Lipschitz constant bounded by $\|A\|_{L^\infty}$. Hence, if $\varrho(\theta_0,Q)\geq C_\#\varepsilon^{2/11}$ (that is, the deviation is large enough) then Q_ε is always $\mathbb{P}_{\mu,\varepsilon}$ regular.*
Also, if $\inf_{\gamma\in Q_\varepsilon^+}\mathscr{I}^-_{\theta_0,\Delta_*}(\gamma)\leq C_{\Delta_*,T}\sqrt{\varepsilon}$, then it must be $\varrho(\theta_0,Q_\varepsilon)\leq C_{\Delta_*,T}\sqrt{\varepsilon}$ (see Lemma 6.16), and the coefficient in front of the rate function in the first of the (2.16) *becomes positive, therefore making the estimate empty.*
Finally, note that, using the results of Section 6 (in particular Lemmata 7.2 and 7.5) one could state the theorem in the case of a small T depending on ε. This is in fact not necessary: indeed any event in $\mathcal{C}^0([0,C_\#\varepsilon^\alpha],\mathbb{R}^d)$, $\alpha\in[0,1)$, can be seen as an event in $\mathcal{C}^0([0,T],\mathbb{R}^d)$. One can then check, using (2.11)*, that times larger than $C_\#\varepsilon^\alpha$ do no contribute to the* inf*, since any such event contains trajectories for which $\gamma'=\bar A$ for all $t\geq C_\#\varepsilon^\alpha$.*

The statement of Theorem 2.4, due to its precise quantitative nature, may feel a bit cumbersome. To help the reader understand its force we spell out few easy consequences in a form of corollaries. Their proof can be found in Section 7.5.

We already mentioned that Theorem 2.4 implies Proposition 2.2; yet the *finite size* version provided by Theorem 2.4 implies much more. Also note that, although the statement of Proposition 2.2 looks very clean, it is not very easy to use since the inf involved is often very hard to compute, even for a simple event like $Q=\{\gamma\in C^0([0,T],\mathbb{R}^d) : \|\gamma(s)-\bar\gamma(s,\theta_0)\|\geq Cs, s\in[0,T]\}$. For deviations that are not too large, one can get some more explicit estimates using the expansion of $\mathscr{I}_{\theta_0}$ stated in (2.11). The following corollary provides precise asymptotic estimates for paths that deviate from the average by at most $C_\#\varepsilon^\beta$, where $\beta\in(0,1/2)$.

Corollary 2.6 (Moderate deviations)**.** *Let $T>0$ and ε_0 small enough. For each $\varepsilon\le\varepsilon_0$, $\theta_0\in\mathbb{T}$, $\beta\in(0,\frac12)$ and Lipschitz bounded set $Q\subset\mathcal{C}^0([0,T],\mathbb{R}^d)$, i.e. the Lipschitz constant is uniformly bounded, define*[12]

$$Q_\varepsilon=\{\varepsilon^\beta\gamma(\cdot)+(1-\varepsilon^\beta)\bar\gamma(\cdot,\theta_0)\}_{\gamma\in Q}.$$

Then, for all $\mu\in\mathcal{P}_\varepsilon(\theta_0)$, $\mathbb{P}_{\mu,\varepsilon}=(\gamma_\varepsilon)_*\mu$, we have

$$\limsup_{\varepsilon\to0}\varepsilon^{1-2\beta}\log\mathbb{P}_{\mu,\varepsilon}(Q_\varepsilon)\le-\inf_{\gamma\in\overline{Q}}\mathscr{I}_{\mathrm{Lin},\theta_0}(\gamma),$$

where

$$\mathscr{I}_{\mathrm{Lin},\theta_0}(\gamma)=\frac12\int_0^T\langle\gamma'(s)-\bar A(\bar\theta(s,\theta_0)),\Sigma^2(\bar\theta(s,\theta_0))^{-1}\left[\gamma'(s)-\bar A(\bar\theta(s,\theta_0))\right]\rangle ds.$$

If, additionally, $\beta<\frac1{16}$ then[13]

$$\liminf_{\varepsilon\to0}\varepsilon^{1-2\beta}\log\mathbb{P}_{\mu,\varepsilon}(Q_\varepsilon)\ge-\inf_{\gamma\in\operatorname{int}Q}\mathscr{I}_{\mathrm{Lin},\theta_0}(\gamma).$$

In fact, Theorem 2.4 allows to estimate the probability of even smaller deviations, up to the scale of the Central Limit Theorem, whereby providing a strong refinement of Proposition 2.3.

Corollary 2.7 (Small deviations)**.** *For each $T>0$ and $\vartheta\in(0,1)$ there exists $\varepsilon_0,C_*>0$ such that, for each $\varepsilon\in(0,\varepsilon_0)$, $\mu\in\mathcal{P}_\varepsilon(\theta_0)$, $\mathbb{P}_{\mu,\varepsilon}=(\gamma_\varepsilon)_*\mu$, Lipschitz bounded event $Q\subset C^0([0,T],\mathbb{R}^d)$ such that $\varrho(\theta_0,Q)\ge C_*$ and setting $Q_\varepsilon=\{\varepsilon^{\frac12}\gamma(\cdot)+(1-\varepsilon^{\frac12})\bar\gamma(\cdot,\theta_0)\}_{\gamma\in Q}$, we have*

$$\mathbb{P}_{\mu,\varepsilon}(Q_\varepsilon)\le e^{-\vartheta\inf_{\gamma\in\hat Q^+}\mathscr{I}_{\mathrm{Lin},\theta_0}(\gamma)},$$

where $\hat Q^+=\bigcup_{\gamma\in Q}B(\gamma,\vartheta\|\gamma-\bar\gamma(\cdot,\theta_0)\|_\infty)$.

2.4. **Local Central Limit Theorem.** Given that in many cases we have seen that the upper bound in Proposition 2.3 is sharp, one expects that typical deviations are of order $\sqrt{\varepsilon T}$. It is then natural to wonder about their distribution. It is possible to prove (see e.g. [20, Theorem 5], where a related class of systems is investigated, or [13] for a pedagogical exposition of the present case) that the deviation of z_ε from the average, when rescaled by $\varepsilon^{-\frac12}$, converges towards a diffusion process. To simplify matters we will discuss only the case $d=1$, but similar results hold for any d.

Let us describe the above statement more precisely. Once again fix θ_0^*, let x be random and define $\Delta^\varepsilon(t)=\varepsilon^{-1/2}\left[\theta_\varepsilon(t)-\bar\theta(t,\theta_0^*)\right]$. Then, as $\varepsilon\to0$, the deviation $\Delta^\varepsilon(t)$ converges weakly to $\Delta(t)$, the solution of:

$$\begin{aligned}&d\Delta(t)=\bar\omega'(\bar\theta(t,\theta_0^*))\Delta_\ell(t)dt+\hat{\boldsymbol\sigma}(\bar\theta(t,\theta_0^*))dB(t)\\&\Delta(0)=0,\end{aligned}\tag{2.17}$$

where $B(t)$ is a standard Brownian motion and[14]

$$\hat{\boldsymbol\sigma}^2(\theta)=\mu_\theta\left(\hat\omega^2(\cdot,\theta)+2\sum_{m=1}^\infty\hat\omega(f_\theta^m(\cdot),\theta)\hat\omega(\cdot,\theta)\right).\tag{2.18}$$

Our next result provides a dramatic sharpening of the above statement.

---

[12] Hence there exists $C>0$ such that, if $\gamma\in Q_\varepsilon$, then $\|\gamma-\bar\gamma\|_{L^\infty}\le C\varepsilon^\beta$.

[13] Our techniques should allow to establish a similar lower bound also for $\beta\in[1/2,\frac1{16}]$, but at the price of further work. As is, if Q is open, we have only $\log\mathbb{P}_{\mu,\varepsilon}(Q_\varepsilon)\ge-C_\#\varepsilon^{-7/8}$, for $\beta\le\frac12$.

[14] Observe that this is nothing else that the matrix element $\Sigma^2_{1,1}$, which appeared in the moderate deviations.

Theorem 2.8. *For any $T > 0$, there exists $\varepsilon_0 > 0$ so that the following holds. For any $\beta > 0$, compact interval $I \subset \mathbb{R}$, real numbers $\kappa > 0$, $\varepsilon \in (0, \varepsilon_0)$ and $t \in [\varepsilon^{1/2000}, T]$, any fixed $\theta_0^* \in \mathbb{T}^1$ and $\mu \in \mathcal{P}_\varepsilon(\theta_0^*)$, we have:*

$$\left| \frac{\mathbb{P}_{\mu,\varepsilon}(\Delta^\varepsilon(t) \in \varepsilon^{1/2} I + \kappa)}{\sqrt{\varepsilon}} - \frac{e^{-\kappa^2/2\boldsymbol{\sigma}_t^2(\theta_0^*)}}{\boldsymbol{\sigma}_t(\theta_0^*)\sqrt{2\pi}} \mathrm{Leb}\, I \right| \le C_{T,\beta}\varepsilon^{1/2-7\beta}\mathrm{Leb}\, I + C_{T,\beta}\varepsilon^{1/2-\beta}, \tag{2.19}$$

where $\mathbb{P}_{\mu,\varepsilon} = (\gamma_\varepsilon)_*\mu$ and the variance $\boldsymbol{\sigma}_t^2(\theta)$ is given by

$$\boldsymbol{\sigma}_t^2(\theta) = \int_0^t e^{2\int_s^t \bar{\omega}'(\bar{\theta}(r,\theta))dr} \hat{\boldsymbol{\sigma}}^2(\bar{\theta}(s,\theta))ds. \tag{2.20}$$

Note that the Gaussian in equation (2.19) is indeed the solution of (2.17).[15] The proof of Theorem 2.8 is given in Section 8.2.

Remark 2.9. *If an Edgeworth expansion for (2.19) would hold, then one would expect the next term to be $\mathcal{O}(\varepsilon^{1/2}\mathrm{Leb} I)$, see [25]. Thus our error term is just slightly bigger than the expected first term in the Edgeworth expansion. In fact, with the technology put forward in this paper it should be possible to obtain such a correction to the CLT at the price of explicitly computing the main contribution of some terms that we have just estimated and considered errors. Unfortunately this, although feasible, is computationally heavy and we decided to avoid it to keep the length and readability of the paper (somewhat) under control.*

3. Standard pairs and families

In this section we introduce standard pairs and families for our system. As mentioned in the introductory section, this tool proved quite powerful in obtaining quantitative statistical results in systems with some degree of hyperbolicity. The first step is thus to establish some hyperbolicity result.

3.1. Dominated splitting.

Let us start with a preliminary inspection of the geometry of our system: for $\gamma^{\mathrm{u}}, \gamma^{\mathrm{c}} > 0$ to be specified later, let us define the *unstable cone* and the *center cone* as, respectively:

$$\mathfrak{C}^{\mathrm{u}} = \{(\xi, \eta) \in \mathbb{R}^2 \ : \ |\eta| \le \varepsilon\gamma^{\mathrm{u}}|\xi|\} \qquad \mathfrak{C}^{\mathrm{c}} = \{(\xi, \eta) \in \mathbb{R}^2 \ : \ |\xi| \le \gamma^{\mathrm{c}}|\eta|\}. \tag{3.1}$$

We claim that there exist $\gamma^{\mathrm{u}}, \gamma^{\mathrm{c}}$ such that, if ε is small enough, $dF_\varepsilon\mathfrak{C}^{\mathrm{u}} \subset \mathfrak{C}^{\mathrm{u}}$ and $dF_\varepsilon^{-1}\mathfrak{C}^{\mathrm{c}} \subset \mathfrak{C}^{\mathrm{c}}$. In fact, let us compute the differential of F_ε:

$$dF_\varepsilon = \begin{pmatrix} \partial_x f & \partial_\theta f \\ \varepsilon\partial_x\omega & 1 + \varepsilon\partial_\theta\omega \end{pmatrix}; \tag{3.2}$$

consequently, if we consider the vector $(1, \varepsilon u)$

$$\begin{aligned} d_pF_\varepsilon(1, \varepsilon u) &= (\partial_x f(p) + \varepsilon u\partial_\theta f(p), \varepsilon\partial_x\omega(p) + \varepsilon u + \varepsilon^2 u\partial_\theta\omega(p)) \\ &= \partial_x f(p)\left(1 + \varepsilon\frac{\partial_\theta f(p)}{\partial_x f(p)}u\right) \cdot (1, \varepsilon\Xi_p(u)) \end{aligned} \tag{3.3}$$

where

$$\Xi_p(u) = \frac{\partial_x\omega(p) + (1 + \varepsilon\partial_\theta\omega(p))u}{\partial_x f(p) + \varepsilon\partial_\theta f(p)u}, \tag{3.4}$$

from which we obtain our claim, choosing for instance

$$\gamma^{\mathrm{u}} = 2\|\partial_x\omega\|_\infty \qquad \text{and} \qquad \gamma^{\mathrm{c}} = 2\|\partial_\theta f\|_\infty. \tag{3.5}$$

---

[15] If in doubt, see [13] for details.

In fact, for any λ' so that $\lambda > \lambda' > 3/2$, we can choose ε so small that if $|u| < \gamma^{\mathrm{u}}$:

$$|\Xi_p(u)| < \frac{\|\partial_x\omega\|_\infty + |u|}{\lambda'} < \gamma^{\mathrm{u}},$$

which proves invariance of $\mathfrak{C}^{\mathrm{u}}$ under dF_ε. Invariance of $\mathfrak{C}^{\mathrm{c}}$ can be similarly established. Hence, for any $p \in \mathbb{T}^2$ and $n \in \mathbb{N}$, we can define the quantities $\upsilon_n^+, u_n, s_n, \upsilon_n^{\mathrm{c}}$ as follows:

$$d_pF_\varepsilon^n(1,0) = \upsilon_n^+(1, \varepsilon u_n) \qquad\qquad d_pF_\varepsilon^n(s_n,1) = \upsilon_n^{\mathrm{c}}(0,1) \tag{3.6}$$

with $|u_n| \le c$ and $|s_n| \le K^{-1}$. Notice that $d_pF_\varepsilon(s_n(p),1) = \upsilon_n^{\mathrm{c}}/\upsilon_{n-1}^{\mathrm{c}}(s_{n-1}(F_\varepsilon p),1)$; therefore, there exists a constant b such that:

$$\exp[-b\varepsilon] \le \frac{\upsilon_n^{\mathrm{c}}}{\upsilon_{n-1}^{\mathrm{c}}} \le \exp[b\varepsilon]\,. \tag{3.7}$$

Furthermore, define $\Gamma_n = \prod_{k=0}^{n-1} \partial_x f \circ F_\varepsilon^k$, and let

$$a = c\left\|\frac{\partial_\theta f}{\partial_x f}\right\|_\infty.$$

Clearly

$$\Gamma_n \exp[-a\varepsilon n] \le \upsilon_n^+ \le \Gamma_n \exp[a\varepsilon n]\,. \tag{3.8}$$

3.2. **Standard pairs: definition and properties.**

We now proceed to define standard pairs for our system: we begin by introducing real standard pairs (which are just special probabilities measures), and then proceed to extend our definitions to complex standard pairs.

3.2.1. *Real standard pairs.* Let us fix a small $\delta > 0$, and $D_1, D_1' > 0$ large to be specified later; for any $c_1 > 0$ let us define the set of functions

$$\Sigma_{c_1} = \{G \in \mathcal{C}^3([a,b],\mathbb{T}^1)\ : a,b \in \mathbb{T}^1, b-a \in [\delta/2,\delta], \\ \|G'\| \le \varepsilon c_1,\ \|G''\| \le \varepsilon D_1 c_1,\ \|G'''\| \le \varepsilon D_1' c_1\}.$$

Let us associate to each $G \in \Sigma_{c_1}$ the map $\mathbb{G}(x) = (x, G(x))$ whose image is a curve –the graph of G– which will be denoted by $\gamma_{\mathbb{G}}$; such curves are called *standard curves*. For any $c_2, c_3 > 0$ define the set of (c_2,c_3)-*standard* probability densities on the standard curve $\gamma_{\mathbb{G}}$ as

$$D^{\mathbb{R}}_{c_2,c_3}(G) = \left\{\rho \in \mathcal{C}^2([a,b],\mathbb{R}_{>0})\ :\ \int_a^b \rho(x)dx = 1,\ \left\|\frac{\rho'}{\rho}\right\| \le c_2,\ \left\|\frac{\rho''}{\rho}\right\| \le c_3\right\}.$$

A *real* (c_1,c_2,c_3)-*standard pair* ℓ is given by $\ell = (\mathbb{G},\rho)$ where $G \in \Sigma_{c_1}$ and $\rho \in D^{\mathbb{R}}_{c_2}(G)$. A real standard pair $\ell = (\mathbb{G},\rho)$ induces a probability measure μ_ℓ on $\mathbb{T}^2$ defined as follows: for any Borel-measurable function g on $\mathbb{T}^2$ let

$$\mu_\ell(g) := \int_a^b g(x, G(x))\rho(x)dx.$$

We define[16] a *standard family* $\mathfrak{L} = (\{\ell_j\}, \nu)$ as a (finite or) countable collection of standard pairs $\{\ell_j\}$ endowed with a finite factor measure ν, i.e. we associate to each standard pair ℓ_j a positive weight ν_{ℓ_j} so that $\sum_{\ell\in\mathfrak{L}} \nu_\ell < \infty$. A standard family $\mathfrak{L}$ naturally induces a finite measure $\mu_{\mathfrak{L}}$ on $\mathbb{T}^2$ defined as follows: for any Borel-measurable function g on $\mathbb{T}^2$ we let

$$\mu_{\mathfrak{L}}(g) := \sum_{\ell\in\mathfrak{L}} \nu_\ell \mu_\ell(g).$$

[16] We remark that this is not the most general definition of standard family, yet it suffices for our purposes and it allows to greatly simplify our notations.

A standard family is a standard probability family if the induced measure is a probability measure (i.e. if ν is itself a probability measure). Let us denote by $\sim$ the equivalence relation induced by the above correspondence i.e. we let $\mathfrak{L} \sim \mathfrak{L}'$ if and only if $\mu_{\mathfrak{L}} = \mu_{\mathfrak{L}'}$.The key property of the above objects is that the pushforward of a standard family is a standard family [12, Proposition 5.2].

Unfortunately, to study large deviations we will need to consider a more general pushforward in which the density is first multiplies by some real positive function (called *weight*, which logarithm is called *potential*) and then pushforwarded (see equation (3.11)). This is analogous to the use of twisted transfer operators so useful in the analytic approach to the statistical properties of dynamical systems [5]. Yet, for the study of the CLT not even this suffices: we need to multiply the density by a complex phase. It is then necessary to generalize the above concepts to the complex setting. As the proofs for complex and real weights are essentially the same, we proceed directly in introducing complex potentials and prove the needed generalization of [12, Proposition 5.2] : Proposition 3.3.

3.2.2. *Complex standard pairs.* We now proceed to introduce *complex standard pairs*. Let us first define the set of complex standard densities:

$$D^{\mathbb{C}}_{c_2,c_3}(G) = \left\{ \rho \in \mathcal{C}^2([a,b],\mathbb{C}^*) \;:\; \int_a^b \rho(x)dx = 1,\; \left\| \frac{\rho'}{\rho} \right\| \leq c_2,\; \left\| \frac{\rho''}{\rho} \right\| \leq c_3 \right\}, \tag{3.9}$$

where we denote $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Yet, this time, for technical reasons, we cannot chose the length fixed once an for all. So we will consider standard curves $\Sigma^{\mathbb{C}}_{c_1}$ made of curves of length $b - a \in [\delta_{\mathbb{C}}/2, \delta_{\mathbb{C}}]$ for some $\delta_{\mathbb{C}} \in (0, \delta)$. We then require $c_2\delta_{\mathbb{C}} \leq \pi/10$. A *complex standard pair* is then given by $\ell = (\mathbb{G}, \rho)$ where $G \in \Sigma^{\mathbb{C}}_{c_1}$ and $\rho \in D^{\mathbb{C}}_{c_2,c_3}(G)$; a complex standard pair induces a natural complex measure on $\mathbb{T}^2$. A complex standard family $\mathfrak{L}$ is defined as its real counterpart, but now we allow ℓ_j's to be complex standard pairs and ν to be a complex measure so that $\sum_{\ell\in\mathfrak{L}} |\nu_\ell| < \infty$. Clearly, a complex standard family naturally induces a complex measure on $\mathbb{T}^2$.

We will say that $b - a$ is the *length* of the standard pair and we will say that a family $\mathfrak{L}$ has length $\delta_{\mathbb{C}}$ if each $\ell \in \mathfrak{L}$ has length $b_\ell - a_\ell \in [\delta_{\mathbb{C}}/2, \delta_{\mathbb{C}}]$.

Lemma 3.1 (Variation)**.** *Let $G \in \Sigma^{\mathbb{C}}_{c_1}$ be a standard curve and $\rho \in D^{\mathbb{C}}_{c_2,c_3}(G)$; if δ is sufficiently small we have:*

$$\operatorname{Range}\rho \subset \{z = re^{i\vartheta} \in \mathbb{C} \;:\; e^{-2c_2\delta_{\mathbb{C}}} < r(b-a) < e^{2c_2\delta_{\mathbb{C}}},\ |\vartheta| < c_2\delta_{\mathbb{C}}\}.$$

Proof. Observe that, by definition of standard density we have $\|(\log\rho)'\| \leq c_2$; since we are assuming $c_2\delta_{\mathbb{C}} \leq \pi/10$, we can unambiguously define the function $\log\rho$, which is contained in a square $S \subset \mathbb{C}$ of side $c_2\delta_{\mathbb{C}}$. Thus, $\operatorname{Range}\rho \subset \exp S$, which is an annular sector. The normalization condition $\int \rho = 1$ and the mean value theorem imply that $\exp S$ must non-trivially intersect the sets $\{\operatorname{Re} z = (b-a)^{-1}\}$ and $\{\operatorname{Im} z = 0\}$; these two conditions immediately imply that $\exp S \subset \{re^{i\vartheta} \in \mathbb{C} :$ $|\vartheta| < c_2\delta_{\mathbb{C}}\}$. It is then immediate to show that

$$\exp S \subset \left\{ re^{i\vartheta} \;:\; e^{-c_2\delta_{\mathbb{C}}} < r(b-a) < \frac{1}{\cos(c_2\delta_{\mathbb{C}})} e^{c_2\delta_{\mathbb{C}}} \right\},$$

which concludes the proof. □

Remark 3.2. *The above lemma also implies a uniform $\mathcal{C}^2$ bound on standard densities given by $\|\rho\|_{\mathcal{C}^2} \leq e^{2c_2\delta_{\mathbb{C}}} c_3 \delta_{\mathbb{C}}^{-1}$. Moreover, we have*

$$|\mu_\ell| < e^{2c_2\delta_{\mathbb{C}}}, \tag{3.10}$$

where $|\mu_\ell|$ is the standard total variation norm.

The key property of the class of real standard pairs is its invariance under push-forward by the dynamics; we are now going to prove a more general result. Let $\mathfrak{P} \subset \mathcal{C}^2(\mathbb{T}^2, \mathbb{C})$ be a family of smooth functions with uniformly bounded $\mathcal{C}^2$-norm; we denote by $\|\mathfrak{P}\|_{\mathcal{C}^r} = \sup_{\Omega \in \mathfrak{P}} \|\Omega\|_{\mathcal{C}^r}$. For any $\Omega \in \mathfrak{P}$ define the operator $F_{\varepsilon*,\Omega}$ acting on a complex measure μ as follows: for any measurable function g of $\mathbb{T}^2$

$$[F_{\varepsilon*,\Omega}\mu](g) := \mu(e^{\Omega} \cdot g \circ F_\varepsilon). \tag{3.11}$$

We call $F_{\varepsilon*,\Omega}$ the *weighted push-forward operator with potential* Ω; observe that $F_{\varepsilon*,0} = F_{\varepsilon*}$ is the usual push-forward.

Proposition 3.3 (Invariance). *Given a family of complex potentials $\mathfrak{P}$, there exist c_1, c_2, c_3 and δ such that the following holds. For any $\Omega \in \mathfrak{P}$ and complex (c_1, c_2, c_3)-standard family $\mathfrak{L}$ of length $\delta_{\mathbb{C}} \le \min\{\delta, \pi/(10\, c_2)\}$, the complex measure $F_{\varepsilon*,\Omega}\mu_{\mathfrak{L}}$ can be decomposed in complex (c_1, c_2, c_3)-standard pairs, i.e. there exists a complex (c_1, c_2, c_3)-standard family $\mathfrak{L}'_\Omega$, of length $\delta_{\mathbb{C}}$, such that $F_{\varepsilon*,\Omega}\mu_{\mathfrak{L}} = \mu_{\mathfrak{L}'_\Omega}$. We say that $\mathfrak{L}'_\Omega$ is a (c_1, c_2, c_3)-*standard decomposition *of $F_{\varepsilon*,\Omega}\mu_{\mathfrak{L}}$ and we write –with a little abuse of notation– $\mathfrak{L}'_\Omega \sim F_{\varepsilon*,\Omega}\mathfrak{L}$. Moreover, the constant c_1 does not depend on $\mathfrak{P}$, whereas the constants c_2 and c_3 (and consequently δ) can be chosen as follows:*

$$c_2 \ge C_\#(1 + \|\mathfrak{P}\|_{\mathcal{C}^1}) \qquad\qquad c_3 \ge C_\#(1 + \|\mathfrak{P}\|_{\mathcal{C}^2} + \|\mathfrak{P}\|^2_{\mathcal{C}^1}). \tag{3.12}$$

Proof. For simplicity, let us assume that $\mathfrak{L}$ is given by a single complex standard pair ℓ; the general case does not require any additional ideas and it is left to the reader.

Let then $\ell = (\mathbb{G}, \rho)$ be a complex (c_1, c_2, c_3)-standard pair. For any sufficiently smooth function A on $\mathbb{T}^2$, by the definition of standard curve, it is trivial to check that:

$$\|(A \circ \mathbb{G})'\| \le \|dA\|(1 + \varepsilon c_1) \tag{3.13a}$$

$$\|(A \circ \mathbb{G})''\| \le \varepsilon\|dA\|D_1 c_1 + \|dA\|_{\mathcal{C}^1}(1 + \varepsilon c_1)^2 \tag{3.13b}$$

$$\|(A \circ \mathbb{G})'''\| \le \varepsilon\|dA\|D'_1 c_1 + \|dA\|_{\mathcal{C}^2}(1 + \varepsilon(1 + D_1)c_1)^3. \tag{3.13c}$$

Let us then introduce the maps $f_{\mathbb{G}} = f \circ \mathbb{G}$, $\omega_{\mathbb{G}} = \omega \circ \mathbb{G}$ and $\Omega_{\mathbb{G}} = \Omega \circ \mathbb{G}$. We will assume ε to be small enough (depending on our choice of c_1) so that $f'_{\mathbb{G}} \ge \lambda - \varepsilon c_1\|\partial_\theta f\| > 3/2$; in particular, $f_{\mathbb{G}}$ is an expanding map. Provided δ has been chosen small enough, $f_{\mathbb{G}}$ is invertible. Let $\varphi(x) = f_{\mathbb{G}}^{-1}(x)$. Differentiating we obtain

$$\varphi' = \frac{1}{f'_{\mathbb{G}}} \circ \varphi \qquad \varphi'' = -\frac{f''_{\mathbb{G}}}{f'^3_{\mathbb{G}}} \circ \varphi \qquad \varphi''' = \frac{3f''^2_{\mathbb{G}} - f'''_{\mathbb{G}} f'_{\mathbb{G}}}{f'^5_{\mathbb{G}}} \circ \varphi. \tag{3.14}$$

Then, by definition, for any measurable function g:

$$\begin{aligned} F_{\varepsilon*,\Omega}\mu_\ell(g) &= \mu_\ell(e^{\Omega} \cdot g \circ F_\varepsilon) \\ &= \int_a^b g(f_{\mathbb{G}}(x), \bar{G}(x)) \cdot e^{\Omega_{\mathbb{G}}(x)} \rho(x) dx, \end{aligned}$$

where $\bar{G}(x) = G(x) + \varepsilon\omega_{\mathbb{G}}(x)$. Then, fix a partition (mod 0) $[f_{\mathbb{G}}(a), f_{\mathbb{G}}(b)] = \bigcup_{j \in \mathcal{J}}[a_j, b_j]$, with $b_j - a_j \in [\delta_{\mathbb{C}}/2, \delta_{\mathbb{C}}]$ and $b_j = a_{j+1}$. We can thus write:

$$F_{\varepsilon*,\Omega}\mu_\ell(g) = \sum_j \int_{a_j}^{b_j} g(x, G_j(x))\tilde{\rho}_j(x) dx, \tag{3.15}$$

provided that $G_j = \bar{G} \circ \varphi_j$ and $\tilde{\rho}_j(x) = e^{\Omega_{\mathbb{G}} \circ \varphi_j} \cdot \rho \circ \varphi_j \cdot \varphi'_j$, where $\varphi_j = \varphi|_{[a_j, b_j]}$.

In order to conclude our proof it suffices to show that (i) there exists c_1 large enough so that if $G \in \Sigma_{c_1}$, then $G_j \in \Sigma_{c_1}$, and (ii) there exist c_2, c_3 large enough and δ small enough so that if $\rho \in D^{\mathbb{C}}_{c_2,c_3}(G)$, $\tilde\rho_j$ can be normalized to a complex standard density belonging to $D^{\mathbb{C}}_{c_2,c_3}(G_j)$.

Item (i) follows from routine computations: differentiating the above definitions and using (3.14) we obtain

$$G_j' = \frac{\bar G'}{f_{\mathbb{G}}'}\circ\varphi_j \tag{3.16a}$$

$$G_j'' = \frac{\bar G''}{f_{\mathbb{G}}'^2}\circ\varphi_j - G_j'\cdot\frac{f_{\mathbb{G}}''}{f_{\mathbb{G}}'^2}\circ\varphi_j \tag{3.16b}$$

$$G_j''' = \frac{\bar G'''}{f_{\mathbb{G}}'^3}\circ\varphi_j - 3G_j''\cdot\frac{f_{\mathbb{G}}''}{f_{\mathbb{G}}'^2}\circ\varphi_j - G_j'\cdot\frac{f_{\mathbb{G}}'''}{f_{\mathbb{G}}'^3}\circ\varphi_j \tag{3.16c}$$

Using (3.16a), the definition of $\bar G$ and (3.13a) we obtain, for small enough ε:

$$\|G_j'\| \le \left\|\frac{G' + \varepsilon\omega_{\mathbb{G}}'}{f_{\mathbb{G}}'}\right\| \le \frac{2}{3}(1+\varepsilon\|d\omega\|)\varepsilon c_1 + \frac{2}{3}\varepsilon\|d\omega\| \le \frac{3}{4}\varepsilon c_1 + \varepsilon D_1$$

where $D_1 = \frac{2}{3}\|d\omega\|$. We can then fix c_1 large enough so that the right hand side of the above inequality is less than c_1. Next we will use C_* for a generic constant depending on c_1, D_1, D_1' and $C_\#$ for a generic constant depending only on F_ε. Then, we find[17]

$$\|G_j''\| \le \frac{3}{4}\varepsilon[c_1 D_1 + C_\#] + \varepsilon^2 C_*\,;$$
$$\|G_j'''\| \le \frac{3}{4}\varepsilon\left[c_1(D_1' + D_1C_\# + C_\#) + C_\#\right] + \varepsilon^2 C_*.$$

We can then fix c_1, D_1' sufficiently large and then ε sufficiently small to ensure that the $\mathbb{G}_j$'s are c_1-standard pairs. We now proceed with item (ii); by differentiating the definition of $\tilde\rho_j$ we obtain

$$\frac{\tilde\rho_j'}{\tilde\rho_j} = \frac{\rho'}{\rho\cdot f_{\mathbb{G}}'}\circ\varphi_j - \frac{f_{\mathbb{G}}''}{f_{\mathbb{G}}'^2}\circ\varphi_j + \frac{\Omega_{\mathbb{G}}'}{f_{\mathbb{G}}'}\circ\varphi_j \tag{3.17a}$$

$$\begin{aligned}\frac{\tilde\rho_j''}{\tilde\rho_j} = {}& \frac{\rho''}{\rho\cdot f_{\mathbb{G}}'^2}\circ\varphi_j - 3\frac{\tilde\rho_j'}{\tilde\rho_j}\cdot\frac{f_{\mathbb{G}}''}{f_{\mathbb{G}}'^2}\circ\varphi_j - \frac{f_{\mathbb{G}}'''}{f_{\mathbb{G}}'^3}\circ\varphi_j +\\ &+ 2\frac{\rho'\Omega_{\mathbb{G}}'}{\rho f_{\mathbb{G}}'^2}\circ\varphi_j + \frac{\Omega_{\mathbb{G}}'^2}{f_{\mathbb{G}}'^2}\circ\varphi_j + \frac{\Omega_{\mathbb{G}}''^2}{f_{\mathbb{G}}'^2}\circ\varphi_j.\end{aligned} \tag{3.17b}$$

From the first of the above expressions and (3.13a) we gather:

$$\left\|\frac{\tilde\rho_j'}{\tilde\rho_j}\right\|_{\mathcal{C}^0} \le \frac{2}{3}\left\|\frac{\rho'}{\rho}\right\|_{\mathcal{C}^0} + D + C_\#\|\Omega\|_{\mathcal{C}^1},$$

where D is a uniform constant related to the *distortion* of the maps $f(\cdot,\theta)$, which can be obtained using our uniform bounds on $\|G_j'\|$ and $\|G_j''\|$. The above expression implies that we can choose $c_2 = \mathcal{O}(1+\|\Omega\|_{\mathcal{C}^1})$ so that if $\|\rho'/\rho\|_{\mathcal{C}^0} \le c_2$, then $\|\tilde\rho_j'/\tilde\rho_j\|_{\mathcal{C}^0} \le c_2$. A similar computation, using (3.17b), yields:

$$\left\|\frac{\tilde\rho_j''}{\tilde\rho_j}\right\| \le \frac{4}{9}\left\|\frac{\rho''}{\rho}\right\| + C_\#(\|\Omega\|_{\mathcal{C}^2} + \|\Omega\|_{\mathcal{C}^1}^2 + c_2(\|\Omega\|_{\mathcal{C}^1} + D) + D'),$$

where, once again, D' is uniformly bounded thanks to our bounds on $\|G_j'\|$, $\|G_j''\|$ and $\|G_j'''\|$. As before, this implies the existence of $c_3 = \mathcal{O}(1+\|\Omega\|_{\mathcal{C}^2}+\|\Omega\|_{\mathcal{C}^1}^2)$ so that if $\|\rho''/\rho\|_{\mathcal{C}^0} \le c_3$, then $\|\tilde\rho_j''/\tilde\rho_j\|_{\mathcal{C}^0} \le c_3$.

[17] The reader can easily fill in the details of the computations.

We are now left to show that, using our requirement on $\delta_{\mathbb{C}}$:

$$\nu_j := \int_{a_j}^{b_j} \tilde\rho_j dx \neq 0; \tag{3.18}$$

this implies that $\rho_j := \nu_j^{-1}\tilde\rho_j \in D^{\mathbb{C}}_{c_2,c_3}(G_j)$, which concludes our proof: in fact, define the standard family $\mathfrak{L}'_\Omega$ given by $(\{\ell_j\},\nu)$, where $\nu_{\ell_j} = \nu_j$; then we can rewrite (3.15) as follows:

$$F_{\varepsilon*,\Omega}\mu_\ell(g) = \sum_{\tilde\ell\in\mathfrak{L}'_\Omega} \nu_{\tilde\ell}\mu_{\tilde\ell}(g) = \mu_{\mathfrak{L}'_\Omega}(g).$$

The proof of (3.18) follows from arguments similar to the ones used in the proof of Lemma 3.1: in fact $\delta_{\mathbb{C}}$ is sufficiently small so that the function $\log\tilde\rho_j$ can be defined and it is contained in a square of side $c_2\delta_{\mathbb{C}}$. Therefore, $\operatorname{Range}\tilde\rho_j$ is contained in an annular sector of small aperture, whose convex hull is bounded away from 0; this implies that $\nu_j \neq 0$. □

Remark 3.4. *Assume $\mathfrak{L}$ to be a standard probability family and $\Omega \in \mathcal{C}^2(\mathbb{T}^2,\mathbb{R})$: then $\mathfrak{L}'_\Omega$ is also a real standard family. Moreover, $\mathfrak{L}'_0$ is a standard probability family.*

Remark 3.5. *A quick inspection to the proof of Proposition 3.3 shows that we can choose the standard family $\mathfrak{L}'_\Omega$ to be of length $\frac{3}{2}\delta_{\mathbb{C}}$, provided $\frac{3}{2}\delta_{\mathbb{C}} \le \min\{\delta,\pi/(10\,c_2)\}$.*

Remark 3.6. *Note that if $\ell = (\mathbb{G},\rho)$ is a complex standard pair, then $\ell_{\mathbb{R}} = (\mathbb{G},|\rho|)$ is also a complex standard pair, and indeed is a regular standard pair if $\delta_{\mathbb{C}} = \delta$. Moreover form the arguments in the proof of Proposition 3.3 it follows that, for δ small enough, if $\mathfrak{L}_\Omega = \{\ell\}$, calling $\mathfrak{L}'_\Omega = \{\{\ell'\}_{\ell'\in\mathfrak{L}'_\Omega},\nu_{\ell'}\}$ and $\mathfrak{L}'_{\mathrm{Re}(\Omega)} = \{\{\ell'_{\mathbb{R}}\}_{\ell'\in\mathfrak{L}'_\Omega},\nu_{\mathbb{R},\ell'}\}$ the family obtained applying Proposition 3.3 to $\{\ell_{\mathbb{R}}\}$ we have $|\nu_{\ell'}| \ge c_\#\nu_{\mathbb{R},\ell'}$.*

We say that ℓ is a $\mathfrak{P}$*-standard pair* if c_1, c_2, c_3 and $\delta_{\mathbb{C}}$ are so that Proposition 3.3 holds with respect to the family $\mathfrak{P}$. Given a $\mathfrak{P}$-standard pair ℓ and a sequence of potentials $(\Omega_k)_{k\in\mathbb{N}} = \boldsymbol{\Omega} \in \mathfrak{P}^{\mathbb{N}}$, we denote (again with an abuse of notation) by $\mathfrak{L}^{(n)}_{\ell,\boldsymbol{\Omega}}$ a standard decomposition of $F^{(n)}_{\varepsilon*,\boldsymbol{\Omega}}\mu_\ell = F_{\varepsilon*,\Omega_{n-1}}\cdots F_{\varepsilon*,\Omega_0}\mu_\ell$, which we obtain by iterating the above proposition. By definition, therefore, we have, for any sufficiently smooth function g of $\mathbb{T}^2$:

$$F^{(n)}_{\varepsilon*,\boldsymbol{\Omega}}\,\mu_\ell(g) = \sum_{\tilde\ell\in\mathfrak{L}^{(n)}_{\ell,\boldsymbol{\Omega}}} \nu_{\tilde\ell}\mu_{\tilde\ell}(g) = \mu_\ell\left(e^{S_n\boldsymbol{\Omega}} g\circ F^n_\varepsilon\right), \tag{3.19}$$

where we have defined the "Birkhoff sum" $S_n\boldsymbol{\Omega} = \sum_{k=0}^{n-1}\Omega_k\circ F^k_\varepsilon$. In particular, (3.19) implies that $\mu_\ell\left(e^{S_n\boldsymbol{\Omega}}\right) = \sum_{\tilde\ell\in\mathfrak{L}^{(n)}_{\ell,\boldsymbol{\Omega}}}\nu_{\tilde\ell}$.

Remark 3.7. *The proof of Proposition 3.3 allows to define, for any $\tilde\ell \in \mathfrak{L}^{(n)}_{\ell,\boldsymbol{\Omega}}$ the corresponding characteristic function $\mathbb{1}_{\tilde\ell}$, that is a random variable on ℓ which equals 1 on points which are mapped to $\tilde\ell$ by F^n_ε and 0 elsewhere. This allows to write:*

$$\nu_{\tilde\ell} = \mu_\ell\left(e^{S_n\boldsymbol{\Omega}}\mathbb{1}_{\tilde\ell}\right) \tag{3.20a}$$

$$\mu_{\tilde\ell}(g) = \nu_{\tilde\ell}^{-1}\mu_\ell\left(e^{S_n\boldsymbol{\Omega}}\mathbb{1}_{\tilde\ell}\cdot g\circ F^n_\varepsilon\right), \tag{3.20b}$$

Observe that (3.20a) and (3.10) immediately implies that, for any $n\in\mathbb{N}$:

$$\sum_{\tilde\ell\in\mathfrak{L}^{(n)}_{\ell,\boldsymbol{\Omega}}} |\nu_{\tilde\ell}| \le \exp\left[\sum_{k=0}^{n-1}\max\operatorname{Re}\Omega_k + 2c_2\delta_{\mathbb{C}}\right]. \tag{3.21}$$

Moreover,

$$\sum_{\ell_1\in\mathfrak{L}^{(n)}_{\ell,\boldsymbol{\Omega}}}\sum_{\ell_2\in\mathfrak{L}^{(n)}_{\ell_1,\mathbf{s}^n\boldsymbol{\Omega}}}\cdots\sum_{\ell_m\in\mathfrak{L}^{(n)}_{\ell_{m-1},\mathbf{s}^{n(m-1)}\boldsymbol{\Omega}}}\prod_{j=1}^{m}|\nu_{\ell_j}|\leq e^{\sum_{k=0}^{nm-1}\max\operatorname{Re}\Omega_k}e^{2c_2\delta_{\mathbb{C}}}, \tag{3.22}$$

where $\mathbf{s}$ is the one-sided shift acting naturally on $\mathfrak{P}^{\mathbb{N}}$. In fact, the above is just a special choice of standard decomposition for $F^{(nm)}_{\varepsilon*,\boldsymbol{\Omega}}\mu_\ell$, indexed by a m-tuple of standard pairs $(\ell_1,\cdots,\ell_{m-1})$ selected at intermediate steps of length n.

Remark 3.8. *Given a standard pair $\ell=(\mathbb{G},\rho)$, we will interpret (x_k,θ_k) as random variables defined as $(x_k,\theta_k)=F^k_\varepsilon(x,G(x))$, where x is distributed according to ρ. We would like do the same for complex standard pairs. Of course, in this case (x_k,θ_k) will be random variables under $\operatorname{Re}(\rho)$ only, so we will simply say that they are functions distributed according to ρ, or, for brevity, functions on ℓ.*

Finally, let us define the set "good probability measures" mentioned in Section 2. Fix $C_\star>0$ large enough; given $\theta_0^*\in\mathbb{T}$, we define

$$\mathcal{P}_\varepsilon(\theta_0^*)=\{\mu_{\mathfrak{L}}\;:\;\mu_{\mathfrak{L}}(G_\ell)=\theta_0^*,\ \sup_{\ell\in\mathfrak{L}}|G_\ell-\theta_0^*|\leq C_\star\varepsilon,\ \mathfrak{L}\in\text{standard families}\}, \tag{3.23}$$

where $\mu_{\mathfrak{L}}(G_\ell):=\sum_{\ell\in\mathfrak{L}}\nu_\ell\int_{\mathbb{T}}G_\ell(x)\rho_\ell(x)dx$.

4. Averaging

This section is devoted to the proof of Theorem 2.1. The aim of this section is mostly notational and didactic; therefore, we keep things as simple as possible we provide the proof only for the variable θ since the argument for z is exactly the same.

In the following, given a standard pair $\ell=(\mathbb{G},\rho)$, we will use the notation

$$\theta_\ell^*=\mu_\ell(\theta_0)\quad\text{and}\quad\bar\theta^*_{\ell,k}=\bar\theta(\varepsilon k,\theta^*_\ell) \tag{4.1}$$

where, according to Remark 3.8, we consider $\theta_0=G$ to be a random variable on the standard pair ℓ and we denote with $\bar\theta(t,\theta^*)$ the unique solution of (2.3) for initial condition $\bar\theta(0)=\theta^*$.

Remark 4.1. *We find it convenient to prove the theorem for slightly more general initial condition: standard pairs ℓ_ε s.t. $\mu_{\ell_\varepsilon}\in\mathcal{P}_\varepsilon(\theta_0)$. In the following we will drop the subscript ε in the standard pair since this does not create confusion.*

4.1. Deterministic approximation.

First, we provide a preliminary useful approximation result, which allows to compare the true dynamics with a fixed one for times of order $\varepsilon^{-1/2}$: for fixed $\theta^*\in\mathbb{T}^1$, let us introduce[18] the map $F_*(x,\theta)=(f_*(x),\theta)$, where $f_*(x)=f(x,\theta^*)$.

Lemma 4.2. *Consider a standard pair $\ell=(\mathbb{G},\rho)$ and fix $\theta^*\in\mathbb{T}^1$ at a distance at most ϵ from the range of G.[19] For any $n\in\mathbb{N}$ so that $\epsilon n+n^2\varepsilon\leq C_\#$, there exists a diffeomorphism $Y_n:[a,b]\to[a^*,b^*]$ such that $(x_n,\theta_n)=F^n_\varepsilon(x,G(x))=(f^n_*(Y_n(x)),\theta_n)=F^n_*(Y_n(x),\theta_n)$. In addition, for all $k\in\{0,\cdots,n\}$ and setting*

[18] The reader should not confuse the notation F_*, which is a map of $\mathbb{T}^2$, with the push-forward $F_{\varepsilon*}$ introduced in the previous section.

[19] We will typically apply this Lemma to the case $\theta^*=\theta^*_\ell$.

$x_k^* = f_*^k(Y_n(x))$,

$$\left\| \theta_k - \theta^* - \varepsilon \sum_{j=0}^{k-1} \omega(x_j^*, \theta^*) \right\|_{\mathcal{C}^0} \le C_\#[\epsilon + \varepsilon^2 k^2],$$

$$\left\| x_k^* - x_k + \varepsilon \sum_{j=k}^{n-1} \Lambda_{k,j}^* \partial_\theta f(x_j^*, \theta^*) \sum_{l=0}^{j-1} \omega(x_l^*, \theta^*) \right\|_{\mathcal{C}^0} \le C_\#[\epsilon + \varepsilon^2 k^2],$$

$$\left\| 1 - Y_n' \prod_{k=0}^{n-1} \frac{f_*'(x_k^*)}{\partial_x f(x_k, \theta_k)} \right\|_{\mathcal{C}^0} \le C_\# n\varepsilon,$$

where we defined $\Lambda_{k,j}^* = \prod_{l=k}^j f_*'(x_l^*)^{-1} = (f_*^{j-k+1})'(x_k^*)^{-1} \le \lambda^{k-j-1}$ *and* $\| \cdot \|_{\mathcal{C}^0}$ *denotes the usual* sup*-norm of the random variables seen as functions of* $x \in [a, b]$.

Proof. Let us denote with $\pi_x : \mathbb{T}^2 \to \mathbb{T}$ the canonical projection on the x coordinate; then, for $x, z \in \mathbb{T}$ and $\varrho \in [0, 1]$, define

$$\mathcal{H}_n(x, z; \varrho) = \pi_x F_{\varrho\varepsilon}^n(x, \theta^* + \varrho(G(x) - \theta^*)) - f_*^n(z).$$

Note that, $\mathcal{H}_n(x, x; 0) = 0$, in addition, for any x, ϱ:

$$\partial_z \mathcal{H}_n(x, z; \varrho) = -(f_*^n)'(z) \neq 0.$$

Accordingly, by the implicit function theorem, for any $n \in \mathbb{N}$ and $\varrho \in [0, 1]$, there exists a diffeomorphism $Y_n(\cdot; \varrho)$ such that $\mathcal{H}_n(x, Y_n(x; \varrho); \varrho) = 0$; from now on $Y_n(x)$ stands for $Y_n(x; 1)$. Observe moreover that

$$Y_n' = \frac{(\pi_x F_\varepsilon^n \circ \mathbb{G})'}{(f_*^n)' \circ Y_n} = \frac{(1 - G' s_n) v_n^+}{(f_*^n)' \circ Y_n}, \tag{4.2}$$

where we have used the notations introduced in (3.6).

Next, we want to estimate to which degree $\{(x_k^*, \theta^*)\}_{k=0}^n$ shadows the true trajectory. Observe that

$$\theta_k = \varepsilon \sum_{j=0}^{k-1} \omega(x_j, \theta_j) + \theta_0$$

thus $|\theta_k - \theta^*| \le C_\# \varepsilon k + \epsilon$. Accordingly, let us set $\xi_k = x_k^* - x_k$; then by the mean value theorem we obtain, for some $x, \theta \in \mathbb{T}$:

$$|\xi_{k+1}| = |f_*'(x) \cdot \xi_k + \partial_\theta f(x_k, \theta) \cdot (\theta_k - \theta^*)| \ge \lambda |\xi_k| - C_\#(\theta_k - \theta^*)$$

which, by backward induction, using the fact that $\xi_n = 0$ and our previous estimates on $|\theta_k - \theta^*|$, yields $|\xi_k| \le C_\#(\epsilon + \varepsilon k)$. We thus obtain:

$$\theta_k - \theta^* = \theta_0 - \theta^* + \varepsilon \sum_{j=0}^{k-1} \omega(x_j^*, \theta^*) + \mathcal{O}(\varepsilon(\epsilon k + \varepsilon k^2))$$

$$\xi_k = - \sum_{j=k}^{n-1} \Lambda_{k,j}^* \partial_\theta f(x_j^*, \theta^*) \left(\varepsilon \sum_{l=0}^{j-1} \omega(x_l^*, \theta^*) + \mathcal{O}(\epsilon + \varepsilon^2 j^2) \right).$$

Finally, recalling (4.2), (3.8) and using invariance of the center cone, we have

$$e^{-c_\# \varepsilon n} \prod_{k=0}^{n-1} \frac{\partial_x f(x_k, \theta_k)}{f_*'(x_k^*)} \le \left| \frac{(1 - G' s_n) v_n^+}{(f_*^n)'} \right| \le e^{c_\# \varepsilon n} \prod_{k=0}^{n-1} \frac{\partial_x f(x_k, \theta_k)}{f_*'(x_k^*)}.$$

Accordingly, Y_n is invertible with uniformly bounded derivative, since we assume $n\epsilon + n^2 \varepsilon \le C_\#$.[20] □

[20] On the contrary, the reader can easily check that $\|Y_n''\|_\infty \sim \lambda^n$.

4.2. **Proof of the Averaging Theorem.**
Let us now ready to prove our first result.

Proof of Theorem 2.1. Let ℓ be a standard pair; recall that we defined $\hat\omega(x,\theta)=\omega(x,\theta)-\bar\omega(\theta)$; for any $t,h>0$ define $H=H(t,h)=\lfloor (t+h)\varepsilon^{-1}\rfloor-\lfloor t\varepsilon^{-1}\rfloor$; observe that $|H(t,h)-\lfloor h\varepsilon^{-1}\rfloor|\leq 1$. Let us start by computing

$$\mu_\ell\left(\left[\varepsilon\sum_{k=\lfloor t\varepsilon^{-1}\rfloor}^{\lfloor (t+h)\varepsilon^{-1}\rfloor-1}\hat\omega(x_k,\theta_k)\right]^2\right)=\sum_{\ell_1\in\mathfrak{L}_\ell^{(\lfloor t\varepsilon^{-1}\rfloor)}}\sum_{k=0}^{H-1}\varepsilon^2\nu_{\ell_1}\mu_{\ell_1}(\hat\omega^2\circ F_\varepsilon^k)+$$

$$+\sum_{\ell_1\in\mathfrak{L}_\ell^{(\lfloor t\varepsilon^{-1}\rfloor)}}2\sum_{j=0}^{H-1}\sum_{k=j+1}^{H-1}\sum_{\ell_2\in\mathfrak{L}_{\ell_1}^{(j)}}\varepsilon^2\nu_{\ell_1}\nu_{\ell_2}\mu_{\ell_2}(\hat\omega\circ F_\varepsilon^{k-j}\cdot\hat\omega),\tag{4.3}$$

where we repeatedly used Proposition 3.3 and the notation introduced before (3.19) without $\boldsymbol{\Omega}$, since in this case $\boldsymbol{\Omega}=0$.

Next, using Lemma 4.2 we introduce, for any standard pair $\tilde\ell=(\tilde{\mathbb{G}},\tilde\rho)$, the diffeomorphisms $Y=Y_H$ and let $[a^*,b^*]=Y([a,b])$. Let us call $\rho^*=\frac{\tilde\rho\circ Y^{-1}}{Y'\circ Y^{-1}}$ the push-forward of $\tilde\rho$ by Y, also let $\theta^*_{\tilde\ell}=\mu_{\tilde\ell}(\theta)$. For any functions $\varphi,g\in\mathcal{C}^1(\mathbb{T}^2)$ and $k\in\mathbb{N}$, Lemma 4.2 implies

$$\begin{aligned}\mu_{\tilde\ell}(g\circ F_\varepsilon^k\cdot\varphi)&=\int_{a^*}^{b^*}\rho^*(x)\varphi(Y^{-1}(x),\theta^*_{\tilde\ell})\cdot g(f^k_{\theta^*_{\tilde\ell}}(x),\theta^*_{\tilde\ell})dx+\mathcal{O}(k\varepsilon\|g\|_{\mathcal{C}^1}\|\varphi\|_{\mathcal{C}^1})\\&=\int_a^b\tilde\rho(x)\varphi(x,\theta^*_{\tilde\ell})\cdot g(f^k_{\theta^*_{\tilde\ell}}(x),\theta^*_{\tilde\ell})dx+\mathcal{O}(k^2\varepsilon\|g\|_{\mathcal{C}^1}\|\varphi\|_{\mathcal{C}^1}).\end{aligned}$$

To continue we introduce one of the main tools in the study of hyperbolic systems: the transfer operator (for now, without potential). Let

$$\mathcal{L}_\theta g(x)=\sum_{y\in f_\theta^{-1}(x)}\frac{g(y)}{f'_\theta(y)}.$$

The basic properties of these operators are well known (see e.g. [5]) but in the following we need several quite sophisticated facts that are either not easily found or absent altogether in the literature. To help the reader we have collected all the needed properties in Appendix A.[21] We can thus estimate the quantity in the second line of (4.3) as

$$\begin{aligned}\mu_{\ell_2}(\hat\omega\circ F_\varepsilon^l\cdot\hat\omega)&=\int[\mathcal{L}^l_{\theta^*_{\ell_2}}(\mathbb{1}_{[a,b]}\rho\hat\omega)](x,\theta^*_{\ell_2})\cdot\hat\omega(x,\theta^*_{\ell_2})dx+\mathcal{O}(l^2\varepsilon)=\\&=\int_{\mathbb{T}^1}h_{\theta^*_{\ell_2}}(x)\hat\omega(x,\theta^*_{\ell_2})dx\int_a^b\rho(x)\hat\omega(x,\theta^*_{\ell_2})dx+\mathcal{O}(l^2\varepsilon+\tau^l)=\\&=\mathcal{O}(l^2\varepsilon+\tau^l),\end{aligned}$$

where we used the fact that $\mu_\theta(\hat\omega(\cdot,\theta))=0$ by construction and $\tau\in(0,1)$ where $1-\tau$ is a lower bound on the spectral gap of $\mathcal{L}_\theta$ for any $\theta\in\mathbb{T}$

[21] For the time being we need only that $\int g\mathcal{L}_\theta\phi=\int g\circ f_\theta\phi$ and that, seen as an operator acting on BV, $\mathcal{L}_\theta$ has 1 as a maximal eigenvalue, a spectral gap, and h_θ (the eigenfunction associated to the eigenvalue 1) is the $\mathcal{C}^{r-1}$ density of the unique absolutely continuous invariant measure of f_θ. In other words $\mathcal{L}_\theta$ has the spectral decomposition $\mathcal{L}_\theta g=h_\theta\int g+Rg$, where the spectral radius of R is smaller that some $\tau\in(0,1)$.

Collecting all the above considerations we obtain

$$\mu_\ell\left(\left[\varepsilon\sum_{k=\lfloor t\varepsilon^{-1}\rfloor}^{\lfloor (t+h)\varepsilon^{-1}\rfloor-1}\hat\omega(x_k,\theta_k)\right]^2\right)\leq C_\#\varepsilon^2\sum_{k=0}^{H-1}\left[1+\sum_{j=1}^{H-k-1}\{\tau^j+j^2\varepsilon\}\right]$$
$$\leq C_\#[\varepsilon h+\varepsilon^{-1}h^4]\leq C_\#\varepsilon^{5/3}, \tag{4.4}$$

where at the very last step we have chosen $h=\varepsilon^{2/3}$, which optimizes the estimate. Recall now the definition of the random element $\theta_\varepsilon\in C^0([0,T],\mathbb{R})$, defined in (2.7). As previously observed, the functions θ_ε are uniformly Lipschitz of constant $\|A\|_{\mathcal{C}^0}$. Using the Cauchy–Schwarz inequality and (4.4):

$$\mu_\ell\left(\left|\theta_\varepsilon(t)-\theta_\varepsilon(0)-\int_0^t\bar\omega(\theta_\varepsilon(s))ds\right|^2\right)$$
$$\leq\lfloor th^{-1}\rfloor\sum_{r=0}^{\lfloor th^{-1}\rfloor-1}\mu_\ell\left(\left|\theta_\varepsilon((r+1)h)-\theta_\varepsilon(rh)-\int_{rh}^{(r+1)h}\bar\omega(\theta_\varepsilon(s))ds\right|^2\right)$$
$$\leq\lfloor th^{-1}\rfloor\sum_{r=0}^{\lfloor th^{-1}\rfloor-1}\mu_\ell\left(\left|\varepsilon\sum_{k=\lfloor rh\varepsilon^{-1}\rfloor}^{\lfloor (r+1)h\varepsilon^{-1}\rfloor-1}\hat\omega(x_k,\theta_k)+\mathcal{O}(\varepsilon h)\right|^2\right)$$
$$\leq C_\# t^2[\varepsilon h^{-1}+\varepsilon^{-1}h^2+\varepsilon]\leq C_\# t^2\varepsilon^{\frac13},$$

where at the very last step we have chosen $h=\varepsilon^{2/3}$. Chebyshev inequality then implies, for any $t\leq T$:

$$\mu_\ell\left(\left\{\left|\theta_\varepsilon(t)-\theta_\varepsilon(0)-\int_0^t\bar\omega(\theta_\varepsilon(s))ds\right|\geq C_\#\varepsilon^{1/8}\right\}\right)\leq C_\# T^2\varepsilon^{1/3}\varepsilon^{-1/4}. \tag{4.5}$$

Let us partition the interval $[0,T]$ in $N=\lfloor T\varepsilon^{-1/24}\rfloor$ intervals of endpoints

$$0=t_0<t_1<\cdots<t_N=T,$$

where for any $k\in\{0,\cdots,N-1\}$ we have $\varepsilon^{1/24}\leq t_{k+1}-t_k<2\varepsilon^{1/24}$. Since θ_ε is uniformly Lipschitz and using (4.5), we conclude:

$$\mu_\ell\left(\left\{\sup_{t\in[0,T]}\left|\theta_\varepsilon(t)-\theta_\varepsilon(0)-\int_0^t\bar\omega(\theta_\varepsilon(s))ds\right|\geq C_\#\varepsilon^{1/24}\right\}\right)$$
$$\leq\mu_\ell\left(\bigcup_{k=0}^{N-1}\left\{\left|\theta_\varepsilon(t_k)-\theta_\varepsilon(0)-\int_0^{t_k}\bar\omega(\theta_\varepsilon(s))ds\right|\geq C_\#\varepsilon^{1/24}\right\}\right) \tag{4.6}$$
$$\leq C_\# T^3\varepsilon^{1/24}.$$

Since θ_ε are a uniformly Lipschitz family of paths, they form a compact set by Ascoli–Arzelà Theorem. Consider then any converging subsequence θ_{ε_j}; choosing $\varepsilon=\varepsilon_j$ and taking the limit of (4.6) for $j\to\infty$ it follows that all accumulation points of θ_ε are solutions of the integral version of (2.3). Since such differential equation admits a unique solution, we conclude that the limit exists and it is given by the solution of (2.3).

If we consider now the initial conditions of the Lemma, which allow to consider all random variables on the same probability space, we immediately have the result for θ. The results for z is more of the same. □

Remark 4.3. *Note that it may be possible to obtain this result* almost surely *rather than in probability. We do not push this venue since it is irrelevant for our purposes.*[22]

Remark 4.4. *The bound (4.6) was obtained by estimating the second moment. This gave us a simple argument, but not sufficient for our later needs. To get sharper bounds we will need to estimate the exponential moment, which is tantamount to studying large deviations.*

5. Moment generating function

We now begin the study of deviations from the average behavior described in the previous section. We start with the problem of investigating large and moderate deviations. It is well known that such information can be obtained from precise estimates of the exponential moment generating function. Hence our next goal is the study of this object. In order to do so it turns out to be helpful to have an approximate description of the dynamics that is more refined than the one presented in Lemma 4.2. This is achieved in the next subsection.

5.1. **Random approximation.**

In order to obtain our main results Theorem 2.4 and Theorem 2.8, we will need to control deviations from the average with resolution up to order ε; this requires very fine bounds which we proceed to obtain in this subsection.

For later reference, we find convenient to state such estimates in a slightly more general form than needed for our immediate purposes; we introduce two different notions of *deviation from the average*: let ℓ be a standard pair; recall the notation $\bar\theta^*_{\ell,k} = \bar\theta(\varepsilon k, \theta^*_\ell)$ introduced in (4.1), where $\theta^*_\ell = \mu_\ell(G_\ell)$. Let us also define the functions $\bar\theta_k(\theta) = \bar\theta(\varepsilon k, \theta)$ (observe that $\bar\theta^*_{\ell,k} = \bar\theta_k(\theta^*_\ell)$). Then we define two corresponding notions of deviation:

$$\Delta^*_{\ell,k}(x,\theta) = \theta_k(x,\theta) - \bar\theta^*_{\ell,k} \tag{5.1a}$$

$$\Delta_k(x,\theta) = \theta_k(x,\theta) - \bar\theta_k(\theta). \tag{5.1b}$$

Since $|\theta_k - \theta_0| \le C_\#\varepsilon k$, $|\bar\theta^*_{\ell,k} - \theta^*_\ell| \le C_\#\varepsilon k$, $|\bar\theta_k - \theta| \le C_\#\varepsilon k$ and $|\theta_0 - \theta^*_\ell| \le C_\#\varepsilon$, we trivially find

$$|\Delta^*_{\ell,k}| \le C_\#\varepsilon(k+1) \qquad\qquad |\Delta_k| \le C_\#\varepsilon k. \tag{5.2}$$

Moreover, observe that

$$\begin{aligned}
\theta_{k+1} - \theta_k &= \varepsilon\bar\omega(\theta_k) + \varepsilon\hat\omega(x_k,\theta_k)\\
\bar\theta^*_{\ell,k+1} - \bar\theta^*_{\ell,k} &= \varepsilon\bar\omega(\bar\theta^*_{\ell,k}) + \frac12\varepsilon^2\bar\omega'(\bar\theta^*_{\ell,k})\bar\omega(\bar\theta^*_{\ell,k}) + \mathcal{O}(\varepsilon^3),\\
\bar\theta_{k+1} - \bar\theta_k &= \varepsilon\bar\omega(\bar\theta_k) + \frac12\varepsilon^2\bar\omega'(\bar\theta_k)\bar\omega(\bar\theta_k) + \mathcal{O}(\varepsilon^3),
\end{aligned}$$

where, recall $\bar\omega(\theta) = \mu_\theta(\omega(\cdot,\theta))$ and $\hat\omega(\cdot,\theta) = \omega(\cdot,\theta) - \bar\omega(\theta)$. The above equations yield the difference equations:

$$\begin{aligned}
\Delta^*_{\ell,k+1} - \Delta^*_{\ell,k} &= \varepsilon\hat\omega(x_k,\theta_k) + \varepsilon\bar\omega'(\bar\theta^*_{\ell,k})\Delta^*_{\ell,k} +\\
&+ \frac{\varepsilon}{2}\bar\omega''(\bar\theta^*_{\ell,k})(\Delta^*_{\ell,k})^2 - \frac{\varepsilon^2}{2}\bar\omega'(\bar\theta^*_{\ell,k})\bar\omega(\bar\theta^*_{\ell,k}) + \mathcal{O}(\varepsilon(\Delta^*_{\ell,k})^3 + \varepsilon^3).
\end{aligned} \tag{5.3a}$$

$$\begin{aligned}
\Delta_{k+1} - \Delta_k &= \varepsilon\hat\omega(x_k,\theta_k) + \varepsilon\bar\omega'(\bar\theta_k)\Delta_k +\\
&+ \frac{\varepsilon}{2}\bar\omega''(\bar\theta_k)(\Delta_k)^2 - \frac{\varepsilon^2}{2}\bar\omega'(\bar\theta_k)\bar\omega(\bar\theta_k k) + \mathcal{O}(\varepsilon(\Delta_k)^3 + \varepsilon^3).
\end{aligned} \tag{5.3b}$$

---

[22] But see [38] for a discussion of possible counterexamples.

Define now the auxiliary functions:

$$H^*_{\ell,k} = \sum_{j=0}^{k-1} \Xi^*_{\ell,j,k}\left[\hat\omega(x_j,\theta_j) - \frac{\varepsilon}{2}\bar\omega'(\bar\theta^*_{\ell,j})\bar\omega(\bar\theta^*_{\ell,j})\right], \tag{5.4a}$$

$$H_k = \sum_{j=0}^{k-1} \Xi_{j,k}\left[\hat\omega(x_j,\theta_j) - \frac{\varepsilon}{2}\bar\omega'(\bar\theta_j)\bar\omega(\bar\theta_j)\right], \tag{5.4b}$$

where

$$\Xi^*_{\ell,j,k} = \prod_{l=j+1}^{k-1}\left[1+\varepsilon\bar\omega'(\bar\theta^*_{\ell,l})\right] \qquad\qquad \Xi_{j,k} = \prod_{l=j+1}^{k-1}\left[1+\varepsilon\bar\omega'(\bar\theta_l)\right]. \tag{5.5}$$

We are now finally ready to state and prove the needed approximation. The following lemma is a refinement of Lemma 4.2.

Lemma 5.1. *For any $T>0$, $0\le k\le T\varepsilon^{-1}$ and standard pair ℓ, we have*

$$\begin{aligned}\Delta^*_{\ell,k} - \varepsilon H^*_{\ell,k} &= \varepsilon\sum_{j=0}^{k-1}\Xi^*_{\ell,j,k}\left[\frac{\bar\omega''(\bar\theta^*_{\ell,j})}{2}(\Delta^*_{\ell,j})^2 + \mathcal{O}((\Delta^*_{\ell,j})^3+\varepsilon^2)\right]\\ &\quad + \Xi^*_{\ell,-1,k}\Delta^*_{\ell,0}\end{aligned} \tag{5.6a}$$

$$\Delta_k - \varepsilon H_k = \varepsilon\sum_{j=0}^{k-1}\Xi_{j,k}\left[\frac{\bar\omega''(\bar\theta_j)}{2}(\Delta_j)^2 + \mathcal{O}((\Delta_j)^3+\varepsilon^2)\right]. \tag{5.6b}$$

Proof. Observe that (5.4a) implies $|H^*_{\ell,k}|\le C_\# k$. Also, it is immediate to check that, by definition, $H^*_{\ell,k}$ satisfies the following recurrence equation:

$$H^*_{\ell,k+1} = \hat\omega(x_k,\theta_k) + (1+\varepsilon\bar\omega'(\bar\theta^*_{\ell,k}))H^*_{\ell,k} - \frac{\varepsilon}{2}\bar\omega'(\bar\theta^*_{\ell,k})\bar\omega(\bar\theta^*_{\ell,k}).$$

Hence, by (5.3a),

$$\begin{aligned}\Delta^*_{\ell,k+1} - \varepsilon H^*_{\ell,k+1} &= (1+\varepsilon\bar\omega'(\bar\theta^*_{\ell,k}))\left[\Delta^*_{\ell,k} - \varepsilon H^*_{\ell,k}\right] + \frac{\varepsilon}{2}\bar\omega''(\bar\theta^*_{\ell,k})(\Delta^*_{\ell,k})^2\\ &\quad + \mathcal{O}(\varepsilon(\Delta^*_{\ell,k})^3+\varepsilon^3).\end{aligned}$$

The first statement of the lemma then follows by induction, since $H^*_{\ell,0}=0$. The second statement follows by identical computations and the observation that, by definition, we have $\Delta_0 = 0$. □

5.2. Computation of the exponential moment.

We can now proceed to the main result of this section, which is the precise computation of the exponential moment. The goal is to compute it with an error much smaller than currently available in the literature. This will allow to obtain precise information not only on large, but also on moderate deviations, as will be shown in the next two sections.

In this section, given a (c_1,c_2,c_3)-standard pair $\ell=(\mathbb{G},\rho)$, we will call it simply a c_2-standard pair, since c_1 will be always fixed (as in the rest of the paper) and c_3 is irrelevant for the estimates in this section. Recall that we fixed $A=(A_1,\cdots,A_d)\in\mathcal{C}^3(\mathbb{T}^2,\mathbb{R}^d)$, with $A_1=\omega$; we also introduced the notation $\bar A(\theta)=\mu_\theta(A(\cdot,\theta))$ and $\hat A = A-\bar A$. Recall that $\theta^*_\ell = \mu_\ell(G) = \int_a^b\rho(x)G(x)dx$ (hence it belongs to the range of the standard pair ℓ). Moreover, recall that we are under the standing assumption (A1').

Remark 5.2. *In this section we will use the notation $\mathrm{BV}([0,T],\mathbb{R}^d)$ to denote the space of functions in $\mathbb{R}^d$ whose components are bounded variation functions. Recall* *that, given a* L^1 *function* $\varphi : I \to \mathbb{R}^d$*, its* BV*-norm is defined as:*

$$\|\varphi\|_{\mathrm{BV}(I)} = \|\varphi\|_{L^1(I)} + V_I(\varphi),$$

where $V_I(\varphi)$ *is the* total variation *of* φ *on the interval* I*, given by:*

$$V_I(\varphi) = \sup_{\substack{\psi \in \mathcal{C}^1_c(I,\mathbb{R}^d) \\ \|\psi\|_\infty = 1}} \int_I \langle \varphi(x), \psi'(x) \rangle dx,$$

where $\mathcal{C}^1_c(I)$ *is the space of* $\mathcal{C}^1$ *functions that are* 0 *in a neighborhood of the boundary of* I*. Moreover in this section, given* $I \subset \mathbb{R}$*, we will denote* $\|f\|_{L^\infty} = \sup_{x \in I} |f(x)|$*. As usual, if the set* I *is not specified, it is understood to be the domain of the function.*

Remark 5.3. *Before giving the main result of this section (an estimate for the* logarithmic moment generating functional*), as an attempt to illustrate its statement, let us consider the following simple example. Let us fix* $\varepsilon > 0$*; consider a (non-stationary) Markov chain on the state space* $\mathscr{S}$ *described at time* n *by the transition matrix* P_n*; assume that (in an appropriate sense)* P_{n+1} *is* ε*-close to* P_n*. Let us fix an arbitrary observable* $A \in \mathbb{R}^{\mathscr{S}}$ *(which we identify with a column vector* $A(x) = A^x$*) and* $x \in \mathscr{S}$ *we can define the* logarithmic moment generating functional of A associated to the Markov chain with initial state x *(denoted by* Λ*) as follows: for any function* $\sigma \in \mathrm{BV}([0,T],\mathbb{R})$

$$\Lambda_x(\sigma) = \varepsilon \log \mathbb{E}_x \left[\exp \varepsilon^{-1} \int_0^T \sigma(s) A(X_{\lfloor \varepsilon^{-1} s \rfloor}) ds \right].$$

where X_n *is a realization of the Markov chain with initial state* $X_0 = x$ *and* $\mathbb{E}_x$ *denotes the expectation conditioned to having initial state* x*. If* σ *were a constant and* $P_n = P$ *for all* n*, then it would be possible to express this expectation as follows: let* $P_{\sigma A}$ *be the transition matrix* twisted with potential σA*, that is* $[P_{\sigma A}]^{xy} = [P]^{xy} \exp(\sigma A^y)$*. Then*

$$\mathbb{E}_x \left[\exp \left(\sum_{n=0}^{\lfloor T\varepsilon^{-1} \rfloor - 1} \sigma A(X_n) \right) \right] = \sum_{y \in \mathscr{S}} \left[P_{\sigma A}^{\lfloor T\varepsilon^{-1} \rfloor} \right]^{xy}.$$

The leading contribution to the logarithmic moment generating functional is thus given by the spectral radius of the matrix $P_{\sigma A}$*, that is, its leading eigenvalue* $e^{\chi_A(\sigma)}$*. We then obtain*

$$\Lambda_x(\sigma) = \lfloor T\varepsilon^{-1} \rfloor \chi_A(\sigma) + \mathcal{R}_x$$

where $\mathcal{R}$ *is a remainder term that hopefully can be neglected. If, on the other hand,* σ *and* P *are not constant, then, heuristically, we can choose* $\varepsilon \ll h \ll T$ *and assume* σ *and* P *to be constant in each block of length* h *in* $[0,T]$*. Arguing in this way we can expect*

$$\Lambda_x(\sigma) = \int_0^T \chi_A(\sigma(s), \lfloor s\varepsilon^{-1} \rfloor) ds + \mathcal{R}_x$$

where $e^{\chi_A(\sigma,n)}$ *is the leading eigenvalue of the transition matrix* P_n *twisted with the potential* σA *and* $\mathcal{R}_x$ *is a remainder term which remains to be estimated.*

The main result of this section is the proof of a formula similar to the above, for our deterministic system. The Markov chain will be replaced by the fast dynamics, which changes in time according to the evolution of the slow variable.

The first object that we need to define is the class of transfer operators associated to the function $A \in \mathcal{C}^2(\mathbb{T}^2, \mathbb{R}^d)$ and a parameter $\sigma \in \mathbb{R}^d$. These will play the role of the $P_{\sigma A}$ in the example of Remark 5.3. For any $\mathcal{C}^1$ (or BV) function g, define

$$[\mathcal{L}_{\theta,\langle\sigma,A\rangle}\, g](x) := \sum_{f_\theta(y)=x} \frac{e^{\langle\sigma,A(y,\theta)\rangle}}{f'_\theta(y)} g(y) = e^{\langle\sigma,\bar{A}(\theta)\rangle}[\mathcal{L}_{\theta,\langle\sigma,\hat{A}\rangle}\, g](x), \tag{5.7}$$

where, recall, we have defined $\bar{A}(\theta) = \mu_\theta(A(\cdot,\theta))$, $\hat{A} = A - \bar{A}$ and μ_θ denotes the unique absolutely continuous invariant probability of f_θ. The above operators are of Perron–Frobenius type when acting on $\mathcal{C}^1$ (see Lemma A.1), and the same is true for sufficiently small σ when acting on BV (see Remark A.11). In other words, they have a simple maximal eigenvalue and a spectral gap. Let $e^{\chi_A(\sigma,\theta)}$ and $e^{\hat{\chi}_A(\sigma,\theta)}$ be their maximal eigenvalues, respectively. By (5.7) it follows

$$\chi_A(\sigma,\theta) = \langle\sigma,\bar{A}(\theta)\rangle + \hat{\chi}_A(\sigma,\theta). \tag{5.8}$$

Moreover, it is well known (see e.g. [5, Remark 2.5]) that

$$\chi_A(\sigma,\theta) = P_{\text{top}}(f_\theta, \langle\sigma,A\rangle - \log f'_\theta),$$

where P_{top} denotes the topological pressure. Also the results of Appendices A.2 and A.3 imply $\chi_A \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{T}, \mathbb{R})$.

Given $\sigma \in \text{BV}([0,T],\mathbb{R}^d)$, and $n \in \mathbb{N}$, we introduce the notation

$$\sigma_n = \varepsilon^{-1}\int_{\varepsilon n}^{\varepsilon(n+1)} \sigma(s)ds; \tag{5.9}$$

observe that for any $s \in [\varepsilon n, \varepsilon(n+1)]$ we have $|\sigma(s) - \sigma_n| \le \|\sigma\|_{\text{BV}([\varepsilon n,\varepsilon(n+1)])}$.

For any standard pair ℓ, $\varepsilon > 0$, $T > 0$ and $\sigma \in \text{BV}([0,T],\mathbb{R}^d)$, we now proceed to obtain some information on the logarithmic moment generating functional

$$\Lambda_{\ell,\varepsilon}(\sigma) = \varepsilon \log \mu_\ell \left[\exp\left(\sum_{n=0}^{\lfloor T\varepsilon^{-1}\rfloor - 1} \langle\sigma_n, A\circ F_\varepsilon^n\rangle\right)\right]. \tag{5.10}$$

Remark 5.3 suggests that $\Lambda_{\ell,\varepsilon}(\sigma) \sim \int_0^T \chi_A(\sigma(s),\bar{\theta}(s,\theta^*_\ell))ds$; it is therefore natural to define the quantity:

$$\mathcal{R}_{\ell,\varepsilon}(\sigma) = \Lambda_{\ell,\varepsilon}(\sigma) - \int_0^T \chi_A(\sigma(s),\bar{\theta}(s,\theta^*_\ell))ds. \tag{5.11}$$

The main result of this section is a bound on the remainder term defined above.

Proposition 5.4. *There exists $\varepsilon_0 > 0$, such that, for any $\varepsilon \in (0,\varepsilon_0)$, $L \in \left[\frac{1}{\varepsilon_0}, \frac{\varepsilon_0}{\sqrt{\varepsilon}}\right]$ and $T \in [\varepsilon L, T_{\max}]$,*

(a) *for any $\sigma \in \text{BV}([0,T],\mathbb{R}^d)$ we have*

$$\begin{aligned}|\mathcal{R}_{\ell,\varepsilon}(\sigma)| \le C_\#\big(&\varepsilon L\|\sigma\|_{\text{BV}} + \varepsilon LT \\ &+ \left[L^{-1} + \min\{T, \|\sigma\|_{L^1} + \varepsilon L\|\sigma\|_{\text{BV}}\}\right]\|\sigma\|_{L^1}\big);\end{aligned}$$

(b) *there exists $\sigma_* = \sigma_*(T_{\max}) > 0$ so that, if $\|\sigma\|_{L^\infty} < \sigma_*$, then*

$$|\mathcal{R}_{\ell,\varepsilon}(\sigma)| \le C_\#\left(\varepsilon\|\sigma\|_{\text{BV}} + \varepsilon LT + \varepsilon(L+T^{-1})\|\sigma\|_{L^1} + \|\sigma\|^2_{L^1} + L^{-1}\|\sigma\|^2_{L^2}\right).$$

The proof of the Proposition 5.4 relies on the spectral properties of the transfer operators (5.7). It is then natural that our ability to bound the size of the remainder term $\mathcal{R}$ depends on the size of σ. Without any assumption on σ we cannot use perturbation theory of the associated transfer operators. This allows only a rough bound, which is stated in item (a). On the other hand, if $\|\sigma\|_{L^\infty}$ is sufficiently small, then the corresponding transfer operators are guaranteed to be of uniform

Perron–Frobenius type and can be treated using perturbation theory. This enables us to give the much sharper bounds stated in (b).

As hinted in Remark 5.3, the main (quite standard) idea of the proof is to introduce a block decomposition: consider a partition of the set $\{0, \cdots, \lfloor T\varepsilon^{-1}\rfloor\}$ in K blocks of length L, where $K = \lfloor T\varepsilon^{-1}\rfloor L^{-1}$ (in Remark 5.3 we have $h \sim L\varepsilon$).[23]

By (3.19), we have that, for any $g \in L^\infty(\mathbb{T}^2, \mathbb{R})$,[24]

$$\mu_\ell\left(e^{\sum_{n=0}^{\lfloor T\varepsilon^{-1}\rfloor-1}\langle \sigma_n, A\circ F_\varepsilon^n\rangle} g\circ F_\varepsilon^{\lfloor T\varepsilon^{-1}\rfloor}\right) = \sum_{\ell_1\in\mathfrak{L}_\ell^L}\cdots\sum_{\ell_K\in\mathfrak{L}_{\ell_{K-1}}^L}\prod_{i=1}^K \nu_{\ell_i}\mu_{\ell_K}(g), \tag{5.12}$$

where, to ease the notation, we dropped the subscript potentials $\langle \sigma_k, \hat{A}\circ F_\varepsilon^k\rangle$ from the symbols for standard families. To further shorten notation, given a standard pair ℓ, we use ρ_ℓ, G_ℓ, a_ℓ and b_ℓ to denote the corresponding data.

Recall from Section 3.2 that ρ_ℓ is a $\mathcal{C}^2$ probability density over $[a_\ell, b_\ell]$; yet for our future purposes it is more convenient to deal with functions that are defined on the whole $\mathbb{T}^1$; to this end we introduce the extension $\mathring{\rho}_\ell$ of ρ_ℓ to $\mathbb{T}^1$ which we indicate by the (slightly abusing) notation $\mathring{\rho}_\ell = \mathbb{1}_{[a_\ell,b_\ell]}\rho_\ell$.

Remark 5.5. *Observe that if ρ_ℓ is a c_*-standard density, then $\mathring{\rho}_\ell$ is a* BV *function and its* BV *norm is bounded by:*

$$\begin{aligned}\|\mathring{\rho}_\ell\|_{\mathrm{BV}} &\le \|\mathring{\rho}_\ell\|_{L^1} + \sup_{\substack{\psi\in\mathcal{C}^1(\mathbb{T},\mathbb{R})\\ \|\psi\|_\infty=1}}\left|\int \psi'(x)\mathring{\rho}_\ell(x)dx\right| \le 1 + 2\|\rho_\ell\|_{L^\infty} + c_*\\ &\le (1+2|b_\ell - a_\ell|^{-1})e^{c_*}.\end{aligned}$$

The next lemma, whose proof we briefly postpone, is our basic computational tool: it contains an estimate of the contribution of each of the blocks of length L appearing in (5.12).[25]

Lemma 5.6. *There exists $\varepsilon_0 > 0$, such that, for any $\varepsilon \in (0, \varepsilon_0)$, any standard pair ℓ, $L \in [\varepsilon_0^{-1}, \varepsilon_0\varepsilon^{-1/2}]$, $\sigma \in \mathrm{BV}([0, \varepsilon L], \mathbb{R}^d)$ and $\Phi \in \mathcal{C}^2(\mathbb{T}, \mathbb{R})$:*

(a) *the following bound holds*

$$\mathrm{Leb}\left[\sum_{\ell'\in\mathfrak{L}_\ell^L}\nu_{\ell'}\mathring{\rho}_{\ell'}(\cdot)e^{\varepsilon^{-1}\Phi\circ G_{\ell'}}\right] = e^{\varepsilon^{-1}\Phi(\bar\theta(\varepsilon L,\theta_\ell^*))+\sum_{j=0}^{L-1}\chi_A(\sigma_j,\bar\theta(\varepsilon j,\theta_\ell^*))}e^{\varepsilon^{-1}\mathcal{S}(\sigma)}$$

where

$$\begin{aligned}|\mathcal{S}(\sigma)| \le\,& C_\#\Big(\varepsilon L\|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + L^{-1}\|\sigma\|_{L^1([0,\varepsilon L])} + \varepsilon^2L^2\\ &+ \varepsilon\|\Phi\|_{\mathcal{C}^2}\left[1 + L\min\{1, \varepsilon^{-1}L^{-1}\|\sigma\|_{L^1([0,\varepsilon L])} + \|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + \|\Phi\|_{\mathcal{C}^2}\}\right]\Big).\end{aligned}$$

Recall that $e^{\chi_A(\sigma,\theta)}$ denotes the maximal eigenvalue of $\mathcal{L}_{\theta,\langle\sigma,A(\cdot,\theta)\rangle}$.

[23] For simplicity of notation we ignore that K may not be an integer, as such a problem can be fixed trivially.

[24] In this section we will use (5.12) only in the case $g = 1$; yet in Section 7.3 this more general formulation will be needed.

[25] The $\{\sigma_0, \cdots, \sigma_{L-1}\}$ in Lemma 5.6 correspond to an arbitrary block $\{\sigma_{jL}, \cdots, \sigma_{(j+1)L-1}\}$ in equation (5.12). Recall that σ_j is defined in (5.9).

(b) *There exists $\bar\sigma_* > 0$ so that if $\|\sigma\|_{L^\infty} + 2\|\Phi\|_{\mathcal{C}^2} \le \bar\sigma_*$, then*

$$\sum_{\ell'\in\mathfrak{L}^L_\ell} \nu_{\ell'}\mathring{\rho}_{\ell'}(x)e^{\varepsilon^{-1}\Phi(G_{\ell'}(x))} = h^*(x) \cdot m^*\left(\mathring{\rho}_\ell(\cdot)e^{\varepsilon^{-1}\Phi(\bar\theta(\varepsilon L, G_\ell(\cdot)))+\sum_{j=0}^{L-1}\chi_A(\sigma_j,\bar\theta(\varepsilon j, G_\ell(\cdot)))}\right)e^{\varepsilon^{-1}\mathcal{S}(\sigma,x)}$$

where

$$\begin{aligned}\|\mathcal{S}(\sigma,\cdot)\|_{L^\infty} \le & C_\#\Big(\varepsilon\|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + \varepsilon^2L^2(1+\|\Phi\|_{\mathcal{C}^2}) + L\varepsilon\|\sigma\|_{L^1([0,\varepsilon L])} \\ & + L^{-1}\|\sigma\|^2_{L^2([0,\varepsilon L])} + \|\Phi\|_{\mathcal{C}^2}\left(\varepsilon L\|\Phi\|_{\mathcal{C}^2} + \|\sigma\|_{L^1([0,\varepsilon L])}\right)\Big),\end{aligned}$$

where h^* is the right eigenvector of the transfer operator $\mathcal{L}_{\theta_1^*,\langle\sigma_1^*,\hat A(\cdot,\theta_1^*)\rangle}$ and m^* is the left eigenvector of $\mathcal{L}_{\theta_2^*,\langle\sigma_2^*,\hat A(\cdot,\theta_2^*)\rangle}$, where θ_1^*,θ_2^* can be chosen arbitrarily with $|\theta_\ell^*-\theta_i^*| \le C_\#\varepsilon L$ and σ_1^*,σ_2^* can be chosen arbitrarily in the essential range[26] *of σ.*

Proof of Proposition 5.4. For $k\in\{0,\cdots,K\}$ and $\theta\in\mathbb{T}$, $T=K\varepsilon L$, define

$$\Phi_k(\theta) = \int_{\varepsilon kL}^{\varepsilon KL}\chi_A(\sigma(s),\bar\theta(s-\varepsilon kL,\theta))ds.$$

It follows from equation (A.22d) that $\Phi_k\in C^2(\mathbb{T},\mathbb{R})$. Also (A.11a) implies that, for any k, we have $\|\Phi_k\|_{\mathcal{C}^2} \le C_\#\|\sigma\|_{L^1([0,T])}$. First of all notice that, by definition

$$\Lambda_{\ell,\varepsilon}(\sigma) = \Phi_0(\theta_\ell^*) + \mathcal{R}_{\ell,\varepsilon}(\sigma).$$

Moreover, let us define $J_k = [\varepsilon kL, \varepsilon(k+1)L]$ and recall (A.22a), (A.11) together with Remark 5.2; then

$$\begin{aligned}\Phi_k(\theta) - \Phi_{k+1}(\bar\theta(\varepsilon L,\theta)) &= \int_{J_k}\chi_A(\sigma(s),\bar\theta(s-\varepsilon kL,\theta)) \\ &= \sum_{j=kL}^{(k+1)L-1}\int_{\varepsilon j}^{\varepsilon(j+1)}\chi_A(\sigma(s),\bar\theta(\varepsilon(j-kL),\theta)) + \mathcal{O}(\varepsilon\|\sigma\|_{L^1(J_k)}) \\ &= \varepsilon\sum_{j=0}^{L-1}\chi_A(\sigma_{j+kL},\bar\theta(\varepsilon j,\theta)) + \mathcal{O}(\varepsilon\|\sigma\|_{\mathrm{BV}(J_k)}),\end{aligned}$$

which is the quantity appearing in Lemma 5.6.

We first proceed to prove item (a): let us fix conventionally $\ell_0=\ell$. Consider (5.12) with $g=1$ and isolate the last term:

$$\mu_\ell\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n, A\circ F_\varepsilon^n\rangle}\right) = \sum_{\ell_1\in\mathfrak{L}^L_{\ell_0}}\cdots\sum_{\ell_{K-1}\in\mathfrak{L}^L_{\ell_{K-2}}}\prod_{k=1}^{K-1}\nu_{\ell_k}\mathrm{Leb}\left[\sum_{\ell'\in\mathfrak{L}^L_{\ell_{K-1}}}\nu_{\ell'}\mathring{\rho}_{\ell'}\right].$$

We then apply Lemma 5.6-(a) with $\Phi=\Phi_K=0$ to the term in brackets and obtain:

$$\begin{aligned}\mu_\ell\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n, A\circ F_\varepsilon^n\rangle}\right) &= \sum_{\ell_1\in\mathfrak{L}^L_{\ell_0}}\cdots\sum_{\ell_{K-1}\in\mathfrak{L}^L_{\ell_{K-2}}}\prod_{k=1}^{K-1}\nu_{\ell_k}\mathrm{Leb}(\mathring{\rho}_{\ell_{K-1}})e^{\varepsilon^{-1}\Phi_{K-1}(\theta^*_{\ell_{K-1}})}e^{\varepsilon^{-1}\tilde{\mathcal{S}}(\sigma)} \\ &= \sum_{\ell_1\in\mathfrak{L}^L_{\ell_0}}\cdots\sum_{\ell_{K-2}\in\mathfrak{L}^L_{\ell_{K-3}}}\prod_{k=1}^{K-2}\nu_{\ell_k}\mathrm{Leb}\left[\sum_{\ell'\in\mathfrak{L}^L_{\ell_{K-2}}}\nu_{\ell'}\mathring{\rho}_{\ell'}e^{\varepsilon^{-1}\Phi_{K-1}\circ G_{\ell'}}\right]e^{\varepsilon^{-1}\tilde{\mathcal{S}}(\sigma)}\end{aligned}$$

[26] Recall that the essential range of a function σ is the "range modulo null sets", i.e. the intersection of the closure of the image of all functions which agree a.e. with σ.

where $\tilde{\mathcal{S}}$ stands for an arbitrary function on $\mathrm{BV}([0,T])$ satisfying the bound

$$|\tilde{\mathcal{S}}(\sigma)| \leq C_\# \left(\varepsilon L\|\sigma\|_{\mathrm{BV}(J_{K-1})} + L^{-1}\|\sigma\|_{L^1(J_{K-1})} + \varepsilon^2 L^2\right).$$

Also, we used the fact that $\mathrm{Leb}(\mathring{\rho}_{\ell_{K-1}}) = 1$ and Lemma A.10 to change the argument in Φ_{K-1}. We now apply Lemma 5.6-(a) to the term in brackets and iterate. This proves item (a) since, recalling that $K = T\varepsilon^{-1}L^{-1}$, we obtain

$$\mu_{\ell_0}\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n, A\circ F_\varepsilon^n\rangle}\right) = e^{\varepsilon^{-1}\Phi_0(\theta^*_{\ell_0}) + \mathcal{O}\left(L\|\sigma\|_{\mathrm{BV}} + \varepsilon^{-1}L^{-1}\|\sigma\|_{L^1} + LT\right)} \cdot e^{\|\sigma\|_{L^1}\mathcal{O}\left(T\varepsilon^{-1}L^{-1} + \min\{T\varepsilon^{-1}, (1+T)\varepsilon^{-1}\|\sigma\|_{L^1} + L\|\sigma\|_{\mathrm{BV}}\}\right)}.$$

To prove item (b) note that, if we assume $(1 + C_\# T)\sigma_* < \bar{\sigma}_*$, it possible to obtain a sharper estimate using Lemma 5.6-(b) and carefully keeping track of the error terms. More precisely: for any $k \in \{0, \cdots, K-1\}$ define $m_{\ell_k} = m_{\theta^*_{\ell_k}, \langle\sigma_{(k+1)L}, \hat{A}(\cdot, \theta^*_{\ell_k})\rangle}$. For each standard pair ℓ, (A.31) and Remark 5.5 yield

$$|m_{\ell_k}(\mathring{\rho}_\ell) - \mathrm{Leb}(\mathring{\rho}_\ell)| \leq C_\#|\sigma_{(k+1)L}| \leq C_\#(T^{-1}\|\sigma\|_{L^1} + \|\sigma\|_{\mathrm{BV}}). \tag{5.13}$$

Then we claim that for any $k \in \{1, \cdots, K\}$:

$$\mu_\ell\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n, A\circ F_\varepsilon^n\rangle}\right) = \sum_{\ell_1 \in \mathfrak{L}^L_{\ell_0}} \cdots \sum_{\ell_k \in \mathfrak{L}^L_{\ell_{k-1}}} \prod_{i=1}^k \nu_{\ell_i} m_{\ell_{k-1}}\left(\mathring{\rho}_{\ell_k} e^{\varepsilon^{-1}\Phi_k(G_{\ell_k}(\cdot))}\right) \cdot e^{\varepsilon^{-1}\mathcal{S}_k(\sigma)} \tag{5.14}$$

where

$$\begin{aligned}|\mathcal{S}_k(\sigma)| \leq & C_\#\big(\varepsilon\|\sigma\|_{\mathrm{BV}([\varepsilon kL, \varepsilon KL])} + (K-k)L^2\varepsilon^2(1 + \|\sigma\|_{L^1}) + L^{-1}\|\sigma\|^2_{L^2([\varepsilon kL, \varepsilon KL])} \\ & + L\varepsilon\|\sigma\|_{L^1([\varepsilon kL, \varepsilon KL])} + \varepsilon\|\sigma\|^2_{L^1} L(K-k) + \|\sigma\|_{L^1}\|\sigma\|_{L^1([\varepsilon kL, \varepsilon KL])} \\ & + \varepsilon T^{-1}\|\sigma\|_{L^1} + \varepsilon\|\sigma\|_{\mathrm{BV}}\big).\end{aligned}$$

Let us give an inductive proof of (5.14). The base case is $k = K$: choosing $g = 1$ in (5.12) yields

$$\mu_\ell\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n, A\circ F_\varepsilon^n\rangle}\right) = \sum_{\ell_1 \in \mathfrak{L}^L_{\ell_0}} \cdots \sum_{\ell_K \in \mathfrak{L}^L_{\ell_{K-1}}} \prod_{i=1}^K \nu_{\ell_i} \mathrm{Leb}\left(\mathring{\rho}_{\ell_K}\right)$$

and (5.14), for $k = K$, follows by (5.13) (i.e. $\|\mathcal{S}_K\|_{L^\infty} \leq C_\#(\varepsilon T^{-1}\|\sigma\|_{L^1} + \varepsilon\|\sigma\|_{\mathrm{BV}})$).

Next, we proceed by backward induction to prove (5.14) for $k < K$. Suppose that the estimate holds for $k + 1 \leq K$, then we need to compute

$$\sum_{\ell_{k+1} \in \mathfrak{L}^L_{\ell_k}} \nu_{\ell_{k+1}} m_{\ell_k}\left(\mathring{\rho}_{\ell_{k+1}} e^{\varepsilon^{-1}\Phi_{k+1}(G_{\ell_{k+1}}(\cdot))}\right) = \tag{5.15}$$

$$= m_{\ell_k}\left(\sum_{\ell_{k+1} \in \mathfrak{L}^L_{\ell_k}} \nu_{\ell_{k+1}} \mathring{\rho}_{\ell_{k+1}} e^{\varepsilon^{-1}\Phi_{k+1}(G_{\ell_{k+1}}(\cdot))}\right).$$

Apply Lemma 5.6-(b) with $\Phi = \Phi_{k+1}$, $m^* = m_{\ell_{k-1}}$ and $h^* = h_{\theta^*_{\ell_k}, \langle\sigma_{(k+1)L}, \hat{A}(\cdot, \theta^*_{\ell_k})\rangle}$. Since by design $m_{\ell_k}(h^*) = 1$, we obtain (5.14) at step k, which concludes the proof of (5.14) for any $k \in \{1, \cdots, K\}$.

In particular, choosing $k = 1$ we have:

$$\mu_\ell\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n, A\circ F_\varepsilon^n\rangle}\right) = m_{\ell_0}\left(\sum_{\ell_1 \in \mathfrak{L}^L_{\ell_0}} \nu_{\ell_1}\mathring{\rho}_{\ell_1} e^{\varepsilon^{-1}\Phi_1(G_{\ell_1}(\cdot))}\right) \cdot e^{\varepsilon^{-1}\mathcal{S}_1(\sigma)}. \tag{5.16}$$

We now apply once again Lemma 5.6-(b) with $\Phi = \Phi_1$, $m^* = m_{\ell_0}$ and $h^* = h_{\theta^*_{\ell_0},\langle\sigma_L,\hat A(\cdot,\theta^*_{\ell_0})\rangle}$. We conclude that

$$\mu_\ell\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n,A\circ F^n_\varepsilon\rangle}\right) = m_{\ell_0}(\mathring\rho_{\ell_0}(\cdot)e^{\varepsilon^{-1}\Phi_0(G_{\ell_0}(\cdot))})e^{\varepsilon^{-1}\mathcal{S}_0(\sigma)}.$$

Since $\|\Phi_0\|_{\mathcal{C}^2} = C_\#\|\sigma\|_{L^1}$ and ℓ_0 is a standard pair, we conclude that $\|\Phi_0(G_\ell(\cdot)) - \Phi_0(\theta^*_\ell)\| < \varepsilon\|\sigma\|_{L^1}$; hence, using once again (5.13) to estimate $m_{\ell_0}(\mathring\rho_{\ell_0})$ we obtain:

$$\mu_\ell\left(e^{\sum_{n=0}^{KL-1}\langle\sigma_n,A\circ F^n_\varepsilon\rangle}\right) = e^{\varepsilon^{-1}(\Phi_0(\theta^*_\ell)+\mathcal{S}_0(\sigma))}$$

which concludes the proof of item (b). □

Proof of Lemma 5.6. For any standard pair $\ell = (G,\rho)$ supported on $[a,b]$, recall that we consider x_j and θ_j to be random variables on ℓ. Let $g \in \mathcal{C}^0(\mathbb{T}^1,\mathbb{R}_{\geq 0})$ be an arbitrary non-negative test function; using (3.19), we can write:

$$\sum_{\ell'\in\mathfrak{L}^L_\ell}\nu_{\ell'}\mu_{\ell'}\left(g\,e^{\varepsilon^{-1}\Phi\circ G_{\ell'}}\right) = \mu_\ell\left(g(x_L)e^{\varepsilon^{-1}\Phi(\theta_L)}e^{\sum_{j=0}^{L-1}\langle\sigma_j,A(x_j,\theta_j)\rangle}\right)$$

$$= \int_a^b g(x_L(x))\rho(x)e^{\varepsilon^{-1}\Phi(\theta_L(x))+\sum_{j=0}^{L-1}\langle\sigma_j,A(x_j(x),\theta_j(x))\rangle}dx. \tag{5.17}$$

First of all, observe that, if θ_0 is distributed according to ℓ, then

$$|\Phi(\theta_0) - \Phi(\theta^*_\ell)| \leq \|\Phi\|_{\mathcal{C}^1}\varepsilon. \tag{5.18}$$

Next, let us define the random variable $\bar\theta_j = \bar\theta(\varepsilon j,\theta_0)$; and recall the notation $\bar\theta^*_{\ell,j} = \bar\theta(\varepsilon j,\theta^*_\ell)$. Observe that

$$\|\bar\theta_j - \bar\theta^*_{\ell,j}\|_{\mathcal{C}^0} \leq C_\#\varepsilon$$

Also, by Lemma 5.1 (more precisely (5.6b)) we have, for any $j\in\{0,\cdots,L-1\}$,

$$\|\theta_j - \bar\theta_j - \varepsilon H_j\| \leq C_\# j^3\varepsilon^3. \tag{5.19}$$

Hence, we conclude that (recall the definition of Ξ^*_ℓ given in (5.5):

$$\begin{aligned}
\varepsilon^{-1}\Phi(\theta_L(x)) &= \varepsilon^{-1}\Phi(\bar\theta_L(x)) + \varepsilon^{-1}\Phi'(\bar\theta_L(x))\cdot(\theta_L-\bar\theta_L) + \mathcal{O}(\|\Phi\|_{\mathcal{C}^2}\varepsilon L^2)\\
&= \varepsilon^{-1}\Phi(\bar\theta_L(x)) + \Phi'(\bar\theta_L)\cdot\sum_{j=0}^{L-1}\Xi_{j,L}\hat\omega(x_j,\theta_j) + \mathcal{O}(\|\Phi\|_{\mathcal{C}^2}\varepsilon L^2).\\
&= \varepsilon^{-1}\Phi(\bar\theta_L(x)) + \Phi'(\bar\theta^*_{\ell,L})\cdot\sum_{j=0}^{L-1}\Xi^*_{\ell,j,L}\hat\omega(x_j,\theta_j) + \mathcal{O}(\|\Phi\|_{\mathcal{C}^2}\varepsilon L^2).
\end{aligned}$$

We now proceed to incorporate the first term of the above expression in the density; the second term will be incorporated as a potential and the third term is small enough to be considered as an error term. Let us introduce the notation

$$\Gamma_{\Phi,j} = \Phi'(\bar\theta^*_{\ell,L})(\Xi^*_{\ell,j,L},0)\in\mathbb{R}^d$$

and let:

$$\rho_\Phi(x) = \rho(x)e^{\varepsilon^{-1}\left[\Phi(\bar\theta_L(x))-\Phi^*\right]} \qquad \text{where } \Phi^* = \varepsilon\log\left[\int_a^b\rho(x)e^{\varepsilon^{-1}\Phi(\bar\theta_L(x))}\right].$$

Observe that ρ_Φ is a $(c_2 + C_\#\|\Phi\|_{\mathcal{C}^2})$-standard probability density and that

$$|\Phi^* - \Phi(\bar\theta^*_{\ell,L})| \leq C_\#\|\Phi\|_{\mathcal{C}^2}\varepsilon. \tag{5.20}$$

We can then rewrite (5.17) as

$$\sum_{\ell'\in\mathfrak{L}^L_\ell} \nu_{\ell'}\mu_{\ell'}\left(g\, e^{\varepsilon^{-1}\Phi\circ G_{\ell'}}\right) = e^{\varepsilon^{-1}\Phi^*+\mathcal{O}(\|\Phi\|_{\mathcal{C}^2}\varepsilon L^2)} \times \int_a^b \rho_\Phi(x)g(x_L)e^{\sum_{j=0}^{L-1}\left[\langle\sigma_j,A(x_j,\theta_j)\rangle+\langle\Gamma_{\Phi,j},\hat{A}(x_j,\theta_j)\rangle\right]}dx. \tag{5.21}$$

It is then convenient to defined

$$\begin{aligned}\Omega_{\Phi,j}(x,\theta) &= \langle\sigma_j, A(x,\theta)\rangle + \langle\Gamma_{\Phi,j},\hat{A}(x,\theta)\rangle\\ &= \langle\sigma_j,\bar{A}(\theta)\rangle + \langle\sigma_j+\Gamma_{\Phi,j},\hat{A}(x,\theta)\rangle.\end{aligned} \tag{5.22}$$

To estimate the integral in (5.21) we use Lemma 4.2 and write, using the notations introduced there,[27] for some θ^*, $|\theta^*-\theta^*_\ell| < C_\# \varepsilon L$, to be chosen later:[28]

$$\begin{aligned}&\int_a^b \rho_\Phi(x)g(x_L)e^{\sum_{j=0}^{L-1}\left[\langle\sigma_j,A(x_j,\theta_j)\rangle+\langle\Gamma_{\Phi,j},\hat{A}(x_j,\theta_j)\rangle\right]}dx\\ &= \int_a^b \rho_\Phi(x)g(x_L)e^{\sum_{j=0}^{L-1}\Omega_{\Phi,j}(x_j^*,\theta^*)+\mathcal{O}(L\|\Phi\|_{\mathcal{C}^2}+\varepsilon^{-1}\|\sigma\|_{L^1})\varepsilon L}dx\\ &= \int_{a^*}^{b^*}\frac{\rho_\Phi\circ Y^{-1}(x)}{Y'\circ Y^{-1}(x)} g\circ f_*^L(x)e^{\sum_{j=0}^{L-1}\Omega_{\Phi,j}(f_*^j(x),\theta^*)+\mathcal{O}(\varepsilon L^2\|\Phi\|_{\mathcal{C}^2}+L\|\sigma\|_{L^1})}dx.\end{aligned} \tag{5.23}$$

Next, we let $\sigma_{\Phi,j} = \sigma_j + \Gamma_{\Phi,j}$ and write $\Omega_{\Phi,j}(x,\theta) = \bar{\Omega}_j(\theta) + \widehat{\Omega}_{\Phi,j}(x,\theta)$, where

$$\bar{\Omega}_j(\theta) = \langle\sigma_j,\bar{A}(\theta)\rangle \qquad \widehat{\Omega}_{\Phi,j}(x,\theta) = \langle\sigma_{\Phi,j},\hat{A}(x,\theta)\rangle. \tag{5.24}$$

It is then natural to introduce the BV-function

$$\sigma_\Phi(s) = \sigma(s) + \Gamma_{\Phi,\lfloor s\varepsilon^{-1}\rfloor}$$

so that $\sigma_{\Phi,j} = \varepsilon^{-1}\int_{\varepsilon j}^{\varepsilon(j+1)}\sigma_\Phi(s)ds$. Observe that the definition of $\Gamma_{\Phi,j}$ and our upper bound on L imply, if ε is sufficiently small:

$$\|\sigma_\Phi - \sigma\|_{L^\infty} \le \|\Phi\|_{\mathcal{C}^2}(1+\varepsilon L) \le 2\|\Phi\|_{\mathcal{C}^2} \tag{5.25a}$$

$$\|\sigma_\Phi\|_{\mathrm{BV}([\varepsilon j,\varepsilon j'])} \le \|\sigma\|_{\mathrm{BV}([\varepsilon j,\varepsilon j'])} + C_\#\|\Phi\|_{\mathcal{C}^2}\varepsilon|j'-j| \tag{5.25b}$$

$$\|\sigma_\Phi\|_{L^1([\varepsilon j,\varepsilon j'])} \le \|\sigma\|_{L^1([\varepsilon j,\varepsilon j'])} + C_\#\|\Phi\|_{\mathcal{C}^2}\varepsilon|j'-j|. \tag{5.25c}$$

Combining (5.23) and (5.21) and using the above definitions we can thus write:

$$\begin{aligned}\sum_{\ell'\in\mathfrak{L}^L_\ell} \nu_{\ell'}\mu_{\ell'}\left(g\, e^{\varepsilon^{-1}\Phi\circ G_{\ell'}}\right) &= e^{\varepsilon^{-1}\Phi^*+\sum_{j=0}^{L-1}\bar{\Omega}_j(\bar{\theta}^*_{\ell,j})}\\ &\cdot\int_{a^*}^{b^*}\frac{\rho_\Phi\circ Y^{-1}(x)}{Y'\circ Y^{-1}(x)}g(f^L(x,\theta^*))e^{\sum_{j=0}^{L-1}\widehat{\Omega}_{\Phi,j}(f^j(x,\theta^*),\theta^*)}dx\\ &\cdot e^{\mathcal{O}\left(\varepsilon L^2\|\Phi\|_{\mathcal{C}^2}+L\|\sigma\|_{L^1}\right)},\end{aligned} \tag{5.26}$$

where, in the above estimate, we also used:

$$\sum_{j=0}^{L-1}\bar{\Omega}_j(\theta^*) = \sum_{j=0}^{L-1}\bar{\Omega}_j(\bar{\theta}^*_{\ell,j}) + \mathcal{O}(L\|\sigma\|_{L^1}).$$

The problem with expression (5.26) is that Y' has a very large derivative (see footnote 20) and hence it cannot be effectively treated as a BV function. In Section 11

[27] Choosing $n = L$ and setting $Y = Y_L$.

[28] To ease notation, for the duration of the proof L^1 will denote $L^1([0,\varepsilon L])$ (and similarly for BV, L^∞ and L^2) unless a different domain is explicitly written.

we will deal with this problem in a more sophisticated way; here it suffices the following rough estimate based, again, on Lemma 4.2:

$$\begin{aligned}\frac{1}{Y'} &= e^{\mathcal{O}(\varepsilon L)}\prod_{j=0}^{L-1}\frac{\partial_x f(x_j^*,\theta^*)}{\partial_x f(x_j,\theta_j)}\\ &= e^{\mathcal{O}(\varepsilon L)+\sum_{j=0}^{L-1}\left[\log \partial_x f(x_j^*,\theta^*)-\log\partial_x f(x_j,\theta_j)\right]} = e^{\mathcal{O}_{L^\infty}(\varepsilon L^2)}.\end{aligned}\tag{5.27}$$

Also note that, setting

$$\tilde\rho_\Phi = \frac{\rho_\Phi\circ Y^{-1}}{\rho_\Phi^*}\mathbb{1}_{[a^*,b^*]} = \frac{\rho_\Phi\,\mathbb{1}_{[a,b]}}{\rho^*_\Phi}\circ Y^{-1}\qquad\text{where } \rho^*_\Phi = \int_{a^*}^{b^*}\rho_\Phi\circ Y^{-1},\tag{5.28}$$

we have that $\tilde\rho_\Phi$ is a $C_\#(c_2+\|\Phi\|_{\mathcal{C}^2})$-standard probability density and $\|\tilde\rho_\Phi\|_{\mathrm{BV}}\leq C_\#\|\rho_\Phi\|_{\mathrm{BV}}$. Observe moreover that (5.27) implies that $\rho^*_\Phi = e^{\mathcal{O}(\varepsilon L^2)}$. Collecting the above estimate together with (5.27) and (5.26) we have

$$\begin{aligned}\sum_{\ell'\in\mathfrak{L}^L_\ell}\nu_{\ell'}\mu_{\ell'}\left(g\,e^{\varepsilon^{-1}\Phi\circ G_{\ell'}}\right) =& e^{\varepsilon^{-1}\Phi^*+\sum_{j=0}^{L-1}\bar\Omega_j(\bar\theta^*_{\ell,j})}\\ &\cdot\int_{\mathbb{T}}\tilde\rho_\Phi(x)g(f^L(x,\theta^*))e^{\sum_{j=0}^{L-1}\widehat\Omega_{\Phi,j}(f^j(x,\theta^*),\theta^*)}dx\\ &\cdot e^{\mathcal{O}\left(\varepsilon L^2(1+\|\Phi\|_{\mathcal{C}^2})+L\|\sigma\|_{L^1}\right)}.\end{aligned}\tag{5.29}$$

Such integrals can be computed by introducing the weighted transfer operators

$$[\mathcal{L}_{\theta,\widehat\Omega_{\Phi,j}}g](x) = \sum_{f_\theta(y)=x}\frac{e^{\widehat\Omega_{\Phi,j}(y,\theta)}}{f'_\theta(y)}g(y),$$

which allow to rewrite the integral in (5.29) as

$$\begin{aligned}&\int_{\mathbb{T}}\tilde\rho_\Phi(x)g(f^L(x,\theta^*))\exp\left[\sum_{j=0}^{L-1}\widehat\Omega_{\Phi,j}(f^j(x,\theta^*),\theta^*)\right]dx\\ &\qquad= \mathrm{Leb}\left(g\mathcal{L}_{\theta^*,\widehat\Omega_{\Phi,L-1}}\cdots\mathcal{L}_{\theta^*,\widehat\Omega_{\Phi,0}}[\tilde\rho_\Phi]\right).\end{aligned}\tag{5.30}$$

Such a quantity can be computed in terms of $\hat\chi_{\theta,\Omega_{\Phi,j}} = \chi_{\theta,\widehat\Omega_{\Phi,j}}$, the logarithm of the maximal eigenvalue of $\mathcal{L}_{\theta,\widehat\Omega_{\Phi,j}}$ when acting on $\mathcal{C}^1$. Observe that by definition, remembering (5.8), (5.24), and by (A.19a), Lemma A.1 we can write[29]

$$\begin{aligned}\hat\chi_{\theta,\Omega_{\Phi,j}} &= \hat\chi_A(\sigma_{\Phi,j},\theta) = \hat\chi_A(\sigma_j,\theta)+\int_0^1 m_{\theta,\langle\sigma_j+s\Gamma_{\Phi,j},\hat A\rangle}(\langle\Gamma_{\Phi,j},\hat A\rangle h_{\theta,\langle\sigma_j+s\Gamma_{\Phi,j},\hat A\rangle})ds\\ &= \hat\chi_A(\sigma_j,\theta)+\mathcal{O}(\|\Phi\|_{\mathcal{C}^2}).\end{aligned}$$

Also, by Lemma A.7 and since $m_{\theta,0}(\hat A(\cdot,\theta)h_{\theta,0})=0$, we have, for $\|\sigma_{\Phi,j}\|$ small,

$$m_{\theta,\langle\sigma_j+s\Gamma_{\Phi,j},\hat A\rangle}(\langle\Gamma_{\Phi,j},\hat A\rangle h_{\theta,\langle\sigma_j+s\Gamma_{\Phi,j},\hat A\rangle}) = \mathcal{O}(\|\sigma_{\Phi,j}\|\|\Phi\|_{\mathcal{C}^2}).$$

Collecting the above facts, yields

$$\hat\chi_{\theta,\Omega_{\Phi,j}} = \hat\chi_A(\sigma_j,\theta)+\mathcal{O}(\min\{1,\|\sigma_{\Phi,j}\|\}\|\Phi\|_{\mathcal{C}^2}).\tag{5.31}$$

Remark. *We will now adopt the following strategy: we first obtain a rather crude bound for (5.30) (see (5.34)): this bound will be valid for arbitrary σ and Φ. We then proceed to obtain a sharper bound, which is however valid only for σ with a relatively small L^∞ norm; the sharper bound will enable us to improve the previously found rough bound to obtain item (a) and to prove (b).*

[29] Also recall the normalization $m_{\theta,\langle\sigma_j+s\Gamma_{\Phi,j},\hat A\rangle}(h_{\theta,\langle\sigma_j+s\Gamma_{\Phi,j},\hat A\rangle})=1$.

We obtain the rough bound by replacing the potential $\widehat{\Omega}_{\Phi,j}$ for $j \in \{0, \cdots, L-1\}$ with a fixed $\widehat{\Omega}^*_\Phi = \widehat{\Omega}_{\Phi,j^*}$ for some $j^* \in \{0, \cdots, L-1\}$ chosen arbitrarily. Notice in fact that, for $g \geq 0$ and any $j \in \{0, \cdots, L-1\}$:

$$\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,j}} g = e^{\mathcal{O}_{L^\infty}(\|\sigma_{\Phi,j}-\sigma_{\Phi,j^*}\|)} \mathcal{L}_{\theta^*,\widehat{\Omega}^*_\Phi} g = e^{\mathcal{O}_{L^\infty}(\|\sigma_\Phi\|_{\mathrm{BV}})} \mathcal{L}_{\theta^*,\widehat{\Omega}^*_\Phi} g, \tag{5.32}$$

whence:

$$\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,L}} \cdots \mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,0}} [\tilde{\rho}_\Phi] = e^{\mathcal{O}_{L^\infty}(L\|\sigma_\Phi\|_{BV([0,\varepsilon L])})} \mathcal{L}^L_{\theta^*,\widehat{\Omega}^*_\Phi} [\tilde{\rho}_\Phi]. \tag{5.33}$$

Since $\tilde{\rho}_\Phi$ is a BV function and σ can be arbitrarily large, we cannot guarantee that $\mathcal{L}_{\theta^*,\widehat{\Omega}^*_\Phi}$ is of Perron–Frobenius type (see Remark A.11). We thus proceed as follows: recall that $\tilde{\rho}_\Phi$ is supported on an interval $[a^*, b^*]$ of size at least $\delta/4$; since $f(\cdot, \theta^*)$ is uniformly expanding there exists $q_0 \sim \log \delta = \mathcal{O}(1)$ so that $f([a^*, b^*], \theta^*) \supset \mathbb{T}^1$; by definition of $\mathcal{L}_{\theta^*,\widehat{\Omega}^*_\Phi}$:

$$e^{-C_\#(1+\|\Phi\|_{\mathcal{C}^2}+q_0\|\widehat{\Omega}^*_\Phi\|_{\mathcal{C}^0})} \leq \mathcal{L}^{q_0}_{\theta^*,\widehat{\Omega}^*_\Phi} \tilde{\rho}_\Phi \leq e^{C_\#(1+\|\Phi\|_{\mathcal{C}^2}+q_0\|\widehat{\Omega}^*_\Phi\|_{\mathcal{C}^0})}.$$

By positivity of the transfer operator, and since it is of Perron–Frobenius type when acting on $\mathcal{C}^1$ densities (here we want to apply it to the constant functions), we can apply $\mathcal{L}^{L-q_0}_{\theta^*,\widehat{\Omega}^*_\Phi}$ to the previous inequalities and, by (A.4), obtain

$$\mathcal{L}^L_{\theta^*,\widehat{\Omega}^*_\Phi} [\tilde{\rho}_\Phi] = e^{L\,\hat{\chi}_{\theta^*,\Omega^*_\Phi} + \mathcal{O}(1+\|\Phi\|_{\mathcal{C}^2}+\|\widehat{\Omega}^*_\Phi\|_{\mathcal{C}^1})} \bar{h}_{\theta^*,\widehat{\Omega}^*_\Phi},$$

where $\bar{h}_{\theta^*,\widehat{\Omega}^*_\Phi}$ is the eigenfunction associated to the maximal eigenvalue and normalized so that $\mathrm{Leb}(\bar{h}_{\theta^*,\widehat{\Omega}^*_\Phi}) = 1$. Thus, using (5.33),

$$\mathrm{Leb}\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,L}} \cdots \mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,0}} [\tilde{\rho}_\Phi] = e^{L\,\hat{\chi}_{\theta^*,\Omega^*_\Phi} + \mathcal{O}(L\|\sigma_\Phi\|_{\mathrm{BV}}+\|\sigma_\Phi\|_{L^\infty}+1+\|\Phi\|_{\mathcal{C}^2})}, \tag{5.34}$$

where we used that $\|\widehat{\Omega}^*_\Phi\|_{\mathcal{C}^1} \leq \|\sigma_\Phi\|_{L^\infty}$. This is our announced preliminary rough bound, which holds for any σ and Φ.

In order to obtain a sharper bound we need to subdivide $\{0, \cdots, L-1\}$ into smaller sub-blocks and replace $\widehat{\Omega}_{\Phi,j}$ on each sub-block with a potential that is constant on the corresponding sub-block.

Let us now assume $\|\sigma\|_{L^\infty} + 2\|\Phi\|_{\mathcal{C}^2} < \bar{\sigma}_*$ (hence $\|\sigma_\Phi\|_{L^\infty} < \bar{\sigma}_*$) for some fixed $\bar{\sigma}_* \leq \sigma_2$ (from Lemma A.13) sufficiently small to be chosen shortly.

Observe that, by definition (5.24), we have that $\|\widehat{\Omega}_{\Phi,j}\| < C_\#\bar{\sigma}_*$ for any $j \in \{0, \cdots, L-1\}$ and thus each $\mathcal{L}_{\theta,\widehat{\Omega}_{\Phi,j}}$ is a perturbation of the Perron–Frobenius operator $\mathcal{L}_{\theta,0}$.

Lemma A.1 implies that can fix $Q \in \mathbb{N}$ such that $q = \mathcal{O}(1)$ and $\mathcal{L}^Q_{\theta,0} = \tilde{\mathcal{P}}_\theta + \tilde{\mathcal{Q}}_\theta$ where $\tilde{\mathcal{P}}_\theta$ is a projector, $\tilde{\mathcal{P}}_\theta\tilde{\mathcal{Q}}_\theta = \tilde{\mathcal{Q}}_\theta\tilde{\mathcal{P}}_\theta = 0$ and $\|\tilde{\mathcal{Q}}_\theta\|_{\mathcal{C}^1} \leq \frac{1}{4}$.

As announced, we now partition $\{0, \cdots, L-1\}$ in $L' = LQ^{-1}$ sub-blocks[30] of length Q. Let us fix arbitrarily $q^* \in \{0, \cdots, Q-1\}$; for any $l \in \{1, \cdots, L'\}$ define $\tilde{\Omega}^*_{\Phi,l} = \widehat{\Omega}_{\Phi,(l-1)Q+q^*}$ and let $\tilde{\mathcal{L}}_{\theta,l} = \mathcal{L}^Q_{\theta,\tilde{\Omega}^*_{\Phi,l}}$. Then in each sub-block, for any $g \geq 0$, similarly to (5.32):

$$\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,lQ-1}} \cdots \mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,(l-1)Q}} [g] = e^{\mathcal{O}_{L^\infty}(\|\sigma_\Phi\|_{\mathrm{BV}([\varepsilon(l-1)Q,\varepsilon lQ-1])})} \tilde{\mathcal{L}}_{\theta,l}[g] \tag{5.35}$$

By Lemma A.1, each $\tilde{\mathcal{L}}_{\theta,l}$ has a simple maximal eigenvalue, which we denote $e^{\tilde{\chi}_{\theta,l}}$. Observe that by definition $\tilde{\chi}_{\theta,l} = Q\chi_{\theta,\tilde{\Omega}^*_{\Phi,l}}$. Moreover, we can write $\tilde{\mathcal{L}}_{\theta,l} = e^{\tilde{\chi}_{\theta,l}}\tilde{\mathcal{P}}_{\theta,l} + \tilde{\mathcal{Q}}_{\theta,l}$ where $\tilde{\mathcal{P}}^2_{\theta,l} = \tilde{\mathcal{P}}_{\theta,l}$, $\tilde{\mathcal{P}}_{\theta,l}\tilde{\mathcal{Q}}_{\theta,l} = \tilde{\mathcal{Q}}_{\theta,l}\tilde{\mathcal{P}}_{\theta,l} = 0$ and the theory of

[30] Once again we ignore the issue that L' may not be an integer.

Section A.2 implies that $\|\tilde{\mathcal{Q}}_{\theta,l}\|_{\mathcal{C}^1} \leq \frac{1}{2}e^{\tilde{\chi}_{\theta,l}}$, provided that $\bar{\sigma}_*$ is sufficiently small and Q has been chosen large enough. Let us write $\tilde{\mathcal{P}}_{\theta,l}g = h_{\theta,l}m_{\theta,l}(g)$, normalized as in Lemma A.6.

The main advantage in defining the iterated operators $\tilde{\mathcal{L}}_{\theta,l}$ is that the bound on the *norm* of $\tilde{\mathcal{Q}}$ (as opposed to the bound on the mere spectral radius which is available for the operator $\mathcal{Q}$ acting on a single iterate) makes them well behaved under composition.

Sub-lemma 5.7. *Using the above notation, if $\bar{\sigma}_* > 0$ is sufficiently small and $\|\sigma_\Phi\|_{L^\infty} < \bar{\sigma}_*$:*

$$\tilde{\mathcal{L}}_{\theta,L'} \cdots \tilde{\mathcal{L}}_{\theta,1}[\tilde{\rho}_\Phi] = h_{\theta,L'}(x)m_{\theta,1}(\tilde{\rho}_\Phi)e^{\sum_{l=1}^{L'}\tilde{\chi}_{\theta,l}}e^{\mathcal{O}_{L^\infty}(\|\sigma_\Phi\|_{\mathrm{BV}})}.$$

The above sub-lemma, whose proof is postponed after the end of the current proof, allows to refine the rough estimate (5.34). Observe that, using (A.17a) and (A.31):

$$\begin{aligned}
&\mathrm{Leb}(h_{\theta,L'}) = e^{\mathcal{O}(\|\sigma_\Phi\|_{L^\infty})}\\
&|m_{\theta,1}\tilde{\rho}_\Phi| = \mathrm{Leb}\,\tilde{\rho}_\Phi + \mathcal{O}(\|\sigma_\Phi\|_{L^\infty}(1+\|\Phi\|_{\mathcal{C}^2})) = e^{\mathcal{O}(\|\sigma_\Phi\|_{L^\infty}(1+\|\Phi\|_{\mathcal{C}^2}))};
\end{aligned}$$

moreover, by (A.11a):

$$\sum_{l=1}^{L'}\tilde{\chi}_{\theta,l} = L\hat{\chi}_{\theta,\Omega_\Phi^*} + \mathcal{O}(L\|\sigma_\Phi\|_{\mathrm{BV}}).$$

We thus conclude that if $\|\sigma_\Phi\|_{L^\infty} < \bar{\sigma}_*$, using (5.35):

$$\mathrm{Leb}\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,L}}\cdots\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,0}}[\tilde{\rho}_\Phi] = e^{L\hat{\chi}_{\theta^*,\Omega_\Phi^*}+\mathcal{O}(L\|\sigma_\Phi\|_{\mathrm{BV}}+\|\sigma_\Phi\|_{L^\infty}(1+\|\Phi\|_{\mathcal{C}^2}))}.$$

Combining the above equation with (5.34), we conclude that for arbitrary σ_Φ:

$$\mathrm{Leb}\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,L}}\cdots\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,0}}[\tilde{\rho}_\Phi] = e^{L\hat{\chi}_{\theta^*,\Omega_\Phi^*}+\mathcal{O}(L\|\sigma_\Phi\|_{\mathrm{BV}}+\|\sigma_\Phi\|_{L^\infty}+\min\{1,\|\sigma_\Phi\|_{L^\infty}\}\|\Phi\|_{\mathcal{C}^2})}.$$

Applying (5.31) we thus obtain

$$\mathrm{Leb}\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,L}}\cdots\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,0}}[\tilde{\rho}_\Phi] = e^{L\hat{\chi}_A(\sigma_{j^*},\theta^*)+\mathcal{O}(L\min\{1,\|\sigma_\Phi\|_{L^\infty}\}\|\Phi\|_{\mathcal{C}^2}+L\|\sigma_\Phi\|_{\mathrm{BV}}+\|\sigma_\Phi\|_{L^\infty})}.$$

At last, setting $g = 1$ and substituting the latter equation in (5.30) and (5.29),

$$\begin{aligned}
&\left[\sum_{\ell'\in\mathfrak{L}_\ell^L}\nu_{\ell'}\mu_{\ell'}\left(e^{\varepsilon^{-1}\Phi\circ G_{\ell'}}\right)\right]e^{-\varepsilon^{-1}\Phi^*-\sum_{j=0}^{L-1}\bar{\Omega}_j(\bar{\theta}_{\ell,j}^*)} = e^{L\hat{\chi}_A(\sigma_{j^*},\theta^*)}\\
&\cdot e^{\mathcal{O}(L\min\{1,\|\sigma_\Phi\|_{L^\infty}\}\|\Phi\|_{\mathcal{C}^2}+L\|\sigma_\Phi\|_{\mathrm{BV}}+\|\sigma_\Phi\|_{L^\infty}+\varepsilon L^2(1+\|\Phi\|_{\mathcal{C}^2})+L\|\sigma\|_{L^1})}.
\end{aligned}$$

Choosing $\theta^* = \bar{\theta}_{\ell,j^*}^*$ and taking the geometric mean of the above expressions for $j^* \in \{0,\cdots,L-1\}$, we conclude

$$\begin{aligned}
&\sum_{\ell'\in\mathfrak{L}_\ell^L}\nu_{\ell'}\mu_{\ell'}\left(e^{\varepsilon^{-1}\Phi\circ G_{\ell'}}\right) = e^{\varepsilon^{-1}\Phi^*+\sum_{j=0}^{L-1}\chi_A(\sigma_j,\bar{\theta}_{\ell,j}^*)}\\
&\cdot e^{\mathcal{O}(L\min\{1,\|\sigma_\Phi\|_{L^\infty}\}\|\Phi\|_{\mathcal{C}^2}+L\|\sigma_\Phi\|_{\mathrm{BV}}+\|\sigma_\Phi\|_{L^\infty}+\varepsilon L^2(1+\|\Phi\|_{\mathcal{C}^2})+L\|\sigma\|_{L^1})}.
\end{aligned}$$

Item (a) then follows using (5.25), (5.20) and $\|\sigma\|_{L^\infty} \leq (\varepsilon L)^{-1}\|\sigma\|_{L^1} + \|\sigma\|_{\mathrm{BV}}$.

We now proceed to the proof of item (b), which follows from a more careful application of Sub-Lemma 5.7.

Recall that, by definition, $h_{\theta,L'} = h_{\theta,\widehat{\Omega}_{\Phi,L'}^*}$, where $\widehat{\Omega}_{\Phi,L'}^* = \langle\sigma_{\Phi,(L'-1)Q+q^*},\hat{A}(\cdot,\theta)\rangle$. Then, notice that for any σ_1^* in the essential range of σ we have, using bounds (5.25), that $\sigma_1^* - \sigma_{\Phi,(L'-1)Q+q^*} \leq \|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + \|\Phi\|_{\mathcal{C}^2}\varepsilon L^2$. By (A.17a) and (A.22b) we thus conclude that for any θ_1^* so that $|\theta_1^* - \theta| < C_\# \varepsilon L$:

$$h^* = h_{\theta_1^*,\langle\sigma_1^*,\hat{A}(\cdot,\theta_1^*)\rangle} = h_{\theta,L'} e^{\mathcal{O}_{L^\infty}(\|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + \|\Phi\|_{\mathcal{C}^2}\varepsilon L^2 + \varepsilon L)}. \tag{5.36}$$

Likewise, for any σ_2^* in the essential range of σ, using (A.30) we gather,

$$m^*(\tilde{\rho}_\Phi) = m_{\theta,\langle\sigma_2^*,\hat{A}(\cdot,\theta)\rangle}(\tilde{\rho}_\Phi) = m_{\theta,1}(\tilde{\rho}_\Phi) e^{\mathcal{O}(\|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + \|\Phi\|\varepsilon L^2)}. \tag{5.37}$$

Also, by (A.11a) and using (5.31) and (5.25) we obtain

$$\begin{aligned}\sum_{l=1}^{L'} \tilde{\chi}_{\theta,l} = Q\sum_{l=1}^{L'} \hat{\chi}_{\theta,\tilde{\Omega}^*_{\Phi,l}} &= \sum_{j=0}^{L-1} \hat{\chi}_{\theta,\Omega_{\Phi,j}} + \mathcal{O}(\|\sigma_\Phi\|_{\mathrm{BV}([0,\varepsilon L])}) \\ &= \sum_{j=0}^{L-1} \hat{\chi}_A(\sigma_j,\theta) + \mathcal{O}(\|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + L\|\Phi\|_{\mathcal{C}^2}(\varepsilon + \varepsilon^{-1}L^{-1}\|\sigma\|_{L^1} + \|\Phi\|_{\mathcal{C}^2}))\end{aligned}$$

and using (A.22a):

$$\sum_{j=0}^{L-1} \hat{\chi}_A(\sigma_j,\theta) = \sum_{j=0}^{L-1} \hat{\chi}_A(\sigma_j,\bar{\theta}_j^*) + \mathcal{O}(\|\sigma\|_{L^2}^2 L).$$

Hence, using Sub-Lemma 5.7 and equations (5.35), (5.36), (5.37), we conclude:

$$\begin{aligned}\mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,L-1}} \cdots \mathcal{L}_{\theta^*,\widehat{\Omega}_{\Phi,0}}[\tilde{\rho}_\Phi] = h^* m^*(\tilde{\rho}_\Phi) e^{\sum_{j=0}^{L-1}\hat{\chi}_A(\sigma_j,\bar{\theta}_j^*)} \\ \cdot e^{\mathcal{O}_{L^\infty}(\|\sigma\|_{\mathrm{BV}} + \varepsilon L^2\|\Phi\|_{\mathcal{C}^2} + \|\Phi\|_{\mathcal{C}^2}(\|\Phi\|_{\mathcal{C}^2}L + \varepsilon^{-1}\|\sigma\|_{L^1}) + \varepsilon L + \|\sigma\|_{L^2}^2 L)}.\end{aligned} \tag{5.38}$$

In order to proceed we need to compare $m^*(\tilde{\rho}_\Phi)$ with $m^*(\mathring{\rho}_\Phi)$. Recall that by (5.28), $\tilde{\rho}_\Phi = (\mathring{\rho}_\Phi \circ Y^{-1})/\rho_\Phi^*$, where $\mathring{\rho}_\Phi = \rho_\Phi\,\mathbb{1}_{[a,b]}$ and $\rho_\Phi^* = 1 + \mathcal{O}(\varepsilon L^2)$. We claim that

$$m^*(\mathring{\rho}_\Phi \circ Y^{-1}) = m^*(\mathring{\rho}_\Phi) e^{\mathcal{O}(\varepsilon L^2 + \|\sigma\|_{\mathrm{BV}([0,\varepsilon L])} + \varepsilon^{-1}L^{-1}\|\sigma\|_{L^2}^2)}. \tag{5.39}$$

Observe that substituting (5.39) into (5.38), item (b) follows by (5.30) and (5.29) since g is arbitrary. In order to conclude, we therefore only need to prove (5.39). First of all, recall that ρ_Φ is a $(c_2 + C_\#\|\Phi\|_{\mathcal{C}^2})$-standard probability density and that, by hypotheses, $\|\Phi\|_{\mathcal{C}^2} \leq C_\#\bar{\sigma}_*$; hence Remark 5.5 implies that $\|\mathring{\rho}_\Phi\|_{\mathrm{BV}} \leq C_\#$. Hence, if $\bar{\sigma}_*$ is sufficiently small and since $\mathrm{Leb}\,\mathring{\rho}_\Phi = 1$, Lemma A.14 yields:

$$m^*(\mathring{\rho}_\Phi) = e^{\mathcal{O}(\bar{\sigma}_*|\log\bar{\sigma}_*|)}.$$

Let us proceed to estimate $m^*(\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi)$: if $\bar{\sigma}_*$ is sufficiently small, Lemma A.14 ensures that

$$\begin{aligned}m^*(\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi) &= \mathrm{Leb}(\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi) + \|\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi\|_{L^1}\mathcal{O}(\|\sigma_2^*\||\log\|\sigma_2^*\||) \\ &\quad + \mathcal{O}(\|\sigma_2^*\|^2\|\mathring{\rho}_\Phi\|_{\mathrm{BV}}) \\ &\leq \|\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi\|_{L^1}\,(1 + \mathcal{O}(\|\sigma_2^*\||\log\|\sigma_2^*\||)) \\ &\quad + \mathcal{O}(\|\sigma_2^*\|^2\|\rho_\Phi\|_{\mathrm{BV}})\end{aligned}$$

Next, we proceed to estimate the L^1 norm; recall that for any bounded ψ:

$$\|\psi\|_{L^1} = \sup_{\substack{\varphi\in L^\infty \\ \|\varphi\|_{L^\infty}=1}} \int \psi\varphi = \sup_{\substack{\varphi\in\mathcal{C}^0 \\ \|\varphi\|_{\mathcal{C}^0}=1}} \int \psi\varphi,$$

where in the last equality we used the fact that continuous functions are dense in L^1. For any $\varphi \in \mathcal{C}^0$ we have

$$\begin{aligned}\left|\text{Leb}(\varphi\left[\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi\right])\right| &= \left|\int_{\mathbb{T}} \left[\varphi(Y(x)) \cdot Y'(x) - \varphi(x)\right] \rho_\Phi(x) dx\right| \\ &= \left|\int_{\mathbb{T}} [\overline{\varphi} \circ Y - \overline{\varphi}]'(x) \mathring{\rho}_\Phi(x) dx\right| \\ &\leq \|\mathring{\rho}_\Phi\|_{\text{BV}} \|\overline{\varphi} \circ Y - \overline{\varphi}\|_{\mathcal{C}^0},\end{aligned}$$

where $\overline{\varphi}' = \varphi$ on $[0,1]$. Since $|Y - \text{Id}|_\infty \leq C_\# \varepsilon L^2$ (by (5.27)), and $\|\mathring{\rho}_\Phi\|_{\text{BV}} \leq C_\#$, we conclude that $|\text{Leb}(\varphi[\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi])| \leq C_\# \|\varphi\|_{\mathcal{C}^0} \varepsilon L^2$, which implies

$$\|\mathring{\rho}_\Phi \circ Y^{-1} - \mathring{\rho}_\Phi\|_{L^1} \leq C_\# \varepsilon L^2.$$

Accordingly, putting together the above estimates:

$$m^*(\mathring{\rho}_\Phi \circ Y^{-1}) = m^*(\mathring{\rho}_\Phi) e^{\mathcal{O}(\varepsilon L^2 + \|\sigma_2^*\|^2)}.$$

Observe that

$$\|\sigma_2^*\|^2 = \frac{1}{\varepsilon L} \int_0^{\varepsilon L} \|\sigma(s) + \sigma_2^* - \sigma(s)\|^2 \leq \frac{1}{\varepsilon L} \|\sigma\|_{L^2}^2 + 4\bar{\sigma}_* \|\sigma\|_{\text{BV}([0,\varepsilon L])},$$

and thus we have

$$m^*(\mathring{\rho}_\Phi \circ Y^{-1}) = m^*(\mathring{\rho}_\Phi) e^{\mathcal{O}(\varepsilon L^2 + \|\sigma\|_{\text{BV}([0,\varepsilon L])} + \varepsilon^{-1} L^{-1} \|\sigma\|_{L^2}^2)}.$$

which gives (5.39) and concludes the proof of the Lemma. □

Proof of Sub-lemma 5.7. First of all observe that, using (A.17a) and (A.30),

$$\|h_{\theta,l} - h_{\theta,l+1}\| \leq \mathcal{R}_{\theta,l+1} \tag{5.40a}$$

$$|m_{\theta,l}(g) - m_{\theta,l+1}(g)| \leq \mathcal{R}_{\theta,l+1} \|g\|_{\text{BV}} \tag{5.40b}$$

where[31]

$$\mathcal{R}_{\theta,l+1} = C_Q \min\left\{\|\sigma_\Phi\|_{\text{BV}([\varepsilon(l-1)Q,\varepsilon(l+1)Q])}, \varepsilon^{-1}\|\sigma_\Phi\|_{L^1([\varepsilon(l-1)Q,\varepsilon(l+1)Q])}\right\} \tag{5.41}$$

where the first term in the min comes from comparing the potential in one block to the potential in the next one and the second term comes from comparing the potential in each block with the zero potential. We assume conventionally $\mathcal{R}_{\theta,0} = 1$.

Let $\rho_{(0)} = \tilde{\rho}_\Phi$ and define for $l \in \{1, \cdots, L'\}$:

$$\rho_{(l)} := \tilde{\mathcal{L}}_{\theta,l} \rho_{(l-1)};$$

observe in particular that $\rho_{(L')} = \tilde{\mathcal{L}}_{\theta,L'} \cdots \tilde{\mathcal{L}}_{\theta,1}[\tilde{\rho}_\Phi]$.

Let us now define $\gamma_l = m_l(\rho_{(l)}) \geq 0$ and $\varphi_l = (1 - \tilde{\mathcal{P}}_l)\rho_{(l)}$ so that $\rho_{(l)} = \gamma_l h_l + \varphi_l$; in particular $\|\rho_{(l)}\|_{\text{BV}} \leq C_\# \gamma_l + \|\varphi_l\|_{\text{BV}}$ and $m_l(\varphi_l) = 0$. Then

$$\begin{aligned}\gamma_{l+1} &= m_{l+1}(\rho_{l+1}) = e^{\tilde{\chi}_{\theta,l+1}} m_{l+1}(\rho_l) \\ &= e^{\tilde{\chi}_{\theta,l+1}} \left(\gamma_l - m_l(\rho_{(l)}) + m_{l+1}(\rho_{(l)})\right) \\ &= e^{\tilde{\chi}_{\theta,l+1}} \gamma_l + e^{\tilde{\chi}_{\theta,l+1}} [m_{l+1} - m_l](\rho_{(l)})\ ;\end{aligned} \tag{5.42a}$$

$$\begin{aligned}\varphi_{l+1} &= \rho_{(l+1)} - \gamma_{l+1} h_{l+1} = \tilde{\mathcal{L}}_{\theta,l+1}(\rho_{(l)} - m_{l+1}(\rho_{(l)}) h_{l+1}) \\ &= \tilde{\mathcal{Q}}_{\theta,l+1}(\rho_{(l)} - m_{l+1}(\rho_{(l)}) h_{l+1}).\end{aligned} \tag{5.42b}$$

[31] The proposed estimate of $\mathcal{R}_{\theta,l}$ may seem a bit cumbersome. The reason is that the second possibility is good locally to verify the condition $\mathcal{R}_l \leq 2Q\bar{\sigma}_*$ below but is otherwise a bad choice since it gives a too large cumulative mistake.

By (5.40b) we have $|[m_{l+1} - m_l](\rho_{(l+1)})| \le C_\# \mathcal{R}_{l+1} \|\rho_{(l+1)}\|_{\mathrm{BV}}$. Accordingly, since $\|\tilde{\mathcal{Q}}_{\theta,l}\|_{\mathrm{BV}} < \frac{1}{2} e^{\tilde{\chi}_{\theta,l}}$ and setting $\alpha = \log 2$:

$$\begin{aligned} \gamma_{l+1} &= e^{\tilde{\chi}_{\theta,l+1}} \gamma_l + \mathcal{O}(e^{\tilde{\chi}_{\theta,l+1}} \mathcal{R}_{l+1} \|\rho_{(l)}\|_{\mathrm{BV}}) \\ &= e^{\tilde{\chi}_{\theta,l+1} + \mathcal{O}(\mathcal{R}_{l+1})} \gamma_l + \mathcal{O}(e^{\tilde{\chi}_{\theta,l+1}} \mathcal{R}_{l+1} \|\varphi_l\|_{\mathrm{BV}}) \, ; \end{aligned} \tag{5.43a}$$

$$\begin{aligned} \|\varphi_{l+1}\|_{\mathrm{BV}} &\le e^{\tilde{\chi}_{\theta,l+1} - \alpha} \|\rho_{(l)} - m_{l+1}(\rho_{(l)}) h_{l+1}\|_{\mathrm{BV}} \\ &\le e^{\tilde{\chi}_{\theta,l+1} - \alpha} \|m_l(\rho_{(l)}) h_l + \varphi_l - m_{l+1}(\rho_{(l)}) h_{l+1}\|_{\mathrm{BV}} \\ &\le e^{\tilde{\chi}_{\theta,l+1} - \alpha} \left[\|\varphi_l\|_{\mathrm{BV}} + C_\# \mathcal{R}_{l+1} \|\rho_{(l)}\|_{\mathrm{BV}} \right] \\ &\le e^{\tilde{\chi}_{\theta,l+1} - \alpha} \left[e^{C_\# \mathcal{R}_{l+1}} \|\varphi_l\|_{\mathrm{BV}} + C_\# \mathcal{R}_{l+1} \gamma_l \right] . \end{aligned} \tag{5.43b}$$

Next, we prove, by induction, that there exists $C_* > 1$ such that, for any $l \in \{1, \cdots, L'\}$

$$\|\varphi_l\|_{\mathrm{BV}} \le C_* \gamma_l \sum_{j=0}^{l} e^{-(l-j)\alpha/2} \mathcal{R}_j. \tag{5.44}$$

Since $\rho_{(0)}$ is a standard density we have $\gamma_0 = m_0(\rho_{(0)}) \ge C_\# \|\rho_{(0)}\|_{\mathrm{BV}} \ge C_\# \|\varphi_0\|_{\mathrm{BV}}$, thus the relation is satisfied for $l = 0$ provided C_* is chosen large enough. Next, combining (5.43a) with (5.44) and observing that for any j we have by definition $\mathcal{R}_j \le 2Q\bar{\sigma}_*$, we obtain:

$$\begin{aligned} \gamma_{l+1} &\ge e^{\tilde{\chi}_{\theta,l+1} - C_\# \mathcal{R}_{l+1}} \gamma_l - C_\# e^{\tilde{\chi}_{\theta,l+1}} \mathcal{R}_{l+1} \left[C_* \gamma_l \sum_{j=0}^{l} e^{-(l-j)\alpha/2} Q \bar{\sigma}_* \right] \\ &\ge e^{\tilde{\chi}_{\theta,l+1} - C_\# Q \bar{\sigma}_* - C_\# C_* Q^2 \bar{\sigma}_*^2} \gamma_l. \end{aligned}$$

Plugging the above estimate into (5.44) and combining with (5.43b) yields,

$$\begin{aligned} \|\varphi_{l+1}\|_{\mathrm{BV}} &\le e^{-\alpha + C_\# Q \bar{\sigma}_* + C_\# Q^2 C_* \bar{\sigma}_*^2} \gamma_{l+1} C_* \sum_{j=0}^{l} e^{-(l-j)\alpha/2} \mathcal{R}_j + C_\# \mathcal{R}_{l+1} \gamma_{l+1} \\ &\le C_* \gamma_l \sum_{j=0}^{l+1} e^{-(l+1-j)\alpha/2} \mathcal{R}_j \end{aligned}$$

provided C_* is chosen large enough and $C_\# Q \bar{\sigma}_* + C_\# C_* Q^2 \bar{\sigma}_*^2 \le \frac{\alpha}{2}$, which can always be satisfied by choosing $\bar{\sigma}_*$ small enough. This concludes the proof of (5.44). Combining this estimate with (5.41) we conclude that

$$\|\varphi_{L'}\|_{\mathrm{BV}} \le C_\# \gamma_{L'} \|\sigma_\Phi\|_{\mathrm{BV}}.$$

Finally, substituting again (5.44) into (5.43a) implies

$$\gamma_{L'} = e^{\tilde{\chi}_{\theta,L'} + \mathcal{O}(\mathcal{R}_{L'})} \gamma_l = e^{\sum_{l=1}^{L'} [\tilde{\chi}_{\theta,l} + \mathcal{O}(\mathcal{R}_l)]} \gamma_0.$$

since $\gamma_0 = m_0(\rho_{(0)}) = m_0(\tilde{\rho}_\Phi)$, we gather that

$$\rho_{(L')}(x) = e^{\sum_{l=1}^{L'} [\tilde{\chi}_{\theta,l} + \mathcal{O}(\mathcal{R}_l)]} m_0(\tilde{\rho}_\Phi)(h_{L'}(x) + C_\# \|\sigma_\Phi\|_{\mathrm{BV}}).$$

Hence, recalling, from Lemma A.1 (or more precisely (A.2)) that $h_{L'} \ge e^{c_\#}$, we can conclude the proof of the sub-lemma, since $\sum_{l=1}^{L'} \mathcal{R}_{\theta,l} \le C_\# \|\sigma_\Phi\|_{\mathrm{BV}}$. □

5.3. **Regularizing moment generating functional.**

The discussion of the previous section tells us that, provided the error term $\mathcal{R}_{\ell,\varepsilon}(\sigma)$ is somewhat under control, the logarithmic moment generating function $\Lambda_{\ell,\varepsilon}(\sigma)$ is well described by $\int_0^T \chi_A(\sigma(s),\bar\theta(s,\theta_\ell^*))ds$. Proposition 5.4, however, shows that any kind of control on the remainder term $\mathcal{R}_{\ell,\varepsilon}(\sigma)$ might fail if the BV-norm of σ is much larger than its L^1-norm (i.e. for rapidly oscillating functions). We will then need to consider regularizations of σ whose BV-norm is controlled by their L^1-norm.

Given a step size $h = T/N_h$, for $N_h \in \mathbb{N}$ suitably large, define the projector $\Pi^{(h)}$ given by averaging on each interval of size h:

$$[\Pi^{(h)}\sigma](t) = h^{-1}\int_{h\lfloor th^{-1}\rfloor}^{h(\lfloor th^{-1}\rfloor+1)} \sigma(s)ds. \tag{5.45}$$

We collect in the following sub-lemma the basic properties of $\Pi^{(h)}$; their proof is elementary and it is left to the reader.

Sub-lemma 5.8. *The operator $\Pi^{(h)}$ satisfies the following properties*

(a) $\Pi^{(h)}1 = 1$
(b) $\int f\cdot\Pi^{(h)}\sigma = \int \Pi^{(h)} f\cdot\sigma$;
(c) $\Pi^{(h)}$ *is a contraction in the* BV *and* L^p*-norms if* $p\in[1,\infty]$;
(d) $\|\Pi^{(h)}\sigma\|_{\mathrm{BV}} \le C_\# h^{-1}\|\sigma\|_{L^1}$.

We then proceed to define the *regularized moment generating functional* as

$$\Lambda^{(h)}_{\ell,\varepsilon} = \Lambda_{\ell,\varepsilon}\circ\Pi^{(h)} \tag{5.46}$$

and as in the previous section we can define

$$\mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma) = \Lambda^{(h)}_{\ell,\varepsilon}(\sigma) - \int_0^T \chi_A(\sigma(s),\bar\theta(s,\theta_\ell^*))ds \tag{5.47}$$

Lemma 5.9. *There exists $\varepsilon_0>0$, such that if $\varepsilon\in(0,\varepsilon_0)$, $L\in\left[\varepsilon_0^{-1},\varepsilon_0\varepsilon^{-1/2}\right]$ and $T\in[\varepsilon L, T_{\max}]$:*

(a) *for any $\sigma\in\mathrm{BV}([0,T],\mathbb{R}^d)$, the following upper bound holds:*

$$\mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma) \le C_\#(\varepsilon LT + \\ + \left[\varepsilon Lh^{-1} + h + L^{-1} + \min\{T,(1+\varepsilon Lh^{-1})\|\sigma\|_{L^1}\}\right]\|\sigma\|_{L^1});$$

(b) *there exists $\sigma_* = \sigma_*(T_{\max})>0$ so that if $\|\sigma\|_{L^\infty}<\sigma_*$, the following upper bound holds:*

$$\mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma) \le C_\#\left(\varepsilon LT + \varepsilon[L+T^{-1}+h^{-1}+h\varepsilon^{-1}]\|\sigma\|_{L^1} + \|\sigma\|^2_{L^1} + L^{-1}\|\sigma\|^2_{L^2}\right).$$

Proof. Observe that, by definition

$$\mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma) = \mathcal{R}_{\ell,\varepsilon}(\Pi^{(h)}\sigma) + \int_0^T \chi_A(\Pi^{(h)}\sigma(s),\bar\theta(s,\theta^*_\ell))ds - \int_0^T \chi_A(\sigma(s),\bar\theta(s,\theta^*_\ell))ds.$$

But since $\chi_A(\cdot,\theta)$ is convex, Jensen inequality yields:

$$\int_0^T \chi_A(\Pi^{(h)}\sigma(s),\bar\theta(s,\theta^*_\ell))ds \le \int_0^T ds\int_{h\lfloor s\varepsilon^{-1}\rfloor}^{h(\lfloor s\varepsilon^{-1}\rfloor+1)} \frac{\hat\chi_A(\sigma(r),\bar\theta(s,\theta^*_\ell))}{h}dr. \tag{5.48}$$

Next, by (A.21a), (A.5) and since, by Lemma A.1, m_θ is a measure, holds the normalization $m_\theta(h_\theta)=1$ and $|h'_\theta|\le C_\#\|\sigma\|h_\theta$, we have

$$|\hat\chi_A(\sigma,\bar\theta(s,\theta)) - \hat\chi_A(\sigma,\bar\theta(r,\theta))| \le C_\# h\|\sigma\|.$$

Hence

$$\int_0^T \chi_A(\Pi^{(h)}\sigma(s), \bar\theta(s,\theta_\ell^*))ds \le \int_0^T [\Pi^{(h)}\chi_A(\sigma(\cdot), \bar\theta(\cdot,\theta_\ell^*))](s)ds + C_\# h\|\sigma\|_{L^1}$$
$$\le \int_0^T \chi_A(\sigma(s), \bar\theta(s,\theta_\ell^*))ds + C_\# h\|\sigma\|_{L^1},$$

where we used items (a-b) of Sub-lemma 5.8 to conclude that $\int \Pi^{(h)} f = \int f$. The lemma readily follows from items (c-d) of Sub-lemma 5.8 and Proposition 5.4. □

6. Deviations from the average: the rate function

Here we study the deviations from the average behavior described in Section 4.

Remark 6.1. *The results in Sections 6 and 7 are in the spirit of [20] although more precise, insofar in [20] only a rough upper bound on the rate function is provided. Regarding the classical Large Deviations Principle, the exact rate function was derived in [38], but with an estimate of the error largely insufficient to handle moderate deviations, as the function was computed with a mistake of order* $o(1)$*. Here we estimate the error much more precisely and we are therefore able to study accurately also deviations of order* ε^α*, with* $\alpha < 1/2$*. In addition, contrary to [20], we derive not only an upper bound but a lower bound as well, at least for deviations larger than* $\varepsilon^{1/96}$*. We refrain from obtaining completely optimal results (which may be obtained using the techniques developed later in this paper) only to keep the length of the paper (somewhat) under control.*

Recall that we fixed $d \in \mathbb{N}$ and $A = (A_1, \cdots, A_d) \in \mathcal{C}^2(\mathbb{T}^2, \mathbb{R}^d)$, with $A_1(x,\theta) = \omega(x,\theta)$. Recall moreover that we are always under the standing assumption (A1'). Finally, note that, for convenience, we will often implicitly lift $\theta \in \mathbb{T}$ to its universal cover $\mathbb{R}$.

The fundamental object in the theory of large deviations is the *rate function*. Because its definition is a bit involved, we start by discussing it in some detail. The reader that is not familiar with the meaning and the use of such a function may want to review the discussion in Sections 2.1, 2.3 and have a preliminary look at Section 7 where it is made clear the role of the rate function in the statements of the various large and moderate deviations results.

6.1. Definition and properties: the preliminary rate function.

We start by discussing a rate function that is expressed in terms of the averaged trajectory of θ and therefore turns out to be accurate only for short times. However its discussion entails all the quantities and ideas needed for the general case.

Recall that $e^{\chi_A(\sigma,\theta)}$ (resp. $e^{\hat\chi_A(\sigma,\theta)}$) denotes the maximal eigenvalue of the transfer operator $\mathcal{L}_{\theta,\langle\sigma,A\rangle}$ (resp. $\mathcal{L}_{\theta,\langle\sigma,\hat A\rangle}$) which has been introduced in Section 5.2. Recall also that $\chi_A(\sigma,\theta) = \langle\sigma, \bar A(\theta)\rangle + \hat\chi_A(\sigma,\theta)$; finally, observe that (A.12a), (A.12b) and Lemma A.16, together with assumption (A1') imply that $\chi_A(\cdot,\theta)$ is a strictly convex function.

For any $\sigma, b \in \mathbb{R}^d$ and $\theta \in \mathbb{T}^1$, define

$$\kappa(\sigma,b,\theta) = \langle\sigma,b\rangle - \chi_A(\sigma,\theta) = \langle\sigma, b - \bar A(\theta)\rangle - \hat\chi_A(\sigma,\theta), \tag{6.1}$$

and define the function $\mathcal{Z} : \mathbb{R}^d \times \mathbb{T} \to \mathbb{R} \cup \{+\infty\}$ as

$$\mathcal{Z}(b,\theta) = \sup_{\sigma\in\mathbb{R}^d} \kappa(\sigma,b,\theta). \tag{6.2}$$

Observe that $\mathcal{Z}(\cdot,\theta)$ is the Legendre transform of $\chi_A(\cdot,\theta)$. We are now able to give a first preliminary definition of the *rate function*; for any $\theta^* \in \mathbb{T}$, let

$$I_{\text{pre},\theta^*} : \mathcal{C}^0([0,T],\mathbb{R}^d) \to \mathbb{R}\cup\{+\infty\}$$

$$I_{\text{pre},\theta^*}(\gamma) = \begin{cases} +\infty & \text{if } \gamma \text{ is not Lipschitz, or } \gamma(0)\neq 0 \\ \int_0^T \mathcal{Z}(\gamma'(s),\bar\theta(s,\theta^*))\,ds & \text{otherwise.}\end{cases} \tag{6.3}$$

Our next task is to investigate the properties of I_{pre,θ^*} or, equivalently, of $\mathcal{Z}$. Let $\mathbb{D}(\theta) = \{b\in\mathbb{R}^d : \mathcal{Z}(b,\theta) < +\infty\}$ be the *effective domain* of $\mathcal{Z}(\cdot,\theta)$,

Lemma 6.2. *Assume (A1') (i.e. for all $\sigma\in\mathbb{R}^d$ and $\theta\in\mathbb{T}^1$, $\langle\sigma,\hat A(\cdot,\theta)\rangle$ is not an f_θ-coboundary). Then the following properties hold:*

(0) $\mathcal{Z}(\cdot,\theta)$ is a convex lower semi-continuous function; in particular $\mathbb{D}(\theta)$ is convex.
(1) let $\mathbb{D}_*(\theta) = \partial_\sigma\chi_A(\mathbb{R}^d,\theta)$; then $\mathbb{D}_*(\theta) = \operatorname{int}\mathbb{D}(\theta)$; in particular $\mathbb{D}_*(\theta)$ is convex;
(2) let $U = \{(b,\theta) : \theta\in\mathbb{T},\, b\in\mathbb{D}_*(\theta)\}$; $\mathcal{Z}\in\mathcal{C}^2(U,\mathbb{R}_{\geq 0})$ and it is analytic in b;
(3) $\mathbb{D}(\theta)$ contains a neighborhood of $\bar A(\theta)$;
(4) $\mathcal{Z}(\bar A(\theta),\theta) = 0$, $\partial_b\mathcal{Z}(\bar A(\theta),\theta) = 0$, $\partial_\theta\mathcal{Z}(\bar A(\theta),\theta) = 0$, and $\mathcal{Z}\geq 0$;
(5) $\partial_b^2\mathcal{Z}(b,\theta) > 0$, and setting $\left[\partial_b^2\mathcal{Z}(\bar A(\theta),\theta)\right]^{-1} = \Sigma^2(\theta)$ we have

$$\Sigma^2(\theta) = \mu_\theta\left(\hat A(\cdot,\theta)\otimes\hat A(\cdot,\theta)\right) + 2\sum_{m=1}^{\infty}\mu_\theta\left(\hat A(f_\theta^m(\cdot),\theta)\otimes\hat A(\cdot,\theta)\right).$$

Proof. Item (0) follows since, for each θ, $\mathcal{Z}(\cdot,\theta)$ is the (convex) conjugate function of a proper function, hence a convex lower semi-continuous function (see [48, Theorems 10.1, 12.2]).

Since χ_A is a strictly convex function, $\partial_\sigma\chi_A$ in an injective map and hence, by the theorem of Invariance of Domain, we conclude that $\mathbb{D}_*(\theta)$ is open. The equality $\operatorname{int}\mathbb{D}(\theta) = \mathbb{D}_*(\theta)$ follows then from [48, Theorem 23.4, Corollary 26.4.1]. We have thus proved item (1).

Observe now that if $b\in\mathbb{D}_*(\theta)$, then $\mathcal{Z}(b,\theta) = \kappa(\bar\sigma,b,\theta)$ where $\bar\sigma = \bar\sigma(b,\theta)$ is the unique solution of $b = \partial_\sigma\chi_A(\bar\sigma(b,\theta),\theta)$. Item (2) follows by the implicit function theorem and the perturbation theory results collected in Appendix A.2 and A.3.

By (A.11a) $\partial_\sigma\chi_A(\sigma,\theta) = \nu_{\theta,\langle\sigma,A\rangle}(A(\cdot,\theta))$, where $\nu_{\theta,\langle\sigma,A\rangle}$ is the invariant probability measure associated to the operator (5.7); in particular $\nu_{0,\theta} = \mu_\theta$. Then $\partial_\sigma\chi_A(0,\theta) = \bar A(\theta)$, which implies that $\bar A(\theta)\in\mathbb{D}_*(\theta)$, hence proving item (3).

Next, let us prove item (4). First $\mathcal{Z}(\bar A(\theta),\theta) = 0$ and $\mathcal{Z}(b,\theta)\geq -\hat\chi_A(0,\theta) = 0$. Also, a direct computation shows that

$$\partial_b\mathcal{Z}(b,\theta) = \bar\sigma(b,\theta); \tag{6.4}$$

in particular $\partial_b\mathcal{Z}(\bar A(\theta),\theta) = 0$. Next, $(\partial_\theta\mathcal{Z})(\bar A(\theta),\theta) = -\partial_\theta\hat\chi_A(0,\theta) = 0$ by (A.22a).

Finally, by [48, Theorem 26.5], $\partial_b^2\mathcal{Z}(b,\theta) = [\partial_\sigma^2\chi_A(a(b,\theta),\theta)]^{-1}$. This and (A.12a) imply item (5). □

Remark 6.3. *Arguing as in (A.11a) it follows that $\|\partial_\sigma\chi_A\|_\infty \leq \|A\|_\infty$, thus $\mathbb{D}_*(\theta)$ is uniformly (in θ) bounded. It follows that if $b\in\mathbb{D}_*(\theta)$ and σ_b is the solution of*

$$b = \partial_\sigma\chi_A(\sigma,\theta), \tag{6.5}$$

then

$$\mathcal{Z}(b,\theta) = \kappa(\sigma_b,b,\theta), \tag{6.6}$$

and $\partial_b\mathcal{Z}(b,\theta) = \sigma_b$.

Remark 6.4. *Using the above facts, it would be possible to show that $I_{\mathrm{pre},\theta^*}$ is lower semi-continuous with respect to the uniform topology. We refrain from proving it here because the proof will be given later in Lemma 6.11.*

We conclude this subsection with a useful estimate.

Lemma 6.5. *Fix $\theta \in \mathbb{T}$ and let $b, \sigma \in \mathbb{R}^d$ so that $b = \partial_\sigma \hat{\chi}_A(\sigma, \theta)$; then there exists $C_{**} > 1$ so that*

(a) $\|b\| \le C_{**}\|\sigma\|$
(b) $C_{**}^{-1} \min\{\|\sigma\|, \|\sigma\|^2\} \le \langle \sigma, b\rangle \le C_{**} \min\{\|\sigma\|, \|\sigma\|^2\}$.

Proof. The first item follows by the definition, equations (A.11) and the fact that $\partial_{\sigma^2}\hat{\chi}_A(\cdot, \theta)$ is bounded (as a quadratic form). We proceed to prove the second item. Let $\hat{\sigma} = \|\sigma\|^{-1}\sigma$; then, by definition:

$$\langle \sigma, b\rangle = \int_0^{\|\sigma\|} \langle \sigma, \partial_\sigma^2 \hat{\chi}_A(\lambda\hat{\sigma}, \theta)\hat{\sigma}\rangle d\lambda. \tag{6.7}$$

Moreover, (A.12b) and (A1) imply that $\partial_\sigma^2 \hat{\chi}_A(\cdot, \theta) \ge C_\#^{-1}\mathbb{1}$ (as quadratic forms) for $\|\sigma\| \le \sigma_\#$ for some $\sigma_\#$ sufficiently small. Observe that $\sigma_\#$ depends on f and A only. Hence by (6.7) we gather

$$\langle \sigma, b\rangle \ge C_\#^{-1} \int_0^{\min\{\|\sigma\|, \sigma_\#\}} \langle \sigma, \hat{\sigma}\rangle d\lambda \ge C_\#^{-1} \min\{\|\sigma\|, \|\sigma\|^2\},$$

which gives the lower bound.

On the other hand, since $b = \partial_\sigma \hat{\chi}_A(\sigma, \theta)$, we have $b \in \mathbb{D}_*(\theta)$; hence, by Remark 6.3, b is uniformly bounded and thus we obtain $\langle \sigma, b\rangle \le C_\# \|\sigma\|$.

Moreover, using once again (6.7) and since $\partial_\sigma^2 \hat{\chi}_A$ is locally bounded from above (see (A.12)), $\langle \sigma, b\rangle \le C_\# \|\sigma\|^2$, for all $\sigma \le \sigma_\#$, which concludes the proof. □

6.2. Entropy characterization.

As already mentioned, $\mathcal{Z}$ can also be expressed in terms of entropy (see e.g. [38]).

Lemma 6.6. *For any $\theta \in \mathbb{T}^1$ and $b \in \mathbb{R}^d$, let $\mathcal{M}_\theta(b) = \{\nu \in \mathcal{M}_\theta \,:\, \nu(A(\cdot, \theta)) = b\}$, where $\mathcal{M}_\theta$ denotes the set of f_θ-invariant probability measures. Then:*

$$\mathcal{Z}(b, \theta) = - \sup_{\nu \in \mathcal{M}_\theta(b)} \{h_{\mathrm{KS},\theta}(\nu) - \nu(\log f_\theta')\}, \tag{6.8}$$

where $h_{\mathrm{KS},\theta}(\nu)$ is the Kolmogorov-Sinai metric entropy of the measure ν with respect to the map f_θ. In particular[32]*, $\mathbb{D}(\theta) = \{b \in \mathbb{R}^d \,:\, \mathcal{M}_\theta(b) \ne \emptyset\}$.*

Proof. It is well known (see e.g. [5, Remark 2.5]), that

$$\begin{aligned} \chi_A(\sigma, \theta) &= \sup_{\nu \in \mathcal{M}_\theta} \{h_{\mathrm{KS},\theta}(\nu) + \nu(\langle \sigma, A\rangle - \log f_\theta')\} \\ &= h_{\mathrm{KS},\theta}(\nu_{\theta,\langle\sigma,A\rangle}) + \nu_{\theta,\langle\sigma,A\rangle}(\langle \sigma, A\rangle - \log f_\theta') \end{aligned} \tag{6.9}$$

where $\nu_{\theta,\langle\sigma,A\rangle}(g) = m_{\theta,\langle\sigma,A\rangle}(g\, h_{\theta,\langle\sigma,A\rangle})$ and $m_{\theta,\langle\sigma,A\rangle}$ and $h_{\theta,\langle\sigma,A\rangle}$ are respectively the left and right eigenvectors of $\mathcal{L}_{\theta,\langle\sigma,A\rangle}$ corresponding to the eigenvalue $e^{\chi_A(\sigma,\theta)}$, normalized so that $\nu_{\theta,\langle\sigma,A\rangle}$ is a probability measure. We record, for future use, some properties of the entropy: since each f_θ is expanding, $h_{\mathrm{KS},\theta}$ is (as a function of ν) an upper-semicontinuous function with respect to the weak topology (see [35, Theorem 4.5.6]). Also, $h_{\mathrm{KS},\theta}$ is a convex affine function[33] by [35, Theorem 3.3.2].

[32] Recall that we adopt the convention $\sup \emptyset = -\infty$.
[33] i.e. it is both convex and concave.

Incidentally, this implies that the sup in (6.9) would be the the same if taken only on ergodic measures, see [35, Theorem 4.3.7]. Then, using the definition (6.2):

$$\begin{aligned}\mathcal{Z}(b,\theta) &= \sup_{\sigma\in\mathbb{R}^d}\left\{\langle\sigma,b\rangle - \sup_{\nu\in\mathcal{M}_\theta}[h_{\mathrm{KS},\theta}(\nu)+\nu(\langle\sigma,A\rangle-\log f'_\theta)]\right\}\\ &\leq \sup_{\sigma\in\mathbb{R}^d}\left\{\langle\sigma,b\rangle - \sup_{\nu\in\mathcal{M}_\theta(b)}[h_{\mathrm{KS},\theta}(\nu)+\nu(\langle\sigma,A\rangle-\log f'_\theta)]\right\}\\ &\leq -\sup_{\nu\in\mathcal{M}_\theta(b)}\{h_{\mathrm{KS},\theta}(\nu)-\nu(\log f'_\theta)\}.\end{aligned} \tag{6.10}$$

In particular, the above implies that if $\mathcal{Z}(b,\theta)=\infty$, then (6.8) holds. We may thus assume that $\mathcal{Z}(b,\theta)<\infty$. Observe that:

$$\mathcal{Z}(b,\theta) = \sup_{\sigma\in\mathbb{R}^d}\left\{-h_{\mathrm{KS},\theta}(\nu_{\theta,\langle\sigma,A\rangle})+\nu_{\theta,\langle\sigma,A\rangle}(\langle\sigma,b-A\rangle)+\nu_{\theta,\langle\sigma,A\rangle}(\log f'_\theta)\right\}.$$

Note that the first term on the right hand side is bounded by the topological entropy [35, Theorem 4.2.3], while the last term is bounded because $f'_\theta>1$. Thus, since Lemma 6.2(4) implies that $\mathcal{Z}\geq 0$ and we assume $\mathcal{Z}(b,\theta)<\infty$, we conclude that

$$\sup_{\sigma\in\mathbb{R}^d}|\nu_{\theta,\langle\sigma,A\rangle}(\langle\sigma,b-A\rangle)|<\infty. \tag{6.11}$$

For any $\lambda\in\mathbb{R}$ consider the function $K_\lambda\in\mathcal{C}^0(\mathbb{R}^d,\mathbb{R}^d)$ defined by $K_\lambda(\sigma)=\nu_{\theta,\lambda\langle\sigma,A\rangle}(b-A)$. Since $\nu_{\theta,\lambda\langle\sigma,A\rangle}$ is a probability measure, we have $K_\lambda(\mathbb{R}^d)\subset B=\{x\in\mathbb{R}^d : \|x\|\leq\|b-A\|_\infty\}$. By Brouwer fixed-point theorem it follows that there exists $\sigma_\lambda\in B$ such that $K_\lambda(\sigma_\lambda)=\sigma_\lambda$. Accordingly, for any non-negative sequence (λ_j) with $\lambda_j\to+\infty$:

$$\langle\lambda_j\sigma_{\lambda_j},\nu_{\theta,\langle\lambda_j\sigma_{\lambda_j},A\rangle}(b-A)\rangle=\lambda_j\|\nu_{\theta,\langle\lambda_j\sigma_{\lambda_j},A\rangle}(b-A)\|^2\geq 0.$$

Since the left hand side is bounded, see (6.11), $\lim_{j\to\infty}\nu_{\theta,\langle\lambda_j\sigma_{\lambda_j},A\rangle}(b-A)=0$. By passing to a subsequence $\{j_k\}$ we can assume, setting $\bar\sigma_k=\lambda_{j_k}\sigma_{\lambda_{j_k}}$, that $\nu_{\theta,\langle\bar\sigma_k,A\rangle}$ weakly converges to a measure ν_*. Moreover, for any $k\in\mathbb{N}$

$$\mathcal{Z}(b,\theta)\geq -h_{\mathrm{KS},\theta}(\nu_{\theta,\langle\bar\sigma_k,A\rangle})+\nu_{\theta,\langle\bar\sigma_k,A\rangle}(\log f'_\theta).$$

Since $h_{\mathrm{KS},\theta}$ is upper-semicontinuous, we conclude that

$$\mathcal{Z}(b,\theta)\geq -\limsup_{k\to\infty}\left[h_{\mathrm{KS},\theta}(\nu_{\theta,\langle\bar\sigma_k,A\rangle})-\nu_{\theta,\langle\bar\sigma_k,A\rangle}(\log f'_\theta)\right]\geq -h_{\mathrm{KS},\theta}(\nu_*)+\nu_*(\log f'_\theta).$$

Finally, notice that $\nu_*\in\mathcal{M}_\theta$ and $\nu_*(b-A)=0$, hence $\nu_*\in\mathcal{M}_\theta(b)$. Thus we have

$$\mathcal{Z}(b,\theta)\geq -\sup_{\nu\in\mathcal{M}_\theta(b)}\{h_{\mathrm{KS},\theta}(\nu)-\nu(\log f'_\theta)\},$$

which together with (6.10) concludes the proof of the lemma. □

The entropy characterization allows to add two useful properties to those listed in Lemma 6.2.

Lemma 6.7. *The following properties hold:*

- *(6) $\mathbb{D}(\theta)$ is a compact set for all $\theta\in\mathbb{T}$;*
- *(7) The map $\theta\mapsto\mathbb{D}(\theta)$ is Lipschitz in the Hausdorff metric.*

Proof. If $\{b_n\}\subset\mathbb{D}(\theta)$, then there exists $\{\nu_n\}\subset\mathcal{M}_\theta$ such that $\nu_n(A)=b_n$. Since $\mathcal{M}_\theta$ is compact in the weak topology, by extracting a convergent subsequence, item (6) follows.

To prove item (7), note that all the maps f_θ are topologically conjugated to f_0 by a homeomorphism $\xi(\cdot,\theta)=\xi_\theta(\cdot)$ with the property[34]

$$\|\xi_\theta-\xi_{\theta'}\|_{\mathcal{C}^0}+\|\xi_\theta^{-1}-\xi_{\theta'}^{-1}\|_{\mathcal{C}^0}\leq C_\#|\theta-\theta'|.$$

Accordingly, using the notation of Lemma 6.6, $\mathcal{M}(\theta)=(\xi_\theta)_*\mathcal{M}(0)$. Hence, for each $\theta,\theta'\in\mathbb{T}$ and $b\in\mathbb{D}(\theta)$ there exist $\nu\in\mathcal{M}(0)$ such that $b=\nu(\bar A(\xi_\theta(\cdot),\theta))$ and $b'=\nu(\bar A(\xi_{\theta'}(\cdot),\theta'))\in\mathbb{D}(\theta')$. Then

$$\|b-b'\|\leq\|\nu(\bar A(\xi_\theta(\cdot),\theta))-\nu(\bar A(\xi_{\theta'}(\cdot),\theta'))\|\leq C_\#|\theta-\theta'|.$$

Thus b must belong to a $C_\#|\theta-\theta'|$ neighborhood of $\mathbb{D}(\theta')$ and exchanging the role of θ,θ', the item follows. □

Lemma 6.6 allows to specify exactly the effective domain of I_{pre,θ^*}:

$$\mathfrak{D}(I_{\text{pre},\theta^*})=\{\gamma\in C([0,T])\ :\ I_{\text{pre},\theta^*}(\gamma)<\infty\}.$$

In fact $I_{\text{pre},\theta^*}(\gamma)<\infty$ if and only if $\gamma(0)=0$, γ is Lipschitz and $\mathcal{M}_{\bar\theta(t,\theta^*)}(\gamma'(t))\neq\emptyset$ for almost all $t\in[0,T]$. Having fixed $\theta^*\in\mathbb{T}$, we will call *s-admissible* the paths such that $\gamma\in\mathfrak{D}(I_{\text{pre},\theta^*})$. To use effectively this definition, it would be convenient if one could characterize s-admissibility in terms of periodic orbits. To this end, given a periodic orbit p, let ν_p the measure determined by the average along the orbit of p.

Lemma 6.8. *Given $\theta\in\mathbb{T}$ and $b\in\mathbb{R}^d$, $b\in\operatorname{int}\mathbb{D}(\theta)$ if and only if there exist $d+1$ periodic orbits $\{p_i\}$ of f_θ such that the convex hull of $\nu_{p_i}(b-A(\cdot,\theta))$ contains a neighborhood of zero. Also if there exists $n\in\mathbb{N}$ such that*

$$\inf_{x\in\mathbb{T}}\left|\langle b,b-\frac{1}{n}\sum_{k=0}^{n-1}A(f_\theta^k(x),\theta)\rangle\right|>0,$$

then $b\notin\mathbb{D}(\theta)$.

Proof. If the convex hull contains a neighborhood of zero, then there exists $\delta>0$ such that, for all $b'\in\mathbb{R}^d$, $\|b-b'\|<\delta$, there exists $\{\alpha_i\}_{i=1}^{d+1}\subset\mathbb{R}_{\geq 0}$, $\sum_{i=1}^{d+1}\alpha_i=1$ such that $\sum_{i=1}^{d+1}\alpha_i\nu_{p_i}(A(\cdot,\theta))=b'$, hence $b'\in\mathbb{D}(\theta)$ and $b\in\mathbb{D}_*(\theta)$. On the other hand if $b\in\mathbb{D}_*(\theta)$ then there are $\{b_i\}_{i=1}^{d+1}\subset\mathbb{D}_*(\theta)$ such that b belongs to the interior of their convex hull. Hence there exists $\nu_i\in\mathcal{M}_\theta(b_i)$ such that their convex combination gives an element of $\mathcal{M}_\theta(b)$. Since the measures supported on periodic orbits are weakly dense in $\mathcal{M}_\theta$ (see[35] [46]) it is possible to find periodic orbits $\{p_i\}$ such that the convex hull of $\nu_{p_i}(A(\cdot,\theta))$ contains a neighborhood of b, hence the necessity of the condition.

To prove the other necessary condition, note that, by (A.11a), (6.5) reads

$$b=\nu_{\theta,\langle\sigma,A\rangle}(A(\cdot,\theta)). \tag{6.12}$$

Thus $b\in\mathbb{D}_*(\theta)$ if and only if (6.12) has a solution. If the second condition in the lemma is satisfied, then for each invariant measure ν we have $|\langle b,b-\nu(A(\cdot,\theta))\rangle|>0$, hence equation (6.12) cannot be satisfied. Moreover, the same conclusion holds for any b' in a small neighborhood of b, hence the claim. □

[34] This is folklore, e.g. it can be proven using shadowing and keeping track of the constants.

[35] In fact the proof in [46] is for the invertible case but it applies almost verbatim to the present one.

6.3. **An equivalent definition.**
Unfortunately, in our subsequent discussion, the rate function will appear first in a much less transparent form, a priori different from the definition (6.3). Namely, recall the notation (6.1); then for any $\sigma \in \mathbb{R}^d$ and any Lipschitz path $\gamma \in \mathcal{C}^0([0,T],\mathbb{R}^d)$, we introduce the shorthand notation

$$\kappa_{\gamma,\theta^*}(\sigma,s) = \kappa(\sigma,\gamma'(s),\bar\theta(s,\theta^*)). \tag{6.13}$$

Remark 6.9. *For further use remark that Lemmata 6.10, 6.11 and 6.12 hold verbatim if in the above definition of $\kappa_{\gamma,\theta^*}(\sigma,s)$ one substitutes $\bar\theta(s,\theta^*)$ with some other continuous function of s.*

Also let us fix $C \geq 2\|A\|_{\mathcal{C}^0}$ and define:

$$\mathrm{Lip}_{C,*} = \{\gamma \in \mathcal{C}^0([0,T],\mathbb{R}^d) : \ \gamma(0) = 0, \|\gamma(t)-\gamma(s)\| \leq C|t-s| \ \forall\, t,s \in [0,T]\},$$

Then, the functional $I_{\theta^*} : \mathcal{C}^0([0,T],\mathbb{R}^d) \to \mathbb{R}\cup\{+\infty\}$ will appear naturally, where

$$I_{\theta^*}(\gamma) = \begin{cases} +\infty & \text{if } \gamma \notin \mathrm{Lip}_{C,*} \\ \sup\limits_{\sigma\in\mathrm{BV}} \int_0^T \kappa_{\gamma,\theta^*}(\sigma(s),s)ds & \text{otherwise.}\end{cases} \tag{6.14}$$

It is the task of this subsection to show that the two definitions (6.3) and (6.14) coincide. At a superficial level, it amounts to prove that we can bring the sup inside the integral. This will be proved essentially via a compactness argument.

First, observe that I_{θ^*} is convex, because it is the conjugate function of a proper function. Moreover, $I_{\theta^*} \geq 0$ (just consider $\sigma = 0$ in the sup) and since $\hat\chi_A \geq 0$ we obtain $I_{\theta^*}(\bar\gamma(\cdot,\theta^*)) = 0$, where recall that $\bar\gamma(s,\theta)$ (defined in (2.9)) satisfies the equation $\bar\gamma'(t,\theta) = \bar A(\bar\theta(t,\theta))$. Our first task is to show that we can replace the sup on $\sigma \in \mathrm{BV}$ with the sup on $\sigma \in L^1$.

Lemma 6.10. *Let $\gamma \in \mathrm{Lip}_{C,*}$; then:*

$$I_{\theta^*}(\gamma) = \sup_{\sigma\in L^1} \int_0^T \kappa_{\gamma,\theta^*}(\sigma(s),s)ds.$$

Proof. First, notice that (A.11a) implies that $\|\partial_\sigma\bar\kappa_{\gamma,\theta^*}(\sigma,s)\| \leq C_\#(C+1)$ (and consequently $\|\kappa_{\gamma,\theta^*}(\sigma,s)\| \leq C_\#(C+1)\|\sigma\|$) for all $s \in [0,T]$. It follows that, for all $\gamma \in \mathrm{Lip}_{C,*}$ the functional $\sigma \mapsto \int_0^T \kappa_{\gamma,\theta^*}(\sigma(s),s)ds$ is continuous in the L^1 topology.

Let $\sigma \in L^1$; since BV is dense in L^1, [40, Theorem 2.16], for any $\epsilon > 0$ there exists $\sigma_\epsilon \in \mathrm{BV}$ such that $\|\sigma-\sigma_\epsilon\|_{L^1} < \epsilon$ and thus

$$\begin{aligned}\int_0^T \kappa_{\gamma,\theta^*}(\sigma(s),s)ds \leq{}& C_\#\epsilon + \int_0^T \kappa_{\gamma,\theta^*}(\sigma_\epsilon(s),s)ds \\ &\leq C_\#\epsilon + \sup_{\bar\sigma\in\mathrm{BV}} \int_0^T \kappa_{\gamma,\theta^*}(\bar\sigma(s),s)ds.\end{aligned}$$

Taking the limit $\epsilon \to 0$ first and then sup on $\sigma \in L^1$ we have that the sup on BV equals the sup on L^1, proving the lemma. □

Lemma 6.11. *The functional I_{θ^*} is lower semi-continuous on $\mathcal{C}^0([0,T],\mathbb{R}^d)$.*

Proof. Consider a sequence $\{\gamma_n\} \subset \mathcal{C}^0([0,T],\mathbb{R}^d)$ converging uniformly to γ. If $\liminf_{n\to\infty} I_{\theta^*}(\gamma_n) = +\infty$, then obviously $\liminf_{n\to\infty} I_{\theta^*}(\gamma_n) \geq I_{\theta^*}(\gamma)$. Otherwise, there exists a subsequence $\{\gamma_{n_j}\}$, $M > 0$ and $j_0 \in \mathbb{N}$ such that

$$\liminf_{n\to\infty} I_{\theta^*}(\gamma_n) = \lim_{j\to\infty} I_{\theta^*}(\gamma_{n_j}),$$

and $I_{\theta^*}(\gamma_{n_j}) \le M$ for all $j \ge j_0$. This implies that if $j \ge j_0$, then $\gamma_{n_j} \in \mathrm{Lip}_{C,*}$; hence, we also conclude that $\gamma \in \mathrm{Lip}_{C,*}$. This implies that, for any $\sigma \in L^1$,

$$\lim_{j\to\infty} \int_0^T \langle \sigma, \gamma'_{n_j}\rangle = \int_0^T \langle \sigma, \gamma'\rangle.$$

In fact, for any $\epsilon > 0$ there exists $\sigma_\epsilon \in \mathcal{C}^1$, such that $\|\sigma - \sigma_\epsilon\|_{L^1} \le \epsilon$, [40, Theorem 2.16]. Then

$$\left|\int_0^T \langle \sigma, \gamma'_{n_j}\rangle - \int_0^T \langle \sigma, \gamma'\rangle\right| \le 2C\epsilon + \left|\int_0^T \langle \sigma'_\epsilon, \gamma_{n_j} - \gamma\rangle\right| + |\langle \sigma_\epsilon(T), \gamma_{n_j}(T) - \gamma(T)\rangle|.$$

We conclude that, for any $\sigma \in L^1$,

$$\begin{aligned}\liminf_{n\to\infty} I_{\theta^*}(\gamma_n) &\ge \lim_{j\to\infty} \int_0^T \left[\langle \sigma(s), \gamma'_{n_j}(s))\rangle - \chi_A(\sigma(s), \bar\theta(s,\theta^*))\right] ds \\ &= \int_0^T \left[\langle \sigma(s), \gamma'(s)\rangle - \chi_A(\sigma(s), \bar\theta(s,\theta^*))\right] ds.\end{aligned}$$

The proof follows by taking the sup on σ. □

We can finally show that the definition of I_{θ^*} given in the current section coincides with (6.3).

Lemma 6.12. *For any $\theta^* \in \mathbb{T}$, let $\gamma \in \mathcal{C}^0([0,T],\mathbb{R}^d)$, then:*

$$I_{\theta^*}(\gamma) = I_{\mathrm{pre},\theta^*}(\gamma).$$

Proof. If $\gamma(0) \ne 0$ or γ is not Lipschitz, we have $I_{\mathrm{pre},\theta^*}(\gamma) = \infty = I_{\theta^*}(\gamma)$; we can thus assume γ to be a Lipschitz function so that $\gamma(0) = 0$. Recall that in this case

$$I_{\mathrm{pre},\theta^*}(\gamma) = \int_0^T \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta^*))ds.$$

If $\gamma \notin \mathrm{Lip}_{C,*}$, then, provided C has been chosen large enough, there is a positive measure set in which $\gamma'(t) \notin \mathbb{D}(\bar\theta(t,\theta^*))$ and hence $\int_0^T \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta^*))ds = \infty$, which coincides with I_{θ^*}. We can then assume $\gamma \in \mathrm{Lip}_{C,*}$.

Observe that by definition we have $I_{\theta^*}(\gamma) \le I_{\mathrm{pre},\theta^*}(\gamma)$; it just suffices to prove the reverse inequality.

Suppose first that $I_{\mathrm{pre},\theta^*}(\gamma) = \infty$: we want to show that $I_{\theta^*}(\gamma) = \infty$. Define $z_*(s) = \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta^*))$; by assumption $z_* \notin L^1[0,T]$. Let us fix arbitrarily $M > 0$; by Lusin Theorem and Lebesgue monotone convergence Theorem there exists $\lambda > 0$ and a compact set E such that γ' and $\min\{\lambda, z_*(t)\}$ are continuous on E and $\int_E \min\{\lambda, z_*(t)\}dt \ge M$. Then for $t \in E$ let $\sigma_\lambda(t)$ be such that $\kappa_{\gamma,\theta^*}(\sigma_\lambda(t), t) \ge \frac{1}{2}\min\{\lambda, z_*(t)\}$. Since $\kappa_{\gamma,\theta^*}(\sigma_\lambda(t), s)$ is continuous in $s \in E$, it follows that, for all $t \in E$, there exists and open set $U(t) \ni t$ such that $\kappa_{\gamma,\theta^*}(\sigma_\lambda(t), s) \ge \frac{1}{4}\min\{\lambda, z_*(s)\}$ for all $s \in U(t) \cap E$. We can then extract a finite sub cover $\{U(t_i)\}$ of E and define

$$\bar\sigma_\lambda(s) = \begin{cases} \sigma_\lambda(t_{k(s)}) & \text{if } s \in E, \text{ where } k(s) = \inf\{i : s \in U(t_i)\} \\ 0 & \text{if } s \notin E. \end{cases}$$

By construction $\bar\sigma_\lambda \in L^\infty$ and $\kappa_{\gamma,\theta^*}(\bar\sigma_\lambda(t), t) \ge \frac{1}{4}\min\{\lambda, z_*(t)\}$ for each $t \in E$. Accordingly, setting $z_\lambda(t) = \mathbb{1}_E(t) \cdot \min\{\lambda, z_*(t)\}$, by Lemma 6.10 we have

$$I_{\theta^*}(\gamma) \ge \int_0^T \kappa_{\gamma,\theta^*}(\bar\sigma_\lambda(s), s)ds \ge \frac{1}{4}\int_0^T z_\lambda(s)ds \ge \frac{M}{4}.$$

By the arbitrariness of M it follows $I_{\theta^*}(\gamma) = +\infty$.

On the other hand, if $z_* \in L^1$, then by Lemma 6.10

$$I_{\theta^*}(\gamma) = \sup_{\sigma \in L^1} \int_0^T \kappa(\sigma(s), \gamma'(s), \bar\theta(s,\theta^*))ds \le \int_0^T z_*(s)ds < +\infty, \tag{6.15}$$

and $\gamma'(s) \in \mathbb{D}(\bar\theta(s,\theta^*))$ for almost every $s \in [0,T]$. For $\varrho \in (0,1)$ and $s \in [0,T]$ let us define the convex combination

$$\gamma_\varrho(s) = (1-\varrho)\gamma(s) + \varrho\bar\gamma(s,\theta^*).$$

Since $\bar\gamma'(s,\theta^*) = A(\bar\theta(s,\theta^*)) \in \operatorname{int}\mathbb{D}(\bar\theta(s,\theta^*))$ it follows that, for any $\varrho \in (0,1)$ and $s \in [0,T]$, there exists a compact set $K(\varrho,s) \subset \mathbb{D}_*(\bar\theta(s,\theta))$, such that $\gamma_\varrho'(s) \in K(\varrho,s)$ for almost all $s \in [0,T]$. By Lemma 6.7 such compacts depend continuously on s. Since the inverse of $\partial_\sigma\chi_A(\cdot,\theta)$ is a continuous function with depends continuously on θ, it follows that the preimages of $K(\varrho,s)$ are all contained in a fixed compact set K_ϱ. Hence, there exists $\sigma_\varrho \in L^\infty$ such that $\gamma_\varrho'(s) = \partial_\sigma\chi_A(\sigma_\varrho(s), \bar\theta(s,\theta^*))$ for almost all $s \in [0,T]$.

$$\begin{aligned}
I_{\theta^*}(\gamma) &= \lim_{\varrho\to 0}(1-\varrho)I_{\theta^*}(\gamma) \ge \liminf_{\varrho\to 0} I_{\theta^*}(\gamma_\varrho) \ge \liminf_{\varrho\to 0}\int_0^T \kappa(\sigma_\varrho(s), \gamma_\varrho'(s), \bar\theta(s,\theta^*))ds\\
&= \liminf_{\varrho\to 0}\int_0^T \mathcal{Z}(\gamma_\varrho'(s), \bar\theta(s,\theta^*))ds \ge \int_0^T \liminf_{\varrho\to 0}\mathcal{Z}(\gamma_\varrho'(s), \bar\theta(s,\theta^*))ds\\
&\ge \int_0^T \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta^*))ds = I_{\text{pre},\theta^*}(\gamma),
\end{aligned}$$

where we have used the convexity of I_{θ^*} first, then Lemma 6.10, then equation (6.6), then Fatou's Lemma and finally Lemma 6.2(0). The above concludes the proof. □

6.4. **Definition and properties: the rate function.**

Lemma 6.8 tells us that it might be difficult to exactly determine the boundary of the effective domain of the rate function, i.e. to distinguish between impossible and almost impossible paths. To circumvent this problem it is convenient to thicken the boundary of the effective domain by slightly modifying the rate function.

For any $\epsilon > 0$ let $\partial_\epsilon\mathbb{D}(\theta) = \{b \in \mathbb{R}^d : \operatorname{dist}(b, \partial\mathbb{D}(\theta)) < \epsilon\}$ and define

$$\begin{aligned}
\mathcal{Z}_\epsilon^+(b,\theta) &= \begin{cases}\mathcal{Z}(b,\theta) & \text{if } b \notin \partial_\epsilon\mathbb{D}(\theta)\\ +\infty & \text{otherwise,}\end{cases}\\
\mathcal{Z}_\epsilon^-(b,\theta) &= \begin{cases}\mathcal{Z}(b,\theta) & \text{if } b \notin \overline{\partial_\epsilon\mathbb{D}(\theta)}\\ \mathcal{Z}(\bar A(\theta) + \varrho_{b,\theta}(b - \bar A(\theta)), \theta) & \text{otherwise,}\end{cases}
\end{aligned} \tag{6.16}$$

where $\varrho_{b,\theta} = \sup\{\varrho > 0 : \bar A(\theta) + \varrho(b - \bar A(\theta)) \in \mathbb{D}(\theta)\setminus\partial_\epsilon\mathbb{D}(\theta)\}$. We will conveniently assume that ϵ is so small that for any $\theta \in \mathbb{T}^1$ and $b \in \partial_\epsilon\mathbb{D}(\theta)$ we have $\varrho_{b,\theta} > 1/2$. Note that, by Lemma 6.8, the set $\partial_\epsilon\mathbb{D}(\theta)$ can be explicitly determined for arbitrarily small ϵ by computing longer and longer periodic orbits and ergodic averages (see [11] for a discussions on the speed of such approximation). Moreover, for any $\epsilon' < \epsilon$ we have:

$$\mathcal{Z}_\epsilon^- < \mathcal{Z}_{\epsilon'}^- < \mathcal{Z} < \mathcal{Z}_{\epsilon'}^+ < \mathcal{Z}_\epsilon^+.$$

Remark 6.13. *Note that Lemmata 6.2 and 6.6 show that $\mathbb{D}(\theta)$ is a convex compact non-empty set (in fact, (6.8) implies $\sup_{\theta\in\mathbb{T}^1}\sup_{b\in\mathbb{D}(\theta)}\mathcal{Z}(b,\theta) < \infty$). Moreover, they characterize $\partial\mathbb{D}(\theta)$ as those values that can be attained as averages of A with respect to an invariant measure which is not associated to a transfer operator of type (5.7). Hence, again by Lemma 6.2, there exists $\Sigma^+ > \Sigma^- > 0$ (as quadratic* *forms) such that, for any ϵ small enough,*

$$\mathcal{Z}_\epsilon^+(b,\theta) \geq \langle b - \bar{A}(\theta), \Sigma^-(b - \bar{A}(\theta))\rangle$$
$$\mathcal{Z}_\epsilon^-(b,\theta) \leq \langle b - \bar{A}(\theta), \Sigma^+(b - \bar{A}(\theta))\rangle.$$

The rate function defined in the previous sections would suffice to describe deviations from the average behavior for relatively short times. If we want to study longer times, then we must consider slightly different rate functions. That is, for any $\theta^* \in \mathbb{T}$ and $\gamma \in \mathcal{C}^0([0,T];\mathbb{R}^d)$, let $\theta^\gamma(s,\theta^*) = \theta^* + (\gamma(s))_1$ (where recall that $(\gamma(s))_1$ denotes the first component of the vector $\gamma(s)$). Then for any $\epsilon > 0$:

$$\begin{aligned} &\mathscr{I}_{\theta^*}(\gamma) = \begin{cases} +\infty & \text{if } \gamma \notin \text{ Lipschitz, or } \gamma(0) \neq 0 \\ \int_0^T \mathcal{Z}(\gamma'(s), \theta^\gamma(s,\theta^*))\, ds & \text{otherwise;} \end{cases} \\ &\mathscr{I}^\pm_{\theta^*,\epsilon}(\gamma) = \begin{cases} +\infty & \text{if } \gamma \notin \text{ Lipschitz, or } \gamma(0) \neq 0 \\ \int_0^T \mathcal{Z}^\pm_\epsilon(\gamma'(s), \theta^\gamma(s,\theta^*))\, ds & \text{otherwise.} \end{cases} \end{aligned} \tag{6.17}$$

Finally, we define

$$\mathscr{I}^\pm_{\theta^*}(\gamma) = \lim_{\epsilon\to 0} \mathscr{I}^\pm_{\theta^*,\epsilon}(\gamma). \tag{6.18}$$

We stress the important difference with the definition of I_θ which comes from the fact that the function $\mathcal{Z}$ is now calculated along the actual path θ^γ rather than the averaged path $\bar{\theta}$.

Remark 6.14. *Note that, in general, $\mathscr{I}_{\theta^*}$ is not convex. The effective domain $\mathfrak{D}(\mathscr{I}_{\theta^*}) = \{\gamma \in \mathcal{C}^0 \ : \ \mathscr{I}_{\theta^*}(\gamma) < \infty\}$ is given by those paths γ that are C-Lipschitz, $\gamma(0) = 0$ and such that $\mathcal{M}_{\gamma_1(t)}(\gamma'(t)) \neq \emptyset$ for almost all $t \in [0,T]$. By lower semicontinuity (which we prove shortly) $\mathfrak{D}(\mathscr{I}_{\theta^*})$ is closed, and hence compact by Ascoli-Arzelà, in $\mathcal{C}^0$; however it has empty interior. It is therefore more convenient to consider $\mathfrak{D}(\mathscr{I}_{\theta^*})$ as a subset of the Lipschitz functions with the associated topology. Then the interior and the boundary are non trivial and this is the topology we will always consider for the effective domains otherwise differently stated. The effective domain of $\mathscr{I}^+_{\theta^*}$ is given by $\operatorname{int}\mathfrak{D}(\mathscr{I}_{\theta^*})$. If $\gamma \in \mathfrak{D}(\mathscr{I}^+_{\theta^*})$ we say that γ is* admissible. *Note that the two functionals only differ on $\partial\mathfrak{D}(\mathscr{I}_{\theta^*})$.*

Lemma 6.15. *For any $\theta^* \in \mathbb{T}$ and ϵ sufficiently small, the rate functions $\mathscr{I}_{\theta^*}, \mathscr{I}^-_{\theta^*,\epsilon}$ are lower-semicontinuous and $\mathscr{I}_{\theta^*} = \mathscr{I}^-_{\theta^*}$. Also, $\partial\mathfrak{D}(\mathscr{I}_{\theta^*})$ has empty interior and, for each $\gamma \in \partial\mathfrak{D}(\mathscr{I}_{\theta^*})$, there exists a sequence $\{\gamma_n\} \subset \operatorname{int}\mathfrak{D}(\mathscr{I}_{\theta^*})$ such that $\lim_{n\to\infty} \mathscr{I}_{\theta^*}(\gamma_n) = \mathscr{I}_{\theta^*}(\gamma)$.*

Proof. Let $\{\gamma_n\} \in \mathcal{C}^0([0,T],\mathbb{R}^d)$ be a converging sequence and call γ_* its limit. We start proving that:

$$\liminf_{n\to\infty} \mathscr{I}_{\theta^*}(\gamma_n) \geq \mathscr{I}_{\theta^*}(\gamma_*).$$

If the left hand side equals $+\infty$, then the statement is obviously true; if not, then there exists a subsequence such that $\{\gamma_{n_j}\} \subset \text{Lip}_{C,*}$, hence $\gamma_* \in \text{Lip}_{C,*}$. From now on the proof follows very closely the argument in Section 6.3 (recall Remark 6.9): for $\gamma \in \text{Lip}_{C,*}$ define

$$\mathscr{I}_{\theta_*,\text{pre}}(\gamma) = \sup_{\sigma\in L^1} \int_0^T \langle \sigma(s), \gamma'(s)\rangle - \chi_A(\sigma(s), \theta^\gamma(s))ds$$

Then arguing as in Lemma 6.11 it follows that $\mathscr{I}_{\theta_*,\text{pre}}$ is lower semicontinuous. The only difference being in the last display of the proof, since now the second argument of χ_A depends on γ, which can be controlled using Lemma A.10. Then the equivalent of Lemma 6.12 holds verbatim whereby establishing $\mathscr{I}_{\theta_*,\text{pre}} = \mathscr{I}_{\theta_*}$. The argument for $\mathscr{I}^-_{\theta^*,\epsilon}$ (for arbitrary $\epsilon > 0$) is more of the same.

Next, by the convexity of $\mathcal{Z}(\cdot,\theta)$, $\mathscr{I}^-_{\theta^*,\epsilon} \leq \mathscr{I}^-_{\theta^*}$ and $\mathscr{I}^-_{\theta^*,\epsilon}(\gamma) = \int_0^T \mathcal{Z}(\gamma'_\epsilon, \theta^\gamma)$, where $\lim_{\epsilon\to 0} \|\gamma'_\epsilon - \gamma'\|_{\mathcal{C}^0} = 0$. Fatou Lemma and the lower semicontinuity of $\mathcal{Z}(\cdot,\theta)$ then imply that $\mathscr{I}_{\theta^*} = \mathscr{I}^-_{\theta^*}$.

Further, suppose $\gamma \in \partial\mathfrak{D}(\mathscr{I}_{\theta^*})$. Fix an arbitrary $\delta > 0$ and, for all $\alpha > 1$ we define the path γ_α as the unique solution of the ODE

$$\begin{aligned}\gamma'_\alpha(s) &= (1-\delta e^{-\alpha(T-s)})\gamma'(s) + \delta e^{-\alpha(T-s)}\bar{A}(\theta^{\gamma_\alpha}(s))\\ \gamma_\alpha(0) &= 0.\end{aligned}$$

Then $\|\gamma_\alpha - \gamma\|_{\mathcal{C}^0} \leq 2C\delta\alpha^{-1}e^{-\alpha(T-s)}$ and $\|\gamma'_\alpha - \gamma'\|_{\mathcal{C}^0} \leq 2C\delta e^{-\alpha(T-s)}$. Also, by Lemma 6.7,

$$d(\gamma_\alpha, \partial\mathbb{D}(\theta^{\gamma_\alpha}(s))) \geq d(\gamma_\alpha, \partial\mathbb{D}(\theta^{\gamma}(s)) - c_4 2C\delta\alpha^{-1}e^{-\alpha(T-s)}.$$

On the other hand by Lemma 6.2-(3) and Lemma 6.7 the distance between $\bar{A}(\theta^\gamma(s))$ and $\partial\mathbb{D}(\theta^\gamma(s))$ is continuous and strictly positive, hence it has a minimum $\tau > 0$. Accordingly, provided δ is small enough,[36]

$$\begin{aligned}d(\gamma_\alpha, \partial\mathbb{D}(\theta^\gamma(s))) \geq & d((1-\delta e^{-\alpha(T-s)})\gamma'(s) + \delta e^{-\alpha(T-s)}\bar{A}(\theta^\gamma(s)), \partial\mathbb{D}(\theta^\gamma(s)))\\ & - C_\# \delta^2\alpha^{-1}e^{-\alpha(T-s)} \geq C_\#\tau\delta e^{-\alpha(T-s)}.\end{aligned}$$

Thus, by choosing α large enough, we have $d(\gamma_\alpha, \partial\mathbb{D}(\theta^{\gamma_\alpha}(s))) \geq c_T\delta$. Accordingly, $\gamma_\alpha \in \mathfrak{D}(\mathscr{I}_{\theta^*}) \setminus \partial\mathfrak{D}(\mathscr{I}_{\theta^*})$, thus $\operatorname{int}\partial\mathfrak{D}(\mathscr{I}_{\theta^*}) = \emptyset$.

Finally, by (6.6), (6.1), Lemma A.10 and Lemma 6.5, setting $\lambda(s) = \delta e^{-\alpha(T-s)}$ and choosing α large enough, we have

$$\begin{aligned}\mathscr{I}_{\theta^*}(\gamma_\alpha) &= \int_0^T \langle\sigma_\alpha, \hat{\gamma}'_\alpha\rangle - \hat{\chi}_A(\sigma_\alpha, \theta^{\gamma_\alpha})\\ &\leq \int_0^T (1-\lambda)\langle\sigma_\alpha, \hat{\gamma}'\rangle - \hat{\chi}_A(\sigma_\alpha, \theta^\gamma) + C_\# \min\{\|\sigma_\alpha^2\|, \|\sigma_\alpha\|\}\,\|\gamma - \gamma_\alpha\|\\ &\leq \int_0^T \langle\sigma_\alpha, \hat{\gamma}'\rangle - \hat{\chi}_A(\sigma_\alpha, \theta^\gamma) - C_\# \min\{\|\sigma_\alpha^2\|, \|\sigma_\alpha\|\}\,\left[\lambda - C_\#\lambda\alpha^{-1}\right]\\ &\leq \sup_{\sigma\in L^1}\int_0^T \langle\sigma, \gamma'\rangle - \chi_A(\sigma, \theta^\gamma) = \mathscr{I}_{\theta^*}(\gamma).\end{aligned}$$

The then lemma follows by the lower semicontinuity of $\mathscr{I}_{\theta^*}$. □

We conclude this section with a useful estimate:

Lemma 6.16. *For any $\theta \in \mathbb{T}$ and $\gamma \in \operatorname{Lip}_{C,*}([0,T], \mathbb{R}^d)$:*

$$\mathscr{I}^\pm_{\theta,\epsilon}(\gamma) \geq C_\#\|\gamma' - \bar{A}(\theta^\gamma(\cdot,\theta))\|^2_{L^2}. \tag{6.19}$$

Proof. Let us fix γ and introduce the shorthand notation $\hat{\chi}_A(\sigma) = \hat{\chi}_A(\sigma(\cdot), \theta^\gamma(\cdot))$; let $\check{\gamma}$ be so that $\check{\gamma}' = \gamma'(s) - \bar{A}(\theta^\gamma(s))$. Observe that if $\mathscr{I}^\pm_{\theta,\epsilon}(\gamma) = \infty$ then the statement trivially holds; hence let us assume that this is not the case and fix $s \in [0,T]$ so that $\mathcal{Z}^-_\epsilon(\gamma'(s), \theta^\gamma(s)) < \infty$. Then, for any $\varrho \in [0, 1/2)$, by the definition (6.16) of $\mathcal{Z}^-_\epsilon$ and the smallness condition on ϵ, we can define $\bar{\sigma}_\varrho(s)$ to be the solution of $\varrho\check{\gamma}'(s) = \partial_\sigma\hat{\chi}_A(\bar{\sigma}_\varrho(s))$. Define, moreover

$$\varphi(s,\varrho) = \langle\bar{\sigma}_\varrho(s), \varrho\check{\gamma}'(s)\rangle - \hat{\chi}_A(\bar{\sigma}_\varrho(s)).$$

Observe that $\bar{\sigma}_0 = 0$ and $\partial_\varrho\varphi(s,\varrho) = \langle\bar{\sigma}_\varrho(s), \check{\gamma}'(s)\rangle$, finally:

$$\partial^2_\varrho\varphi(s,\varrho) = \langle\partial_\varrho\bar{\sigma}_\varrho(s), \check{\gamma}'(s)\rangle = \langle\partial_\varrho\bar{\sigma}_\varrho(s), \partial^2_\sigma\hat{\chi}_A(\bar{\sigma}_\varrho)\partial_\varrho\bar{\sigma}_\varrho(s)\rangle \geq 0.$$

[36] Just note that $\mathbb{D}(\theta^\gamma(s))$ must contain a right triangle with vertexes $\gamma'(s)$ and $\bar{A}(\theta^\gamma(s))$ and base of length at least τ.

In particular $\varphi(s,\cdot)$ is increasing; hence, using once again the definition (6.16) of $\mathcal{Z}_\epsilon^-$ and the smallness condition on ϵ, we conclude that $\mathcal{Z}_\epsilon^-(\gamma(s),\theta^\gamma(s)) \geq \varphi(s,1/2)$.

Recall moreover there exists $c>0$ such that $\inf_{\|\sigma\|\leq 1}\partial_\sigma^2\hat{\chi}_A(\sigma)\geq c\mathbb{1}$ (as quadratic forms). Let

$$\varrho_0 = \max\{\varrho\in[0,1/2]\;:\;\|\bar\sigma_{\varrho'}(s)\|\leq 1 \text{ for all } \varrho'\in[0,\varrho]\}.$$

Since $\gamma\in\mathrm{Lip}_{C,*}([0,T],\mathbb{R}^d)$ we have $c\|\partial_\varrho\bar\sigma_\varrho(s)\|\leq\|\check\gamma'(s)\|\leq 2C$ for all $\varrho\in[0,\varrho_0]$. Hence, either $\varrho_0=1/2$ or, otherwise, $1=\|\bar\sigma_{\varrho_0}(s)\|\leq 2Cc^{-1}\varrho_0$. In any case we have $\varrho_0\geq C_\#$. We thus conclude:

$$\varphi(s,1/2)\geq\varphi(s,\varrho_0)=\int_0^{\varrho_0}d\varrho\int_0^{\varrho}d\nu\langle\check\gamma'(s),\partial_\sigma^2\hat\chi_A(\bar\sigma_\nu(s))^{-1}\check\gamma'(s)\rangle\geq C_\#\|\check\gamma'(s)\|^2.$$

Since $\mathscr{I}_{\theta,\epsilon}^-(\gamma)<\infty$ we conclude that s can be chosen in a full-measure set in $[0,T]$; hence the above estimate holds a.e., which concludes the proof of our lemma since $\mathscr{I}_{\theta,\epsilon}^+\geq\mathscr{I}_{\theta,\epsilon}^-$. □

Note that a similar, but simpler, argument shows that

$$I_\theta(\gamma)\geq C_\#\|\gamma'-\bar A(\bar\theta(\cdot,\theta))\|_{L^2}^2. \tag{6.20}$$

7. Deviations from the average: Large Deviations

We are now at last ready to precisely state and prove our Large Deviations results.

Let us recall the definition (2.8) of the random element $\gamma_\varepsilon(t)$; observe that equivalently, we have:

$$\gamma_\varepsilon(t)=\varepsilon\sum_{j=0}^{\lfloor t\varepsilon^{-1}\rfloor-1}A\circ F_\varepsilon^j(x,\theta)+(t-\varepsilon\lfloor t\varepsilon^{-1}\rfloor)A\circ F_\varepsilon^{\lfloor t\varepsilon^{-1}\rfloor}(x,\theta). \tag{7.1}$$

Recall that $\gamma_\varepsilon\in\mathcal{C}^0_*([0,T],\mathbb{R}^d):=\{\gamma\in\mathcal{C}^0([0,T],\mathbb{R}^d)\;:\;\gamma(0)=0\}$; moreover $\gamma_\varepsilon(t)=(\theta_\varepsilon(t)-\theta,\zeta_\varepsilon(t))$ and the family $\{\gamma_\varepsilon\}$ is uniformly Lipschitz of constant $\|A\|_{\mathcal{C}^0}$: in fact it is differentiable at all $t\notin\varepsilon\mathbb{Z}$.

Given a standard pair ℓ we can consider γ_ε as a random element of $\mathcal{C}^0_*([0,T],\mathbb{R}^d)$ by assuming that (x,θ) are distributed according to ℓ. In fact, as already mentioned in Section 2, it is more convenient to work directly in the probability space $\mathcal{C}^0([0,T],\mathbb{R}^d)$ endowed with the probability measure $\mathbb{P}_{\ell,\varepsilon}$ determined by the law of γ_ε under ℓ, that is $\mathbb{P}_{\ell,\varepsilon}=(\gamma_\varepsilon)_*\mu_\ell$. In particular, for any function $g\in\mathcal{C}^0(\mathbb{R}^d,\mathbb{R})$, $k\in\mathbb{N}$ and standard pair ℓ:[37]

$$\mathbb{E}_{\ell,\varepsilon}(g\circ\gamma(k\varepsilon))=\mu_\ell(g(\gamma_\varepsilon(k\varepsilon))=\mu_\ell(g\circ\mathbb{F}_\varepsilon^k)$$

where $\mathbb{E}_{\ell,\varepsilon}$ is the expectation associated to the probability $\mathbb{P}_{\ell,\varepsilon}$ and $\mathbb{F}_\varepsilon$ is defined in (2.6).

Remark 7.1. *By the above mentioned Lipschitz property of the paths γ_ε and since $\gamma_\varepsilon\in\mathcal{C}^0_*([0,T];\mathbb{R}^d)$ we conclude that the support of $\mathbb{P}_{\ell,\varepsilon}$ is contained in a compact set that is independent on ε and ℓ; in particular the family $\{\mathbb{P}_{\ell,\varepsilon}\;:\;\varepsilon>0,\ell$ standard pair$\}$ is tight. More precisely, for any $C>\|A\|_{\mathcal{C}^0}$ we have that $\mathbb{P}_{\ell,\varepsilon}(\mathrm{Lip}_{C,*})=1$ where*

$$\mathrm{Lip}_{C,*}=\{\gamma\in\mathcal{C}^0_*([0,T],\mathbb{R}^d):\;\|\gamma(t)-\gamma(s)\|\leq C|t-s|\;\forall\,t,s\in[0,T]\}.$$

[37] According to the usual probabilistic notation $\gamma(t)$ stands both for the numerical value of the path γ at time t and for the evaluation functional $\gamma(t):\mathcal{C}^0([0,T],\mathbb{R}^d)\to\mathbb{R}^d$ defined by $\gamma(t)(\tilde\gamma)=\tilde\gamma(t)$, for all $\tilde\gamma\in\mathcal{C}^0([0,T],\mathbb{R}^d)$.

For any standard pair ℓ recall (see (4.1)) that we defined $\theta_\ell^* = \int_a^b \rho(x)G(x)dx$, to be the average of the random variable θ_0. If we let $\varepsilon \to 0$ and consider standard pairs ℓ_ε (each standard with respect to the corresponding ε) with fixed θ_ℓ^*, Theorem 2.1 implies that $\mathbb{P}_{\ell_\varepsilon,\varepsilon}$ converges, as $\varepsilon \to 0$, to a measure supported on the single path $\bar\gamma(t)$ defined in (2.9). The goal of this section is to establish estimates for the deviations from this path.

We begin with Sections 7.1 and 7.2 where we establish large deviations results that are optimal only for relatively short times. Then in Section 7.3 we use such preliminary results to prove Theorem 2.4; finally in Section 7.5 we prove the remaining propositions stated in Section 2.

7.1. **Upper bound for arbitrary sets (short times).**

For any $\gamma \in C^0([0,T];\mathbb{R}^d)$, let $B(\gamma, r)$ denote the $\mathcal{C}^0$-ball of radius r centered at γ.[38] For each measurable set (event) $Q \subset \mathcal{C}^0([0,T],\mathbb{R}^d)$ define $Q_{C,*} = Q \cap \mathrm{Lip}_{C,*}$. By Remark 7.1, we conclude that for any C sufficiently large:

$$\mathbb{P}_{\ell,\varepsilon}(Q \setminus Q_{C,*}) = 0. \tag{7.2}$$

Lemma 7.2 (Upper bound)**.** *There exist $C_0 > 0$ and $\varepsilon_0, T_{\max} \in (0,1]$ such that, for all $\varepsilon \le \varepsilon_0$, $T \in [\varepsilon_0^{-4}\varepsilon, T_{\max}]$, and $Q \subset \mathcal{C}^0([0,T],\mathbb{R}^d)$, for any $\theta \in \mathbb{T}$ and standard pair ℓ with $\theta_\ell^* = \theta$.*

$$\varepsilon \log \mathbb{P}_{\ell,\varepsilon}(Q) \le -\inf_{\gamma \in Q_{\varepsilon,+}} I_\theta(\gamma),$$

where I_θ is defined in (6.14), $Q_{\varepsilon,+} = \bigcup_{\gamma\in\overline{Q}} B(\gamma, R_\varepsilon(\gamma))$ with

$$R_\varepsilon(\gamma) = C_0 \max\left\{(\varepsilon^{1/4}T^{-1/4} + T)\|\hat\gamma\|_{L^\infty}, \min\left\{\varepsilon^{1/4}T^{3/4}, (\varepsilon T)^{1/6}\|\hat\gamma\|_{L^\infty}^{2/3}\right\}, \sqrt{\varepsilon T}\right\}$$

and $\hat\gamma = \gamma - \bar\gamma(\cdot,\theta)$, the latter being defined in (2.9).

Proof. For any linear functional $\varphi \in \mathcal{M}^d([0,T]) = \mathcal{C}^0([0,T],\mathbb{R}^d)' = \left[\mathcal{C}^0([0,T],\mathbb{R})'\right]^d$, recalling (7.2), we have

$$\begin{aligned}\mathbb{P}_{\ell,\varepsilon}(Q) &\le \mathbb{E}_{\ell,\varepsilon}\left(\mathbb{1}_{Q_{C,*}} e^{\varphi - \inf_{\gamma\in Q_{C,*}}\varphi(\gamma)}\right)\\ &\le \exp\left[-\inf_{\gamma\in\overline{Q_{C,*}}}\varphi(\gamma)\right]\mathbb{E}_{\ell,\varepsilon}\left(e^{\varphi}\right),\end{aligned} \tag{7.3}$$

where $\mathbb{E}_{\ell,\varepsilon}(e^\varphi)$ denotes the expectation of $\gamma \mapsto \exp(\varphi(\gamma))$ with respect to the probability $\mathbb{P}_{\ell,\varepsilon}$. Let $\Lambda_{\ell,\varepsilon}$ be the *logarithmic moment generating function* and $\Lambda_{\ell,\varepsilon}^*$ be its convex conjugate function, i.e.:[39]

$$\Lambda_{\ell,\varepsilon}(\varphi) = \varepsilon \log \mathbb{E}_{\ell,\varepsilon}\left(e^\varphi\right); \qquad \Lambda_{\ell,\varepsilon}^*(\gamma) = \sup_{\varphi\in\mathcal{M}^d([0,T])}(\varepsilon\varphi(\gamma) - \Lambda_{\ell,\varepsilon}(\varphi)). \tag{7.4}$$

Note that $|\Lambda_{\ell,\varepsilon}(\varphi)| \le \varepsilon C_\# \|\varphi\| < \infty$, hence $\Lambda_{\ell,\varepsilon}$ is a proper convex function.[40] Since $\Lambda_{\ell,\varepsilon}(0) = 0$, we have $\Lambda_{\ell,\varepsilon}^* \ge 0$. Moreover $\Lambda_{\ell,\varepsilon}^* : \mathcal{C}^0([0,T],\mathbb{R}^d) \to \mathbb{R}\cup\{+\infty\}$ is

---

[38] We prefer not to write the explicit dependence of B on T since this can be recovered by the fact that $\gamma \in \mathcal{C}^0([0,T],\mathbb{R}^d)$.

[39] We will see shortly, in (7.8), that $\Lambda_{\ell,\varepsilon}$ agrees with the previous definition (5.10), hence justifying the abuse of notations (in one case we have a functional on measures, in the other a functional on BV).

[40] The first assertion follows by (7.2) which implies $\|\gamma\|_\infty \le CT$, $\mathbb{P}_{\ell,\varepsilon}$-a.s.. The second follows from the Hölder inequality since, for all $t \in [0,1]$, and $\varphi, \varphi' \in \mathcal{M}([0,T])$,

$$\begin{aligned}\Lambda_{\ell,\varepsilon}(t\varphi + (1-t)\varphi') &= \varepsilon \log \mathbb{E}_{\ell,A,\varepsilon}\left([e^\varphi]^t[e^{\varphi'}]^{1-t}\right)\\ &\le \varepsilon\log\left[\mathbb{E}_{\ell,A,\varepsilon}\left(e^\varphi\right)^t \mathbb{E}_{\ell,A,\varepsilon}\left(e^{\varphi'}\right)^{1-t}\right] = t\Lambda_{\ell,\varepsilon}(\varphi) + (1-t)\Lambda_{\ell,\varepsilon}(\varphi').\end{aligned}$$

convex as well and lower semi-continuous (with respect to the C^0 topology), since it is the conjugate function of a proper function. We can then follow the strategy of [15, Exercise 4.5.5]. Note that $(\varphi,\gamma)\mapsto \varepsilon\varphi(\gamma)-\Lambda_{\ell,\varepsilon}(\varphi)$ is a function concave in φ, continuous in γ with respect to the $\mathcal{C}^0$ topology, also it is convex in γ for any $\varphi\in\mathcal{M}([0,T])$. Finally, $\overline{Q_{C,*}}$ is compact in $\mathcal{C}^0$. Thus, the Minimax Theorem ([49], but see [39] for an elementary proof) guarantees that

$$\sup_{\varphi\in\mathcal{M}^d}\inf_{\gamma\in\overline{Q_{C,*}}}[\varepsilon\varphi(\gamma)-\Lambda_{\ell,\varepsilon}(\varphi)]=\inf_{\gamma\in\overline{Q_{C,*}}}\sup_{\varphi\in\mathcal{M}^d}[\varepsilon\varphi(\gamma)-\Lambda_{\ell,\varepsilon}(\varphi)].$$

The above implies, taking the inf on φ in (7.3),

$$\begin{aligned}\varepsilon\log\mathbb{P}_{\ell,\varepsilon}(Q)&\leq-\inf_{\gamma\in\overline{Q_{C,*}}}\sup_{\varphi\in\mathcal{M}^d}\varepsilon\varphi(\gamma)-\Lambda_{\ell,\varepsilon}(\varphi)\\&\leq-\inf_{\gamma\in\overline{Q_{C,*}}}\Lambda^*_{\ell,\varepsilon}(\gamma).\end{aligned}\tag{7.5}$$

The above estimate looks indeed quite promising, but unfortunately it is completely useless without sharp information on $\Lambda^*_{\ell,\varepsilon}$.

We are thus left with the task of computing $\Lambda^*_{\ell,\varepsilon}$. It turns out to be convenient to associate to φ the function σ defined as:

$$\sigma(s)=\varepsilon\varphi((s,T]),\tag{7.6}$$

where the right hand side is interpreted by applying the Jordan Decomposition to φ. Note that, by definition, $\sigma=(\sigma_1,\cdots,\sigma_d)$ with $\sigma_i\in\mathrm{BV}$, thus $\sigma\in\mathrm{BV}^d$, that we will simply call BV to ease notation. By definition $\|\varphi\|=\varepsilon^{-1}\|\sigma\|_{\mathrm{BV}}$ and, for any $\gamma\in\mathrm{Lip}_{C,*}$,

$$\varphi(\gamma)=\varepsilon^{-1}\int_0^T\langle\sigma(s),\gamma'(s)\rangle ds.\tag{7.7}$$

On the other hand, for each $\sigma\in\mathrm{BV}$ there exists $\varphi\in\mathcal{M}^d([0,T])$ such that (7.7) holds, see [21, Section 5.1, Theorem 1]. Moreover (recall definition (7.1)):[41]

$$\begin{aligned}\varphi(\gamma_\varepsilon)&=\varepsilon^{-1}\int_0^T\langle\sigma(s),\gamma_\varepsilon'(s)\rangle ds\\&=\sum_{k=0}^{\lfloor T\varepsilon^{-1}\rfloor-1}\langle\varepsilon^{-1}\int_{k\varepsilon}^{(k+1)\varepsilon}\sigma(s)ds,A\circ F_\varepsilon^k\rangle=\sum_{k=0}^{\lfloor T\varepsilon^{-1}\rfloor-1}\langle\sigma_k,A\circ F_\varepsilon^k\rangle.\end{aligned}$$

where, as in Section 5.2, we introduced the notation

$$\sigma_n=\varepsilon^{-1}\int_{n\varepsilon}^{(n+1)\varepsilon}\sigma(s)ds.$$

Hence, we conclude that for any $\varphi\in\mathcal{M}^d$ and for the corresponding $\sigma\in\mathrm{BV}$:

$$\begin{aligned}\Lambda_{\ell,\varepsilon}(\varphi)&=\varepsilon\log\mu_\ell\left(e^{\sum_{k=0}^{\lfloor T\varepsilon^{-1}\rfloor-1}\langle\sigma_k,A\circ F_\varepsilon^k\rangle}\right)=\Lambda_{\ell,\varepsilon}(\sigma)\\&=\int_0^T\chi_A(\sigma(s),\bar\theta(s,\theta_\ell^*))ds+\mathcal{R}_{\ell,\varepsilon}(\sigma).\end{aligned}\tag{7.8}$$

where $\mathcal{R}_{\ell,\varepsilon}(\sigma)$ (which is defined by the equation above, see (5.11)) satisfies the estimates given in Proposition 5.4.

[41] As we often do in this work, we are neglecting the contribution of the fact that $T\varepsilon^{-1}$ may not be an integer.

The above implies (recall (7.4), (7.7) and the definition (6.13) of $\kappa_{\gamma,\theta}$):

$$
\begin{aligned}
\Lambda^*_{\ell,\varepsilon}(\gamma) &= \sup_{\sigma\in\mathrm{BV}}\left[\int_0^T\langle\sigma(s),\gamma'(s)\rangle-\Lambda_{\ell,\varepsilon}(\sigma)\right]\\
&= \sup_{\sigma\in\mathrm{BV}}\left[\int_0^T\kappa_{\gamma,\theta^*_\ell}(\sigma(s),s)ds-\mathcal{R}_{\ell,\varepsilon}(\sigma)\right].
\end{aligned}\tag{7.9}
$$

Formula (7.9) closely resembles the definition of rate function given in (6.14). Unfortunately there is an obvious obstacle: we need to ensure that the first term dominates $\mathcal{R}_{\ell,\varepsilon}(\sigma)$. As already observed in Subsection 5.3, this can be taken care of by some regularization procedure for σ; we will now describe the dual procedure, i.e. a regularization procedure for the paths γ.

Given γ and $h=T/N_h$, for $N_h\in\mathbb{N}$ suitably large to be chosen later, we denote with $\gamma_h=\Pi_{(h)}\gamma$ the polygonalization of γ over a mesh of size h. In other words, we define $\gamma_h\in\mathcal{C}^0_*([0,T],\mathbb{R}^d)$ so that $\gamma_h'=\Pi^{(h)}(\gamma')$ where $\Pi^{(h)}$ has been defined in (5.45). Recall (see (7.2)) that it suffices to consider paths $\gamma\in\mathrm{Lip}_C$ and thus, since $\Pi^{(h)}$ is a contraction in L^∞ (see Sub-lemma 5.8), we conclude that $\gamma_h\in\mathrm{Lip}_{C,*}$. Moreover, for $t\in[nh,(n+1)h]$, we have

$$
|\gamma_h(t)-\gamma(t)|\leq\int_{nh}^{(n+1)h}|\gamma_h'(s)-\gamma'(s)|ds\leq 2\int_{nh}^{(n+1)h}|\gamma'(s)|ds,\tag{7.10}
$$

which implies that $\gamma_h\in Q_h=\bigcup_{\gamma\in\overline{Q_{C,*}}}B(\gamma,2Ch)$. Then, by (7.4) and recalling the definition of $\Lambda^{(h)}_{\ell,\varepsilon}$ given in (5.46):

$$
\begin{aligned}
\Lambda^*_{\ell,\varepsilon}(\gamma) &\geq \sup_{\sigma\in\mathrm{BV}}\left[\int_0^T\langle\Pi^{(h)}\sigma(s),\gamma'(s)\rangle-\Lambda^{(h)}_{\ell,\varepsilon}(\sigma)\right]\\
&\geq \sup_{\sigma\in\mathrm{BV}}\left[\int_0^T\langle\sigma(s),\gamma_h'(s)\rangle-\Lambda^{(h)}_{\ell,\varepsilon}(\sigma)\right]=\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma_h)
\end{aligned}
$$

where $\Lambda^{(h)*}_{\ell,\varepsilon}$ is the Legendre transform of the regularized moment generating functional $\Lambda^{(h)}_{\ell,\varepsilon}$. In particular, we have $\Lambda^{(h)*}_{\ell,\varepsilon}\geq 0$ because $\Lambda^{(h)}_{\ell,\varepsilon}(0)=0$.

Hence, we conclude that

$$
\inf_{\gamma\in\overline{Q_{C,*}}}\Lambda^*_{\ell,\varepsilon}(\gamma)\geq\inf_{\gamma\in Q_h}\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma).\tag{7.11}
$$

Observe that, by (5.47),

$$
\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma)=\sup_{\sigma\in\mathrm{BV}}\left[\int_0^T\kappa_{\gamma,\theta^*_\ell}(\sigma(s),s)ds-\mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma)\right],\tag{7.12}
$$

where $\mathcal{R}^{(h)}_{\ell,\varepsilon}$ satisfies the estimates obtained in Lemma 5.9. The aim of the above regularization is to gain control on $\mathcal{R}^{(h)}_{\ell,\varepsilon}$ even for very rough σ. This is the content of the next sub-lemma.

Sub-lemma 7.3. *There exists $\tilde{C}_0>0$ and $\varepsilon_0,T_{\max}\in(0,1]$, such that, for all $\varepsilon\in[0,\varepsilon_0]$, $T\in[\varepsilon_0^{-4}\varepsilon,T_{\max}]$ and $\gamma\in\mathrm{Lip}_{C,*}([0,T],\mathbb{R}^d)$, we have for any $\theta\in\mathbb{T}$ and standard pair ℓ with $\theta^*_\ell=\theta$, setting $h=\sqrt{\varepsilon T}$:*

$$
\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma)\geq\inf_{\tilde{\gamma}\in B(\gamma,\tilde{R}_\varepsilon(\gamma))}I_\theta(\tilde{\gamma})\tag{7.13}
$$

where we define

$$
\tilde{R}_\varepsilon(\gamma)=\tilde{C}_0\max\left\{(\varepsilon^{1/4}T^{-1/4}+T)\|\hat{\gamma}\|_{L^\infty},\min\left\{\varepsilon^{1/4}T^{3/4},(\varepsilon T)^{1/6}\|\hat{\gamma}\|_{L^\infty}^{2/3}\right\},\sqrt{\varepsilon T}\right\}
$$

and recall, $\hat{\gamma}=\gamma-\bar{\gamma}(\cdot,\theta)$.

Observe that, together with (7.5) and (7.11), the above sub-lemma immediately allows to conclude the proof of Lemma 7.2 choosing $C_0 = \tilde{C}_0 + 2C$. □

Proof of Sub-Lemma 7.3. Let us fix $\tilde{C}_0 > 1$ large enough to be specified later; we begin by observing that if $\|\hat{\gamma}\|_{L^\infty} < \tilde{C}_0\sqrt{\varepsilon T}$, then $B(\gamma, \tilde{R}_\varepsilon(\gamma)) \ni \bar{\gamma}$, which implies that $\inf_{\tilde{\gamma}\in B(\gamma,\tilde{R}_\varepsilon(\gamma))} I_\theta(\tilde{\gamma}) = I_\theta(\bar{\gamma}) = 0$ and the sub-lemma holds trivially since $\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) \geq 0$. In the rest of the proof, we will therefore always assume that $\|\hat{\gamma}\|_{L^\infty} \geq \tilde{C}_0\sqrt{\varepsilon T}$ provided

$$\tilde{R}_\varepsilon(\gamma) \geq \tilde{C}_0\sqrt{\varepsilon T}. \tag{7.14}$$

Recall (7.12):

$$\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) = \sup_{\sigma\in\mathrm{BV}}\left[\int_0^T \kappa_{\gamma,\theta^*_\ell}(\sigma(s),s)ds - \mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma)\right],$$

and that, by definition (6.14) (since $\gamma \in \mathrm{Lip}_{C,*}$):

$$I_\theta(\gamma) = \sup_{\sigma\in\mathrm{BV}}\int_0^T \kappa_{\gamma,\theta}(\sigma(s),s)ds,$$

where, by definition (6.13),

$$\begin{aligned}\kappa_{\gamma,\theta}(\sigma(s),s) &= \langle\sigma(s),\gamma'(s)\rangle - \chi_A(\sigma(s),\bar{\theta}(s,\theta))\\ &= \langle\sigma(s),\hat{\gamma}'(s)\rangle - \hat{\chi}_A(\sigma(s),\bar{\theta}(s,\theta)).\end{aligned} \tag{7.15}$$

We will proceed as follows: by convexity of the rate function, in any ball $B(\gamma)$ around γ, we can find paths which are more likely than γ itself; in particular $\inf I_\theta(B(\gamma)) < I_\theta(\gamma)$ (the inequality is strict since $\gamma \neq \bar{\gamma}$). The idea is then to take the ball B to be so large that the decrease in the rate function compensates for the remainder term $\mathcal{R}^{(h)}_{\ell,\varepsilon}$. Of course, by choosing larger balls, we obtain worse bounds for the error in the final estimate: the key technical point of the sub-lemma rests exactly in finding a good compromise for the size of B.

For any $\varrho \in [0,1]$, let us define the convex interpolation $\gamma_\varrho = (1-\varrho)\gamma + \varrho\bar{\gamma}$; since $I_\theta(\gamma_{\varrho=1}) = I_\theta(\bar{\gamma}) = 0$ and by convexity of I_θ, we conclude that $I_\theta(\gamma_\varrho)$ is decreasing in ϱ. Hence we want to find ϱ sufficiently large so that the decrease compensates for the remainder term. The choice of ϱ will in fact depend on the distance of γ from $\bar{\gamma}$, that is on $\|\hat{\gamma}\|_{L^\infty}$. We carry out the estimate using two different strategies as they yield optimal bounds in different regimes.

Case I: non-perturbative estimate

Note that, since $\hat{\gamma}'_\varrho = (1-\varrho)\hat{\gamma}'$, we have

$$\kappa(\sigma,\gamma',\theta) = \kappa(\sigma,\gamma'_\varrho,\theta) + \varrho\langle\sigma,\hat{\gamma}'\rangle. \tag{7.16}$$

Collecting (7.12), (7.15) and (7.16) we obtain, for any $\sigma \in \mathrm{BV}$ and $\varrho \in [0,1]$,

$$\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) \geq \int_0^T \kappa(\sigma(s),\gamma'_\varrho(s),\bar{\theta}(s,\theta))ds + \varrho\int_0^T\langle\sigma(s),\hat{\gamma}'(s)\rangle ds - \mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma). \tag{7.17}$$

We then use the estimate for $\mathcal{R}^{(h)}_{\ell,\varepsilon}$ given by Lemma 5.9-(a) with the choice

$$L = T^{1/4}\varepsilon^{-1/4} \in [\varepsilon_0^{-1}, \min\{\varepsilon^{-1}T, \varepsilon_0\varepsilon^{-1/2}\}],$$

where the inclusion follows from our conditions on ε, T and noticing that $[\varepsilon_0^{-4}\varepsilon, T_{\max}]$ is empty for $\varepsilon > T_{\max}\varepsilon_0^4$. Recalling that $h = \sqrt{\varepsilon T}$ we obtain:

$$\mathcal{R}^{(h)}_{\ell,\varepsilon}(\sigma) \leq C_{12}\left[T^{5/4}\varepsilon^{3/4} + [T^{-1/4}\varepsilon^{1/4} + \min\{T,\|\sigma\|_{L^1}\}]\|\sigma\|_{L^1}\right]. \tag{7.18}$$

Let us fix $C_{10} > 0$ large to be specified later and let

$$\varrho_0 = C_{10}\{\varepsilon^{1/4}T^{-1/4} + 2T\}. \tag{7.19}$$

To fix ideas, we may assume $\varepsilon_0, T_{\max}$ to be so small that $\varrho_0 \leq \frac{1}{3}$. Let us first examine the possibility $I(\gamma_{\varrho_0}) = \infty$; in this case we claim that $\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) = \infty$, which trivially implies the sub-lemma. In fact: let us fix arbitrarily $M > 1$; by (6.14), since $\gamma \in \mathrm{Lip}_{C,*}$, there exists $\bar\sigma_M \in \mathrm{BV}$ such that

$$\begin{aligned} M &\leq \int_0^T \kappa(\bar\sigma_M(s), \gamma'_{\varrho_0}(s), \bar\theta(s,\theta))ds \\ &= \int_0^T \langle \bar\sigma_M(s), \hat\gamma'_{\varrho_0}(s)\rangle - \hat\chi_A(\bar\sigma_M(s), \bar\theta(s,\theta))ds. \end{aligned} \tag{7.20}$$

By the properties of $\hat\chi_A$ it follows[42] that if $\|\sigma\| \geq \sigma_*$, then

$$\hat\chi_A(\sigma,\theta) \geq C_\#\|\sigma\|.$$

Define the set $\Sigma_M = \Sigma_M(\varrho_0, M) = \{s \in [0,T] \;:\; \|\bar\sigma_M(s)\| \geq \sigma_*\}$. Then

$$M \leq \int_0^T \langle \bar\sigma_M, \hat\gamma'_{\varrho_0}\rangle - C_\# \int_{\Sigma_M} \|\bar\sigma_M\| \leq \int_0^T \langle \bar\sigma_M, (1-\varrho_0)\hat\gamma'\rangle - C_\#(\|\bar\sigma_M\|_{L^1} - T\sigma_*).$$

Assuming $M \geq C_\#\sqrt{T}$ to be sufficiently large, it follows

$$\int_0^T \langle \bar\sigma_M, \hat\gamma'\rangle > \frac{M}{2} + C_\#\|\bar\sigma_M\|_{L^1}. \tag{7.21}$$

Using (7.17), (7.18) together with (7.20) and (7.21) yields

$$\begin{aligned} \Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) - M &\geq \varrho_0 \int_0^T \langle \bar\sigma_M, \hat\gamma'\rangle - C_{12}\left[\varepsilon^{3/4}T^{5/4} + \left\{\frac{\varepsilon^{1/4}}{T^{1/4}} + T\right\}\|\bar\sigma_M\|_{L^1}\right] \\ &\geq 0, \end{aligned} \tag{7.22}$$

provided ε is small enough, M is large enough, and C_{10}, thus ϱ_0, is large enough. By the arbitrariness of M it follows $\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) = \infty$.

We can therefore assume that $I(\gamma_{\varrho_0}) < \infty$: since $I(\gamma_\varrho)$ is decreasing in ϱ, this implies that $I(\gamma_\varrho) < \infty$ for any $\varrho \in [\varrho_0, 1]$. In particular, recalling also Lemma 6.7, for each $\varrho \in (\varrho_0, 1]$ and $\theta \in \mathbb{T}$, $s \in [0,T]$ there exists an open set $s \in U \subset [0,T]$ and a compact set $\mathcal{K} = \mathcal{K}(\theta,\varrho,s) \subset \cap_{s'\in U}\mathbb{D}_*(\bar\theta(s',\theta))$ such that $\gamma'_\varrho(s') \in \mathcal{K}$ for almost all $s' \in U$ and thus that $(1-\varrho)\hat\gamma' = \hat\gamma'_\varrho(s) = \partial_\sigma\hat\chi_A(\sigma)$ has a (unique) solution for almost all $s \in [0,T]$, which we denote with $\bar\sigma_\varrho$. Since $\partial_\sigma\hat\chi_A$ is a homeomorphism, see the proof of Lemma 6.2, we have $\bar\sigma_\varrho \in L^\infty$.

By construction, $\bar\sigma_\varrho$ realizes the sup in (6.14); in particular (7.17) reads:

$$\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) - I_\theta(\gamma_\varrho) \geq \varrho \int_0^T \langle \bar\sigma_\varrho(s), \gamma'(s)\rangle ds - \mathcal{R}^{(h)}_{\ell,\varepsilon}(\bar\sigma_\varrho).$$

Let us assume $\varrho \in (\varrho_0, \frac{1}{2}]$. Define the set $\Sigma_* = \Sigma_*(\varrho) = \{s \in [0,T] \;:\; \|\bar\sigma_\varrho(s)\| \geq \sigma_*\}$ and let us denote by $\Sigma_*^c = [0,T] \setminus \Sigma_*$ its complement. Then, plugging (7.18) into the above inequality and recalling Lemma 6.5, we find:

$$\begin{aligned} \Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) - I_\theta(\gamma_\varrho) &\geq C_\#\varrho\left[\int_{\Sigma_*}\|\bar\sigma_\varrho\| + \int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right] \\ &\quad - C_{12}\left[T^{5/4}\varepsilon^{3/4} + \left\{\varepsilon^{1/4}T^{-1/4} + \min\{T, \|\bar\sigma_\varrho\|_{L^1}\}\right\}\|\bar\sigma_\varrho\|_{L^1}\right]. \end{aligned} \tag{7.23}$$

---

[42] Equations (A.11) and (A.12b) imply that $\hat\chi_A(\cdot,\theta)$ is convex and has its minimum in $\sigma = 0$ where $\hat\chi_A(0,\theta) = 0$

First, we claim that:

$$\min\{T, \|\bar\sigma_\varrho\|_{L^1}\}\|\bar\sigma_\varrho\|_{L^1} \le 2T\int_{\Sigma_*}\|\bar\sigma_\varrho\| + 2T\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2. \tag{7.24}$$

In fact, assume that $\int_{\Sigma_*}\|\bar\sigma_\varrho\| \ge \int_{\Sigma_*^c}\|\bar\sigma_\varrho\|$; then $\|\bar\sigma_\varrho\|_{L^1} \le 2\int_{\Sigma_*}\|\bar\sigma_\varrho\|$ and

$$\min\{T, \|\bar\sigma_\varrho\|_{L^1}\}\|\bar\sigma_\varrho\|_{L^1} \le 2T\int_{\Sigma_*}\|\bar\sigma_\varrho\|.$$

If, on the other hand, $\int_{\Sigma_*}\|\bar\sigma_\varrho\| \le \int_{\Sigma_*^c}\|\bar\sigma_\varrho\|$, we have

$$\min\{T, \|\bar\sigma_\varrho\|_{L^1}\}\|\bar\sigma_\varrho\|_{L^1} \le T\int_{\Sigma_*}\|\bar\sigma_\varrho\| + 2\left[\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|\right]^2 \le T\int_{\Sigma_*}\|\bar\sigma_\varrho\| + 2T\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2.$$

Which proves (7.24). Plugging it into (7.23) we obtain:

$$\begin{aligned}
\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) - I(\gamma_\varrho) \ge{}& C_\# \varrho\left[\int_{\Sigma_*}\|\bar\sigma_\varrho\| + \int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right] - C_{12}\Bigg[\varepsilon^{3/4}T^{5/4} +\\
&+ \{\varepsilon^{1/4}T^{-1/4} + 2T\}\int_{\Sigma_*}\|\bar\sigma_\varrho\| + (\varepsilon T)^{1/4}\left[\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right]^{1/2} + 2T\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\Bigg]\\
\ge{}& C_\#\varrho\left[\int_{\Sigma_*}\|\bar\sigma_\varrho\| + \int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right]\\
&- C_{12}\left[\varepsilon^{3/4}T^{5/4} + (\varepsilon T)^{1/4}\left[\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right]^{1/2}\right],
\end{aligned}$$

provided C_{10} has been chosen large enough (recall (7.19) and that $\varrho \ge \varrho_0$).

We conclude that,

$$\begin{aligned}
&\text{if } \int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2 \ge C_\#\varrho^{-2}\varepsilon^{1/2}T^{1/2}\\
&\text{or } \int_{\Sigma_*}\|\bar\sigma_\varrho\| \ge C_\#\varrho^{-1}\max\left\{\varepsilon^{3/4}T^{5/4}, (\varepsilon T)^{1/4}\left[\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right]^{1/2}\right\},
\end{aligned}$$

then $\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) - I(\gamma_\varrho) \ge 0$. Otherwise,

$$\begin{aligned}
&\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2 \le C_\#\varrho^{-2}(\varepsilon T)^{1/2}\\
&\int_{\Sigma_*}\|\bar\sigma_\varrho\| \le C_\#\varrho^{-1}\max\left\{\varepsilon^{3/4}T^{5/4}, (\varepsilon T)^{1/4}\left[\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right]^{1/2}\right\}\\
&\qquad\le C_\#\varrho^{-2}(\varepsilon T)^{1/2},
\end{aligned}\tag{7.25}$$

provided ε is small enough. Note that Lemma 6.5 implies that $\|\hat\gamma'_\varrho(s)\| \le C_{**}\|\bar\sigma_\varrho(s)\|$; then estimates (7.25) and definition (7.19) imply, since $\varrho > \varrho_0$,

$$\begin{aligned}
\|\hat\gamma\|_{L^\infty} &\le 2\|\hat\gamma'_\varrho\|_{L^1} \le 2C_{**}\|\bar\sigma_\varrho\|_{L^1}\\
&\le 2C_{**}\left[\int_{\Sigma_*}\|\bar\sigma_\varrho\| + \left[T\int_{\Sigma_*^c}\|\bar\sigma_\varrho\|^2\right]^{1/2}\right] \le C_\#\varrho^{-1}\varepsilon^{1/4}T^{3/4}.
\end{aligned}\tag{7.26}$$

Hence, if we choose C_{10} sufficiently large and

$$\varrho = \max\{2\varrho_0, 2C_{10}\varepsilon^{1/4}T^{3/4}\|\hat\gamma\|^{-1}_{L^\infty}\},\tag{7.27}$$

then (7.26) cannot hold true and necessarily $\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) \geq I(\gamma_\varrho)$. Thus, we obtain (7.13), provided that $\tilde{R}_\varepsilon(\gamma) > \varrho\|\hat{\gamma}\|_{L^\infty}$.

Case II: perturbative estimate

In this case we plan to apply the more refined estimate given in Lemma 5.9-(b); yet it may not be possible to do so with the path γ_ϱ since in general $\Sigma_* \neq \emptyset$. In order to circumvent this problem we define another path that is sufficiently close to γ_ϱ and to which we can apply the mentioned estimate. Let $\underline{\gamma} \in \mathrm{Lip}_{C,*}$ so that $\underline{\gamma}(0) = 0$ and $\underline{\gamma}'(s) = \bar{\gamma}'(s)$ for $s \in \Sigma_*$ and $\underline{\gamma}'(s) = \gamma'(s)$ otherwise. Define $\underline{\gamma}_\varrho = (1-\varrho)\underline{\gamma} + \varrho\bar{\gamma}$, hence

$$\hat{\underline{\gamma}}'_\varrho = (1-\varrho)\hat{\underline{\gamma}}' = \mathbb{1}_{\Sigma_*^c}\hat{\gamma}'_\varrho. \tag{7.28}$$

Let us now choose

$$\varrho = C_{10}\max\{(\varepsilon T)^{1/6}\|\hat{\gamma}\|_{L^\infty}^{-1/3}, \varepsilon^{1/4}T^{-1/4} + 2T\}. \tag{7.29}$$

Then, either $\Lambda^{(h)*}_{\ell,\varepsilon}(\gamma) \geq I(\gamma_\varrho)$ (and then (7.13) holds provided $\tilde{R}_\varepsilon(\gamma) \geq \varrho\|\hat{\gamma}\|_{L^\infty}$), or estimates (7.25) hold; in the latter case Lemma 6.5 yields

$$\|\gamma_\varrho - \underline{\gamma}_\varrho\|_{L^\infty} \leq \int_{\Sigma_*}\|\hat{\gamma}'_\varrho(s)\|ds \leq C_{**}\int_{\Sigma_*}\|\bar{\sigma}_\varrho(s)\|ds \leq C_\#\varrho^{-2}(\varepsilon T)^{1/2}. \tag{7.30}$$

In particular, $\|\gamma - \underline{\gamma}_\varrho\|_{L^\infty} \leq \varrho\|\hat{\gamma}\|_{L^\infty} + C_\#\varrho^{-2}(\varepsilon T)^{1/2}$. Observe that, assuming C_{10} large enough, (7.29) implies that (this justifies the choice of the first term in (7.29), which optimizes the above inequality)

$$\|\gamma - \underline{\gamma}_\varrho\|_{L^\infty} \leq 2\varrho\|\hat{\gamma}\|_{L^\infty}. \tag{7.31}$$

Let $\underline{\sigma}_\varrho(s)$ be the unique solution of $\hat{\underline{\gamma}}'_\varrho(s) = \partial_\sigma\hat{\chi}_A(\sigma, \bar{\theta}(s,\theta))$. By definition,

$$\underline{\sigma}_\varrho(s) = \mathbb{1}_{\Sigma_*^c}(s)\bar{\sigma}_\varrho(s) \leq \sigma_*; \tag{7.32}$$

in particular, $\underline{\sigma}_\varrho(s)$ satisfies the hypotheses of Lemma 5.9-(b). Also it satisfies the first inequality of (7.25). We can now proceed as before but using the path $\underline{\gamma}_\varrho$: more precisely, equations (7.12), (7.15), (7.28), (7.32) and Lemma 6.5 imply

$$\begin{aligned}
\Lambda^{(h)}_{\ell,\varepsilon}(\gamma) &\geq \int_0^T \kappa_{\gamma,\theta^*_\ell}(\underline{\sigma}_\varrho(s), s)ds - \mathcal{R}^{(h)}_{\ell,\varepsilon}(\underline{\sigma}_\varrho)\\
&= \int_0^T \kappa_{\underline{\gamma}_\varrho,\theta}(\underline{\sigma}_\varrho(s), s)ds + \frac{\varrho}{1-\varrho}\int_0^T \langle\underline{\sigma}_\varrho(s), \hat{\underline{\gamma}}'_\varrho(s)\rangle - \mathcal{R}^{(h)}_{\ell,\varepsilon}(\underline{\sigma}_\varrho)\\
&\geq I(\underline{\gamma}_\varrho) + C_\#\varrho\|\underline{\sigma}_\varrho\|^2_{L^2} - \mathcal{R}^{(h)}_{\ell,\varepsilon}(\underline{\sigma}_\varrho).
\end{aligned}$$

Before continuing note that, choosing $\tilde{R}_\varepsilon(\gamma) \geq 3\varrho\|\hat{\gamma}\|_{L^\infty}$, we can ensure that (7.13) trivially holds if $\varrho > 1/3$. Hence we can assume $\varrho \leq 1/3$. Next, we claim that

$$\frac{\tilde{C}_0 C_{**}}{3}\sqrt{\varepsilon} \leq \|\underline{\sigma}_\varrho\|_{L^2} \leq C_\# C_{10}^{-1}\sqrt{T}. \tag{7.33}$$

Indeed, by the first of (7.25) and recalling the choice (7.29) $\|\underline{\sigma}_\varrho\|_{L^2} \leq C_\# C_{10}^{-1}\sqrt{T}$. Moreover, recall that we are assuming that $\|\hat{\gamma}\|_{L^\infty} \geq \tilde{C}_0\sqrt{\varepsilon T}$ and $\varrho \leq 1/3$, which by (7.31) yields $\|\hat{\underline{\gamma}}_\varrho\|_{L^\infty} \geq (\tilde{C}_0/3)\sqrt{\varepsilon T}$; in turn, by Lemma 6.5, this implies that $\|\underline{\sigma}_\varrho\|_{L^2} \geq \frac{\tilde{C}_0 C_{**}}{3}\sqrt{\varepsilon}$.

We now make the choice $L = C_\#\|\underline{\sigma}_\varrho\|_{L^2}(\varepsilon T)^{-1/2}$. Let us check when it satisfies the hypotheses of Lemma 5.9-(b). By (7.33), choosing $\tilde{C}_0$ and C_{10} sufficiently large, we have $L \in [\varepsilon_0^{-1}, \varepsilon_0\varepsilon^{-1/2}]$. On the other hand the condition $\varepsilon L \leq T$ is satisfied only if $C_\#\|\underline{\sigma}_\varrho\|_{L^2} \leq T^{3/2}\varepsilon^{-1/2}$, which is automatically ensured only if $T \geq \sqrt{\varepsilon}$, provided C_{10} is sufficiently large.

Thus, if $T \leq \sqrt{\varepsilon}$, then our choice it is not good. In such a case we make the choice $L = \varepsilon_0\varepsilon^{-1}T$. Again, we have $L \in [\varepsilon_0^{-1}, \varepsilon_0\varepsilon^{-1/2}]$, and $\varepsilon L \leq T$, thus the hypotheses of Lemma 5.9-(b) are satisfied.

We are now ready to estimate $\mathcal{R}^{(h)}_{\ell,\varepsilon}$ using Lemma 5.9-(b), recall our choices $h = \sqrt{\varepsilon T}$ and $T \in (\varepsilon L, T_{\max})$. We have to treat the two above regimes separately.

If $T \geq \sqrt{\varepsilon}$, then, using Schwarz inequality,

$$\mathcal{R}^{(h)}_{\ell,\varepsilon}(\underline{\sigma}_\varrho) \leq C_{12}\left[5\sqrt{\varepsilon}\,\|\underline{\sigma}_\varrho\|_{L^2} + 2T\|\underline{\sigma}_\varrho\|^2_{L^2}\right]. \tag{7.34}$$

We thus conclude, assuming C_{10} in (7.29) to be sufficiently large:

$$\Lambda^{(h)}_{\ell,\varepsilon}(\gamma) - I(\underline{\gamma}_\varrho) \geq C_\#\varrho\|\underline{\sigma}_\varrho\|^2_{L^2} - 5C_{12}\sqrt{\varepsilon}\|\underline{\sigma}_\varrho\|_{L^2}. \tag{7.35}$$

Since Lemma 6.5 implies $C_{**}\sqrt{T}\|\sigma\|_{L^2} \geq \|\hat{\gamma}\|_{L^\infty} \geq \tilde{C}_0\sqrt{\varepsilon T}$, the first term in the max of (7.29) implies that (7.13) holds by choosing $\tilde{R}_\varepsilon(\gamma) = 3\varrho\|\hat{\gamma}\|_{L^\infty}$.

Next, we consider the case $T \leq \sqrt{\varepsilon}$ and apply again Lemma 5.9-(b):

$$\begin{aligned}\mathcal{R}^{(h)}_{\ell,\varepsilon}(\underline{\sigma}_\varrho) &\leq C_{12}\left[T^2C_{12}^{-1/2} + 4\sqrt{\varepsilon}\|\underline{\sigma}_\varrho\|_{L^2} + 2\varepsilon T^{-1}\|\underline{\sigma}_\varrho\|^2_{L^2}\right]\\ &\leq C_{12}\left[5\sqrt{\varepsilon}\,\|\underline{\sigma}_\varrho\|_{L^2} + 2\varepsilon T^{-1}\|\underline{\sigma}_\varrho\|^2_{L^2}\right],\end{aligned} \tag{7.36}$$

where we have chosen C_{12} large enough and, in the last line, we have used (7.33). Then (7.13) follows again since, on the one hand, $\varepsilon^{1/4}T^{-1/4} \geq \varepsilon T^{-1}$, and, on the other hand, $(\varepsilon T)^{1/6}\|\hat{\gamma}\|^{-1/3}_{L^\infty}\|\underline{\sigma}_\varrho\|_{L^2} \geq C_\#\sqrt{\varepsilon}$ is implied, again, by $C_{**}\sqrt{T}\|\underline{\sigma}_\varrho\|_{L^2} \geq \|\hat{\gamma}\|_{L^\infty}$ and $\|\hat{\gamma}\|_{L^\infty} \geq \tilde{C}_0\sqrt{\varepsilon T}$. The sub-lemma then follows by recalling (7.14), that ϱ must be larger than ϱ_0, defined in (7.19), as well as larger than either (7.27) or (7.29), and provided that we choose $\tilde{C}_0 \geq 3C_{10}$. □

Remark 7.4. *The above Lemma is based on a trade-off: for small deviations (up to the ones predicted by the Central Limit Theorem) it gives very rough estimates, but up to times of order one; while for larger deviations it provides much sharper results although only for short times. Obtaining sharper results for small deviations would entail more work, in particular a sharper version of Lemma 5.4 (which can be achieved by using the techniques that we will employ in the next sections to prove a local CLT). On the other hand, in order to extend the above sharp results to longer times one can simply divide the time interval in shorter ones and use Lemma 7.2 repeatedly (see Theorem 2.4 for an implementation of this strategy).*

7.2. Lower bound for balls (short times).

In this section we proceed to obtain a lower bound for short times. It turns out that in [12] we need this type of estimates only in the case of large deviations; therefore we do not insist here in obtaining optimal bounds for all moderate deviations, since this would require considerable additional work. On the other hand, we can, and will, obtain results for deviations that are not exceedingly small at essentially no extra cost. Also, in an attempt to simplify the exposition, we will consider trajectories that are not arbitrarily close to being impossible, i.e. so that their derivatives belong to the domains $\mathbb{D}_\epsilon(\theta) = \mathbb{D}(\theta) \setminus \partial_\epsilon\mathbb{D}(\theta)$ for some arbitrarily small, but fixed, ϵ.

Lemma 7.5 (Lower bound)**.** *For any $\delta \in (0, \frac{1}{4})$ there exists $\varepsilon_\delta, K_\delta, T_{\max} > 0$ such that, for any $\theta \in \mathbb{T}$, $\varepsilon \leq \varepsilon_\delta$, $\underline{\gamma} \in \mathrm{Lip}_{C,*}([0,T])$, with $\underline{\gamma}'(s) \in \mathbb{D}_\delta(\bar{\theta}(s,\theta))$ for almost* *all $s \in [0,T]$, and for any standard pair ℓ so that $\theta^*_\ell = \theta$:*

$$\varepsilon \log \mathbb{P}_{\ell,\varepsilon}(B(\underline{\gamma},h)) \geq -I_\theta(\underline{\gamma}) - c_\delta(T\sqrt{\varepsilon/h} + T^2 + T\bar{R}_T/h))$$
$$\bar{R}_T = K_\delta[T^{3/2} + \max\{\varepsilon^{1/4}T^{3/4}, K_\delta^{-1/2}T(\varepsilon/h)^{1/4}\}]$$

provided $T \in [\varepsilon_\delta^{-1}\varepsilon, T_{\max}]$, $h \in [\bar{R}_T, T]$.

Proof. Before describing the core of the proof we need some preparation: we must define a new reference path having several special properties. To this end we first perform the same polygonalization done just before (7.10), with step $h = T/N_h$, $N_h \in \mathbb{N}$ and $h \geq 2\varepsilon$. Let $\underline{\gamma}_h$ the resulting path. Clearly $\|\underline{\gamma} - \underline{\gamma}_h\|_{\mathcal{C}^0} \leq Ch \leq \delta/4$, provided h is small enough, but it also has another important property.

Sub-lemma 7.6. *For all $\delta > 0$ and $\underline{\gamma} \in \mathrm{Lip}_{C,*}([0,T])$, with $\underline{\gamma}'(s) \in \mathbb{D}_\delta(\bar{\theta}(s,\theta))$ for almost all $s \in [0,T]$, there exists $h_\delta \in (0,T]$ such that, for all $h \leq h_\delta$, $\underline{\gamma}_h'(s) \in \mathbb{D}_{\delta/4}(\bar{\theta}(s,\theta))$ for almost all $s \in [0,T]$.*

Proof. By Lemma 6.7, $\mathbb{D}(\bar{\theta}(s,\theta))$ varies continuously, in the Hausdorff topology, with respect to s. Hence, for each $\delta > 0$ there exists $h_\delta > 0$ such that $\mathbb{D}_\delta(\bar{\theta}(s',\theta)) \subset \mathbb{D}_{\delta/2}(\bar{\theta}(s,\theta)) \subset \mathbb{D}_{\delta/4}(\bar{\theta}(s',\theta))$ for all $|s'-s| \leq h_\delta$. Accordingly, for all $h \leq h_\delta$, $k \in \{0,\cdots,N_h\}$ and $s \in (kh,(k+1)h)$ we have $\underline{\gamma}'(s) \in \mathbb{D}_{\delta/2}(\bar{\theta}(kh,\theta))$. By the convexity of the set it follows $\underline{\gamma}_h'(s) \in \mathbb{D}_{\delta/2}(\bar{\theta}(kh,\theta)) \subset \mathbb{D}_{\delta/4}(\bar{\theta}(s,\theta))$, hence the sub-lemma. □

The above sub-lemma implies that the equation $\underline{\gamma}_h'(s) = \partial_\sigma \chi_A(\sigma(s), \bar{\theta}(s,\theta))$ has a unique solution $\underline{\sigma}_h(s) \in \mathrm{BV}$. Moreover, from Lemmata 6.7, 6.2 it also follows that there exists $\hat{R}_\delta > 0$ such that $\partial_\sigma\chi_A(B(0,\hat{R}_\delta), \bar{\theta}(s,\theta)) \supset \mathbb{D}_{\delta/4}(\bar{\theta}(s,\theta))$ for all $s \in [0,T]$. This allows to obtain also some regularity of the function $\underline{\sigma}_h \in \mathrm{BV}$. Indeed, for $k \in \{0,\cdots,N_h\}$ and $s \in (kh,(k+1)h)$ we have

$$0 = \partial^2_\sigma \chi_A(\underline{\sigma}_h(s), \bar{\theta}(s,\theta))\underline{\sigma}_h'(s) + \partial_\theta\partial_\sigma\chi_A(\underline{\sigma}_h(s), \bar{\theta}(s,\theta))\bar{A}(\bar{\theta}(s,\theta).$$

Since $\|\underline{\sigma}_h\|_{L^\infty} \leq \hat{R}_\delta$ and $\inf_{\|\sigma\| \leq \hat{R}_\delta, s} \partial^2_\sigma \chi_A > 0$, there exists $C_\delta > 0$ such that $\|\underline{\sigma}_h'(s)\| \leq C_\delta$.

We have thus obtained a path with a controlled regularity. Unfortunately, we have a further problem: it is not obvious how to compare effectively the rate function computed on the regularized path and the rate function computed on the original one. To this end it turns out to be convenient to further modify the path in a special way: consider a path $\gamma_h \in \mathcal{C}^0([0,T],\mathbb{R}^d)$ such that $\gamma_h(kh) = \underline{\gamma}(kh) = \underline{\gamma}_h(kh)$ for all $k \in \{0,\cdots,N_h\}$ with the following extra property: a) γ_h is Lipschitz, b) the equation $\gamma_h'(s) = \partial_\sigma\chi_A(\sigma(s), \bar{\theta}(s,\theta))$, which by (a) is well defined for almost every $s \in [0,T]$, has a solution σ_h which is constant in each interval $[nh,(n+1)h)$.

The reason for the above construction lies in item (c) of the next lemma.

Sub-lemma 7.7. *There exists $h_\delta > 0$ such that, for all $h \leq h_\delta \in (0,T]$, the above defined path γ_h is well defined and unique. In addition,*

(a) $\gamma_h'(s) \in \mathbb{D}_{\delta/8}(\bar{\theta}(s,\theta))$ *for almost all $s \in [0,T]$;*
(b) $\|\underline{\gamma} - \gamma_h\|_{\mathcal{C}^0} \leq 2Ch$ *;*
(c) $I_\theta(\underline{\gamma}) \geq I_\theta(\gamma_h)$.

Proof. We can change the path interval by interval. Assume that we have a path $\gamma_{h,k}$ that has the wanted properties in the interval $[0,kh]$ and that agrees with $\underline{\gamma}_h$ in the interval $(kh,T]$, and let us consider the the interval $J_k = (kh,(k+1)h]$. Define the function $\Xi_k : L^\infty([0,T]) \times \mathbb{R}^{d+1} \to L^\infty(J_k) \times \mathbb{R}$ given by

$$\Xi_k(\eta,\zeta,\beta) = \left(\underline{\gamma}_h'(s) + \eta(s) - \partial_\sigma\chi_A(\zeta + (1-\beta)\underline{\sigma}_h(s), \bar{\theta}(s,\theta)), \int_{J_k} \eta(s)ds\right).$$

By definition we have $\Xi_k(0,0,0) = 0$. We want to apply the implicit function theorem, hence we have to study the differential

$$\partial_{\eta,\zeta}\Xi_k = \begin{pmatrix} \mathbb{1} & -\partial_\sigma^2\chi_A(\zeta + (1-\beta)\underline{\sigma}_h(s), \bar\theta(s,\theta)) \\ \text{Leb} & 0 \end{pmatrix}.$$

Using Lemmata A.1, A.16, A.9, a direct computation shows that, provided $\|\zeta\| \le 2\hat R_\delta$, $\|(\partial_{\eta,\zeta}\Xi_k)^{-1}\| \le C_\delta$. We can then apply the Implicit Function Theorem and obtain a solution $(\eta(\beta,s),\zeta(\beta))$ of $\Xi_k(\eta,\zeta,\beta) = 0$. Such a solution is differentiable and satisfies

$$\partial_\beta\eta(s) - \partial_\sigma^2\chi_A(\zeta + (1-\beta)\underline{\sigma}_h(s), \bar\theta(s,\theta))[\partial_\beta\zeta - \underline{\sigma}_h(s)] = 0$$
$$\int_{J_k}\partial_\beta\eta(s)ds = 0.$$

Integrating the first and using the second equation yields

$$\partial_\beta\eta(s) = B(s,\zeta,\beta)[\partial_\beta\zeta - \underline{\sigma}_h(s)]$$
$$\partial_\beta\zeta = \left[\int_{J_k} B(s',\zeta,\beta)ds'\right]^{-1}\int_{J_k} B(s',\zeta,\beta)\underline{\sigma}_h(s')ds'$$
$$B(s,\zeta,\beta) = \partial_\sigma^2\chi_A(\zeta + (1-\beta)\underline{\sigma}_h(s), \bar\theta(s,\theta)).$$

Recalling that $\|\underline{\sigma}_h'(s)\| \le C_\delta$ (which was the point of introducing the regularized path $\underline{\gamma}_h$ in the first place), we have $\|B(s,\zeta,\beta) - B(kh,\zeta,\beta)\| \le C_\delta h$. Thus,

$$\left\|\partial_\beta\zeta - h^{-1}\int_{J_k}\underline{\sigma}_h(s)ds\right\| \le C_\delta h \le \hat R_\delta$$
$$\|\partial_\beta\eta(s)\| \le C_\delta h \le \delta/8.$$

provided h is small enough. Accordingly, $\underline{\gamma}_h'(s) + \eta(\beta,s) \in \mathbb{D}_{\delta/8}(\bar\theta(s,\theta))$ for all $\beta \le 1$. The above shows that $\gamma_{h,k+1}$ and hence, by induction, γ_h is uniquely defined and point (b) of the lemma follows as well for δ small enough. To prove (c), we use Lagrange multipliers to find the local minimum among all the paths $\gamma \in \mathcal{C}^0([0,T],\mathbb{R}^d)$ such that $\gamma(kh) = \underline{\gamma}(kh)$. This amount to finding the stationary points of the functional

$$I_L(\gamma,\phi_i) = I(\gamma) + \sum_{k=1}^{h^{-1}T}\langle\phi_k, \gamma(kh) - \underline{\gamma}(kh)\rangle.$$

By (6.3) and (6.4) it follows that, for each γ such that $I(\gamma) < \infty$, we have, for each $\alpha \in \text{Lip}$, $\alpha(0) = 0$,

$$\partial_\gamma I_L(\gamma)(\alpha) = \int_0^T\langle\sigma(s),\alpha'(s)\rangle ds + \sum_{k=1}^{h^{-1}T}\langle\phi_k,\alpha(kh)\rangle = \int_0^T\langle(\sigma(s) - \sigma_\phi(s)),\alpha'(s)\rangle ds$$

where $\gamma'(s) = \partial_\sigma\chi_A(\sigma(s),\bar\theta(s,\theta))$ and $\sigma_\phi(s) = -\sum_{j=k}^{h^{-1}T}\phi_j$ for $s \in [(k-1)kh, kh)$. Hence it must be $\sigma = \sigma_\phi$, that is σ is piecewise constant. But only γ_h has such a property, thus the convex function $t : \mathbb{R} \to I_\theta(t\gamma_h + (1-t)\underline{\gamma})$ has a unique stationary point that must be a minimum, hence point (c) of the sub-lemma. □

Note that, by the above sub-lemma, the equation $\gamma_h'(s) = \partial_\sigma\chi_A(\sigma(s),\bar\theta(s,\theta))$ has a unique solution $\sigma_h(s) \in \text{BV}$, $\|\sigma_h\|_{L^\infty} \le C_\delta$, $\|\sigma_h\|_{\text{BV}} \le C_\delta h^{-1}T$. Also, if $Ch > \bar R_T$, then we have

$$B(\underline{\gamma}, 3Ch) \supset B(\gamma_h, \bar R_T). \tag{7.37}$$

This concludes our preparation; the rest of the proof follows the strategy strategy for proving the lower bound. Consider the linear functional $\varphi_h \in \mathcal{M}^d([0,T])$ defined by

$$\varphi_h(\gamma) = \varepsilon^{-1} \int_0^T \langle \sigma_h(s), \gamma'(s) \rangle ds, \tag{7.38}$$

and introduce the measure $\mathbb{P}_{\varphi_h,\ell,\varepsilon}$ on $\mathcal{C}^0([0,T],\mathbb{R}^d)$ defined by

$$\mathbb{E}_{\varphi_h,\ell,\varepsilon}(\psi) = \frac{\mathbb{E}_{\ell,\varepsilon}(e^{\varphi_h}\psi)}{\mathbb{E}_{\ell,\varepsilon}(e^{\varphi_h})},$$

for any continuous functional $\psi \in \mathcal{C}([0,T],\mathbb{R}^d)'$.

Sub-lemma 7.8. *There exists $\varepsilon_\delta, K_\delta, T_{\max} > 0$ such that for any $0 < \varepsilon \le \varepsilon_\delta$:*

$$\mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_h, \bar{R}_T)) \ge \frac{1}{2},$$

provided $T \in [\varepsilon_\delta^{-1}\varepsilon, T_{\max}]$, $\bar{R}_T \ge K_\delta[T^{3/2} + \max\{\varepsilon^{1/4}T^{3/4}, K_\delta^{-1/2}T(\varepsilon/h)^{1/4}\}]$ and $h \in [\bar{R}_T/(3C), T]$.

Proof. The idea is to cover the complement of $B(\gamma_h, \bar{R}_T)$ in the support of $\mathbb{P}_{\varphi_h,\ell,\varepsilon}$ with finitely many sufficiently small balls whose measure we can estimate using the upper bound obtained in Lemma 7.2. In order to do so, let us partition the interval $[0,T]$ in subintervals of length $c_\# r_\varepsilon$, $r_\varepsilon < h$, so that any path in the support of the measure (and hence C-Lipschitz) can vary in any given subinterval by at most $2r_\varepsilon$. This means that there exists a finite set $\Gamma = \{\gamma_i\} \subset \mathcal{C}^0([0,T],\mathbb{R}^d)$ of cardinality $|\Gamma| = \exp(c_\# r_\varepsilon^{-1}T)$ so that[43]

$$\operatorname{supp} \mathbb{P}_{\varphi_h,\ell,\varepsilon} \subset \bigcup_{\gamma_i\in\Gamma} B(\gamma_i, r_\varepsilon).$$

Define $\Gamma_* = \{\gamma_i \in \Gamma \ : \ \|\gamma_h - \gamma_i\|_\infty \ge \bar{R}_T/2\}$: then, by definition,

$$\mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_h, \bar{R}_T)) \ge 1 - \sum_{\gamma_i\in\Gamma_*} \mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_i, r_\varepsilon)).$$

Let $r_\varepsilon \in [K_\delta^2 \bar{R}_T^{-2}\varepsilon T^2, \bar{R}_T/8]$; observe that the interval is not empty provided $\bar{R}_T \ge 2\varepsilon^{1/3}T^{2/3}K_\delta^{2/3}$, which always holds if ε_δ is small enough. We claim that, for all $\gamma_i \in \Gamma_*$,

$$\mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_i, r_\varepsilon)) < e^{-C_\delta\varepsilon^{-1}\bar{R}_T^2 T^{-1}}. \tag{7.39}$$

Observe that the above estimate suffices to prove the sub-lemma: in fact, using our estimate on the cardinality of Γ, we have

$$\sum_{\gamma_i\in\Gamma_*} \mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_i, r_\varepsilon)) \le e^{-C_\#([C_\delta - c_\# K_\delta^{-2}]\varepsilon^{-1}\bar{R}_T^2 T^{-1}} \le \frac{1}{2},$$

provided we choose K_δ large enough.

We thus proceed to prove (7.39). Using (7.4) we gather

$$\begin{aligned}\mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_i, r_\varepsilon)) &= e^{-\varepsilon^{-1}\Lambda_{\ell,\varepsilon}(\varphi_h)}\mathbb{E}_{\ell,\varepsilon}\left(e^{\varphi_h}\mathbb{1}_{B(\gamma_i,r_\varepsilon)}\right) \\ &\le e^{-\varepsilon^{-1}\Lambda_{\ell,\varepsilon}(\varphi_h)+\varphi_h(\gamma_i)+\varepsilon^{-1}r_\varepsilon\|\sigma_h\|_{\mathrm{BV}}}\mathbb{P}_{\ell,\varepsilon}\left(B(\gamma_i, r_\varepsilon)\right),\end{aligned}$$

---

[43] In fact, consider a lattice of size $r_\epsilon\sqrt{4/d}$ in $\mathbb{R}^d$. If a path γ is in the support of the measure and $\gamma(s)$ belongs to the lattice, by the Lipschitz property $\gamma(s + c_\# r_\varepsilon)$ belongs to the union of balls of radius r_ε centered at *finitely many* points of the lattice.

where in the second line we used the fact that, for any $\gamma \in B(\gamma_i, r_\varepsilon)$, we have by definition (7.38) (recall Remark 5.2),

$$
|\varphi_h(\gamma) - \varphi_h(\gamma_i)| = \varepsilon^{-1}\left|\int \langle \sigma_h, \gamma' - \gamma_i'\rangle\right| \leq \varepsilon^{-1}\|\sigma_h\|_{\mathrm{BV}}\|\gamma - \gamma_i\|_{L^\infty} \\ \leq \varepsilon^{-1}\|\sigma_h\|_{\mathrm{BV}}\, r_\varepsilon. \tag{7.40}
$$

Then, using (7.7) and (7.8):

$$
\mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_i, r_\varepsilon)) \leq e^{\varepsilon^{-1}[-\mathcal{R}_{\ell,\varepsilon}(\sigma_h) + r_\varepsilon\|\sigma_h\|_{\mathrm{BV}}]} \\ \cdot e^{\varepsilon^{-1}\left[\int_0^T \langle \sigma_h(s), \gamma_i'(s)\rangle - \chi_A(\sigma_h(s), \bar\theta(s,\theta))ds\right]}\mathbb{P}_{\ell,\varepsilon}\left(B(\gamma_i, r_\varepsilon)\right). \tag{7.41}
$$

Next, let us set $\Xi_i = \inf_{\gamma \in B(\gamma_i, \bar R_T/4)} I_\theta(\gamma)$; we claim that

$$
\mathbb{P}_{\ell,\varepsilon}\left(B(\gamma_i, r_\varepsilon)\right) \leq e^{-\varepsilon^{-1}\Xi_i}. \tag{7.42}
$$

In fact, if $\|\hat\gamma_i\|_\infty \geq 2CT$, then $\|\hat\gamma\|_\infty \geq 3/2CT$ for all $\gamma \in B(\gamma_i, r_\varepsilon)$, hence their Lipschitz constant must be larger than C and hence $\mathbb{P}_{\ell,\varepsilon}\left(B(\gamma_i, r_\varepsilon)\right) = 0$. Otherwise (7.42) follows from Lemma 7.2, provided that $R_\varepsilon(\gamma_i) + r_\varepsilon < \bar R_T/4$. This holds since one can check that $R_\varepsilon(\gamma_i) \leq C_\#\{\varepsilon^{1/4}T^{3/4} + T^2\}$, hence $R_\varepsilon(\gamma_i) \leq \bar R_T/8$ by choosing K_δ sufficiently large. Substituting (7.42) in (7.41), recalling (6.1), (6.2), (6.3), Sub-Lemma 6.12 and Remark 6.3:

$$
\mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_i, r_\varepsilon)) \leq e^{\varepsilon^{-1}[-\mathcal{R}_{\ell,\varepsilon}(\sigma_h) + r_\varepsilon\|\sigma_h\|_{\mathrm{BV}}]} \\ \cdot e^{\varepsilon^{-1}\sup_{\gamma\in B(\gamma_i,\bar R_T/4)\cap \mathrm{Lip}}\int_0^T\left[\langle \sigma_h(s), \gamma_i'(s)\rangle - \chi_A(\sigma_h(s), \bar\theta(s,\theta)) - \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta))\right]ds} \\ \leq e^{\varepsilon^{-1}[-\mathcal{R}_{\ell,\varepsilon}(\sigma_h) + r_\varepsilon\|\sigma_h\|_{\mathrm{BV}}]} \\ \cdot e^{\varepsilon^{-1}\sup_{\gamma\in B(\gamma_i,\bar R_T/4)}\int_0^T\left[\langle \sigma_h(s), \gamma'(s) - \gamma_h'(s)\rangle + \mathcal{Z}(\gamma_h'(s), \bar\theta(s,\theta)) - \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta))\right]ds}. \tag{7.43}
$$

Next, we proceed to estimate the argument of the sup appearing on the last line: note that we only need to consider paths γ so that $s \mapsto \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta))$ is integrable on $[0,T]$, hence $\gamma'(s) \in \mathbb{D}$ for almost all s. Then, for any $\varrho \in [0,1]$, let us define the interpolating function

$$
H(\varrho, s) := \langle \sigma_h(s), \varrho(\gamma'(s) - \gamma_h'(s))\rangle \\ - \mathcal{Z}(\gamma_h'(s) + \varrho(\gamma'(s) - \gamma_h'(s)), \bar\theta(s,\theta)) + \mathcal{Z}(\gamma_h'(s), \bar\theta(s,\theta)).
$$

Let σ_ϱ be the solution of $\gamma_h' + \varrho(\gamma' - \gamma_h') = \partial_\sigma\hat\chi_A(\sigma_\varrho, \bar\theta(\cdot,\theta))$, which exists, for all $\varrho < 1$, by the convexity of $\mathbb{D}$. In addition, note that, if $\varrho \leq 1/2$, then $\gamma_h' + \varrho(\gamma' - \gamma_h') \in \mathbb{D}_{\delta/16}$. Note that $H(0,s) = 0$ and

$$
\partial_\varrho H(0,\cdot) = \langle \sigma_h - \bar\sigma, \gamma' - \gamma_h'\rangle - \langle \bar\sigma, \partial_b\mathcal{Z}(\gamma_h'(s), \bar\theta(s,\theta)), \gamma' - \gamma_h'\rangle = 0, \\
\partial_{\varrho\varrho}H(\varrho,\cdot) = -\langle [\partial_\sigma^2\hat\chi_A(\sigma_\varrho)]^{-1}(\gamma' - \gamma_h'), \gamma' - \gamma_h'\rangle \leq 0, \\
\partial_{\varrho\varrho}H(\varrho,\cdot) \leq -C_\delta\|\gamma' - \gamma_h'\|^2 \quad \text{for all } \varrho \leq 1/2.
$$

Thus

$$
\langle \sigma_h, \gamma' - \gamma_h'\rangle - \mathcal{Z}(\gamma', \bar\theta(\cdot,\theta)) + \mathcal{Z}(\gamma_h', \bar\theta(\cdot,\theta)) = H(1) \leq -C_\delta\|\gamma' - \gamma_h'\|^2.
$$

The term containing $\mathcal{R}_{\ell,\varepsilon}$ can be estimated by Proposition 5.4-(a); by our bounds on σ_h and choosing $L = \varepsilon^{-1/2}h^{1/2}$ we obtain, since $\sqrt{\varepsilon h} \leq T$,

$$
|\mathcal{R}_{\ell,\varepsilon}(\sigma_h)| \leq c_*^{-1}C_\delta\varepsilon Lh^{-1}T + C_\delta(L^{-1} + T + \varepsilon Lh^{-1}T)T \leq C_\delta(T\sqrt{\varepsilon/h} + T^2). \tag{7.44}
$$

Thus,

$$
\mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_i, r_\varepsilon)) \leq e^{-C_\delta\varepsilon^{-1}\left[\inf_{\gamma\in B(\gamma_i,\bar R_T/4)}\|\gamma' - \gamma_h'\|_{L^2}^2 - C_\delta(T\sqrt{\varepsilon/h} + T^2) - r_\varepsilon\|\sigma_h\|_{\mathrm{BV}}\right]}.
$$

To conclude, note that, for $\gamma_i \in \Gamma_*$, if $\gamma \in B(\gamma_i, \bar R_T/4)$ we have

$$\bar R_T/4 \le \|\gamma - \gamma_h\|_{L^\infty} \le \|\gamma' - \gamma_h'\|_{L^1} \le \|\gamma' - \gamma_h'\|_{L^2}\sqrt{T},$$

which, by choosing K_δ large and $r_\varepsilon = K_\delta^2 \bar R_T^{-2}\varepsilon T^2$, proves (7.39). □

To conclude the proof of Lemma 7.5 it suffices to compare the measures $\mathbb{P}_{\varphi_h,\ell,\varepsilon}$ and $\mathbb{P}_{\ell,\varepsilon}$. By Sub-Lemma 7.8, equations (7.4), (7.8), and using (7.44), and arguing like in (7.40) we have

$$\begin{aligned}\frac12 &\le \mathbb{P}_{\varphi_h,\ell,\varepsilon}(B(\gamma_h,\bar R_T)) = \frac{\mathbb{E}_{\ell,\varepsilon}\left(e^{\varphi_h}\mathbb{1}_{B(\gamma_h,\bar R_T)}\right)}{\mathbb{E}_{\ell,\varepsilon}(e^{\varphi_h})}\\ &\le C_\# e^{\varepsilon^{-1}\left[\int_0^T\langle\sigma_h,\gamma_h'\rangle - \chi_A(\sigma_h,\bar\theta(s,\theta))ds + C_\delta(T\sqrt{\varepsilon/h}+T^2)\right]}\,\mathbb{E}_{\ell,\varepsilon}\left(e^{\varphi_h-\varphi_h(\gamma_h)}\mathbb{1}_{B(\gamma_h,\bar R_T)}\right)\\ &\le C_\# e^{\varepsilon^{-1}\left[I_\theta(\gamma_h)+C_\delta(T\sqrt{\varepsilon/h}+T^2)+\|\sigma_h\|_{\mathrm{BV}}\bar R_T\right]}\,\mathbb{P}_{\ell,\varepsilon}(B(\gamma_h,\bar R_T)).\end{aligned}$$

By (7.37), Sub-Lemma 7.7-(c) and $h, \bar R_T$ as in Sub-Lemma 7.8 we have

$$\mathbb{P}_{\ell,\varepsilon}(B(\underline{\gamma},3Ch)) \ge \mathbb{P}_{\ell,\varepsilon}(B(\gamma_h,\bar R_T)) \ge e^{-\varepsilon^{-1}(I(\gamma_h)+c_\delta(T\sqrt{\varepsilon/h}+T^2+T\bar R_T/h))}.$$

The Lemma follows by choosing $\bar R_T$ as small as allowed and renaming $3Ch$ as h. □

7.3. **Large and moderate deviations for balls: long times.**

Lemmata 7.2 and 7.5 are the basic ingredients to prove Theorem 2.4. Their major drawback is that they are really effective only for short times. In order to proceed and obtain a Large Deviation estimate for times of order 1 with the announced small error, we will subdivide a trajectory in shorter subintervals and apply the mentioned lemmata to each subinterval. To this end, some type of Markov-like property is needed. Before stating it in Lemma 7.9, we need to introduce a bit of notation.

For any $J \subset J' \subset [0,T]$, $\underline{\gamma} \in \mathcal{C}^0(J',\mathbb{R}^d)$ and $r > 0$, we introduce the notation

$$B|_J(\underline{\gamma}, r) = \left\{\gamma \in C^0([0,T],\mathbb{R}^d)\ :\ \|\gamma - \underline{\gamma}\|_{L^\infty(J)} < r\right\}, \tag{7.45}$$

where $\|\gamma\|_{L^\infty(J)} = \sup_{s\in J}\|\gamma(s)\|$. In other words, $B|_J(\underline{\gamma}, r)$ is a set of paths in $C^0([0,T],\mathbb{R}^d)$ that are r-close to $\bar\gamma$ on J, but are otherwise arbitrary on $[0,T]\setminus J$. Naturally, if $J = [0,T]$, the set $B(\underline{\gamma}, r) = B|_{[0,T]}(\underline{\gamma}, r)$ is the standard C^0-ball of radius r around $\underline{\gamma}$.

Let us fix $\theta \in \mathbb{T}$; consider a path $\gamma \in \mathcal{C}^0([0,T],\mathbb{R}^d)$ and a number $r > 0$. For any standard pair $\ell = (\mathbb{G},\rho)$ and $t \in [0,T]$ we define $\mathbb{1}^{\pm}_{\theta,\gamma,r}(\ell,t)$ as follows:

$$\mathbb{1}^-_{\theta,\gamma,r}(\ell,t) = \begin{cases}1 & \text{if } \inf_x\{|G_\ell(x) - \theta^\gamma(t)|\} < r\\ 0 & \text{otherwise}\end{cases} \tag{7.46a}$$

$$\mathbb{1}^+_{\theta,\gamma,r}(\ell,t) = \begin{cases}1 & \text{if } \sup_x\{|G_\ell(x) - \theta^\gamma(t)|\} < r\\ 0 & \text{otherwise,}\end{cases} \tag{7.46b}$$

where we used the previously introduced notation $\theta^\gamma(s) = \theta^\gamma(s,\theta) = \theta + \gamma(s)_1$, and $\gamma(s)_1$ denotes the first component of the vector $\gamma(s)$. Observe that, by construction, $\mathbb{1}^+_{\theta,\gamma,r}(\ell,t) \le \mathbb{1}^-_{\theta,\gamma,r}(\ell,t)$. Finally, for any interval $E = [s,t] \subset [0,T]$, let

$$\begin{aligned}P_-(E,\theta,\gamma,r) &= \sup_{\{\ell\,:\,\mathbb{1}^-_{\theta,\gamma,r}(\ell,s)=1\}} \mathbb{P}_{\ell,\varepsilon}(B|_{[0,t-s]}(\gamma(s+\cdot)-\gamma(s),r)),\\ P_+(E,\theta,\gamma,r) &= \inf_{\{\ell\,:\,\mathbb{1}^+_{\theta,\gamma,r}(\ell,s)=1\}} \mathbb{P}_{\ell,\varepsilon}(B|_{[0,t-s]}(\gamma(s+\cdot)-\gamma(s),r));\end{aligned} \tag{7.47}$$

observe that by construction $P_+(E,\theta,\gamma,r) \le P_-(E,\theta,\gamma,r)$.

Lemma 7.9. *For $T > 0$, $k \in \{0, \cdots, K-1\}$ and $\tau = TK^{-1} \geq \varepsilon$ let $E_k = [k\tau, (k+1)\tau]$. Then, for any $\theta^* \in \mathbb{T}$ and standard pair ℓ_0 with $\theta^*_{\ell_0} = \theta^*$ we have:*

$$\prod_{k=0}^{K-1} P_+(E_k, \theta^*, \underline{\gamma}, r - C_\# \varepsilon) \leq \mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma}, r)) \leq \prod_{k=0}^{K-1} P_-(E_k, \theta^*, \underline{\gamma}, r + C_\# \varepsilon),$$

where the first inequality holds for $\underline{\gamma} \in \mathcal{C}^0$, while the second for $\underline{\gamma} \in \operatorname{supp} \mathbb{P}_{\ell_0,\varepsilon}$.[44]

Proof. For each $k \in \mathbb{N}$ let $\mathfrak{L}^{(k)}_{\ell_0,0}$ be a standard family representing $(F^k_\varepsilon)_* \mu_{\ell_0}$, see Proposition 3.3 and Remark 3.7 for exact definitions. Let $\pi_z : \mathbb{T}^2 \times \mathbb{R}^{d-1} \to \mathbb{T} \times \mathbb{R}^{d-1}$ be the projection on the last d (i.e. slow) coordinates. Let $\ell' \in \mathfrak{L}^{(k)}_{\ell,0}$, by Remark 3.7, definition (2.6) and the expansivity of the x dynamics, it follows that, for all $j \leq k$,

$$\begin{aligned}
&\sup_{x,x'} \left\| \mathbb{1}_{\ell'}(x) \mathbb{1}_{\ell'}(x') [\pi_z \mathbb{F}^j(x, G_\ell(x), 0) - \pi_z \mathbb{F}^j(x', G_\ell(x'), 0)] \right\| \\
&= \sup_{x,x'} \left\| \mathbb{1}_{\ell'}(x) \mathbb{1}_{\ell'}(x') \varepsilon \sum_{m=0}^{j-1} [A \circ F^m(x, G_\ell(x)) - A \circ F^m(x', G_\ell(x'))] \right\| \\
&\leq C_\# \varepsilon \lambda^{-k+j}.
\end{aligned} \tag{7.48}$$

Let us define, for $j, k \in \mathbb{N}$, $\ell' \in \mathfrak{L}^{(k)}_{\ell,0}$, $E = [j\varepsilon, (j+k)\varepsilon]$, $\mathbb{1}^-_{B|_E(\underline{\gamma},r)}(\ell, \ell') = 1$ if

$$\sup_{m \leq k} \inf_x \| \mathbb{1}_{\ell'}(x) [\underline{\gamma}(\varepsilon(j+m)) - \underline{\gamma}(\varepsilon j) + (G_\ell(x), 0) - \pi_z \mathbb{F}^m(x, G_\ell(x), 0)] \| \leq r,$$

while $\mathbb{1}^-_{B|_E(\underline{\gamma},r)}(\ell, \ell') = 0$ otherwise. Note that, $\mathbb{1}^-_{B|_{[0,t]}(\underline{\gamma},r)}(\ell, \ell') \leq \mathbb{1}^-_{\theta^*_\ell, \underline{\gamma}, r + C_\# \varepsilon}(\ell', t)$. Also, $\mathbb{1}^-_{B|_{[0,t]}(\underline{\gamma}, r + C_\# \varepsilon)}(\ell, \ell') \geq \mathbb{1}^-_{\theta^*_\ell, \underline{\gamma}, r}(\ell', t)$ since, if $\mathbb{1}^-_{\theta^*_\ell, \underline{\gamma}, r}(\ell', t) = 1$, then, by (7.48),

$$\sup_{m \leq k} \sup_x \| \mathbb{1}_{\ell'}(x) [\underline{\gamma}(\varepsilon(j+m)) - \underline{\gamma}(\varepsilon j) + (G_\ell(x), 0) - \pi_z \mathbb{F}^m(x, G_\ell(x), 0)] \| \leq r + C_\# \varepsilon.$$

Setting $\underline{\gamma}_k(s) = \underline{\gamma}(k\tau + s) - \underline{\gamma}(k\tau)$, by Proposition 3.3 and (7.47) it follows that

$$\begin{aligned}
&\mathbb{P}_{\ell_0,\varepsilon}\left(B(\underline{\gamma}, r)\right) \leq \mu_{\ell_0}\left(\prod_{k=0}^{\varepsilon^{-1}T} \mathbb{1}_{B|_{\{k\varepsilon\}}(\underline{\gamma} + (\theta_0, 0), r)} \circ \mathbb{F}^k_\varepsilon \right) \\
&\leq \sum_{\ell' \in \mathfrak{L}^{K\tau/\varepsilon}_{\ell_0,0}} \nu_{\ell'} \mathbb{1}^-_{B|_{[0,K\tau]}(\underline{\gamma},r)}(\ell_0, \ell') \\
&\leq \sum_{\ell' \in \mathfrak{L}^{(K-1)\tau/\varepsilon}_{\ell_0,0}} \nu_{\ell'} \mathbb{1}^-_{B|_{[0,(K-1)\tau]}(\underline{\gamma},r)}(\ell_0, \ell') \mathbb{P}_{\ell',\varepsilon}(B|_{[0,\tau]}(\underline{\gamma}_{K-1}, r + C_\# \varepsilon)) \\
&\leq \sum_{\ell' \in \mathfrak{L}^{(K-1)\tau/\varepsilon}_{\ell_0,0}} \nu_{\ell'} \mathbb{1}^-_{B|_{[0,(K-1)\tau]}(\underline{\gamma},r)}(\ell_0, \ell') P_-(E_{K-1}, \theta^*, \underline{\gamma}, r + C_\# \varepsilon).
\end{aligned} \tag{7.49}$$

Iterating, and arguing similarly for the lower bound, the lemma follows.[45] □

We are finally ready to prove our main Large Deviations result; the proof is divided in two steps; the first step is to obtain upper and lower bounds for the probability of a small ball around a trajectory of length of order 1 and it is given by Lemma 7.10 below. The second will be carried out in the proof of Theorem 2.4.

---

[44] In particular, $\underline{\gamma}$, is C-Lipschitz and piecewise linear in each interval $[k\varepsilon, (k+1)\varepsilon]$.

[45] Note that if $\underline{\gamma} \in \operatorname{supp} \mathbb{P}_{\ell_0,\varepsilon}$, then in the first line of (7.49) holds equality.

Lemma 7.10. *Let $T > 0$, $\theta \in \mathbb{T}$, $\underline{\gamma} \in \mathcal{C}^0([0,T],\mathbb{R}^d)$ and, recalling the notation in (2.12), $R^+(\gamma) = C_{\epsilon,T}\left\{\varepsilon^{1/7}\|\hat\gamma\|_{L^\infty}^{5/7} + \sqrt{\varepsilon}\right\}$, where $\hat\gamma = \gamma - \bar\gamma(\cdot,\theta)$; let moreover $\varsigma = (T^{1/7}\|\hat{\underline{\gamma}}\|_{L^\infty}^{-1/7}\varepsilon^{4/7})$, $\varsigma_- = \varepsilon^{\frac12}$. Then for any $\epsilon > 0$, if ε is sufficiently small, we have, for any ℓ_0 with $\theta^*_{\ell_0} = \theta^*$:*

$$\varepsilon \log \mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma},\varsigma/3)) \le -\left[1 - C_\epsilon\left(\frac{T^2\varepsilon}{\|\hat{\underline{\gamma}}\|^2_{L^\infty}}\right)^{1/7}\right]\inf_{\gamma\in B(\underline{\gamma},R^+(\underline{\gamma}))}\mathscr{I}^-_{\theta^*,\epsilon}(\gamma). \tag{7.50a}$$

$$\varepsilon \log \mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma},\varsigma_-)) \ge -(1 + C_\epsilon\varepsilon^{1/2})\mathscr{I}^+_{\theta^*,\epsilon}(\underline{\gamma}) - C_{\epsilon,T}\varepsilon^{1/8}. \tag{7.50b}$$

Proof. In accordance with our conventions we will use C_ϵ or C_T to designate an arbitrary constant depending only on the values of ε or T, respectively.

We begin by estimating the upper bound (7.50a). First of all observe that if $\|\hat{\underline{\gamma}}\|_{L^\infty} \le C^4_{\epsilon,T}\sqrt{\varepsilon}$, then $R^+(\underline{\gamma}) > \|\hat{\underline{\gamma}}\|_{L^\infty}$ and therefore $\bar\gamma \in B(\underline{\gamma}, R^+(\underline{\gamma}))$, which implies that (7.50a) holds trivially, since $\mathscr{I}_{\theta^*,\epsilon}(\bar\gamma) = 0$. Hence we can assume that $\|\hat{\underline{\gamma}}\|_{L^\infty} \ge C^4_{\epsilon,T}\sqrt{\varepsilon}$ and, provided $C_{\epsilon,T}$ has been chosen large enough, $\varsigma \le \sqrt{\varepsilon}$.

As mentioned, we will divide $[0,T]$ in time intervals of length $\tau \le c_\epsilon$. Although, for convenience, most of the following argument is done for arbitrary τ satisfying said inequalities, we will eventually choose[46]

$$\varsigma = \sqrt{\varepsilon\tau} \quad \text{that is} \quad \tau = \left(\frac{T^2\varepsilon}{\|\hat{\underline{\gamma}}\|^2_{L^\infty}}\right)^{1/7}. \tag{7.51}$$

Note that if $B(\underline{\gamma},\varsigma)\cap \mathrm{Lip}_{C,*}([0,T],\mathbb{R}^d) = \emptyset$, then $\mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma},\varsigma)) = 0$ and (7.50a) is trivially true. We can then assume that there exists $\tilde{\underline{\gamma}} \in B(\underline{\gamma},\varsigma)\cap\mathrm{Lip}_{C,*}([0,T],\mathbb{R}^d)$. We then consider the piecewise linear path $\overline{\underline{\gamma}}$ such that $\overline{\underline{\gamma}}(k\varepsilon) = \tilde{\underline{\gamma}}(k\varepsilon)$ for all $k \in \mathbb{N}$. Since $\|\tilde{\underline{\gamma}} - \overline{\underline{\gamma}}\|_{L^\infty} \le C\varepsilon$, $B(\underline{\gamma},\varsigma) \subset B(\overline{\underline{\gamma}},3\varsigma)$. We can thus assume, without loss of generality, that $\underline{\gamma} = \overline{\underline{\gamma}}$ by substituting ς to $\varsigma/3$.

Let $E = [0,\tau]$ and recall the notation $E_k = [k\tau,(k+1)\tau)$ introduced in Lemma 7.9. By Lemma 7.9, we gather that, provided ε is sufficiently small,

$$\mathbb{P}_{\ell_0,\varepsilon}\left(B(\underline{\gamma},\varsigma)\right) \le \prod_{k=0}^{K-1} P_-(E_k,\theta^*,\underline{\gamma},2\varsigma) \tag{7.52}$$

where recall that

$$P_-(E,\theta^*,\underline{\gamma},r) = \sup_{\{\ell\,:\,\mathbb{1}^-_{\theta^*,\underline{\gamma},r}(\ell,k\tau)=1\}} \mathbb{P}_{\ell,\varepsilon}(B|_E(\underline{\gamma}_k,r)), \tag{7.53}$$

and $\underline{\gamma}_k \in \mathrm{Lip}_{C,*}(E,\mathbb{R}^d)$ is the translation of the path $\underline{\gamma}$ defined by:

$$\underline{\gamma}_k(s) = \underline{\gamma}(k\tau + s) - \underline{\gamma}(k\tau).$$

Let us fix arbitrarily $k \in \{0,\cdots,K-1\}$ and let $\underline{\theta}_k = \theta^{\underline{\gamma}}(k\tau,\theta^*)$; fix also ℓ so that $\mathbb{1}^-_{\theta^*,\underline{\gamma},\varsigma}(\ell,k\tau) = 1$, i.e. we have $|\theta^*_\ell - \underline{\theta}_k| \le C_\#\varsigma$.

Let us define the shorthand notations $\bar\theta_\ell : E \to \mathbb{T}$ as $\bar\theta_\ell(\cdot) = \bar\theta(\cdot,\theta^*_\ell)$ and correspondingly $\bar z_\ell : E \to \mathbb{R}^d$ as $\bar z_\ell = \bar z(\cdot,\theta^*_\ell)$ (i.e. $\bar z_\ell$ satisfies the differential equation $\bar z'_\ell(s) = \bar A(\bar\theta_\ell(s))$ with initial condition $\bar z_\ell(0) = (\theta^*_\ell, 0)$). For convenience, remembering (2.9), let us also introduce, for any $\gamma \in \mathrm{Lip}_{C,*}(E,\mathbb{R}^d)$,

$$\begin{aligned}\bar\gamma_\ell(s) &= \bar z_\ell(s) - (\theta^*_\ell, 0)\\ \hat\gamma_\ell &= \gamma - \bar\gamma_\ell.\end{aligned} \tag{7.54}$$

[46] We do not claim this choice to be optimal; its motivation is to simplify (7.70).

Our first step is to estimate $\mathbb{P}_{\ell,\varepsilon}(B|_E(\underline{\gamma}_k, 2\varsigma))$. For further use let us introduce the notations (recall definitions (6.2), (6.3), (6.16) and (6.17)):

$$
\begin{aligned}
I_\theta|_E(\gamma) &= \begin{cases} +\infty & \text{if } \gamma(0) \neq 0 \text{ or } \gamma \notin \text{Lip} \\ \int_E \mathcal{Z}(\gamma'(s), \bar\theta(s,\theta))\, ds & \text{otherwise.} \end{cases} \\
\mathscr{I}^{\pm}_{\theta,\epsilon}|_E(\gamma) &= \begin{cases} +\infty & \text{if } \gamma(0) \neq 0 \text{ or } \gamma \notin \text{Lip} \\ \int_E \mathcal{Z}^{\pm}_\epsilon(\gamma'(s), \theta^\gamma(s,\theta))\, ds & \text{otherwise.} \end{cases}
\end{aligned} \tag{7.55}
$$

Of course, we have $I_\theta = I_\theta|_{[0,T]}$ and $\mathscr{I}^{\pm}_{\theta,\epsilon} = \mathscr{I}^{\pm}_{\theta,\epsilon}|_{[0,T]}$.

Next, we apply Lemma 7.2 in the interval E: assume $\sqrt{\varepsilon\tau} \geq \varsigma$, $\tau \geq \varepsilon_0^{-4}\varepsilon$, let

$$
R_{k,\ell} = C_0' \max\left\{((\varepsilon/\tau)^{1/4} + \tau)\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}, \min\left\{\varepsilon^{1/4}\tau^{3/4}, (\varepsilon\tau)^{1/6}\|\hat{\underline{\gamma}}_{k,\ell}\|^{2/3}_{L^\infty}\right\}, \sqrt{\varepsilon\tau}\right\} \tag{7.56}
$$

where $\hat{\underline{\gamma}}_{k,\ell} = \underline{\gamma}_k - \bar\gamma_\ell$, $C_0' = C_{11}C_0$, C_0 was introduced in Lemma 7.2 and the constant C_{11} is chosen so large that $C_0'\sqrt{\varepsilon\tau} \geq C_0\sqrt{\varepsilon\tau} + 2\varsigma$, then

$$
\varepsilon \log \mathbb{P}_{\ell,\varepsilon}(B|_E(\underline{\gamma}_k, 2\varsigma)) \leq -\inf_{\gamma \in B|_E(\underline{\gamma}_k, R_{k,\ell})} I_{\theta^*_\ell}|_E(\gamma). \tag{7.57}
$$

We now proceed to relate the rate function $I_{\theta^*_\ell}|_E$ appearing in (7.57) with the modified rate function $\mathscr{I}^-_{\underline{\theta},\epsilon/2}|_E$.

Sub-lemma 7.11. *For each $\epsilon > 0$ there exists $\varepsilon_0, c_\epsilon > 0$ such that, for all $\varepsilon < \varepsilon_0$, $c_\epsilon \geq \tau > 0$, $E = [0,\tau]$, standard pair ℓ such that $\mathbb{1}^-_{\theta^*,\underline{\gamma},\varsigma}(\ell, k\tau) = 1$, and $\gamma \in \text{Lip}_{C,*}(E, \mathbb{R}^d)$,*

$$
\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) \leq (1 + C_\epsilon\tau) I_{\theta^*_\ell}|_E(\gamma) + C_\epsilon\varsigma^2.
$$

Proof. Since $\|\cdot\|_{L^1(E)} \leq \sqrt{\tau}\|\cdot\|_{L^2(E)}$, for any $\gamma \in \text{Lip}_{C,*}(E, \mathbb{R}^d)$ we have

$$
\begin{aligned}
\|\bar\theta_\ell - \theta^\gamma(\cdot, \underline{\theta}_k)\|_{L^\infty(E)} &\leq \int_E \|\bar A(\bar\theta_\ell(s)) - \gamma'(s)\| ds + C_\# \varsigma \\
&\leq \sqrt{\tau}\|\bar A(\bar\theta_\ell) - \gamma'\|_{L^2(E)} + C_\#\varsigma
\end{aligned}
$$

which yields, recalling (6.20):

$$
\|\bar\theta_\ell - \theta^\gamma(\cdot, \underline{\theta}_k)\|_{L^\infty(E)} \leq C_\# \sqrt{\tau I_{\theta^*_\ell}|_E(\gamma)} + C_\#\varsigma. \tag{7.58}
$$

Similarly, we have:

$$
\begin{aligned}
\|\bar\theta_\ell - \theta^\gamma(\cdot, \underline{\theta}_k)\|_{L^\infty(E)} &\leq \int_E \|\bar A(\bar\theta_\ell(s)) - \bar A(\theta^\gamma(s, \underline{\theta}_k))\| + \|\bar A(\theta^\gamma(s,\underline{\theta}_k)) - \gamma'(s)\| ds + C_\#\varsigma \\
&\leq C_\#\tau\|\bar\theta_\ell - \theta^\gamma(\cdot,\underline{\theta}_k)\|_{L^\infty(E)} + \|\bar A(\theta^\gamma(s,\underline{\theta}_k)) - \gamma'(s)\|_{L^1(E)} + C_\#\varsigma.
\end{aligned}
$$

Then, provided $C_\#\tau \leq 1/2$, Lemma 6.16 implies

$$
\|\bar\theta_\ell - \theta^\gamma(\cdot, \underline{\theta}_k)\|_{L^\infty(E)} \leq C_\# \sqrt{\tau \mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma)} + C_\#\varsigma. \tag{7.59}
$$

Notice that if $I_{\theta^*_\ell}|_E(\gamma) = \infty$, then the sub-lemma holds trivially. We can then assume that, for a.e. $s \in E$ we have $\mathcal{Z}(\gamma'(s), \bar\theta_\ell(s)) < \infty$, i.e. $\gamma'(s) \in \mathbb{D}(\bar\theta_\ell(s))$. Hence, if τ is sufficiently small (with respect to ϵ), since $|\bar\theta_\ell(\cdot) - \theta^\gamma(\cdot, \underline{\theta}_k)| \leq C_\#\sqrt{\tau}$, by Lemma 6.7 we conclude that $\gamma'(s) \in \mathbb{D}(\theta^\gamma(s,\underline{\theta}_k)) \cup \partial_{\epsilon/2}\mathbb{D}(\theta^\gamma(s,\underline{\theta}_k))$, that is, $\mathcal{Z}^-_{\epsilon/2}(\gamma'(s), \theta^\gamma(s,\underline{\theta}_k)) < \infty$. Thus, by definition (6.16), we have that, for any $s \in E$,

$$
\mathcal{Z}^-_{\epsilon/2}(\gamma'(s), \theta^\gamma(s,\underline{\theta}_k)) = \mathcal{Z}(b_\varrho(s), \theta^\gamma(s,\underline{\theta}_k))
$$

where $b_\varrho(s) = \bar A(\theta^\gamma(s,\underline{\theta}_k)) + \varrho(s)(\gamma'(s) - \bar A(\theta^\gamma(s,\underline{\theta}_k)))$ and $\varrho(s) \in [0,1]$ is the largest ϱ such that $b_\varrho \notin \partial_{\epsilon/2}\mathbb{D}(\theta^\gamma(s,\underline{\theta}_k))$. Let $\bar b_\varrho(s) = \bar A(\bar\theta(s,\theta^*_\ell)) + \varrho(s)(\gamma'(s) - \bar A(\bar\theta(s,\theta^*_\ell)))$;

observe that, by definition, $\|\bar b_\varrho - b_\varrho\| < C_\# |\bar\theta_\ell - \theta^\gamma(\cdot, \underline{\theta}_k)|$. By Lemma 6.2-(2) we can expand $\mathscr{Z}$ to second order obtaining:[47]

$$\begin{aligned}\mathscr{Z}(b_\varrho(s), \theta^\gamma(s, \underline{\theta}_k)) - \mathscr{Z}(\bar b_\varrho(s), \bar\theta_\ell(s)) &= \mathscr{Z}(b_\varrho(s), \theta^\gamma(s, \underline{\theta}_k)) - \mathscr{Z}(b_\varrho(s), \bar\theta_\ell(s))) \\ &\quad + \mathscr{Z}(b_\varrho(s), \bar\theta_\ell(s)) - \mathscr{Z}(\bar b_\varrho(s), \bar\theta_\ell(s)) \\ = \partial_\theta \mathscr{Z}(b_\varrho(s), \bar\theta_\ell(s))(\theta^\gamma(s, \underline{\theta}_k) - \bar\theta_\ell(s)) &+ \partial_b \mathscr{Z}(\bar b_\varrho(s), \bar\theta_\ell(s))(b_\varrho - \bar b_\varrho) \\ + C_\epsilon \mathcal{O}(|\theta^\gamma(s, \underline{\theta}_k) - \bar\theta_\ell(s)|^2).\end{aligned}$$

Next, we can expand $\partial_\theta \mathscr{Z}$ and $\partial_b \mathscr{Z}$ around the point $(\bar A(\bar\theta_\ell(s)), \bar\theta_\ell(s))$. Recalling Lemma 6.2-(2),(4) we obtain:

$$\begin{aligned}|\partial_\theta \mathscr{Z}(b_\varrho(s), \bar\theta_\ell(s))| &\le C_\epsilon \|\bar A(\bar\theta_\ell(s)) - b_\varrho\| \le C_{\#\epsilon} \|\bar A(\bar\theta_\ell(s)) - \bar b_\varrho\| + \|\bar b_\varrho - b_\varrho\| \\ &\le C_\epsilon \|\bar A(\bar\theta_\ell(s)) - \gamma'(s)\| + C_\epsilon |\theta^\gamma(s, \underline{\theta}_k) - \bar\theta_\ell(s)|. \\ |\partial_b \mathscr{Z}(\bar b_\varrho(s), \bar\theta_\ell(s))| &\le C_\epsilon \|\bar A(\bar\theta_\ell(s)) - \bar b_\varrho\| \le C_\epsilon \|\bar A(\bar\theta_\ell(s)) - \gamma'(s)\|.\end{aligned}$$

Since Lemma 6.2-(0),(2),(4) imply that $\mathscr{Z}(\bar b_\varrho(s), \bar\theta_\ell(s)) \le \mathscr{Z}(\gamma'(s), \bar\theta_\ell(s))$, we conclude that

$$\begin{aligned}\mathscr{Z}^-_{\epsilon/2}(\gamma'(s), \theta^\gamma(s, \underline{\theta}_k)) &\le \mathscr{Z}(\gamma'(s), \bar\theta_\ell(s)) \\ &+ C_\epsilon \left[\|\bar A(\bar\theta_\ell(s)) - \gamma'(s)\| \|\theta^\gamma(\cdot, \underline{\theta}_k) - \bar\theta_\ell\|_{L^\infty(E)} + \|\theta^\gamma(\cdot, \underline{\theta}_k) - \bar\theta_\ell\|^2_{L^\infty(E)}\right].\end{aligned}$$

Integrating the above expression over E yields

$$\begin{aligned}\mathscr{I}^-_{\underline{\theta}_k, \epsilon/2}|_E(\gamma) \le I_{\theta^*_\ell}|_E(\gamma) + C_\epsilon \Big[&\|\bar A(\bar\theta_\ell(s)) - \gamma'\|_{L^1(E)} \|\theta^\gamma(\cdot, \underline{\theta}_k) - \bar\theta_\ell\|_{L^\infty(E)} \\ &+ \tau \|\theta^\gamma(\cdot, \underline{\theta}_k) - \bar\theta_\ell\|^2_{L^\infty(E)}\Big].\end{aligned}$$

Finally, using (7.58), the relation $\|\cdot\|_{L^1(E)} \le \sqrt{\tau}\|\cdot\|_{L^2(E)}$ and (6.20) we obtain

$$\mathscr{I}^-_{\underline{\theta}_k, \epsilon/2}|_E(\gamma) \le (1 + C_\epsilon \tau) I_{\theta^*_\ell}|_E(\gamma) + C_\epsilon \varsigma \sqrt{\tau I_{\theta^*_\ell}|_E(\gamma)} + C_\epsilon \varsigma^2 \tau.$$

Since $2ab \le a^2 + b^2$, this concludes the proof of the sub-lemma. □

By (7.57) and Sub-Lemma 7.11 it follows that for each standard pair ℓ such that $\mathbb{1}^-_{\theta^*, \underline{\gamma}, \varsigma}(\ell, k\tau) = 1$, and $c_\epsilon \ge \tau \ge \varepsilon^{-1}\varsigma^2$,

$$\varepsilon \log \mathbb{P}_{\ell, \varepsilon}(B|_{E_k}(\underline{\gamma}_k, 2\varsigma)) \le - \inf_{\gamma \in B|_{E_k}(\underline{\gamma}_k, R_{k,\ell})} (1 - C_\epsilon \tau) \mathscr{I}^-_{\underline{\theta}_k, \epsilon/2}|_{E_k}(\gamma) + C_\epsilon \varsigma^2. \tag{7.60}$$

We have now obtained an estimate for each of the terms appearing in the product on the right hand side of (7.52) (see also (7.53)). Next, we must join the above estimates for different time intervals E_k. This can be done in various ways: we choose to control the trajectories at the endpoint of the intervals so that the paths corresponding to different time intervals will join naturally into a continuous path. To this end, for each $\tilde\gamma \in \mathcal{C}^0([0, T], \mathbb{R}^d)$ and $r > 0$, let us define the sets

$$B^*|_E(\tilde\gamma, r) = \{\gamma \in B|_E(\tilde\gamma, r) \;:\; \gamma(0) = 0\,; \gamma(\tau) = \tilde\gamma(\tau)\}.$$

Then, for any $\gamma \in B|_E(\underline{\gamma}_k, R_{k,\ell})$ with $\gamma(0) = 0$, we define $\gamma^* \in B^*|_E(\underline{\gamma}_k, 2R_{k,\ell})$ as:

$$\gamma^*(s) = \gamma(s) + \frac{\underline{\gamma}_k(\tau) - \gamma(\tau)}{\tau} s.$$

Our next goal is to relate $\mathscr{I}^-_{\underline{\theta}_k, \epsilon/2}|_E(\gamma)$ with $\mathscr{I}^-_{\underline{\theta}_k, \epsilon}|_E(\gamma^*)$.

[47] Notice that all second derivatives of $\mathscr{Z}$ are uniformly bounded by some constant that depends on ϵ, which we denote with C_ϵ.

Sub-lemma 7.12. *There exist $c_\epsilon > 0$ such that, for all $c_\epsilon > \tau > c_\epsilon^{-1}\varepsilon$,*

$$\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) \ge \left(1 - C_\epsilon \max\left\{\left(\frac{\varepsilon}{\tau}\right)^{1/4} + \tau, \left(\frac{\varepsilon}{\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)}\right)^{1/6}\right\}\right)\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*) - C_\epsilon\varepsilon,$$

provided ε is small enough.

Proof. Again, it suffices to consider the case $\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) < \infty$ (hence $\gamma'(s) \in \mathbb{D}(\theta^\gamma(s,\underline{\theta}_k))$ for all $s \in E$), the other case being trivial. Observe that $\|\hat{\underline{\gamma}}_{k,\ell}\| \le 2C\tau$ hence, by definition, $R_{k,\ell} \le \epsilon\tau/4$ provided $C_\# \epsilon^{-2}\varepsilon \le \tau \le C_\# \epsilon^2$. Accordingly, for a.e. $s \in E$.

$$\begin{aligned} &\|(\gamma^*)(s) - \gamma(s)\| \le R_{k,\ell} \\ &\|(\gamma^*)'(s) - \gamma'(s)\| = \|\underline{\gamma}_k(\tau) - \gamma(\tau)\|/\tau \le R_{k,\ell}/\tau \le \frac{\epsilon}{4}. \end{aligned} \tag{7.61}$$

Since $\mathbb{D}(\theta)$ varies continuously (see Lemma 6.7) it follows that, for ε small enough, if, for all s, $\gamma'(s) \notin \partial_{3\epsilon/4}\mathbb{D}(\theta^\gamma(s,\underline{\theta}_k))$, then $(\gamma^*)'(s) \notin \partial_{\epsilon/2}\mathbb{D}(\theta^{\gamma^*}(s,\underline{\theta}_k))$. We start to discuss such case as we will see that the other possibility can be essentially reduced to the present one.

We expand $\mathcal{Z}$ to first order at $(\gamma^{*\prime}, \theta^{\gamma^*}(s,\underline{\theta}_k))$ using Lemma 6.2-(2) and obtain

$$\begin{aligned} |\mathcal{Z}(\gamma'(s), \theta^\gamma(s,\underline{\theta}_k)) - \mathcal{Z}((\gamma^*)'(s), \theta^{\gamma^*}(s,\underline{\theta}_k)| &\le \|\partial_b\mathcal{Z}((\gamma^*)', \theta^\gamma(s,\underline{\theta}_k))\|\frac{R_{k,\ell}}{\tau} \\ &+ |\partial_\theta\mathcal{Z}((\gamma^*)', \theta^{\gamma^*}(s,\underline{\theta}_k))|R_{k,\ell} + C_\epsilon(R_{k,\ell}/\tau)^2. \end{aligned}$$

Once again, expanding the derivatives to the first order at $(\bar{A}(\theta^{\gamma^*}(s,\theta)), \theta^{\gamma^*}(s,\theta))$ and using Lemma 6.2-(4) we conclude

$$\begin{aligned} &|\mathcal{Z}(\gamma'(s), \theta^\gamma(s,\underline{\theta}_k)) - \mathcal{Z}((\gamma^*)'(s), \theta^{\gamma^*}(s,\underline{\theta}_k)| \\ &\qquad\qquad \le C_\epsilon\left(\|(\gamma^*)' - \bar{A}(\theta^{\gamma^*}(s,\theta))\|R_{k,\ell}/\tau + (R_{k,\ell}/\tau)^2\right). \end{aligned} \tag{7.62}$$

Integrating over E, it follows that

$$\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) = \mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma^*) + C_\epsilon\mathcal{O}\left(\frac{R_{k,\ell}}{\tau}\int_E \|(\gamma^*)' - \bar{A}(\theta^{\gamma^*}(s,\theta))\| + \frac{R_{k,\ell}^2}{\tau}\right)$$

and, by Lemma 6.16 and Remark 6.13,

$$\ge \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*) - C_\epsilon\left(\frac{R_{k,\ell}}{\sqrt{\tau}}\sqrt{\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)} + \frac{R_{k,\ell}^2}{\tau}\right). \tag{7.63}$$

Then, recalling that $\hat{\underline{\gamma}}_{k,\ell} = \underline{\gamma}_k - \bar{\gamma}_\ell$ and (7.54), (7.61),

$$\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty} \le \|\underline{\gamma}_k - \gamma^*\|_{L^\infty} + \|\gamma^* - \bar{\gamma}_\ell\|_{L^\infty} \le 2R_{k,\ell} + \|\gamma^* - \bar{\gamma}_\ell\|_{L^\infty},$$

using the same argument as in (7.59) and since $R_{k,\ell} \ge \varsigma$, we have

$$\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty} \le 3R_{k,\ell} + C_\#\sqrt{\tau\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)}. \tag{7.64}$$

Before continuing, note that if $R_{k,\ell} \ge \|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}/4$, then the value of $R_{k,\ell}$ cannot be given by $(\varepsilon^{1/4}\tau^{-1/4} + \tau)\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}$. If $R_{k,\ell} = \min\left\{\varepsilon^{1/4}\tau^{3/4}, (\varepsilon\tau)^{1/6}\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}^{2/3}\right\}$, then $(\varepsilon\tau)^{1/6}\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}^{2/3} \ge \|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}/4$ implies $\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty} \le 4^3\sqrt{\varepsilon\tau}$. In turn, the later inequality implies, provided ε is small enough,

$$(\varepsilon\tau)^{1/6}\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}^{2/3} \le 4^2\sqrt{\varepsilon\tau} \le \varepsilon^{1/4}\tau^{3/4}.$$

We have thus seen that $\sqrt{\varepsilon\tau} \le R_{k,\ell} \le C_\# \sqrt{\varepsilon\tau}$ and substituting it in (7.63) yields

$$\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) \ge \left[1 - C_\epsilon\left(\left[\frac{\varepsilon}{\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)}\right]^{1/2}\right)\right]\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*) - C_\#\varepsilon. \tag{7.65}$$

We are thus left considering the case $R_{k,\ell} < \|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty}/4$. By (7.64):

$$\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty} \le C_\# \sqrt{\tau \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)}.$$

Accordingly, we have:

$$R_{k,\ell} \le C_\# \max\left\{(\varepsilon^{1/4}\tau^{-1/4}+\tau)\sqrt{\tau\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)}, \varepsilon^{1/6}\tau^{1/2}\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)^{1/3}, \sqrt{\varepsilon\tau}\right\}.$$

If the first term realizes the max, then by (7.63) we conclude

$$\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2,E}(\gamma) \ge (1 - C_\epsilon\mathcal{O}(\varepsilon^{1/4}\tau^{-1/4}+\tau))\mathscr{I}^-_{\underline{\theta}_k,\epsilon,E}(\gamma^*).$$

Otherwise, if the second term realizes the max, (7.63) gives:[48]

$$\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) \ge \left[1 - C_\epsilon\left(\left[\frac{\varepsilon}{\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)}\right]^{1/6}\right)\right]\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*) - C_\epsilon\varepsilon.$$

Finally, if the third term realizes the max, then we have (7.65) again. This proves the sub-lemma in the case under consideration.

We are thus left with the case that, for some s, we have $\gamma'(s) \in \partial_{3\epsilon/4}\mathbb{D}(\theta^\gamma(s,\underline{\theta}_k))$. Let $S_\gamma \ne \emptyset$ be the collection of such s. Then we define $b_\varrho(s) = \varrho(s)\gamma'(s) + (1-\varrho(s))\bar{A}(\bar\theta(s,\underline{\theta}_k))$ where $\varrho(s) \in [0,1]$ is zero on the complement of S_γ and such that $b_\varrho(s)$ belongs to the boundary of $\mathbb{D}(\bar\theta(s,\underline{\theta}_k)) \setminus \partial_{3\epsilon/4}\mathbb{D}(\bar\theta(s,\underline{\theta}_k))$ otherwise. Also, we define $b^*_\varrho(s) = \varrho(s)(\gamma^*)'(s) + (1-\varrho(s))\bar{A}(\bar\theta(s,\underline{\theta}_k))$. Note that $b_\varrho(s) \notin \partial_{\epsilon/2}\mathbb{D}(\bar\theta^{\gamma_\varrho}(s,\underline{\theta}_k))$ and $b^*_\varrho(s) \notin \partial_{\epsilon/2}\mathbb{D}(\bar\theta^{\gamma^*_\varrho}(s,\underline{\theta}_k))$ but, for $s \in S_\gamma$, $b^*_\varrho(s) \in \partial_\epsilon\mathbb{D}(\bar\theta^{\gamma^*_\varrho}(s,\underline{\theta}_k))$. By (7.61) we have $\|b_\varrho - b^*_\varrho\|_{L^\infty} \le R_{k,\ell}/\tau$ and

$$\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) \ge \int_E \mathcal{Z}(b_\varrho(s),\bar\theta^\gamma(s,\underline{\theta}_k))ds$$

$$\int_E \mathcal{Z}(b^*_\varrho(s),\bar\theta^{\gamma^*}(s,\underline{\theta}_k))ds \ge \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*).$$

We can then conclude by expanding $\mathcal{Z}$ as in (7.62) and, using the above relations, we obtain again

$$\mathscr{I}^-_{\underline{\theta}_k,\epsilon/2}|_E(\gamma) \ge \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*) - C_\epsilon\left(\frac{R_{k,\ell}}{\sqrt{\tau}}\sqrt{\mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)} + \frac{R^2_{k,\ell}}{\tau}\right).$$

The argument is then concluded exactly in the same manner as before. □

Using equations (7.52), (7.53), (7.60) and Sub-Lemma 7.12 we obtain:

$$\begin{aligned}
\varepsilon\log\mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma},\varsigma)) \le &-(1-\mathbb{K}_{\epsilon,\varepsilon,\tau})\left[\sum_k \inf_{\gamma\in B|_E(\underline{\gamma}_k,R_k)} \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)\right] \\
&+ C_\epsilon\left(\frac{T\varepsilon}{\tau}\right)^{1/6}\left[\sum_k \inf_{\gamma\in B|_E(\underline{\gamma}_k,R_k)} \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*)\right]^{5/6} + C_\epsilon\frac{T\varepsilon}{\tau}
\end{aligned} \tag{7.66}$$

$$\mathbb{K}_{\epsilon,\varepsilon,\tau} = C_\epsilon(\varepsilon^{1/4}\tau^{-1/4}+\tau),$$

[48] Just consider the two possibilities $\mathscr{I}^-_{\underline{\theta}_k,\epsilon,E}(\gamma^*) > \varepsilon$ and $\mathscr{I}^-_{\underline{\theta}_k,\epsilon,E}(\gamma^*) < \varepsilon$.

where, in the second line, we have used Hölder inequality, the assumption $\varsigma^2 \le \varepsilon$ and

$$R_k = 2 \sup_{\{\ell\,:\,\mathbb{1}^-_{\theta^*,\underline{\gamma},2\varsigma}(\ell,k\tau)=1\}} R_{k,\ell}\ ; \quad R = \max_k\{R_k\}.$$

Next, we must compute the sum in the square brackets. Let us define the sets $\underline{B}^*_k = B^*|_E(\underline{\gamma}_k, R_k)$. For each set of paths $\{\tilde\gamma_k\}_{k\in\{0,\cdots,K-1\}}$, $\tilde\gamma_k \in \underline{B}^*_k$, we can "glue them together" defining $\tilde\gamma(s) = \tilde\gamma_k(s-k\tau)+\underline{\gamma}(k\tau)$ for $s \in E_k$. Clearly $\tilde\gamma \in B(\underline{\gamma}, R)$. In addition, $\mathscr{I}^-_{\theta^*,\epsilon}(\tilde\gamma) = \sum_k \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\tilde\gamma_k)$, which yields

$$\sum_k \inf_{\gamma\in B|_E(\underline{\gamma}_k,R_k/2)} \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma^*) \ge \sum_k \inf_{\gamma\in\underline{B}^*_k} \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma) \ge \inf_{\gamma\in B(\underline{\gamma},R)} \mathscr{I}^-_{\theta^*,\epsilon}(\gamma) =: J_{*,R}.$$

Note that the right hand side of (7.66) is bounded by $C_\# C_\epsilon^5 \varepsilon\tau^{-1}T$ and the maximum is achieved for $\sum_k \inf_{\gamma\in\underline{B}|_E(\underline{\gamma}_k,R_k)} \mathscr{I}^-_{\underline{\theta}_k,\epsilon}|_E(\gamma)$ proportional to $C_\epsilon\varepsilon\tau^{-1}T$. Hence, for $J_{*,R} \ge C_\epsilon\varepsilon\tau^{-1}T$, the right hand side of (7.66) is a decreasing function of the quantity in square brackets. Accordingly, by eventually increasing the value of C_ϵ,

$$\varepsilon \log \mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma},\varsigma)) \le -(1-\mathbb{K}_{\epsilon,\varepsilon,\tau})J_{*,R^+} + C_\epsilon\left(\frac{T\varepsilon}{\tau}\right)^{1/6} J^{5/6}_{*,R^+} + C_\epsilon\varepsilon\tau^{-1}T, \tag{7.67}$$

provided that $R^+ := R^+(\underline{\gamma}) \ge R = \max_k\{R_k\}$. Further,

$$\begin{aligned}\|\hat{\underline{\gamma}}_{k,\ell}(s)\| &= \|\underline{\gamma}(k\tau+s) - \underline{\gamma}(k\tau) - \bar z(s,\theta^*_\ell) + (\theta^*_\ell,0)\|\\ &\le \|\hat{\underline{\gamma}}(k\tau+s)\| + \|\hat{\underline{\gamma}}(k\tau)\| + \|\bar z(k\tau+s,\theta^*_{\ell_0}) - \bar z(k\tau,\theta^*_{\ell_0}) - \bar z(s,\theta^*_\ell) + (\theta^*_\ell,0)\|\\ &\le 2\|\hat{\underline{\gamma}}\|_{L^\infty} + \left\|\int_0^s \left[\bar A(\bar\theta(s',\bar\theta_{\ell_0}(k\tau))) - \bar A(\bar\theta(s',\theta^*_\ell))\right] ds'\right\|.\end{aligned}$$

By continuity with respect to the initial conditions and recalling the assumption $\|\hat{\underline{\gamma}}\|_{L^\infty} \ge C_{\epsilon,T}\sqrt{\varepsilon}$, hence $\|\hat{\underline{\gamma}}\|_{L^\infty} \ge C_\#\varsigma\tau$, it follows

$$\|\hat{\underline{\gamma}}_{k,\ell}\|_{L^\infty} \le C_\#\|\hat{\underline{\gamma}}\|_{L^\infty}. \tag{7.68}$$

Also, by Lemma 6.16, for each $\gamma \in B(\underline{\gamma}, R^+)$ we have $\|\hat{\underline{\gamma}}\|_{L^\infty} \le 2\|\hat\gamma\|_{L^\infty} \le C_\#\sqrt{T\mathscr{I}^-_{\theta^*,\epsilon}(\gamma)}$. Hence $\|\hat{\underline{\gamma}}\|^2_{L^\infty} \le C_\# T J_{*,R^+}$ and, by (7.67),

$$\varepsilon\log\mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma},\varsigma)) \le -\left[1-\mathbb{K}_{\epsilon,\varepsilon,\tau} - C_\epsilon\left(\frac{T^2\varepsilon}{\tau\|\hat{\underline{\gamma}}\|^2_{L^\infty}}\right)^{1/6} - C_\epsilon\frac{T^2\varepsilon}{\tau\|\hat{\underline{\gamma}}\|^2_{L^\infty}}\right] J_{*,R^+}. \tag{7.69}$$

To validate (7.69) we still need to verify that $R^+ \ge R$. To this end, notice (see the beginning of the proof of Lemma 7.12) that $R_k \le \tau$, provided $\tau \le C_\epsilon$, which is implied by (7.51) when $\|\hat{\underline{\gamma}}\|_{L^\infty} \ge C_{\epsilon,T}\sqrt{\varepsilon}$. Thus, using (7.56), (7.68) and the choice (7.51), we have

$$R_k \le C_T \max\left\{\frac{\|\hat{\underline{\gamma}}\|^{15/14}_{L^\infty}\varepsilon^{3/14}}{T^{1/14}} + \|\hat{\underline{\gamma}}\|^{5/7}_{L^\infty}T^{2/7}\varepsilon^{1/7}, \|\hat{\underline{\gamma}}\|^{13/21}_{L^\infty}T^{1/21}\varepsilon^{4/21}, \frac{\varepsilon^{4/7}T^{1/7}}{\|\hat{\underline{\gamma}}\|^{1/7}_{L^\infty}}\right\}.$$

One can then check that

$$R_k \le C_T\left[\|\hat{\underline{\gamma}}\|^{5/7}_{L^\infty}\varepsilon^{1/7} + \sqrt{\varepsilon}\right].$$

The above implies the claim $R^+ \ge R$ and $R^+ \le \max\{\|\hat{\underline{\gamma}}\|_{L^\infty}, C_{\epsilon,T}\sqrt{\varepsilon}\}$. Substituting the choice (7.51) in equation (7.69), yields

$$\varepsilon\log\mathbb{P}_{\ell_0,\varepsilon}(B(\underline{\gamma},\varsigma)) \le -\left[1 - C_\epsilon\left(\frac{T^2\varepsilon}{\|\hat{\underline{\gamma}}\|^2_{L^\infty}}\right)^{1/7}\right] J_{*,R^+}. \tag{7.70}$$

To obtain the lower bound (7.50b) we argue along the same lines (with a different choice of ς and τ), but the argument turns out to be a bit simpler. To further simplify our discussion we are not going to pursue optimal results. We use Lemma 7.5 with $\varsigma_- = h = K_\delta^{-2}T$ and $T = \tau = \sqrt{\varepsilon}$, and $\delta = \epsilon$ to write

$$\varepsilon \log \mathbb{P}_{\ell,\varepsilon}(B|_E(\underline{\gamma}_k, \varsigma_-)) \geq -I_{\theta^*_\ell}(\underline{\gamma}_k) - c_\delta \varepsilon^{5/8}. \tag{7.71}$$

Next, we claim that, for all ℓ such that $\mathbb{1}^+_{\theta^*,\underline{\gamma},\varsigma_-}(\ell, k\tau) = 1$,

$$\mathscr{I}^+_{\underline{\theta}_k,\epsilon}|_E(\underline{\gamma}_k) \geq (1 - C_\epsilon \tau) I_{\theta^*_\ell}|_E(\underline{\gamma}_k) - C_\epsilon \varsigma_-^2. \tag{7.72}$$

The above relation is trivial if the left hand side is infinite. Otherwise, recalling (6.17) and (6.16), it can be proven along the lines of Sub-Lemma 7.11. Accordingly, Lemma 7.9 implies

$$\begin{aligned}\varepsilon \log \mathbb{P}_{\ell,\varepsilon}(B(\underline{\gamma}, \varsigma_-)) &\geq -(1 + C_\epsilon\sqrt{\varepsilon}) \sum_{k=0}^{K-1} \mathscr{I}^+_{\underline{\theta}_k,\epsilon}|_E(\underline{\gamma}_k) - C_\epsilon T \varepsilon^{1/8} \\ &= -(1 + C_\epsilon\sqrt{\varepsilon})\mathscr{I}^+_{\theta^*,\epsilon}(\underline{\gamma}) - C_\epsilon T \varepsilon^{1/8}. \qquad \square\end{aligned}$$

7.4. **Large and moderate deviations for general sets.**
This subsection contains the second step of our argument that leads to the proof of our main Large Deviations result. Concretely, we show how Theorem 2.4 follows from Lemma 7.10. For the upper bound, we use a relatively standard combinatorial argument which allows to obtain an estimate for the probability of an arbitrary event by covering it with balls; for the lower bound, we simply bound it from below with the measure of a ball contained in the event.

Proof of Theorem 2.4. Let ℓ_0 be a standard pair so that $|\theta^*_{\ell_0} - \theta_0| \leq \varepsilon$ and let $\epsilon = \Delta_*$. Clearly, it suffices to prove the theorem for such standard pairs.

Our goal is to estimate, from above and below, the probability of the event Q_ε.

We start with the lower bound, let $Q_\varepsilon^- = \{\gamma \in Q_\varepsilon \;:\; B|_{[0,T]}(\gamma, \varepsilon^{1/2}) \subset Q_\varepsilon\}$. Then, for each $\underline{\gamma} \in Q_\varepsilon^-$, inequality (7.50b) implies

$$\mathbb{P}_{\ell_0,\varepsilon}(Q_\varepsilon) \geq \mathbb{P}_{\ell_0,\varepsilon}\left(B(\underline{\gamma}, \varepsilon^{1/2})\right) \geq e^{-\varepsilon^{-1}[(1+C_\epsilon\varepsilon^{1/2})\mathscr{I}^+_{\theta_0,\epsilon}(\underline{\gamma}) + C_{\epsilon,T}\varepsilon^{1/8}]}.$$

We can then conclude the argument by taking the sup for $\underline{\gamma} \in Q_\varepsilon^-$.

To obtain the upper bound, first recall that $\operatorname{supp}\mathbb{P}_{\ell_0,\varepsilon} \subset \operatorname{Lip}_{C,*}([0,T], \mathbb{R}^d)$ hence, setting $Q = Q_\varepsilon \cap \operatorname{Lip}_{2C,*}([0,T], \mathbb{R}^d)$ holds

$$\mathbb{P}_{\ell_0,\varepsilon}(Q_\varepsilon) = \mathbb{P}_{\ell_0,\varepsilon}(Q).$$

We will construct a class of coverings of Q and use Lemma 7.10 to estimate the probability of each elements of these coverings. Let us first recap some notations. For any set $\tilde{Q} \subset C^0([0,T], \mathbb{R}^d)$ let $\varrho(\tilde{Q}) = \varrho(\theta_0, \tilde{Q}) = \inf_{\gamma \in \tilde{Q}} \|\gamma - \bar{\gamma}(\cdot, \theta_0)\|_\infty$, and

$$Q^+ := \bigcup_{\gamma \in Q} B(\gamma, R^+(\gamma)) \supset \overline{Q}$$

where $R^+(\gamma) = C_{\epsilon,T} \min\left\{\varepsilon^{1/7}\|\hat{\gamma}\|_{L^\infty}^{5/7} + \varepsilon^{1/2}\right\}$. Note that if $\varrho(Q) \leq C_{\epsilon,T}\sqrt{\varepsilon}$, then $\bar{\gamma}_{\ell_0}(\cdot) = \bar{\gamma}(\cdot, \theta_0) \in Q^+$, hence the statement of the theorem is trivially true. We can thus assume $\varrho(Q) \geq C_{\epsilon,T}\sqrt{\varepsilon}$.

We want to estimate the measure of Q by covering it with balls of the type $B(\underline{\gamma}, \varsigma_\star)$, for some $\varsigma_\star > 0$. To this end we must construct a $(\varsigma_\star/2)$-net. To do so, subdivide the interval $[0,T]$ in sub-intervals of equal lengths $J_j = [s_j, s_{j+1})$, where $s_j = j\varsigma_\star/(1+6C) =: s_{j-1} + \Delta_s$ and recall that C is an upper bound on the Lipschitz constant of all paths that are in the support of $\mathbb{P}_{\ell_0,\varepsilon}$. Denote with $Z = T/\Delta_s$, so that[49] $s_Z = T$. Let $\mathbf{a} = \{a_l\}_{l\in\{0,\cdots,Z-1\}}$ be a (finite) sequence with values in $\frac{1}{2\sqrt{d}}\mathbb{Z}^d$ and let $\gamma_{\mathbf{a}}$ be the unique (Lipschitz) continuous path in $C^0([0,T],\mathbb{R}^d)$ that, for each k, satisfies (for a.e. $s \in [0,T]$) the equation

$$\gamma'_{\mathbf{a}}(s) = \bar{A}(\theta^{\gamma_{\mathbf{a}}}(s,\theta_0)) + a_j \qquad \text{for } s \in J_j$$

with initial condition $\gamma_{\mathbf{a}}(0) = 0$. Let $\mathscr{A} = \{\mathbf{a} : \|a_j\| < 2C \text{ for all } j\}$; observe that, by construction, $\mathscr{A}$ is a finite set (indeed $\#\mathscr{A} < e^{c_\# T\varsigma_\star^{-1}}$) and since $\operatorname{supp}\mathbb{P}_{\ell_0,\varepsilon} \subset \operatorname{Lip}_{C,*}([0,T],\mathbb{R}^d)$ we conclude that if $B(\gamma_{\mathbf{a}},\varsigma_\star)\cap\operatorname{supp}\mathbb{P}_{\ell_0,\varepsilon} \neq \emptyset$, then $\mathbf{a}\in\mathscr{A}$. We now claim that $\bigcup_{\mathbf{a}\in\mathscr{A}}\gamma_{\mathbf{a}}$ is a $\varsigma_\star$-net for the support of $\mathbb{P}_{\ell_0,\varepsilon}$, i.e. $\bigcup_{\mathbf{a}\in\mathscr{A}} B(\gamma_{\mathbf{a}},\varsigma_\star) \supset \operatorname{supp}\mathbb{P}_{\ell_0,\varepsilon}$. In fact, for each $\mathbf{a}\in\mathscr{A}$ and $k\in\{0,\cdots,Z\}$, $\partial_{a_j}\gamma_{\mathbf{a}}(s_{j+1}) = \varsigma_\star\mathbb{1} + \mathcal{O}(\varsigma_\star^2)$, by the smooth dependence of a solution from the vector field. Thus for any path $\gamma\in\operatorname{Lip}_{C,*}([0,T],\mathbb{R}^d)$, provided ε is small enough, there exists $\mathbf{a}\in\mathscr{A}$ so that[50]

$$\|\gamma(s_j) - \gamma_{\mathbf{a}}(s_j)\| < \frac{3}{8}\Delta_s \text{ for any } j\in\{0,\cdots,Z\}.$$

By the Lipschitz property, for any $j\in\{0,\cdots,Z\}$ and $s\in J_j$,

$$\|\gamma(s) - \gamma_{\mathbf{a}}(s)\| < \frac{3}{8}\Delta_s + 3C\Delta_s < \varsigma_\star/2.$$

This proves our claim and concludes the construction of a $\varsigma_\star$-net of paths.

Next, let us define $Q_k = \{\gamma\in Q \mid \|\hat\gamma\|_{L^\infty}\in[2^k\varrho(Q), 2^{k+1}\varrho(Q))$. By our current assumption $\varrho(Q)\geq C_{\epsilon,T}\sqrt{\varepsilon}$ and the fact that $\|\gamma\|_{L^\infty}\leq CT$, we have

$$Q \subset \bigcup_{k=0}^{c_\#\log\varepsilon^{-1}} Q_k.$$

Let us fix some k. Then, by hypothesis, for $\mathbb{P}_{\ell_0,\varepsilon}$-almost all $\gamma\in Q_k$ we have (see (2.14)) $C_{\text{Lip}}(\gamma)\leq T^{-11/7}\varepsilon^{-2/7}\varrho(Q)^{11/7}2^{11k/7} =: C_{\text{Lip}}(k)$, $\varsigma(\gamma)\in[\varsigma_k 2^{-1/7},\varsigma_k]$, $\varsigma_k = \sqrt{\varepsilon}\left(\frac{T^2\varepsilon}{\varrho(Q)^2}\right)^{1/14}2^{-k/7}$ and, for each $|s-s'|\leq\frac{\varsigma_k}{2C_{\text{Lip}}(k)} =: h_\star$,

$$\|\gamma(s)-\gamma(s')\|\leq\varsigma_k/4$$

Hence, if $\max\{\|\gamma_{\mathbf{a}}(jh_\star)-\gamma(jh_\star)\|, \|\gamma_{\mathbf{a}}((j+1)h_\star)-\gamma((j+1)h_\star)\|\}\leq\varsigma_k/2$ then we have, for each $s\in[0,h_\star]$

$$\|(1-sh_\star^{-1})\gamma_{\mathbf{a}}(jh_\star) + sh_\star^{-1}\gamma_{\mathbf{a}}((j+1)h_\star) - \gamma(jh_\star+s)\|\leq 3\varsigma_k/4.$$

Accordingly, we need only $C_\#$ paths to describe all possible behaviors in an interval $[h_\star,(j+1)h_\star]$ with a precision ς_k. This implies that there exists $\mathcal{A}_{Q_k}\subset\mathcal{A}$ such that $\bigcup_{\mathbf{a}\in\mathcal{A}_{Q_k}} B(\gamma_{\mathbf{a}},\varsigma_k)\supset Q_k$ and $\#\mathcal{A}_{Q_k}\leq e^{c_\# T h_\star^{-1}} = e^{c_\# T\varsigma_k^{-1}C_{\text{Lip}}(k)}$.

Accordingly, Lemma 7.10 implies

$$\begin{aligned}\mathbb{P}_{\ell_0,\varepsilon}(Q) &\leq \sum_{k=0}^{c_\#\log\varepsilon^{-1}}\mathbb{P}_{\ell_0,\varepsilon}(Q_k) \leq \sum_{k=0}^{c_\#\log\varepsilon^{-1}}\sum_{\mathbf{a}\in\mathcal{A}_{Q_k}}\mathbb{P}_{\ell_0,\varepsilon}(B(\gamma_{\mathbf{a}},\varsigma_k)) \\ &\leq \sum_{k=0}^{c_\#\log\varepsilon^{-1}}\#(\mathcal{A}_{Q_k})\exp\left[-\varepsilon^{-1}\left(1-C_\epsilon\frac{T^{2/7}\varepsilon^{1/7}}{\varrho(Q_k)^{2/7}}\right)\inf_{\gamma\in Q_k^+}\mathscr{I}^-_{\theta_0,\epsilon}(\gamma)\right].\end{aligned} \tag{7.73}$$

---

[49] Once again we disregard the possibility that Z is not a natural number.

[50] Recall that the lattice $\frac{1}{2\sqrt{d}}\mathbb{Z}^d$ is a r-net for $\mathbb{R}^d$ for any $r\geq 1/4$.

Note that, since $\rho(Q_k) \geq C_{\epsilon,T}\sqrt{\varepsilon}$, we have $\rho(Q_k^+) \geq \frac{1}{2}\rho(Q_k)$. Then, by Lemma 6.16, (6.17) and Remark 6.13, we have

$$\begin{aligned}\rho(Q_k)^2 &\leq C_\epsilon T \inf_{\gamma\in Q_k^+} \mathscr{I}^-_{\theta_0,\epsilon}(\gamma) \leq C_\epsilon T \inf_{\gamma\in Q_k}\int_0^T \|\hat\gamma'(s)\|^2 ds \leq C_\epsilon T^2 C_{\text{Lip}}(k)^2 \\ &\leq C_\epsilon T^{-8/7}\varepsilon^{-4/7}\varrho(Q_k)^{22/7}.\end{aligned} \tag{7.74}$$

Hence, $\varepsilon C_{\text{Lip}}(k) T \varsigma_k^{-1} \leq C_\epsilon \frac{T^{2/7}\varepsilon^{1/7}}{\varrho(Q_k)^{2/7}} \inf_{\gamma\in Q_k^+} \mathscr{I}^-_{\theta_0,\epsilon}(\gamma)$. Accordingly,

$$\mathbb{P}_{\ell_0,\varepsilon}(Q) \leq \sum_{k=0}^{c_\# \log\varepsilon^{-1}} \exp\left[-\varepsilon^{-1}\left(1 - C_\epsilon \frac{T^{2/7}\varepsilon^{1/7}}{\varrho(Q_k)^{2/7}}\right)\inf_{\gamma\in Q_k^+}\mathscr{I}^-_{\theta_0,\epsilon}(\gamma)\right]. \tag{7.75}$$

Next, let us define the sequence $k_0 = 0$, k_{j+1} being the smallest integer k such that $2^k \geq C_\varepsilon(\varrho(Q)^2/(T^2\varepsilon))^{2/7} 2^{\frac{11}{7}k_j}$. By (7.74) it follows $\inf_{\gamma\in Q_{k^+_{j+1}}} \mathscr{I}^-_{\theta_0,\epsilon}(\gamma) \geq 2\inf_{\gamma\in Q^+_{k_j}} \mathscr{I}^-_{\theta_0,\epsilon}(\gamma)$. One can check by induction that $k_j \leq e^{c_\star j}\log(\varrho(Q)^2/\varepsilon)$ for some constant $c_\star > 0$, depending on T. Using again (7.74), we can finally conclude:

$$\begin{aligned}\mathbb{P}_{\ell_0,\varepsilon}(Q) &\leq \sum_{j=0}^\infty e^{c_\star j}\log(\varrho(Q)^2/\varepsilon)\exp\left[-\varepsilon^{-1}\left(1 - C_\epsilon\frac{T^{2/7}\varepsilon^{1/7}}{\varrho(Q)^{2/7}}\right)2^j \inf_{\gamma\in Q^+}\mathscr{I}^-_{\theta_0,\epsilon}(\gamma)\right] \\ &\leq \sum_{j=0}^\infty \exp\left[-\varepsilon^{-1}\left(1 - C_{\epsilon,T}\frac{\varepsilon^{1/7}}{\varrho(Q)^{2/7}}\right)2^j\inf_{\gamma\in Q^+}\mathscr{I}^-_{\theta_0,\epsilon}(\gamma) + c_\# j\right] \\ &\leq \sum_{j=1}^\infty \exp\left[-j\varepsilon^{-1}\left(1 - C_{\epsilon,T}\frac{\varepsilon^{1/7}}{\varrho(Q)^{2/7}}\right)\inf_{\gamma\in Q^+}\mathscr{I}^-_{\theta_0,\epsilon}(\gamma)\right] \\ &\leq \exp\left[-\varepsilon^{-1}\left(1 - C_{\epsilon,T}\frac{\varepsilon^{1/7}}{\varrho(Q)^{2/7}}\right)\inf_{\gamma\in Q^+}\mathscr{I}^-_{\theta_0,\epsilon}(\gamma)\right]. \qquad \square\end{aligned}$$

7.5. **Proof of Propositions 2.2, 2.3 and Corollaries 2.6, 2.7.**

We conclude this section by proving the propositions and corollaries that were stated in Section 2 without a proof.

Proof of Proposition 2.2. We start by proving (2.10). Fix $R > 0$ and $\Delta_* > 0$; by Lemma 6.6 for any $C > \|A\|_{L^\infty}$ if γ is not C-Lipschitz, then $\mathscr{I}_{\theta_0}(\gamma) = \infty$. Hence we can assume that all elements of Q are C-Lipschitz paths; this in particular implies that $R^+(\gamma) < C_{\Delta_*,T}\varepsilon^{1/8}$ (recall that R^+ was defined in (2.12)). Now let $Q_R^+ = \bigcup_{\gamma\in Q} B(\gamma, R)$ and $Q_R^- = \{\gamma \in Q \;:\; B(\gamma,R) \subset Q\}$. For ε small enough, $Q_R^- \subset Q^-$ and, $Q^+ \subset Q_R^+$ (see (2.13) for the definition of Q^-, Q^+) and, by Theorem 2.4, taking first lim inf and lim sup as $\varepsilon \to 0$ and then the liminf for $R \to 0$:

$$-\inf_{\gamma\in\operatorname{int} Q}\mathscr{I}^+_{\theta_0,\Delta_*}(\gamma) \leq \liminf_{\varepsilon\to 0}\varepsilon\log\mathbb{P}_{\mu,\varepsilon}(Q) \leq \limsup_{\varepsilon\to0}\varepsilon\log\mathbb{P}_{\mu,\varepsilon}(Q) \leq -\inf_{\gamma\in\overline{Q}}\mathscr{I}^-_{\theta_0,\Delta_*}(\gamma),$$

the only non-obvious inequality being the last one. To prove it note that if $\rho(\theta_0, Q) = 0$, then the inequality is trivially true, we can then assume $\rho(\theta_0, Q) > 0$, hence, for ε small enough Q is $\mathbb{P}_{\mu,\varepsilon}$-regular (see Remark 2.5). Next, let us define $\beta = \liminf_{R\to 0}\inf_{\gamma\in Q_R^+}\mathscr{I}^-_{\theta_0,\Delta_*}(\gamma)$. Then for each $\delta > 0$ there exists $R_\delta < \delta$ and $\gamma_\delta \in Q^+_{R_\delta}$ such that $\mathscr{I}^-_{\theta_0,\Delta_*}(\gamma_\delta) \leq \beta + \delta$. Since the C-Lipschitz function are compact, there exists a subsequence $\delta_j \to 0$ such that $\gamma_{\delta_j} \to \gamma_* \in \bar{Q}$. The claim follows by the lower semicontinuity of $\mathscr{I}^-_{\theta_0,\Delta_*}$ (see Lemma 6.15).

Next, we want to take the limit $\Delta_* \to 0$ and prove (2.10), that is

$$
\begin{aligned}
(7.76) \qquad -\inf_{\gamma\in\operatorname{int} Q} \mathscr{I}_{\theta_0}(\gamma) &\le \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_{\mu,\varepsilon}(Q) \\
&\le \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_{\mu,\varepsilon}(Q) \le -\inf_{\gamma\in\overline{Q}} \mathscr{I}_{\theta_0}(\gamma).
\end{aligned}
$$

If $\eta = \inf_{\gamma\in\operatorname{int} Q} \mathscr{I}_{\theta_0}(\gamma) = \infty$, then the first inequality is trivially true. Otherwise, by Lemma 6.15, for each $\delta > 0$ there exists $\gamma_\delta \in \operatorname{int} Q \cap \operatorname{int} \mathfrak{D}(\mathscr{I}_\theta)$ such that $\eta + \delta > \mathscr{I}_{\theta_0}(\gamma_\delta)$. Accordingly, there exists Δ_* such that

$$
\eta + \delta \ge \mathscr{I}_{\theta_0}(\gamma_\delta) = \mathscr{I}^+_{\theta_0,\Delta_*}(\gamma_\delta) \ge \inf_{\gamma\in\operatorname{int} Q} \mathscr{I}^+_{\theta_0,\Delta_*}(\gamma)
$$

by the arbitrariness of δ the first inequality of (7.76) follows.

To prove the last inequality of (7.76) let $\eta = \lim_{\Delta_*\to 0} \inf_{\gamma\in\overline{Q}} \mathscr{I}^-_{\theta_0,\Delta_*}(\gamma)$. If $\eta = \infty$, then $\overline{Q} \cap \mathfrak{D}(\mathscr{I}^-_{\theta_0}) = \emptyset$ hence the inequality follows. Otherwise, for each δ there exists $\Delta_\delta > 0$ such that, for all $\Delta_* \le \Delta_\delta$, there exists $\gamma_{\Delta_*} \in \overline{Q} \cap \mathfrak{D}(\mathscr{I}^-_{\theta_0})$ such that $\eta + \delta \ge \mathscr{I}^-_{\theta_0,\Delta_*}(\gamma_{\Delta_*}) \ge \mathscr{I}^-_{\theta_0,\Delta_\delta}(\gamma_{\Delta_*})$, where the last inequality follows form the definition of $\mathscr{I}^-_{\theta_0,\Delta_*}$. By taking a subsequence we can assume that γ_{Δ_*} converges to $\gamma \in \overline{Q} \cap \mathfrak{D}(\mathscr{I}^-_{\theta_0})$. We can then establish (2.10) by taking first the limit $\Delta_* \to 0$ followed by $\delta \to 0$ and applying Lemma 6.15 twice.

Item (a) follows from Lemma 6.6 and Remark 6.14 while item (b) is a direct consequence of the properties of $\mathcal{Z}$ detailed in Lemma 6.2. □

Proof of Proposition 2.3. By Lemma 6.6 , for any $\theta \in \mathbb{T}$, $\mathcal{Z}(\cdot, \theta)$ (defined in (6.2)) is finite only in a compact set on which it is bounded. Then Lemma 6.2 implies that there exists $c > 0$ such that $\mathcal{Z}(b, \theta) \ge c(b - \bar{\omega}(\theta))^2$ for all $\theta \in \mathbb{T}$. Hence,

$$
(7.77) \qquad I_{\theta_0^*}(\gamma) = \int_0^T \mathcal{Z}(\gamma'(s), \bar{\theta}(s, \theta_0^*))ds \ge c \int_0^T \left\| \gamma'(t) - \bar{A}(\bar{\theta}(s, \theta_0^*)) \right\|^2 dt,
$$

where we assumed that γ is Lipschitz (otherwise $I_{\theta_0^*}(\gamma) = \infty$ by definition). Hence, for each $\gamma \in Q_\star = \{\|\hat{\gamma}\|_\infty \ge \frac{1}{2}R\}$,

$$
R \le \int_0^T \|\gamma'(t) - \bar{\gamma}'(t, \theta_0^*)\| dt \le \sqrt{c^{-1} T I_{\theta_0^*}(\gamma)},
$$

We can now apply Lemma 7.2. Note that, for $T_{\max}$ small enough and $\bar{C}$ large enough, $R_\varepsilon(\gamma) \le \|\hat{\gamma}\|_\infty/2$. This implies that $Q_{\varepsilon,+} \subset Q_\star$ and the Lemma follows. □

Proof of Corollary 2.6. Let us start with the upper bound. For any $\gamma \in Q$, let $\gamma_\varepsilon = \varepsilon^\beta \gamma + (1 - \varepsilon^\beta)\bar{\gamma}$. Since Q is bounded, we have $\|\gamma_\varepsilon - \bar{\gamma}\|_{\mathcal{C}^0} < C_Q \varepsilon^\beta$ and in particular (recall the definitions of R^+ given in (2.12) and of $\rho, \varsigma, C_{\text{Lip}}$ in (2.14)) $R^+(\gamma_\varepsilon) \le C_T \varepsilon^{1/7 + 5\beta/7}$, $\rho(\theta_0, Q_\varepsilon) = \varepsilon^\beta \rho(\theta_0, Q)$, $C_{\text{Lip}}(\gamma_\varepsilon) = \|\hat{\gamma}\|_{L^\infty}^{11/7} T^{11/7} \varepsilon^{-2/7 + 11\beta/7}$. Thus $C_{\text{Lip}} \le C$ only if $\beta \ge \varepsilon^{2/11}$, in such a case

$$
\|\gamma_\varepsilon(s) - \gamma_\varepsilon(s')\| \le \varepsilon^\beta \|\gamma(s) - \gamma(s')\| \le C_\# \varepsilon^\beta |s - s'| \le C_Q \varepsilon^{-4\beta/7 + 2/7} \varsigma(\gamma_\varepsilon) \le \frac{\varsigma(\gamma_\varepsilon)}{4}
$$

since $\beta < \frac{1}{2}$, that is the events Q_ε are always $\mathbb{P}_{\mu,\varepsilon}$-regular. In addition, since $\frac{1}{7} + \frac{5}{7}\beta > \beta$, it follows that, for all $R > 0$, for all ε small enough we have $Q^+_\varepsilon \subset \{\varepsilon^\beta \gamma(\cdot) + (1 - \varepsilon^\beta)\bar{\gamma}(\cdot, \theta_0)\}_{\gamma\in Q^+_R} =: Q^+_{\varepsilon,R}$ where $Q^+_R = \bigcup_{\gamma\in\overline{Q}} B(\gamma, R)$. Also, for ε small enough, $Q^+_{\varepsilon,R} \subset \operatorname{int} \mathfrak{D}(\mathscr{I}_{\theta_0})$. In particular, for any $\epsilon > 0$ and sufficiently small ε, $\mathscr{I}^\pm_{\theta_0,\epsilon}(\gamma_\varepsilon) = \mathscr{I}_{\theta_0}(\gamma_\varepsilon)$ for any $\gamma \in Q^+_{\varepsilon,R}$ (recall the definition of $\mathscr{I}^\pm_{\theta_0,\epsilon}$ given in (7.55)). Also by (2.11) and the smoothness of Σ, since Q is Lipschitz bounded,

$$
\mathscr{I}_{\theta_0}(\gamma_\varepsilon) = \frac{\varepsilon^{2\beta}}{2} \int_0^T \langle \gamma'(s) - \bar{A}(\bar{\theta}(s)), \Sigma(\bar{\theta}(s))^{-1} \left[\gamma'(s) - \bar{A}(\bar{\theta}(s))\right] \rangle ds + o(\varepsilon^{2\beta}).
$$

We then apply Theorem 2.4 and the above estimate. Taking the lim sup as $\varepsilon \to 0$ followed by the limits $R \to 0$ yields the wanted result. The lower bound follows by similar arguments. □

Proof of Corollary 2.7. Let C_* large enough and set $\gamma_\varepsilon = \varepsilon^{\frac{1}{2}}\gamma - (1-\varepsilon^{\frac{1}{2}})\bar{z}$. For each $\gamma \in Q$ we have (recall the definition of R^+ given in (2.12)) $R^+(\gamma_\varepsilon) \leq \vartheta \|\hat{\gamma}\|_\infty \sqrt{\varepsilon}$ and that Q_ε is $\mathbb{P}_{\mu,\varepsilon}$-regular. Thus, in the notation of Theorem 2.4, $(\hat{Q}^+)_\varepsilon \supset Q^+_\varepsilon$. Since (2.11) implies

$$\mathscr{I}_{\theta_0}(\gamma_\varepsilon) = \varepsilon \mathscr{I}_{\text{Lin}}(\gamma) + \mathcal{O}(\varepsilon^{\frac{3}{2}})$$

the result follows directly by Theorem 2.4. □

8. Local Limit Theorem

The results of the previous section allow to study deviations $\Delta^*_{\ell,n} = \theta_n - \bar{\theta}(\varepsilon n, \theta^*_\ell)$ from the average of order larger than $\sqrt{\varepsilon}$, but give no information on smaller fluctuations, except for the fact that with very high probability the fluctuations are of order $\sqrt{\varepsilon}$ or smaller. In fact, in [20], it is proven that the fluctuations from the average, once renormalized by the multiplicative factor $\varepsilon^{-1/2}$, converge in law to a diffusion process. Here we go one (long) step forward and we prove Theorem 2.8 which is the equivalent of a Local Central Limit Theorem with error terms for the above convergence.

Remark 8.1. *As already mentioned before the statement of Theorem 2.8, although we will restrict our discussion to fluctuations of the variable θ, the same type of arguments would yield corresponding results for z.*

A standard technique to prove local CLT type results for a dynamical systems leads to the study of the leading eigenvalue of a suitable transfer operator (see, e.g., [31]). While this idea works quite well for uniformly hyperbolic systems, it is much harder to implement for partially hyperbolic systems. Here we will use the standard pair technology to reduce our problem to a slowly varying uniformly hyperbolic system. This will be achieved in several steps.

The first step consists in expressing the fluctuation in terms of a more explicit random variable $\mathbb{A}$: this is done in Section 8.1. Then, in Section 8.2, we first show how Theorem 2.8 follows rather easily once one has computed the characteristic function of the random variable $\mathbb{A}^\dagger$, which is a suitable mollification of $\varepsilon\mathbb{A}$. We then discuss which technical estimates are necessary to compute the Fourier transform defining the characteristic function of $\mathbb{A}^\dagger$ and we use the standard pair formalism to recast them in a form to which, in the next sections, it will be possible to apply the transfer operator technique, effectively reducing the problem to one similar to the skew-product case. The difference being that the fast dynamics is slowly varying rather than a constant. Hence, instead of having a power of a single transfer operator we will have to deal with a product of similar, but different, operators.

Let $T > 0$ be the one appearing in the statement of Theorem 2.8 and consider $t \in [\varepsilon^{1/2000}, T]$ to be fixed. In the following we will find convenient to work with a definition of "deviation" that is independent of the standard pair language. This definition has been already introduced in (5.1b), but we report it here for the reader's convenience. Recall the notation $\bar{\theta}_k = \bar{\theta}(\varepsilon k, \theta)$; then let

$$\Delta_k(x,\theta) = \theta_k(x,\theta) - \bar{\theta}_k(\theta) \tag{8.1}$$

where, as usual, $(x_k, \theta_k) = F^k_\varepsilon(x,\theta)$ and $\bar{\theta}(t,\theta)$ is the unique solution of $\dot{\bar{\theta}} = \bar{\omega}(\bar{\theta})$, with initial condition $\bar{\theta}(0) = \theta$. On the other hand, the deviation $\Delta^\varepsilon(t)$, which appears in the statement of Theorem 2.8 is related to the initial measure μ;[51] the first goal of this section is to obtain an explicit relation between the two definitions.

Remark 8.2. *In the following we will need to iterate complex standard pairs. The basic tool to do so will be a generalization of Proposition 3.3 where the potentials that appear are proportional to σ. This means that we will need $c_2 \geq C_\#|\sigma|$ in order for Proposition 3.3 to apply. Accordingly, by the condition $c_2\delta_{\mathbb{C}} \leq \pi/10$, stated just after (3.9), we will need to consider $\delta_{\mathbb{C}} \leq C_\#|\sigma|^{-1}$. On the other hand we will see shortly that we need worry only about $|\sigma| \leq \varepsilon^{-1/2-2\delta_*}$ for some conveniently chosen small constant $\delta_* > 0$. Due to this, we are going to consider* complex *standard pairs with $\delta \geq \delta_{\mathbb{C}} \geq \delta_{\mathrm{c}} = c_*\varepsilon^{1/2+2\delta_*}$ for some conveniently chosen small constant c_*. We will call* short complex standard pairs *the ones for which $\delta_{\mathbb{C}} = \delta_{\mathrm{c}}$ and* long complex standard pairs *the ones for which $\delta_{\mathbb{C}} = \delta$.*

Due to the above remark it is necessary to write a standard pair ℓ_0 as a family of short complex standard pairs. Recall that $\ell_0 = (\mathbb{G}_{\ell_0}, \rho_{\ell_0})$, $\mathbb{G}_{\ell_0} : [a_{\ell_0}, b_{\ell_0}] \to \mathbb{T}^2$, has length $|b_{\ell_0} - a_{\ell_0}| \in [\delta/2, \delta]$, where, as in the previous sections, δ is some fixed number independent on ε. Hence we must cut $[a_{\ell_0}, b_{\ell_0}]$ in $\delta\delta_{\mathrm{c}}^{-1}$ pieces $[\alpha_i, \alpha_{i+1}]$ of length between $\delta_{\mathrm{c}}/2$ and δ_{c}. We can then define the complex standard pairs $\ell_i^{\mathrm{c}} = (\mathbb{G}_{\ell_0,i}, \rho_i)$, where $\mathbb{G}_{\ell_0,i} = \mathbb{G}_{\ell_0}|_{[\alpha_i,\alpha_{i+1}]}$ and $\rho_i = Z_i^{-1}\rho_{\ell_0}\mathbb{1}_{[\alpha,\alpha_{i+1}]}$, $Z_i = \int_{\alpha_i}^{\alpha_{i+1}} \rho_{\ell_0}$.[52] Remark that $Z_i \sim \varepsilon^{1/2+2\delta_*}$ and $\sum_i Z_i = 1$. Clearly, for each continuous function B,

$$\mu_{\ell_0}(B) = \sum_i Z_i \mu_{\ell_i^{\mathrm{c}}}(B). \tag{8.2}$$

Let (x, θ) be distributed according to a measure in $\mathcal{P}_\varepsilon(\theta_0^*)$, we can apply to each standard pair in the family the decomposition (8.2). We can thus write

$$\Delta^\varepsilon(t) = \varepsilon^{-1/2}\sum_i \mathbb{1}_{\ell_i^{\mathrm{c}}}\left[\theta_\varepsilon(t) - \bar\theta(t, \theta^*_{\ell_i^{\mathrm{c}}})\right] - \mathbb{1}_{\ell_i^{\mathrm{c}}}\left[\bar\theta(t,\theta_0^*) - \bar\theta(t,\theta^*_{\ell_i^{\mathrm{c}}})\right]. \tag{8.3}$$

In addition, for any $\alpha > 0$, except for a set of exponentially small probability, the relation between the random variable in (8.1) and $\Delta^\varepsilon_{\ell_i^{\mathrm{c}}}(t) = \varepsilon^{-1/2}[\theta_\varepsilon(t) - \bar\theta(t, \theta^*_{\ell_i^{\mathrm{c}}})]$, under ℓ_i^{c}, is given by:

$$\begin{aligned}\Delta^\varepsilon_{\ell_i^{\mathrm{c}}}(t) &= \varepsilon^{-1/2}[\theta_\varepsilon(t) - \bar\theta(t,\theta)] + \mathcal{O}(\varepsilon^{1-2\delta_*})\\ &= \varepsilon^{-1/2}\left\{\Delta_{\lfloor t\varepsilon^{-1}\rfloor} + (t\varepsilon^{-1} - \lfloor t\varepsilon^{-1}\rfloor)[\Delta_{\lfloor t\varepsilon^{-1}\rfloor+1} - \Delta_{\lfloor t\varepsilon^{-1}\rfloor}]\right\} + \mathcal{O}(\varepsilon^{1-2\delta_*})\\ &= \varepsilon^{-1/2}\Delta_{\lfloor t\varepsilon^{-1}\rfloor} + \varepsilon^{1/2}(t\varepsilon^{-1} - \lfloor t\varepsilon^{-1}\rfloor)\hat\omega(x_{\lfloor t\varepsilon^{-1}\rfloor}, \theta_{\lfloor t\varepsilon^{-1}\rfloor}) + \mathcal{O}(\varepsilon^{1-2\delta_*}),\end{aligned} \tag{8.4}$$

where we have argued as in (5.3a) and used our large deviation results.[53]

Remark 8.3. *In the following we will consider only values of t such that $\lfloor t\varepsilon^{-1}\rfloor = t\varepsilon^{-1}$, i.e. we will assume $t \in \varepsilon\mathbb{N} \cap [0, T]$. As the formula above shows, the general case can be treated by modifying the last term in the sum defining $H_{0,k}$ in (8.8) below. We refrain from doing so explicitly to alleviate our notation. Note however that if one wanted to compute the first term of the Edgeworth expansion, then one would need to treat explicitly all times and even use a formula slightly more precise than (8.4), which anyhow also follows from the arguments used in (5.3a).*

---

[51] Recall the definition of the random variable $\Delta^\varepsilon(t) = \varepsilon^{-1/2}\left[\theta_\varepsilon(t) - \bar\theta(t,\theta_0^*)\right]$ where $\theta_\varepsilon(t)$ is defined in (2.7) by $\theta_\varepsilon(t) = \theta_{\lfloor t\varepsilon^{-1}\rfloor} + (t\varepsilon^{-1} - \lfloor t\varepsilon^{-1}\rfloor)[\theta_{\lfloor t\varepsilon^{-1}\rfloor+1} - \theta_{\lfloor t\varepsilon^{-1}\rfloor}]$, $\mu \in \mathcal{P}_\varepsilon(\theta_0^*)$.

[52] The reader should not be confused by the fact that the ℓ_i^{c} are real: the adjective "complex" here refers to the fact that they satisfy all the conditions for complex standard pairs, in particular the one stated in Remark 8.2 concerning their length.

[53] See the arguments around equation (8.10) for more details.

8.1. **Reduction to a Birkhoff sum.**
As it is often done in the study of sums of weakly dependent random variables (and already several times in this paper), we need to divide the time interval $[0,t]$ in blocks. For technical reasons it turns out to be convenient to allow such blocks to be of variable length. We thus consider a number R of blocks of length identified by the sequence $\{L_k\}_{k=0}^{R-1}$ and set

$$S_k = \sum_{j=0}^{k} L_j, \qquad\qquad S_{-1} = 0$$

so that $S_{R-1} = t\varepsilon^{-1}$. In our situation, it suffices to consider the case in which all the blocks are equal except the last one. More precisely: let us fix[54]

$$\delta_* \in (1/99, 1/32), \tag{8.5}$$

to be specified later, let $L_* = \varepsilon^{-3\delta_*}$ and define the lengths L_k as follows:

$$\begin{aligned} &L_k = L_* \quad \text{for all } k \in \{0, \cdots, R-2\}.\\ &L_* \leq L_{R-1} \leq 2L_*. \end{aligned} \tag{8.6}$$

Remark 8.4. *The estimates in this section are sharper than needed for our purposes, given our choice of L_*. Yet, they are instructive as they show, at very little extra cost, how to proceed if one wants to obtain a full Edgeworth expansion.*

Lemma 8.5. *For any $\varepsilon > 0$, let $t \in \varepsilon\mathbb{N} \cap [0, T]$ and $\{L_k\}_{k=0}^{R-1} \subset \mathbb{N}$ as above:*

$$\begin{aligned} \Delta_{t\varepsilon^{-1}} &= \sum_{k=0}^{R-1} \left[D_k + \mathcal{O}(\Delta_{L_k}^3)\right] \circ F_\varepsilon^{S_{k-1}}\\ D_k &= \widehat{\Xi}(t - \varepsilon S_k, \bar\theta_{L_k}) \left[\Delta_{L_k} + \frac{1}{2} P(t - \varepsilon S_k, \bar\theta_{L_k}) \Delta_{L_k}^2\right]\\ \widehat{\Xi}(s, \theta) &= \exp\left[\int_0^s \bar\omega'(\bar\theta(\tau, \theta)) d\tau\right]; \quad P(s, \theta) = \int_0^s \widehat{\Xi}(\tau, \theta) \bar\omega''(\bar\theta(\tau, \theta)) d\tau. \end{aligned} \tag{8.7}$$

Proof. Note that

$$\begin{aligned} \Delta_{t\varepsilon^{-1}} &= \theta_{t\varepsilon^{-1}} - \bar\theta_{t\varepsilon^{-1}} = \theta_{t\varepsilon^{-1} - L_0} \circ F_\varepsilon^{L_0} - \bar\theta(t - \varepsilon L_0, \bar\theta_{L_0})\\ &= \Delta_{t\varepsilon^{-1} - L_0} \circ F_\varepsilon^{L_0} - \bar\theta(t - \varepsilon L_0, \bar\theta_{L_0}) + \bar\theta(t - \varepsilon L_0, \theta_{L_0})\\ &= \Delta_{t\varepsilon^{-1} - L_0} \circ F_\varepsilon^{L_0} + \partial_\theta\bar\theta(t - \varepsilon L_0, \bar\theta_{L_0})\Delta_{L_0} + \frac{1}{2}\partial_\theta^2\bar\theta(t - \varepsilon L_0, \bar\theta_{L_0})\Delta_{L_0}^2\\ &\quad + \mathcal{O}(\Delta_{L_0}^3). \end{aligned}$$

Next, note that, by the smooth dependence on initial data of the solutions of ordinary differential equations, the functions $\eta_1 = \partial_\theta\bar\theta, \eta_2 = \partial_\theta^2\bar\theta$ solve, respectively, the differential equations $\dot\eta_1 = \bar\omega'(\bar\theta)\eta_1$, $\eta_1(0) = 1$ and $\dot\eta_2 = \bar\omega'(\bar\theta)\eta_2 + \bar\omega''(\bar\theta)\eta_1^2$, $\eta_2(0) = 0$. That is, $\partial_\theta\bar\theta(s, \theta) = \widehat{\Xi}(s, \theta)$ and $\partial_\theta^2\bar\theta(s, \theta) = \widehat{\Xi}(s, \theta)P(s, \theta)$. Iterating the above formulae yields the lemma. □

Next, we want to write the random variables D_k, associated to the k-th block, in terms of the (more explicit) random variables defined in (5.4b): recall that

[54] The choices of 1/32 and 1/99 are both arbitrary and largely irrelevant; in fact one could work with values of δ_* arbitrarily small (see Footnote 77) .

$H_k = H_{0,k} + H_{1,k}$:

$$H_{0,k} = \sum_{j=0}^{k-1} \Xi_{j,k}\hat\omega(x_j,\theta_j); \qquad H_{1,k} = -\frac{\varepsilon}{2}\sum_{j=0}^{k-1}\Xi_{j,k}\bar\omega'(\bar\theta_j)\bar\omega(\bar\theta_j) \tag{8.8}$$

$$\text{where } \Xi_{j,k} = \prod_{l=j+1}^{k-1}\left[1+\varepsilon\bar\omega'(\bar\theta_l)\right].$$

Lemma 8.6. *There exists ε_0 such that, for all $k \in \{0,\cdots,R-1\}$, $\varepsilon \in [0,\varepsilon_0]$, $j \in \{0,\cdots,L_k\}$, $\alpha \in (0,\delta_*]$ and standard pair ℓ we have*

$$\mu_\ell\left(\left\{|\Delta_j| \geq \varepsilon L_k^{1/2+\alpha}\right\}\right) \leq e^{-c_\# L_k^\alpha}$$
$$\mu_\ell\left(\left\{|\varepsilon H_j - \Delta_j| \geq \varepsilon^3 L_k^{2+2\alpha}\right\}\right) \leq e^{-c_\# L_k^\alpha}$$
$$\mu_\ell\left(\left\{\left|(\varepsilon H_{0,j})^2 - \Delta_j^2\right| \geq \varepsilon^3 L_k^{3/2+\alpha}\right\}\right) \leq e^{-c_\# L_k^\alpha}.$$

Proof. By Lemma 5.1 (or, more precisely, (5.6b))

$$\Delta_j - \varepsilon[H_{0,j}+H_{1,j}] = \varepsilon\sum_{l=0}^{j-1}\Xi_{l,j}\left[\frac{\bar\omega''(\bar\theta_l)}{2}\Delta_l^2 + \mathcal{O}(\Delta_l^3+\varepsilon^2)\right]. \tag{8.9}$$

Next, let us define $\mathcal{B}_{\alpha,j} = \{(x,\theta)\in\mathbb{T}^2 \;:\; |\Delta^*_{\ell,j}| \geq \varepsilon j^{\frac12+\alpha}\}$. By Proposition 2.3

$$\mu_\ell(\mathcal{B}_{\alpha,j}) \leq e^{-c_\# j^{2\alpha}}. \tag{8.10}$$

Hence, for all $j \leq C_\#\sqrt{L_k}$, $|\Delta_j| \leq \varepsilon L_k^{1/2}$, while for $j \in [C_\#\sqrt{L_k}, L_k]$, since we have $|\Delta^*_{\ell,j} - \Delta_j| \leq C_\#\varepsilon$, (8.10) implies

$$\mu_\ell\left(\left\{|\Delta_j| \geq \varepsilon L_k^{1/2+\alpha}\right\}\right) \leq e^{-c_\# j^{2\alpha}} \leq e^{-c_\# L_k^\alpha}$$

from which the first assertion of the Lemma follows. Next, we have

$$\left|\varepsilon\sum_{l=0}^{\min\{j,c_\#\sqrt{L_k}\}-1}\Xi_{l,j}\left[\frac{\bar\omega''(\bar\theta_l)}{2}\Delta_l^2 + \mathcal{O}(\Delta_l^3+\varepsilon^2)\right]\right| \leq \varepsilon^3 L_k^{3/2}.$$

This proves the second assertion for $j \leq C_\#\sqrt{L_k}$, while, for $j \in [C_\#\sqrt{L_k}, L_k]$,

$$\mu_\ell\left(\left\{|\varepsilon[H_{0,j}+H_{1,j}] - \Delta_j| \geq \varepsilon^3 j^{2+2\alpha}\right\}\right) \leq e^{-c_\# L_k^\alpha}, \tag{8.11}$$

which yields the second assertion in the general case, recalling the constraints (8.6) on L_k. The last assertion follows analogously since $\Delta_j^2 = (\varepsilon H_{0,j})^2 + 2(\Delta_j - \varepsilon H_{0,j})\Delta_j - (\Delta_j - \varepsilon H_{0,j})^2$, and, for $j \leq L_k$,

$$\mu_\ell\left(\left\{\sup_{0\leq j\leq\lfloor t\varepsilon^{-1}\rfloor}|(\varepsilon H_{0,j})^2 - \Delta_j^2| \geq \varepsilon^3 j^{3/2+\alpha}\right\}\right) \leq e^{-c_\# L_k^\alpha},$$

where we used the fact that $\Delta_j - H_{0,j} = \mathcal{O}(H_{1,j}) = \mathcal{O}(\varepsilon^2 j)$. □

The above Lemma, which is even sharper than necessary, suggests to define

$$\mathbb{M}_k = \widehat\Xi(t-\varepsilon S_k, \bar\theta_{L_k})\left[H_{0,L_k} + H_{1,L_k} + \frac{\varepsilon}{2}P(t-\varepsilon S_k,\bar\theta_{L_k})(H_{0,L_k})^2\right]. \tag{8.12}$$

Then, for $\alpha \leq \delta_* < 1/32$, Lemmata 8.5, 8.6 and equations (8.7), (8.10) yield

$$\begin{aligned} &\mu_{\ell_0}\left(\{|\Delta_{t\varepsilon^{-1}} - \varepsilon\mathbb{A}| \geq C_\#\varepsilon^2 L_*^{1+2\alpha}\}\right) \leq e^{-c_\#\varepsilon^{-3\alpha\delta_*}} \\ &\mathbb{A} = \sum_{k=0}^{R-1}\mathbb{M}_k\circ F_\varepsilon^{S_{k-1}}. \end{aligned} \tag{8.13}$$

Since $\varepsilon^2 L_*^{1+2\alpha} \leq \varepsilon^{3/2+\delta_*}$, $\Delta_{t\varepsilon^{-1}} - \varepsilon\mathbb{A}$ is $o(\varepsilon^{3/2})$ with probability almost one.

Thanks to (8.13) we have reduced ourselves to computing the distribution of the random variable $\varepsilon\mathbb{A}$. The rest of the paper will therefore mostly deal with the problem of obtaining a local CLT for the variable $\mathbb{A}$.

8.2. **Proof of the Local CLT.**

In this subsection we will obtain a LCLT for the random variable $\varepsilon\mathbb{A}$, defined in (8.13), assuming the validity of several propositions that will be proven later on. Using this result we will be able to prove the LCLT for $\Delta^\varepsilon(t)$.

Our first problem is that the random variable $\mathbb{A}$ may have a very rough density (if it has a density at all): it is then convenient to introduce a regularization procedure.[55] To this end let $\boldsymbol{Z}$ be a bounded, independent, zero average random variable so that $|\boldsymbol{Z}| \leq 1$ with smooth density $\psi \in \mathcal{C}^\infty$. We can then consider the random variable $\mathbb{A}^\dagger = \varepsilon\mathbb{A} + \varepsilon^{\beta_*}\boldsymbol{Z}$, where $\beta_* = \frac{3}{2} + \delta_*$ and recall that $\delta_* \in (1/99, 1/32)$. The random variable $\mathbb{A}^\dagger$ indeed admits a density, which we denote with $N_{\mu,\mathbb{A}^\dagger}$ (where μ denotes the distribution of initial conditions). In fact, denote by $\widehat{\psi}$ the Fourier transform of ψ:

$$
\begin{aligned}
N_{\mu,\mathbb{A}^\dagger}(y) &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\xi y}\mathbb{E}(e^{i\xi\mathbb{A}^\dagger})d\xi \\
&= \frac{1}{2\pi\varepsilon}\int_{\mathbb{R}} e^{-i\sigma\varepsilon^{-1}y}\mu(e^{i\sigma\mathbb{A}})\widehat{\psi}(\varepsilon^{\beta_*-1}\sigma)d\sigma. \qquad (8.14)
\end{aligned}
$$

The above discussion motivates us to prove the following result

Proposition 8.7. *For any $T > 0$ there exists ε_0 so that the following holds. For any real numbers $\varepsilon \in (0, \varepsilon_0)$, $t \in [\varepsilon^{1/1000}, T]$ so that $t\varepsilon^{-1} = \lfloor t\varepsilon^{-1}\rfloor$, any $\theta^* \in \mathbb{T}$ and any short complex standard pair ℓ^{c} so that $\theta^* = \mathrm{Re}(\mu_{\ell^{\mathrm{c}}}(\theta))$,[56] we have:*

$$
N_{\ell^{\mathrm{c}},\mathbb{A}^\dagger}(y) = \frac{e^{-y^2/(2\varepsilon\boldsymbol{\sigma}_t^2(\theta^*))}}{\boldsymbol{\sigma}_t(\theta^*)\sqrt{2\pi\varepsilon}} + \mathcal{O}(\varepsilon^{-7\delta_*}), \qquad (8.15)
$$

where $\boldsymbol{\sigma}_t(\cdot)$ is given by (2.20); in particular it is a differentiable function so that $|\boldsymbol{\sigma}_t'| \leq C_\#$.

Let us postpone the proof of Proposition 8.7 and see immediately how it implies our main result.

Proof of Theorem 2.8. Let us remind once again the reader that we will give the proof only in the case $t\varepsilon^{-1} = \lfloor t\varepsilon^{-1}\rfloor$ (see Remark 8.3). By equations (8.3) and (8.4), given any $I = [a, b]$ and $\kappa \in \mathbb{R}$, we have

$$
\begin{aligned}
\mathbb{P}_{\mu,\varepsilon}(\Delta^\varepsilon(t) \in \varepsilon^{1/2}I + \kappa) &= \sum_i Z_i\mathbb{P}_{\mu_{\ell_i^{\mathrm{c}}},\varepsilon}(\Delta^\varepsilon_{\ell_i^{\mathrm{c}}}(t) \in \varepsilon^{1/2}I + \kappa + \tau_i) \\
&\leq \sum_i Z_i\mathbb{P}_{\mu_{\ell_i^{\mathrm{c}}},\varepsilon}(\varepsilon^{-1/2}\Delta_{t\varepsilon^{-1}} \in \varepsilon^{1/2}I^+ + \kappa + \tau_i) + C_\# e^{-\varepsilon^{-c_\#}}
\end{aligned}
$$

where $\tau_i = \varepsilon^{-1/2}[\bar\theta(t, \theta_0^*) - \bar\theta(t, \theta^*_{\ell_i^{\mathrm{c}}})]$ and $I^+ = [a - C_\#\varepsilon^{1-\delta_*}, b + C_\#\varepsilon^{1-\delta_*}]$. By the same token

$$
\mathbb{P}_{\mu,\varepsilon}(\Delta^\varepsilon(t) \in \varepsilon^{1/2}I + \kappa) \geq \sum_i Z_i\mathbb{P}_{\mu_{\ell_i^{\mathrm{c}}},\varepsilon}(\varepsilon^{-1/2}\Delta_{t\varepsilon^{-1}} \in \varepsilon^{1/2}I^- + \kappa + \tau_i) - C_\# e^{-\varepsilon^{-c_\#}}
$$

where $I^- = [a + C_\#\varepsilon^{1-\delta_*}, b - C_\#\varepsilon^{1-\delta_*}]$. From now on we follow only the upper bound, the lower bound being more of the same.

---

[55] This is not the only way to handle the problem, it is just the one we find more convenient, see Remark 8.14 for a standard alternative.

[56] This generalizes (4.1) to the case of complex standard pairs.

By (8.13) and the definition of $\mathbb{A}^\dagger$ (see the beginning of this subsection) we have

$$\mathbb{P}_{\mu,\varepsilon}(\Delta^\varepsilon(t)\in\varepsilon^{1/2}I+\kappa)\le\sum_i Z_i\mathbb{P}_{\mu_{\ell_i^c},\varepsilon}(\varepsilon^{-1/2}\mathbb{A}^\dagger\in\varepsilon^{1/2}I^+ +\kappa+\tau_i)+C_\# e^{-\varepsilon^{-c_\#}}.$$

We can now use Proposition 8.7 to obtain

(8.16)

$$\begin{aligned}\mathbb{P}_{\mu,\varepsilon}(\Delta^\varepsilon(t)\in\varepsilon^{1/2}I+\kappa)&\le\sum_i Z_i\int_{\kappa+\tau_i+\varepsilon^{1/2}I^+}\left[\frac{e^{-\eta^2/(2\boldsymbol{\sigma}_t^2(\theta^*_{\ell_i^c}))}}{\boldsymbol{\sigma}_t(\theta^*_{\ell_i^c})\sqrt{2\pi}}+C_\#\varepsilon^{1/2-7\delta_*}\right]d\eta\\&=\int_{\mathbb{T}}\int_{\kappa+\varepsilon^{1/2}I^+}\left[\frac{e^{-(\eta-\varepsilon^{-1/2}[\bar\theta(t,\theta_0^*)-\bar\theta(t,\theta)])^2/(2\boldsymbol{\sigma}_t^2(\theta))}}{\boldsymbol{\sigma}_t(\theta)\sqrt{2\pi}}\mathcal{N}_\mu(d\theta)+C_\#\varepsilon^{1/2-7\delta_*}\right]d\eta,\end{aligned}$$

where $\mathcal{N}_\mu$ is the law of θ under μ. The obvious analog holds for the lower bound.

The above formula is valid for any standard family, but if $\mu\in\mathcal{P}_\varepsilon(\theta_0^*)$, since by definition $|\bar\theta(t,\theta_0^*)-\bar\theta(t,\theta)|\le C_\#\varepsilon$, we can obtain the simplified expression:

$$\varepsilon^{-1/2}\mathbb{P}_{\mu,\varepsilon}(\Delta^\varepsilon(t)\in\varepsilon^{1/2}I+\kappa)=\operatorname{Leb}I\cdot\left[\frac{e^{-\kappa^2/2\boldsymbol{\sigma}_t^2(\theta_0^*)}}{\boldsymbol{\sigma}_t(\theta_0^*)\sqrt{2\pi}}+\mathcal{O}(\varepsilon^{1/2-7\delta_*})\right]+\mathcal{O}(\varepsilon^{1/2-\delta_*}).$$

This proves the theorem. □

Our task is then reduced to the proof of Proposition 8.7.

Proof of Proposition 8.7. It suffices to compute the integral (8.14) when $\mu=\mu_{\ell^c}$ is a short complex standard pair. To do so, we find convenient to split the integral in five different regimes: let us fix $\sigma_0>0$ small enough and $C_1>0$ large enough to be determined later; also let $\varkappa=\beta_*-1+\delta_*=\frac12+2\delta_*$.[57] Recall moreover that we have chosen $L_*=\varepsilon^{-3\delta_*}$; we consider then the partition $\mathbb{R}=\bigcup_{k=0}^4\mathcal{J}_k$, where

$$\begin{aligned}&\mathcal{J}_0=\{|\sigma|\le C_1\varepsilon^2L_*\}, &&\mathcal{J}_1=\{C_1\varepsilon^2L_*<|\sigma|\le\varepsilon^{\delta_*}\},\\&\mathcal{J}_2=\{\varepsilon^{\delta_*}<|\sigma|\le\sigma_0\}, &&\mathcal{J}_3=\{\sigma_0<|\sigma|\le\varepsilon^{-\varkappa}\},\\&\mathcal{J}_4=\{\varepsilon^{-\varkappa}<|\sigma|\}.\end{aligned}$$

Correspondingly, we can rewrite (8.14) as

$$N_{\mu,\mathbb{A}^\dagger}=\mathcal{I}_0+\mathcal{I}_1+\mathcal{I}_2+\mathcal{I}_3+\mathcal{I}_4,$$

where each $\mathcal{I}_j$ denotes the contribution of $\mathcal{J}_j$ to the integral on the right hand side of (8.14). Recall that we are allowed to neglect contributions that are of order $\varepsilon^{-7\delta_*}$; we will now show that the main contribution to (8.14) is given by $\mathcal{I}_1$, as the contributions of all other terms are, in fact, negligible. First notice that the contribution of $\mathcal{I}_0$ can be neglected; in fact:

$$|\mathcal{I}_0|\le\frac{1}{2\pi\varepsilon}\int_{\sigma\le C_1\varepsilon^2L_*}\left|\widehat{\psi}(\varepsilon^{\beta_*-1}\sigma)\right|d\sigma\le C_\#\varepsilon L_*\|\psi\|_{L^1}\le C_\#\varepsilon^{1-3\delta_*}\tag{8.17}$$

The contribution of $\mathcal{I}_4$ can also be neglected, since, for each $r\in\mathbb{N}$, by Cauchy–Schwarz:

$$\begin{aligned}|\mathcal{I}_4|&\le\frac{\varepsilon^{-\beta_*}}{2\pi}\int_{\sigma\ge\varepsilon^{\beta_*-1-\varkappa}}\left|\widehat{\psi}(\sigma)\right|d\sigma\\&\le C_\#\varepsilon^{-\beta_*}\|\psi^{(r)}\|_{L^2}\left[\int_{\sigma\ge\varepsilon^{-\delta_*}}\sigma^{-2r}d\sigma\right]^{1/2}\\&\le C_\#\|\psi^{(r)}\|_{L^2}\varepsilon^{2r\delta_*-\delta_*-\beta_*}.\end{aligned}\tag{8.18}$$

[57] Informally, σ_0 specifies the region in which we can use perturbation theory, while C_1 and $\varepsilon^{-\varkappa}$ specifies the regions that can be bounded trivially, see equations (8.17), (8.18).

If we take r large enough, depending on the choice of δ_*, we can thus conclude that $|\mathcal{I}_4| \leq C_\# \|\psi^{(r)}\|_{L^2}\varepsilon^{100} \leq C_\# \varepsilon^{-7\delta^*}$. We are then left with the estimate of the contributions of $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$. We will (impressionistically) call $\mathcal{J}_1$ the *small* (frequencies) regime, $\mathcal{J}_2$ the *intermediate* regime and $\mathcal{J}_3$ the *large* regime.

The basic tool to compute these integrals is described by Lemma 8.9, which will be stated below. Before giving its statement, however, it is convenient to introduce a systematic notation for the many correlation terms that will appear in the sequel. It will turn out that, for the level of precision needed for our current investigation, the exact form of such terms is irrelevant. It will thus suffice to consider the following, very rough, bookkeeping strategy.

Notation 8.8. *Let $C^* > 1$ be some fixed constant sufficiently large. Given a standard pair ℓ, we will use the symbol $\mathfrak{C}^{k,p}_{\ell,l,\bar\imath}$ to denote a coefficient which depends only on the averaged trajectory $\bar\theta(t, \theta^*_\ell)$, indexed by $\bar\imath = (i_1, \cdots, i_l) \in \{-1, \cdots, L_k - 1\}^l$ (or $\bar\imath = \emptyset$ if $k = 0$)*[58] *and which satisfies the estimates $|\mathfrak{C}^{k,p}_{\ell,l,\bar\imath}| \leq (C^*)^l$, and $\sum_{\bar\imath} |\mathfrak{C}^{k,p}_{\ell,l,\bar\imath}| \leq (C^*)^l L_k^p$.*

We will use $A_{j,\bar\imath}$, $i_j \geq 0$, as a placeholder for an arbitrary $\mathcal{C}^2(\mathbb{T}^2, \mathbb{C})$ function possibly explicitly depending on ℓ such that $\|A_{j,\bar\imath}\|_{\mathcal{C}^1(\mathbb{T}^2,\mathbb{C})} \leq C^*$, and we assume conventionally that $A_{j,\bar\imath} = 1$ if $i_j = -1$. Finally, we will use the notation

$$\mathfrak{K}^{k,p}_{\ell,l} = \sum_{\bar\imath} \mathfrak{C}^{k,p}_{\ell,l,\bar\imath} \prod_{j=1}^{l} A_{j,\bar\imath} \circ F_\varepsilon^{i_j}.$$

For obvious reasons we will call such expressions correlation terms. *Note that $\mathfrak{K}^{k,p}_{\ell,l}\mathfrak{K}^{k,p'}_{\ell,l'} = \mathfrak{K}^{k,p+p'}_{\ell,l+l'}$. Finally, observe that $\mathfrak{K}^{k,p}_{\ell,l}$ can also be written as $\mathfrak{K}^{k,p}_{\ell,m}$ for any $m \geq l$ (just set $\mathfrak{C}^{k,p}_{\ell,m,\bar\imath} = 0$ if $i_j \neq -1$ for all $j > l$).*

Also let us introduce the potentials (recall that the value of t is fixed)

$$\varpi^k_{\ell,j}(x, \theta) = \widehat{\Xi}(t - \varepsilon S_k, \bar\theta^*_{\ell,L_k})\Xi^*_{\ell,j,L_k}\hat\omega(x, \theta) \tag{8.19}$$

where $\Xi^*_{\ell,j,k}$ is defined in (5.4a), $\widehat{\Xi}$ is defined in (8.7) and, generalizing (4.1):

$$\theta^*_\ell = \operatorname{Re}(\mu_\ell(\theta)) \; ; \quad \bar\theta^*_{\ell,k} = \bar\theta(\varepsilon k, \theta^*_\ell). \tag{8.20}$$

Let us fix $q \in \mathbb{N}$ sufficiently large to be specified later; associated with the above potentials, choosing a standard pair ℓ, $k \in \{0, \cdots, R-1\}$, $\mathfrak{C}^{k,0}_{\ell,0,\emptyset}$ and $(\mathfrak{K}^{k,2s}_{\ell,3s})_{s=1}^{q-1}$, we define an operator[59] $\mathcal{T}_k$: the operator $\mathcal{T}_k$ acts on complex measures over $\mathbb{T}^2$ as a "weighted L_k-push-forward with correlation terms up to q points", according to the following formula

$$\mathcal{T}_k\mu(g) = e^{i\sigma\mathfrak{C}_\ell(\varepsilon)}\mu\left(e^{i\sigma\sum_{j=0}^{L_k-1}\varpi^k_{\ell,j}\circ F_\varepsilon^j}\left[1 + \sum_{s=1}^{q-1}(i\sigma\varepsilon)^s\mathfrak{K}^{k,2s}_{\ell,3s}\right] g \circ F_\varepsilon^{L_k}\right), \tag{8.21}$$

where $\mathfrak{C}_\ell(\varepsilon)$ is a constant depending only on ℓ and ε. Observe that when $q = 1$ and $\mathfrak{C}_\ell(\varepsilon) = 0$, we recover the push-forward operator with complex potential (8.19) defined in (5.35) . The key fact is that the action of such operators on complex standard families can still be described in the standard pair language, as the following lemma shows.

---

[58] We use the convention that, for any set A, $A^0 = \{\emptyset\}$.

[59] To be precise $\mathcal{T}$ should have a lot of indexes $(\{\ell, k, q, \mathfrak{C}^{k,0}_{\ell,0,\emptyset}, \varpi^k_{\ell,j}, (\mathfrak{K}^{k,2s}_{\ell,3s})_{s=1}^{q-1}\})$, we drop all of them (except k) for readability.

Lemma 8.9. *There exists $\varepsilon_0 > 0$ such that, for each $k \in \{1, \cdots, R\}$, $\sigma \in \mathbb{R}$, short complex standard pair ℓ and $g \in L^\infty(\mathbb{T}^2, \mathbb{C})$ there exist a family of short standard pairs $\mathfrak{L}_\ell^k$ such that, provided $|\sigma| \le \varepsilon^{-1/2-2\delta_*}$ and $L_k \le \varepsilon^{-1/4+\delta_*}$, we have*

$$\mu_\ell\left(e^{i\sigma\mathbb{M}_k} g \circ F_\varepsilon^{L_k}\right) = \sum_{\ell' \in \mathfrak{L}_\ell^k} \nu_{k,\ell,\ell'}\mu_{\ell'}(g) + \mathcal{O}\left(\varepsilon^q\sigma^q L_k^{2q} + \sigma\varepsilon^2\delta_{\mathbb{C}}L_k^2\right) \cdot |\mu_\ell|(|g \circ F_\varepsilon^{L_k}|),$$

where $\displaystyle\sum_{\ell' \in \mathfrak{L}_\ell^k} \nu_{k,\ell,\ell'}\mu_{\ell'}(g) = \mathcal{T}_k\mu_\ell(g)$ *and $\mathcal{T}_k$ is given by (8.21) with $\mathfrak{C}_\ell(\varepsilon) = \varepsilon\mathfrak{C}_{\ell,0,\emptyset}^{k,1}$.*

Moreover, if $|\sigma| \le \sigma_0$, we can take $\{\ell\}$ and/or $\mathfrak{L}_\ell^k$ to consist of long standard pairs. In addition, if we define iteratively the standard families $\mathfrak{L}_{\ell_0}^0 = \{\ell_0\} = \{\ell\}$ and $\mathfrak{L}_{\ell_k}^k$ where, for all $\ell_k \in \mathfrak{L}_{\ell_{k-1}}^{k-1}$, $\mathfrak{L}_{\ell_k}^k$ is defined as above, then, for each $k \in \{0, \cdots, R-1\}$, if $q \ge 4$ and $L_* \le C\varepsilon^{-1/4+\delta_*+(3/4+\delta_*)/(2q-1)}$ for sufficiently small C, we have

$$\sum_{\ell_1 \in \mathfrak{L}_{\ell_0}^0} \cdots \sum_{\ell_{k+1} \in \mathfrak{L}_{\ell_k}^k} \prod_{j=0}^{k} |\nu_{j,\ell_j,\ell_{j+1}}| \le C_\#. \tag{8.22}$$

The proof of the above lemma will be given in the next subsection. We now show how to conclude the proof of Proposition 8.7: Lemma 8.9 and (8.13) allow to write the expectation $\mu(e^{i\sigma\mathbb{A}}) = \mu_{\ell^c}(e^{i\sigma\mathbb{A}})$ appearing in (8.14) as (recall $R = \mathcal{O}(\varepsilon^{-1}L_*^{-1})$):

$$\mu_{\ell^c}(e^{i\sigma\mathbb{A}}) = \sum_{\ell_1 \in \mathfrak{L}_{\ell^c}^0} \cdots \sum_{\ell_R \in \mathfrak{L}_{\ell_{R-1}}^{R-1}} \prod_{j=0}^{R-1} \nu_{j,\ell_j,\ell_{j+1}} + \mathcal{O}(\varepsilon^{q-1}\sigma^q L_*^{2q-1} + \sigma\varepsilon\delta_{\mathbb{C}}L_*). \tag{8.23}$$

Remark 8.10. *Note that the above decomposition depends on the choice of $\delta_{\mathbb{C}}$ which, in turns, depends on σ. From now on we will talk only of "complex standard pairs" and it will be understood that the families $\mathfrak{L}_{\ell_k}^k$ are made of short standard pairs for $\sigma \in \mathcal{J}_3$ and long standard pairs if $\sigma \in \mathcal{J}_1 \cup \mathcal{J}_2$.*

Note that the estimate given by (8.22) is very crude as it completely ignores possible cancellations among complex phases. Our next step are the following – much sharper– results which take into consideration such cancellations.

Proposition 8.11 (Large σ regime)**.** *For any $\delta_* \in (0, 1/32)$, $\sigma \in \mathcal{J}_3$, let $L_* \le L_{R-1} \le 2L_*$. Then, for any complex standard pair $\ell_{R-1} \in \mathfrak{L}_{\ell_{R-2}}^{R-2}$:*

$$\left|\sum_{\ell_R \in \mathfrak{L}_{\ell_{R-1}}^{R-1}} \nu_{R-1,\ell_{R-1},\ell_R}\right| = \mathcal{O}(\varepsilon^{2-9\delta_*}).$$

The proof of the above proposition will be given in Section 10.

Proposition 8.12 (Intermediate σ regime)**.** *For any $\delta_* \in (1/99, 1/32)$ and $\sigma \in \mathcal{J}_2$, let $L_* \le L_{R-1} \le 2L_*$. Then, for any complex standard pair $\ell_{R-1} \in \mathfrak{L}_{\ell_{R-2}}^{R-2}$:*

$$\left|\sum_{\ell_R \in \mathfrak{L}_{\ell_{R-1}}^{R-1}} \nu_{R-1,\ell_{R-1},\ell_R}\right| = \mathcal{O}(\varepsilon^{2-9\delta_*}).$$

The proof of the above proposition can be found in Section 11. As mentioned previously, the above propositions imply that the main contribution to the integral (8.14) is given by $\mathcal{I}_1$. The next proposition estimates precisely this contribution

Proposition 8.13 (Small σ regime). *For $\delta_* \in (1/99, 1/32)$, $\sigma \in \mathcal{J}_1$, $q \geq 5$ and $L_k = L_*$, $0 \leq k < R-1$:*

$$\mu_{\ell^c}(e^{i\sigma\mathbb{A}}) = e^{-\frac{2}{\varepsilon}\sigma^2 \boldsymbol{\sigma}_t^2(\theta_\ell^*)} + \mathcal{E}(\sigma, \varepsilon),$$

where recall (see (2.20) and (2.18)) that

$$\boldsymbol{\sigma}_t(\theta) = \int_0^t e^{2\int_s^t \bar{\omega}'(\bar{\theta}(s',\theta))ds'} \hat{\boldsymbol{\sigma}}^2(\bar{\theta}(s,\theta))ds$$

and $\mathcal{E}$ is a small remainder term in the sense that

$$\frac{1}{\varepsilon}\int_{\mathcal{J}_1} |\mathcal{E}(\sigma,\varepsilon)|d\sigma = \mathcal{O}(L_*^2 \log \varepsilon^{-1}).$$

The proof of Proposition 8.13 can be found in Section 13.

Let us now fix $q = 7$ and recall that $L_* = \varepsilon^{-3\delta_*}$. We can now compute the integral (8.14) by using estimates (8.17), (8.18), Lemma 8.9 and (8.23) in the first line below, while using Propositions 8.11–8.13 and (8.22) in the second line:

$$\begin{aligned}
N_{\ell^c,\mathbb{A}^\dagger}(y) &= \sum_{\ell_1 \in \mathfrak{L}^0_{\ell_0}} \cdots \sum_{\ell_R \in \mathfrak{L}^{R-1}_{\ell_{R-1}}} \int_{|\sigma| \leq \varepsilon^{-\frac{1}{2}-2\delta_*}} \frac{e^{-i\sigma\varepsilon^{-1}y}}{2\pi\varepsilon} \prod_{j=1}^{R-1} \nu_{j-1,\ell_{j-1},\ell_j} \widehat{\psi}(\varepsilon^{\beta_*-1}\sigma) + \mathcal{O}(1) \\
&= \mathcal{O}(\varepsilon^{-6\delta_*}\log\varepsilon^{-1}) + \frac{1}{2\pi\sqrt{\varepsilon}} \int_{\mathbb{R}} \widehat{\psi}(\eta\sqrt{\varepsilon}) e^{-i\varepsilon^{-1/2}\eta y - \frac{1}{2}\eta^2 \boldsymbol{\sigma}_t^2(\theta^*_{\ell_0}))} d\eta \\
&= \mathcal{O}(\varepsilon^{-6\delta_*}\log\varepsilon^{-1}) + \frac{e^{-y^2/(2\varepsilon\boldsymbol{\sigma}_t^2(\theta_\ell^*))}}{\boldsymbol{\sigma}_t(\theta_\ell^*)\sqrt{2\pi\varepsilon}},
\end{aligned}$$

where we have used that $|\widehat{\psi}(s) - 1| \leq C_\# s^2$, since $\boldsymbol{Z}$ has zero average. □

Remark 8.14. *In alternative to the above strategy we can choose $\mathcal{N}$ to be the distribution of a Gaussian random variable with density $\mathcal{N}' = \frac{1}{\sqrt{2\pi\varepsilon}\boldsymbol{\sigma}_t} e^{-x^2/2\boldsymbol{\sigma}_t^2\varepsilon}$ and apply* [25, Lemma 2, Chapter XVI.3] *with $T = \varepsilon^{-\beta_*}$*

$$|N_{\mathbb{A}}(x) - \mathcal{N}(x)| \leq \frac{1}{\pi}\int_{-\varepsilon^{-\beta_*}}^{\varepsilon^{-\beta_*}} \left|\frac{\widehat{N}_{\mathbb{A}}(\xi) - \widehat{\mathcal{N}}(\xi)}{\xi}\right| d\xi + \mathcal{O}(\varepsilon^{\beta_*-\frac{1}{2}}),$$

where $N_{\mathbb{A}}$ is the distribution of the random variable $\mathbb{A}$. The above integral can be computed, and shown to be small, using Propositions 8.11, 8.12 and 8.13 as we have done in the proof of Proposition 8.7. Note however that this would yield weaker results, as far as we are concerned, since the errors in the distribution function translate badly on errors for probability of small intervals (which represent our current interest).

8.3. Standard pairs decomposition.

To complete our argument we need to provide the proofs of the previously stated Propositions. Such proofs turn out to be rather laborious and to them is devoted the rest of the paper.

We start first with a generalization of Proposition 3.3.

Lemma 8.15. *There exists a constant $C_* \in (0,1)$ such that, for each short complex standard pair ℓ, $K \in \mathbb{N}$, $|\sigma| \leq \varepsilon^{-1/2-2\delta_*}$, imaginary potential families $\Omega = (i\sigma\varpi_j)_{j=0}^{K-1}$, ϖ_j defined as in (8.19), finite index set $\mathcal{A}$ and functions $\overline{B} = (B_a)_{a\in\mathcal{A}}$, $\|B_a\|_{\mathcal{C}^1} \leq C_*$, and times $\{k_a\}_{a\in\mathcal{A}} \subset \{0, \cdots, K-1\}$, $k_a = k_{a'} \implies a = a'$, for any $\vartheta \in (0, C_* 3^{-\sharp\mathcal{A}})$ there exists a short complex standard family $\mathfrak{L}_{K,\Omega,\overline{B}}$ such that, for* *all* $A \in L^\infty(\mathbb{T}^2, \mathbb{C})$*:*

$$\nu_{B,\ell}(A) := \mu_\ell\left(A \circ F_\varepsilon^K e^{i\sigma \sum_{j=0}^{K-1} \varpi_j \circ F_\varepsilon^j}\left[1 + \vartheta \prod_{a\in\mathcal{A}} B_a \circ F_\varepsilon^{k_a}\right]\right) = \sum_{\tilde\ell \in \mathfrak{L}_{K,\Omega,\overline{B}}} \nu_{\tilde\ell}\mu_{\tilde\ell}(A).$$

In addition, we have

$$|\nu_{\tilde\ell}| \geq C_\# \exp(-c_\# K) \tag{8.24}$$

for some uniform $C_\#, c_\#$*.*

Finally, if $|\sigma| \leq \sigma_0$ *and* $K \geq C \log \varepsilon^{-1}$*, for some* C *large enough, then the above holds also requiring that the family* $\mathfrak{L}_{K,\Omega,\overline{B}}$ *or/and* ℓ *consists of long complex standard pairs.*

Proof. We will use a baby cluster expansion like strategy. Note that, provided C_* is small enough, $\pi_a = \log(1 + B_a)$ are allowed potentials for both short and long standard pairs. Then, calling $\sharp\mathcal{A}$ the cardinality of $\mathcal{A}$, $\mathcal{P}(\mathcal{A})$ the power set of $\mathcal{A}$ and $S^c = \mathcal{A} \setminus S$, we have

$$\prod_{a\in\mathcal{A}} B_a \circ F_\varepsilon^{k_i} = \prod_{a\in\mathcal{A}} (B_a \circ F_\varepsilon^{k_a} + 1 - 1) = \sum_{S\in\mathcal{P}(\mathcal{A})} (-1)^{\sharp S^c} \prod_{a\in S} e^{\pi_a \circ F_\varepsilon^{k_a}}.$$

Then, if we set $\bar\pi_{S,k} = 0$ if $k \notin \{k_a\}_{a\in S}$ and $\bar\pi_{S,k_a} = \pi_a$ otherwise, we can write

$$\begin{aligned}\nu_{B,\ell}(A) &= \mu_\ell\left(A \circ F_\varepsilon^K e^{i\sigma\sum_{j=0}^{K-1}\varpi_j \circ F_\varepsilon^j}\right)\\ &\quad + \vartheta \sum_{S\in\mathcal{P}(\mathcal{A})} (-1)^{\sharp S^c} \mu_\ell\left(A \circ F_\varepsilon^K e^{\sum_{j=0}^{K-1}[i\sigma\varpi_j + \bar\pi_{S,j}]\circ F_\varepsilon^j}\right).\end{aligned}$$

We can now use Lemma 3.3 on each term of the above sums. Note that the decomposition in complex standard curves does not depend on the details of the potential but only on $|\sigma|$ and the dynamics. In particular, we can write

$$\nu_{B,\ell}(A) = \sum_{\ell'\in\mathfrak{L}} \nu^0_{\ell'}\mu_{\ell'}(A) + \vartheta \sum_{S\in\mathcal{P}(\mathcal{A})} (-1)^{\sharp S^c} \sum_{\ell'\in\mathfrak{L}_S} \nu_{S,\ell'}\mu_{\ell'}(A)$$

where $\mathfrak{L} = \{(\mathbb{G}_j, \rho_j^0), \nu^0\}$ and, for each $S \subset \mathcal{A}$, $\mathfrak{L}_S = \{(\mathbb{G}_j, \rho_{S,j}), \nu_S\}$. Note that, if $\ell'_j = (\mathbb{G}_j, \rho_{S,j})$,

$$\begin{aligned}|\nu_{S,\ell'_j}| &= \left|\mu_\ell\left(\mathbb{1}_{\ell'_j} e^{i\sigma\sum_{j=0}^{K-1}\varpi_j\circ F_\varepsilon^j} \prod_{a\in S}(1 + B_a) \circ F_\varepsilon^{k_a}\right)\right|\\ &\leq (1 + C_*)^{\sharp S}|\mu_\ell|\left(\mathbb{1}_{\ell'_j}\right),\end{aligned}$$

see Remark 3.7 for an explanation of the notation $\mathbb{1}_\ell$. Next, notice that, by the usual distortion arguments

$$|\operatorname{supp} \mathbb{1}_{\ell'_j}| \left|\frac{d}{dx}\sum_{j=0}^{K-1} \varpi_j \circ F_\varepsilon^j(x)\right| \leq C_\# \delta_{\mathbb{C}} \sum_{j=0}^{K-1} \lambda^{-K+j} \leq C_\# \delta_{\mathbb{C}}.$$

Thus

$$|\mu_\ell|\left(\mathbb{1}_{\ell'_j}\right) \leq C_\# \left|\mu_\ell\left(\mathbb{1}_{\ell'_j} e^{i\sigma\sum_{j=0}^{K-1}\varpi_j\circ F_\varepsilon^j}\right)\right|,$$

hence

$$|\nu_{S,j}| \leq C_\#(1 + C_*)^{\sharp S}|\nu_j^0|.$$

The above implies

$$\vartheta \sum_{S\in\mathcal{P}(\mathcal{A})} |\nu_{S,j}| \leq \vartheta C_\# |\nu_j^0| \sum_{S\in\mathcal{P}(\mathcal{A})} (1 + C_*)^{\sharp S} = \vartheta C_\#(2 + C_*)^{\sharp\mathcal{A}}|\nu_j^0| \leq \frac{|\nu_j^0|}{2},$$

provided C_* is small enough.

We can then define the standard family $\mathfrak{L}_{K,\Omega,\overline{B}} = \{(\mathbb{G}_j, \rho_j), \nu_j\}$ where

$$\nu_j = \nu_j^0 + \vartheta \sum_{S\in\mathcal{P}(\mathcal{A})} (-1)^{\sharp S^c} \nu_{S,j} \ ; \quad \rho_j = \nu_j^{-1}\left[\nu_j^0\rho_j^0 + \vartheta \sum_{S\in\mathcal{P}(\mathcal{A})} (-1)^{\sharp S^c} \nu_{S,j}\rho_{S,j}\right],$$

which concludes the first part of Lemma (see also Remark 3.6).

If $|\sigma| \le \sigma_0$, then the above argument works verbatim in the case in which ℓ is a long standard pair. If ℓ is a short complex standard pair, then, by Remark 3.5 we can, at each step, use complex standard pairs of length $\frac{3}{2}$ longer than the ones at the previous step, provided the length stays smaller than δ. Thus, at most after $C_\# \log \varepsilon^{-1}$ steps we have families that consist of long complex standard pairs. □

Proof of Lemma 8.9. Recall that, by (8.12) and (8.8), we have

$$\begin{aligned}
\mathbb{M}_k &= \widehat{\Xi}(t-\varepsilon S_k, \bar\theta_{L_k})\left\{\sum_{j=0}^{L_k-1} \Xi_{j,L_k}\left[\hat\omega(x_j,\theta_j) - \frac{\varepsilon}{2}\bar\omega'(\bar\theta_j)\bar\omega(\bar\theta_j)\right] + \varepsilon\,\mathbb{C}_k\right\}\\
\mathbb{C}_k &= \frac{1}{2}P(t-\varepsilon S_k, \bar\theta_{L_k})\left[\sum_{j=0}^{L_k-1}\Xi_{j,L_k}\hat\omega(x_j,\theta_j)\right]^2.
\end{aligned}$$

We would like to argue by using Proposition 3.3. Unfortunately, the above random variables are not of a form suitable to play the role of a potential since they contain products of functions computed at different times (that is *correlation terms*). We will solve this problem in three steps. First we will express the averaged trajectory $\bar\theta_k$ in terms of one starting from an initial condition that depends only on the standard pair, so that the averaged trajectory becomes deterministic. Then we will develop in series the exponential and finally we will show how to deal, in general, with the type of objects so obtained (using Lemma 8.15).

Arguing as at the end of Lemma 8.5 we have, for any function $\varphi \in \mathcal{C}^2$,[60]

$$\varphi(\bar\theta_k) = \varphi(\bar\theta^*_{\ell,k}) + \varphi'(\bar\theta^*_{\ell,k})\widehat{\Xi}(\varepsilon k, \theta^*_\ell)(\theta_0 - \theta^*_\ell) + \mathcal{O}(\varepsilon^2\delta_{\mathbb{C}}^2).$$

Using (8.20) and Notation 8.8 we can (see Appendix C for a detailed explanation on how to perform these, and similar, computations) rewrite (8.12) as

$$\begin{aligned}
&\mathbb{M}_k = \mathbb{M}^*_{\ell,k} + \varepsilon\mathfrak{K}^{k,2}_{\ell,3} + \mathcal{O}(\varepsilon^2\delta_{\mathbb{C}}L_k^2)\\
&\mathbb{M}^*_{\ell,k} = \widehat{\Xi}(t-\varepsilon S_k, \bar\theta^*_{\ell,L_k})\left\{\sum_{j=0}^{L_k-1}\Xi^*_{\ell,j,L_k}\left[\hat\omega(x_j,\theta_j) - \frac{\varepsilon}{2}\bar\omega'(\bar\theta^*_{\ell,j})\bar\omega(\bar\theta^*_{\ell,j})\right]\right\}.
\end{aligned}\tag{8.25}$$

By equations (8.25), (8.19) and the Taylor expansion

$$\begin{aligned}
e^{i\sigma\mathbb{M}_k} &= e^{i\sigma\left[\mathbb{M}^*_{\ell,k}+\varepsilon\mathfrak{K}^{k,2}_{\ell,3}\right]} + \mathcal{O}(\sigma\varepsilon^2\delta_{\mathbb{C}}L_k^2)\\
&= e^{i\sigma\left[\varepsilon\mathfrak{C}^{k,1}_{\ell,0,\emptyset} + \sum_{j=0}^{L_k-1}\varpi^k_{\ell,j}\circ F^j_\varepsilon\right]}\left[1+\mathbb{C}^*_{\ell,k,q}\right] + \mathcal{E}_{\ell,k,q}\\
\mathfrak{C}^{k,1}_{\ell,0,\emptyset} &= -\frac{1}{2}\widehat{\Xi}(t-\varepsilon S_k, \bar\theta^*_{\ell,L_k})\sum_{j=0}^{L_k-1}\Xi^*_{\ell,j,L_k}\bar\omega'(\bar\theta^*_{\ell,j})\bar\omega(\bar\theta^*_{\ell,j})\\
\mathbb{C}^*_{\ell,k,q} &= \sum_{s=1}^{q-1}(i\sigma\varepsilon)^s\mathfrak{K}^{k,2s}_{\ell,3s}\ ; \quad \mathcal{E}_{\ell,k,q} = \mathcal{O}(\varepsilon^q\sigma^q L_k^{2q} + \sigma\varepsilon^2\delta_{\mathbb{C}}L_k^2),
\end{aligned}\tag{8.26}$$

[60] In this section we use the shorthand notation $\mathcal{O}(\cdot) = \mathcal{O}_{L^\infty}(\cdot)$.

where we used Notation 8.8. Set

$$\vartheta := \sum_{s=1}^{q-1}\sum_{\bar\imath}\left|(\sigma\varepsilon)^s \mathfrak{C}^{k,2s}_{\ell,3s,\bar\imath}\right|$$

and notice that

$$\vartheta \le C_q \sum_{s=1}^{q-1}(|\sigma|\varepsilon L_*^2)^s \le C_q|\sigma|\varepsilon L_*^2 \le C_\# \varepsilon^{\delta_*}.$$

The above shows that, for ε small enough, (8.26) is a sum of terms to which we can apply Lemma 8.15, plus a small remainder; in fact[61]

$$\left[1+\mathbb{C}^*_{\ell,k,q}\right] = \vartheta^{-1}\sum_{s=1}^{q-1}\sum_{\bar\imath}|(\sigma\varepsilon)^s\mathfrak{C}^{2k,s}_{\ell,3s,\bar\imath}|\left[1+\vartheta\prod_{j=1}^{3s}A_{j,\bar\imath}\circ F_\varepsilon^{i_j}\right].$$

We have thus written $e^{i\sigma\mathbb{M}_k}$ as a weighted sum of terms which satisfy the hypotheses of Lemma 8.15, also the analogous of Remark 3.6 applies. Note that, again, the decomposition in standard curves can be chosen to be exactly the same for all terms. We can then define the standard family $\mathfrak{L}^k_\ell$ exactly as it was done at the end of the proof of Lemma 8.15. By (8.24) we conclude that the total weight of each standard pair differs uniformly from zero, which allows to normalize the densities. This proves the first part of the lemma.

To conclude the proof we need to prove (8.22); notice that, by the first part of the lemma and using the same notation as in Remark 3.7,[62] for any $0\le r\le s\le R$, we have[63]

$$\begin{aligned}
&\mu_{\ell_r}\left(e^{i\sigma\sum_{j=r}^s \mathbb{M}_j\circ F_\varepsilon^{S_{j-1}-S_{r-1}}}\mathbb{1}_{\ell_{s+1}}\circ F_\varepsilon^{S_{s-1}-S_{r-1}}\cdots\mathbb{1}_{\ell_{r+2}}\circ F_\varepsilon^{S_r-S_{r-1}}\mathbb{1}_{\ell_{r+1}}\right)\\
&=\mu_{\ell_r}\left(\left[e^{i\sigma\sum_{j=r+1}^s \mathbb{M}_j\circ F_\varepsilon^{S_{j-1}-S_r}}\mathbb{1}_{\ell_{s+1}}\circ F_\varepsilon^{S_{s-1}-S_r}\cdots\mathbb{1}_{\ell_{r+2}}\right]\circ F_\varepsilon^{L_r}\cdot e^{i\sigma\mathbb{M}_r}\mathbb{1}_{\ell_{r+1}}\right)\\
&=\nu_{r,\ell_r,\ell_{r+1}}\mu_{\ell_{r+1}}\left(e^{i\sigma\sum_{j=r+1}^s \mathbb{M}_j\circ F_\varepsilon^{S_{j-1}-S_r}}\mathbb{1}_{\ell_{s+1}}\circ F_\varepsilon^{S_{s-1}-S_r}\cdots\mathbb{1}_{\ell_{r+2}}\right)\\
&\quad+\mathcal{O}(\varepsilon^q\sigma^q L_r^{2q}+\sigma\varepsilon^2\delta_{\mathbb{C}}L_r^2)|\mu_{\ell_r}|(\mathbb{1}_{\ell_{s+1}}\circ F_\varepsilon^{S_{s-1}-S_{r-1}}\cdots\mathbb{1}_{\ell_{r+1}}).
\end{aligned}$$

Iterating the above equation yields, for all $r\le s$,

$$\begin{aligned}
&\mu_{\ell_r}\left(e^{i\sigma\sum_{j=r}^s \mathbb{M}_j\circ F_\varepsilon^{S_{j-1}-S_{r-1}}}\mathbb{1}_{\ell_{s+1}}\circ F_\varepsilon^{S_{s-1}-S_{r-1}}\cdots\mathbb{1}_{\ell_{r+1}}\right)\\
&=\prod_{l=r}^s \nu_{l,\ell_l,\ell_{l+1}}+\mathcal{O}(\varepsilon^q\sigma^q L_*^{2q}+\sigma\varepsilon^2\delta_{\mathbb{C}}L_*^2)\\
&\quad\times\sum_{j=r}^s\prod_{l=r}^{j-1}|\nu_{l,\ell_l,\ell_{l+1}}|\,|\mu_{\ell_j}|\left(\mathbb{1}_{\ell_{s+1}}\circ F_\varepsilon^{S_{s-1}-S_{r-1}}\cdots\mathbb{1}_{\ell_{j+1}}\right).
\end{aligned}\tag{8.27}$$

In particular, choosing $s=r$ we conclude that there exists $C_2>0$ such that:,

$$\sum_{\ell_{r+1}\in\mathfrak{L}^r_{\ell_r}}|\nu_{r,\ell_r,\ell_{r+1}}|\le\sum_{\ell_{r+1}\in\mathfrak{L}^r_{\ell_r}}C_\#|\mu_{\ell_r}|\left(\mathbb{1}_{\ell_{r+1}}\right)\le C_2.\tag{8.28}$$

---

[61] Note that we have absorbed the sign of $\sigma\mathfrak{C}^{k,s}_{\ell,3s,\bar\imath}$ into some $A_{j,\bar\imath}$, which is always possible since the $A_{j,\bar\imath}$ are names for arbitrary functions.

[62] Below we consider $\mathbb{1}_{\ell_j}$ to be a function defined on the standard pair ℓ_{j-1}. Also notice that $\mathbb{1}_{\ell_j}$ can be written, if needed, as the restriction to ℓ_{j-1} of $\varphi_{\ell_{j-1},\ell_j}\circ F_\varepsilon^{L_{j-1}}$, for some function $\varphi_{\ell_{j-1},\ell_j}\in L^\infty(\mathbb{T}^2,\mathbb{R})$.

[63] We use the convention that $\sum_{j=a}^b c_j=0$ and $\prod_{j=a}^b c_j=1$ if $b<a$.

To conclude we prove, by induction on $m = s - r$, that

$$\sum_{\ell_{r+1}\in\mathfrak{L}^r_{\ell_r}}\cdots\sum_{\ell_{s+1}\in\mathfrak{L}^s_{\ell_s}}\prod_{j=r}^{s}|\nu_{j,\ell_j,\ell_{j+1}}| \leq 2C_2 \tag{8.29}$$

Equation (8.28) shows that (8.29) holds for each r and $m = 0$. Let us suppose it holds for each r and $n \leq m$ for some $m \in \{0, \cdots R-2\}$. Let $s = r + m + 1$. Then, recalling (3.10), the fact that $|\sigma|\delta_{\mathbb{C}} \leq C_\#$ and the condition on L_*, we can use (8.27) to write:

$$\begin{aligned}
&\sum_{\ell_{r+1}\in\mathfrak{L}^r_{\ell_r}}\cdots\sum_{\ell_{s+1}\in\mathfrak{L}^s_{\ell_s}}\prod_{j=r}^{s}|\nu_{j,\ell_j,\ell_{j+1}}| \leq |\mu_{\ell_r}| + C_\#\,|\mu_{\ell_r}|\,(\varepsilon^q\sigma^q L_*^{2q} + \sigma\varepsilon^2\delta_{\mathbb{C}}L_*^2)\\
&+ C_\#\sum_{j=r+1}^{s}\sum_{\ell_{r+1}\in\mathfrak{L}^r_{\ell_r}}\cdots\sum_{\ell_j\in\mathfrak{L}^{j-1}_{\ell_{j-1}}}\prod_{l=r}^{j-1}|\nu_{l,\ell_l,\ell_{l+1}}|\,|\mu_{\ell_j}|\,(\varepsilon^q\sigma^q L_*^{2q} + \sigma\varepsilon^2\delta_{\mathbb{C}}L_*^2)\\
&\leq C_2 + (m+1)C_2^2 C_\#(\varepsilon^q\sigma^q L_*^{2q} + \sigma\varepsilon^2\delta_{\mathbb{C}}L_*^2) \leq 2C_2,
\end{aligned}$$

provided ε is small enough and $q \geq 4$. □

The remaining sections of the paper are devoted to the proofs of Propositions 8.11, 8.12 and 8.13 although we first need a preparatory technical section.

9. One block estimate: technical preliminaries

Our next step consists in transforming the sums on the standard pairs associated to each of the R blocks into an expression involving transfer operators related to a cocycle over the (slowly varying) averaged dynamics. This will at last allow us to perform the needed computations by functional analytic means.

Let us start by defining the slowly varying dynamics. Let $\ell = (\mathbb{G}, \rho)$ be a complex standard pair; recall that we introduced the notations $\theta^*_\ell = \mathrm{Re}(\mu_\ell(\theta_0))$, $\bar\theta^*_{\ell,k} = \bar\theta(\varepsilon k, \theta^*_\ell)$ in (8.20), where $\bar\theta(t,\theta)$ is the unique solution of (2.3) with initial condition $\bar\theta(0,\theta) = \theta$. Recall also that we defined (in (5.1a)) $\Delta^*_{\ell,k} = \theta_k - \bar\theta^*_{\ell,k}$ and that, for real standard pairs, we will regard x_k and $\Delta^*_{\ell,k}$ as functions on ℓ (see Remark 3.8).

Let us define the shorthand notations $\bar f_{\ell,k} = f(\cdot, \bar\theta^*_{\ell,k})$, $\bar f^{(n)}_\ell = \bar f_{\ell,n-1}\circ\cdots\circ \bar f_{\ell,0}$; consider the map $\bar F_\varepsilon(x,\theta) = (f(x,\theta), \bar\theta(\varepsilon,\theta))$. Observe that $(\bar f^{(k)}_\ell(x), \bar\theta^*_{\ell,k}) = \bar F^k_\varepsilon(x, \theta^*_\ell)$, i.e. the first component of $\bar F_\varepsilon$ yields our wanted slowly varying dynamics. Finally, let us define the function

$$\bar\Lambda_{j,k} = \prod_{r=j}^{k}(\partial_x f)^{-1}\circ \bar F^r_\varepsilon(\cdot, \theta^*_\ell) \tag{9.1}$$

Notice that, by definition, $\bar\Lambda_{j,k} < \lambda^{-(k-j)-1}$ and for any $x \in \mathbb{T}$:

$$\left|\frac{d}{dx}\bar\Lambda_{0,j}(x)\right| \leq C_\#\bar\Lambda_{0,j}(x)\sum_{l=0}^{j-1}\bar\Lambda_{0,l}(x)^{-1} \leq C_\#. \tag{9.2}$$

9.1. **Error in the slowly varying dynamics approximation.**

Our first task is to obtain sufficiently good estimates on the difference between F^k_ε and $\bar F^k_\varepsilon$ when k is not too large.

Lemma 9.1. *Fix a complex standard pair $\ell = (\mathbb{G}_\ell, \rho_\ell)$ of length δ_c and $L \in \mathbb{N}$ so that $L \le C_\# \varepsilon^{-1/2}$. There exists a diffeomorphism $\Upsilon_{\ell,L} : [a,b] \to [\bar a, \bar b]$ such that $(x_L, \theta_L) = F_\varepsilon^L \circ \mathbb{G}_\ell(x) = \left(\bar f_\ell^{(L)} \circ \Upsilon_{\ell,L}(x), \theta_L\right)$ with*

$$\frac{d\Upsilon_{\ell,L}}{dx} = (1 - G' s_L)\upsilon_L^+ \Lambda_{0,L-1} \tag{9.3}$$

where υ_L^+ was defined in (3.6) and $\Lambda_{k,j} = \bar\Lambda_{k,j} \circ \Upsilon_{\ell,L}$. Moreover $\Upsilon_{\ell,L}$ satisfies the following estimates:

$$\|\Upsilon_{\ell,L} - \mathbb{1}\|_{\mathcal{C}^0} \le C_\# \varepsilon \min\{1, L^2 \delta_c\} \qquad \frac{d\Upsilon_{\ell,L}}{dx} = 1 + \mathfrak{R}_\ell^{1,2} \tag{9.4}$$

where the notation $\mathfrak{R}_\ell^{p,q}$ denotes an arbitrary differentiable function of x that satisfies the bounds

$$\|\mathfrak{R}_\ell^{p,q}\|_{\mathcal{C}^0} \le C_\# \varepsilon^p L^q \qquad \left|\frac{d\,\mathfrak{R}_\ell^{p,q}}{dx}\right| \le C_\# \Lambda_{0,L-1}^{-1}. \tag{9.5}$$

Additionally, for any $k \in \{0, \cdots, L\}$, let us introduce the functions

$$\bar\xi_{\ell,k} = x_k \circ \Upsilon_{\ell,L}^{-1} - \bar x_k \qquad \bar\Delta_{\ell,k}^* = \Delta_{\ell,k}^* \circ \Upsilon_{\ell,L}^{-1}$$

where we introduced the shorthand notation $\bar x_k = \bar f_\ell^{(k)}(\cdot)$. Recall the definition of the quantities $\Xi_{\ell,j,k}^*$ given in (5.5); then let

$$\overline{W}_{\ell,k} = \varepsilon^{-1} \Xi_{\ell,-1,k}^* \bar\Delta_{\ell,0}^* + \sum_{j=0}^{k-1} \Xi_{\ell,j,k}^* \hat\omega(\bar x_l, \bar\theta_{\ell,l}^*) \tag{9.6a}$$

$$\overline{\mathfrak{W}}_{\ell,k} = -\sum_{l=k}^{L-1} \Lambda_{0,l-k} \partial_\theta f(\bar x_l, \bar\theta_{\ell,l}^*) \overline{W}_{\ell,l} \tag{9.6b}$$

Moreover define:

$$\overline{W}_{\ell,k,2} = \sum_{j=0}^{k-1} \left[\partial_x \hat\omega(\bar x_j, \bar\theta_{\ell,j}^*) \overline{\mathfrak{W}}_{\ell,j} + \partial_\theta \hat\omega(\bar x_j, \bar\theta_{\ell,j}^*) \overline{W}_{\ell,j} - \frac{1}{2} \bar\omega'(\bar\theta_{\ell,j}^*) \bar\omega(\bar\theta_{\ell,j}^*)\right]$$

$$\begin{aligned}\overline{\mathfrak{W}}_{\ell,k,2} = -\sum_{l=k}^{L-1} \Lambda_{0,l-k} \Big[&\partial_\theta f(\bar x_l, \bar\theta_{\ell,l}^*) \overline{W}_{\ell,l,2} + \frac{1}{2} \partial_\theta^2 f(\bar x_l, \bar\theta_{\ell,l}^*) \overline{W}_{\ell,l}^2 \\ &+ \frac{1}{2} \partial_\theta \partial_x f(\bar x_l, \bar\theta_{\ell,l}^*) \overline{W}_{\ell,l} \overline{\mathfrak{W}}_{\ell,l} + \frac{1}{2} \partial_x^2 f(\bar x_l, \bar\theta_{\ell,l}^*) \overline{\mathfrak{W}}_{\ell,l}^2\Big].\end{aligned}$$

Then the following bounds hold true

$$\bar\Delta_{\ell,k}^* = \varepsilon \overline{W}_{\ell,k} + \varepsilon^2 \overline{W}_{\ell,k,2} + \overline{\mathfrak{R}}_\ell^{3,3} \qquad \bar\xi_{\ell,k} = \varepsilon \overline{\mathfrak{W}}_{\ell,k} + \varepsilon^2 \overline{\mathfrak{W}}_{\ell,k,2} + \overline{\mathfrak{R}}_\ell^{3,3}, \tag{9.7}$$

where the notation $\overline{\mathfrak{R}}_\ell^{p,q}$ is analogous to $\mathfrak{R}_\ell^{p,q}$ but with Λ replaced by $\bar\Lambda$ in (9.5).

Remark 9.2. *Note that the above lemma is essentially a series expansion in which we only keep the first few terms. More precise formulae can be obtained, if needed, at the price of more work.*

Remark 9.3. *The approximation formulae obtained in the previous lemma are close, in spirit, to the ones obtained earlier in Lemma 5.1, but differ from them because they are written in terms of the averaged dynamics $(\bar x_k, \bar\theta_{\ell,k}^*)$, rather than the real dynamics (x_k, θ_k).*

Remark 9.4. *Observe that the random variables $\bar\xi_{\ell,k}$ and $\bar\Delta^*_{\ell,k}$ (defined in the previous lemma) do depend on L (through $\Upsilon_{\ell,L}$). In order to make the notation precise their symbols should thus have indices L. Since it will not create any confusion, we omit some of the indices to ease notation. Similarly, we will suppress the indices in Υ as well when no confusion arises,*

Proof of Lemma 9.1. The lemma follows from a variation on the proof given for Lemma 4.2. Let us recall that we denote with $\pi_x : \mathbb{T}^2 \to \mathbb{T}$ the projection on the x-coordinate and define, for $\varrho \in [0,1]$,

$$\bar{\mathcal{H}}_{\ell,L}(x,z;\varrho) = \pi_x F^L_{\varrho\varepsilon}(x, \theta^*_\ell + \varrho(G(x) - \theta^*_\ell)) - \pi_x \bar F^L_{\varrho\varepsilon}(z, \theta^*_\ell).$$

As in the proof of Lemma 4.2, observe that $\bar{\mathcal{H}}_{\ell,L}(x,x;0) = 0$, and moreover $\partial_z \bar{\mathcal{H}}_{\ell,L} = -\partial_z(\pi_x \bar F^L_{\varrho\varepsilon}(\cdot,\theta^*_\ell)) < -\lambda^L$. Therefore the implicit function theorem implies that for any $\varrho \in [0,1]$ there exists a diffeomorphism $\Upsilon_{\ell,L}(\cdot;\varrho)$ so that $\bar{\mathcal{H}}_{\ell,L}(x,\Upsilon_{\ell,L}(x;\varrho);\varrho) = 0$. Define $\Upsilon_{\ell,L}(x) = \Upsilon_{\ell,L}(x,1)$; then $\pi_x F^L_\varepsilon \circ \mathbb{G}_\ell = x_L = \bar f^{(L)}_\ell \circ \Upsilon_{\ell,L}$. The expression (9.3) then immediately follows using the notation and discussion of Subsection 3.1. Let us postpone the derivation of (9.4) to the end of the proof and first obtain the bounds (9.7). Using (5.6a) yields

$$
\begin{aligned}
\bar\Delta^*_{\ell,k} &= \Xi^*_{\ell,-1,k}\bar\Delta^*_{\ell,0} + \varepsilon\sum_{j=0}^{k-1}\Xi^*_{\ell,j,k}\left[\hat\omega(x_j,\theta_j) - \frac{\varepsilon}{2}\bar\omega'(\bar\theta^*_{\ell,j})\bar\omega(\bar\theta^*_{\ell,j})\right] + \mathcal{O}\left(\varepsilon\sum_{j=0}^{k-1}(\bar\Delta^*_{\ell,j})^2 + \varepsilon^3 k\right)\\
&= \Xi^*_{\ell,-1,k}\bar\Delta^*_{\ell,0} + \varepsilon\sum_{j=0}^{k-1}\Xi^*_{\ell,j,k}\Big[\hat\omega(\bar x_j,\bar\theta^*_{\ell,j}) + \partial_x\hat\omega(\bar x_j,\bar\theta^*_{\ell,j})\bar\xi_j + \partial_\theta\hat\omega(\bar x_j,\bar\theta^*_{\ell,j})\bar\Delta^*_{\ell,j}\\
&\qquad - \frac{\varepsilon}{2}\bar\omega'(\bar\theta^*_{\ell,j})\bar\omega(\bar\theta^*_{\ell,j})\Big] + \mathcal{O}(\varepsilon^3k^3) + \sum_{j=0}^{k-1}\mathcal{O}(\varepsilon\bar\xi_j^2)
\end{aligned}
\tag{9.8}
$$

where we have used (5.2). In addition, we can consider the Taylor expansion

$$
\begin{aligned}
\bar\xi_{k+1} &= f(x_k,\theta_k) - f(\bar x_k,\bar\theta^*_{\ell,k})\\
&= \partial_x f(\bar x_k,\bar\theta^*_{\ell,k})\bar\xi_k + \partial_\theta f(\bar x_k,\bar\theta^*_{\ell,k})\bar\Delta^*_{\ell,k} + \frac12\partial^2_\theta f(\bar x_k,\bar\theta^*_{\ell,k})(\bar\Delta^*_{\ell,k})^2\\
&\quad + \frac12\partial_\theta\partial_x f(\bar x_k,\bar\theta^*_{\ell,k})\bar\Delta^*_{\ell,k}\bar\xi_k + \frac12\partial^2_x f(\bar x_k,\bar\theta^*_{\ell,k})\bar\xi_k^2 + \mathcal{O}((\bar\Delta^*_{\ell,k})^3 + \bar\xi_k^3).
\end{aligned}
\tag{9.9}
$$

From the first line of (9.9) we have $|\bar\xi_{k+1}| \geq \lambda|\bar\xi_k| - C_\#|\bar\Delta^*_{\ell,k}|$. Recalling that, by definition, $\bar\xi_L = 0$, we can conclude that

$$|\bar\xi_k| \leq C_\#\sum_{j=k}^{L-1}\lambda^{k-j}|\bar\Delta^*_{\ell,j}| \leq C_\#\varepsilon(k+1). \tag{9.10}$$

Moreover, by the above estimates, we have

$$\bar\Delta^*_{\ell,k} = \varepsilon\overline{W}_{\ell,k} + \mathcal{O}(\varepsilon^2k^2). \tag{9.11}$$

A more precise result can now be obtained by (backward) iteration of (9.9):

$$
\begin{aligned}
\bar\xi_k &= -\sum_{j=k}^{L-1}\prod_{l=k}^{j}(\partial_x f(\bar x_l,\bar\theta^*_{\ell,l}))^{-1}\left[\partial_\theta f(\bar x_j,\bar\theta^*_{\ell,j})\bar\Delta^*_{\ell,j} + \mathcal{O}(\bar\xi_j^2 + (\bar\Delta^*_{\ell,j})^2)\right]\\
&= -\sum_{j=k}^{L-1}\Lambda_{k,j}\partial_\theta f(\bar x_j,\bar\theta^*_{\ell,j})\bar\Delta^*_{\ell,j} + \mathcal{O}(\varepsilon^2L^2) = \varepsilon\overline{\mathfrak{W}}_{\ell,k} + \mathcal{O}(\varepsilon^2L^2).
\end{aligned}
\tag{9.12}
$$

Finally, we can get a sharper estimate for $\bar\Delta^*_{\ell,k}$ by substituting (9.12) and (9.11) in (9.8):

$$\begin{aligned}\bar\Delta^*_{\ell,k} &= \varepsilon\overline{W}_{\ell,k} + \varepsilon^2\sum_{j=0}^{k-1}\Big[\partial_x\hat\omega(\bar x_j,\bar\theta^*_{\ell,j})\overline{\mathfrak{W}}_{\ell,j} + \partial_\theta\hat\omega(\bar x_j,\bar\theta^*_{\ell,j})\overline{W}_{\ell,j} \\ &\quad -\frac{1}{2}\bar\omega'(\bar\theta^*_{\ell,j})\bar\omega(\bar\theta^*_{\ell,j})\Big] + \mathcal{O}(\varepsilon^3L^3) = \varepsilon\overline{W}_{\ell,k} + \varepsilon^2\overline{W}_{\ell,k,2} + \mathcal{O}(\varepsilon^3L^3);\end{aligned} \tag{9.13}$$

and a sharper estimate for $\bar\xi_k$ by writing (9.9) as

$$\begin{aligned}\bar\xi_{k+1} &= \partial_x f(\bar x_k,\bar\theta^*_{\ell,k})\bar\xi_k + \varepsilon\partial_\theta f(\bar x_k,\bar\theta^*_{\ell,k})\overline{W}_{\ell,k} + \varepsilon^2\partial_\theta f(\bar x_k,\bar\theta^*_{\ell,k})\overline{W}_{\ell,k,2} \\ &\quad + \frac{\varepsilon^2}{2}\partial^2_\theta f(\bar x_k,\bar\theta^*_{\ell,k})\overline{W}^2_{\ell,k} \\ &\quad + \frac{\varepsilon^2}{2}\partial_\theta\partial_x f(\bar x_k,\bar\theta^*_{\ell,k})\overline{W}_{\ell,k}\overline{\mathfrak{W}}_{\ell,k} + \frac{\varepsilon^2}{2}\partial^2_x f(\bar x_k,\bar\theta^*_{\ell,k})\overline{\mathfrak{W}}^2_{\ell,k} + \mathcal{O}(\varepsilon^3L^3),\end{aligned} \tag{9.14}$$

which, iterating backward as before, yields the wanted result. The bound on the derivatives of the error terms, that is needed to write $\mathcal{O}(\varepsilon^3L^3)$ as $\mathfrak{R}^{3,3}_\ell$, follows by definition of $\bar\Delta^*_{\ell,k}$ and $\bar\xi_k$, (9.2) and the fact that $|\Upsilon'_{\ell,L}| \le C_\#$, which in turn follows from the second bound in (9.4).

In order to conclude the proof we now proceed to prove the two bounds of (9.4), which will be obtained by a careful analysis of (9.3). Recall the definition (3.6) of the quantities υ^+_k and u_k; by the discussion of Subsection 3.1 (see (3.3)) we have

$$\begin{aligned}\upsilon^+_L &= \prod_{k=0}^{L-1}[\partial_x f(x_k,\theta_k) + \varepsilon\partial_\theta f(x_k,\theta_k)u_k] \\ u_{k+1} &= \frac{\partial_x\omega(x_k,\theta_k) + (1+\varepsilon\partial_\theta\omega(x_k,\theta_k))u_k}{\partial_x f(x_k,\theta_k) + \varepsilon\partial_\theta f(x_k,\theta_k)u_k}, \quad \varepsilon u_0 = G'(x).\end{aligned} \tag{9.15}$$

As already noted we have $|u_k| \le c_1$ (one can also see this using Proposition 3.3, since εu_k is the slope of a standard curve). The above immediately implies, using (9.3), (9.10) and (5.2):

$$\begin{aligned}\frac{d\Upsilon_{\ell,L}}{dx} &= (1+\mathcal{O}(\varepsilon))\exp\left[\sum_{j=0}^{L-1}\log\partial_x f(x_j,\theta_j) - \log\partial_x f(\bar x_j,\bar\theta^*_{\ell,j}) + \mathcal{O}(\varepsilon)\right] \\ &= e^{\mathcal{O}(\varepsilon L^2)}.\end{aligned} \tag{9.16}$$

which yields the C^0-bound of the right expression in (9.4). Integrating in dx yields the bound on the left, since by (9.10), we know a priori that $|\Upsilon_{\ell,L}(x_0) - x_0| = |\xi_0| \le C_\#\varepsilon$.

At last we want to estimate the second derivative $\Upsilon''_{\ell,L}$; differentiating (9.3) we obtain

$$\frac{d^2\Upsilon_{\ell,L}}{dx^2} = \frac{d}{dx}(1-G's_L)\upsilon^+_L\Lambda_{0,L-1} + (1-G's_L)\left[\frac{d}{dx}(\upsilon^+_L)\Lambda_{0,L-1} + \upsilon^+_L\frac{d}{dx}(\Lambda_{0,L-1})\right].$$

The last term on the right hand side is bounded by $C_\#\upsilon^+_L$ using (9.2) and (9.16). The second term can be estimated by differentiating the first of (9.15), which gives:

$$\left|\frac{d\upsilon^+_L}{dx}\right| \le C_\#\upsilon^+_L\sum_{k=0}^{L-1}\left[\left|\frac{dx_k}{dx}\right| + \left|\frac{d\theta_k}{dx}\right| + \varepsilon\left|\frac{du_k}{dx}\right|\right]$$

To continue, notice that (3.6) implies $\left|\frac{dx_k}{dx}\right| \le C_\# \upsilon_k^+$ and $\left|\frac{d\theta_k}{dx}\right| \le C_\# \varepsilon \upsilon_k^+$. Moreover, by the second one of (9.15), we gather

$$\left|\frac{du_k}{dx}\right| \le C_\# \upsilon_k^+ + C_\# \left|\frac{du_{k-1}}{dx}\right| \le C_\# \upsilon_k^+.$$

Hence, the second term is also bounded by $C_\# \upsilon_L^+$. To conclude, we need an estimate for the first term on the right hand side.

Sub-lemma 9.5. *We have*

$$\left|\frac{d}{dx}(1 - G' s_L)\right| \le C_\# \varepsilon L e^{c_\# \varepsilon L}$$

Proof. From (3.6) it follows, for all $n \in \mathbb{N}$

$$\begin{aligned}\upsilon_{n+1}^{\mathrm{c}}(0,1) &= d_p F_\varepsilon^n \begin{pmatrix} \partial_x f & \partial_\theta f \\ \varepsilon \partial_x \omega & 1 + \varepsilon \partial_\theta \omega \end{pmatrix} (s_{n+1}, 1) \\ &= (1 + \varepsilon(\partial_\theta \omega + \partial_x \omega s_{n+1})) d_p F_\varepsilon^n \left(\frac{\partial_\theta f + \partial_x f s_{n+1}}{1 + \varepsilon(\partial_\theta \omega + \partial_x \omega s_{n+1})}, 1 \right).\end{aligned}$$

Since $\upsilon_n^{\mathrm{c}}(0,1) = d_p F_\varepsilon^n(s_n, 1)$, it follows that $\upsilon_{n+1}^{\mathrm{c}}(1 + \varepsilon(\partial_\theta \omega + \partial_x \omega s_{n+1}))^{-1} = \upsilon_n^{\mathrm{c}}$ and

$$s_n = \frac{\partial_\theta f + \partial_x f s_{n+1}}{1 + \varepsilon(\partial_\theta \omega + \partial_x \omega s_{n+1})}.$$

Inverting the above formula yields

$$s_{L-k}(x_k) = \frac{s_{L-k-1}(x_{k+1})(1 + \varepsilon \partial_\theta \omega(x_k, \theta_k)) - \partial_\theta f(x_k, \theta_k)}{\partial_x f(x_k, \theta_k) - \varepsilon \partial_x \omega(x_k, \theta_k) s_{L-k-1}(x_{k+1})}.$$

In order to estimate the derivatives of s_j we proceed by induction. Note that $s_0(x_L) = 0$. Next, suppose $|\partial_x s_{L-k-1}| \le C(L - k - 1) e^{C\varepsilon(L-k-1)} \left|\frac{\partial x_{k+1}}{\partial x}\right|$; then

$$|\partial_x s_{L-k}| \le C_\# \left|\frac{\partial x_k}{\partial x}\right| + |\partial_x s_{L-k-1}| e^{C_\# \varepsilon} \left|\frac{\partial x_k}{\partial x_{k+1}}\right| \le C_\# (L-k) e^{c_\# \varepsilon (L-k)} \left|\frac{\partial x_k}{\partial x}\right|,$$

provided $C_\#$ is large enough. Since $\|G'\|_{\mathcal{C}^1} = \mathcal{O}(\varepsilon)$, we conclude the proof by using the above formula with $k = 0$. □

We conclude that

$$\left|\frac{d^2 \Upsilon_{\ell,L}}{dx^2}\right| \le C_\# \upsilon_L^+ \le C_\# \Lambda_{0,L-1},$$

which gives the needed bound on the derivatives of $\mathfrak{R}_\ell^{1,2}$ in (9.4) and concludes the proof of our lemma. □

9.2. **Transfer operator representation.**

We are now ready to write the contribution of the standard pairs belonging to one block in terms of a product of transfer operators. This is made explicit by (9.18) in the statement of the next proposition. Unfortunately, in the following we will need rather detailed information on the error terms present in (9.18) which therefore must be painstakingly reported in the statement of the proposition, making it rather unpleasant. Yet, the reader can skip such details and come back to them later when they are needed, and recalled.

Notation 9.6. *In the sequel we will use notation similar to Notation 8.8 where, in addition, we introduce symbols for correlations terms computed along the averaged dynamics which will be denoted with*

$$\bar{\mathfrak{K}}^{k,p}_{\ell,l} = \sum_{\bar{\imath}} \mathfrak{C}^{k,p}_{\ell,l,\bar{\imath}} \prod_{j=1}^{l} A_{j,\bar{\imath}} \circ \bar{F}^{i_j}_{\varepsilon}.$$

Observe that, according to Notation 9.6, we can write:

$$\overline{W}_{\ell,k} = \bar{\mathfrak{K}}^{k,1}_{\ell,1} \qquad \overline{\mathfrak{W}}_{\ell,k} = \bar{\mathfrak{K}}^{k,1}_{\ell,1} \qquad \overline{W}_{\ell,k,2} = \bar{\mathfrak{K}}^{k,2}_{\ell,3} \qquad \overline{\mathfrak{W}}_{\ell,k} = \bar{\mathfrak{K}}^{k,2}_{\ell,5}$$

hence we gather, for any $0 \le j \le L$:

$$\bar{\Delta}^*_{\ell,j} = \varepsilon \bar{\mathfrak{K}}^{k,1}_{\ell,1} + \varepsilon^2 \bar{\mathfrak{K}}^{k,2}_{\ell,3} + \overline{\mathfrak{R}}^{3,3}_{\ell_k} \qquad \bar{\xi}_j = \varepsilon \bar{\mathfrak{K}}^{k,1}_{\ell,2} + \varepsilon^2 \bar{\mathfrak{K}}^{k,2}_{\ell,5} + \overline{\mathfrak{R}}^{3,3}_{\ell_k} \tag{9.17}$$

Proposition 9.7. *For any complex standard pair ℓ_0, let $\{\mathfrak{L}^k_{\ell_k}\}_{i=1}^{R-1}$ be the complex standard families obtained in Lemma 8.9, of length δ_{c}, and assume $|\sigma| \le \varepsilon^{-1/2-2\delta_*}$. For any $\Phi \in \mathcal{C}^2(\mathbb{T}, \mathbb{C})$, so that[64] $\mathrm{Re}(\Phi) \le C_\#$, $\varepsilon \|\Phi'\|_{\mathcal{C}^0} L_* \le C_\#$, any $k \in \{0, \cdots, R-1\}$ and $\varrho > 0$, we have*

$$\sum_{\ell_{k+1} \in \mathfrak{L}^k_{\ell_k}} \nu_{k,\ell_k,\ell_{k+1}} e^{\Phi \circ G_{\ell_{k+1}}} \mathring{\rho}_{\ell_{k+1}} = \mathcal{E}^*_{\ell_k} + e^{i\varepsilon\sigma \mathfrak{C}^{k,1}_{\ell_k,0,\emptyset}} \mathcal{L}_{\ell_k,k,L_k-1} \cdots \mathcal{L}_{\ell_k,k,0} \left[\Psi_{\ell_k,q} e^{\Phi(\bar{\theta}_{\ell_k,L_k})} \tilde{\rho}_\varrho \right], \tag{9.18}$$

where $\mathring{\rho}_\ell = \mathbb{1}_{[a_\ell,b_\ell]} \rho_\ell$ (as introduced in Section 5.2) $\bar{\theta}_{\ell,j}(x) = \bar{\theta}(\varepsilon j, G_\ell(x))$, and

(a) *$\mathcal{L}_{\ell,k,j}$ is the weighted transfer operator defined by*

$$\begin{aligned} [\mathcal{L}_{\ell,k,j} g](x) &= \sum_{y \in \bar{f}^{-1}_{\ell,j}(x)} \frac{e^{\Omega^{k,\Phi}_{\ell,j}(\sigma,y,\bar{\theta}^*_{\ell,j})}}{\bar{f}'_{\ell,j}(y)} g(y), \\ \Omega^{k,\Phi}_{\ell,j}(\sigma,x,\theta) &= i\sigma \varpi^k_{\ell,j}(x,\theta) + \varepsilon \Phi'(\bar{\theta}^*_{\ell_k,L_k}) \Xi^*_{\ell,j,L_k} \hat{\omega}(x,\theta) \end{aligned} \tag{9.19}$$

with $\bar{\theta}^*_{\ell,j} = \bar{\theta}(\varepsilon j, \theta^*_\ell)$, $\bar{f}_{\ell,j}(\cdot) = f(\cdot, \bar{\theta}^*_{\ell,j})$ and $\varpi^k_{\ell,j}$ defined in (8.19);

(b) *$\Psi_{\ell_k,q}$ is defined by*

$$\Psi_{\ell_k,q} = \left[1 + \sum_{s=1}^{q-1} (i\sigma\varepsilon)^s \bar{\mathfrak{K}}^{k,2s}_{\ell_k,3s} + (i\sigma\varepsilon)^s \varepsilon \bar{\mathfrak{K}}^{k,2(s+1)}_{\ell_k,3(s+1)} \right] \times \exp\left[i\sigma\varepsilon \bar{\mathfrak{K}}^{k,2}_{\ell_k,3} + i\sigma\varepsilon^2 \bar{\mathfrak{K}}^{k,3}_{\ell_k,6} + \varepsilon \mathcal{K}_0 \right],$$

where $\mathcal{K}_0$ is a $\bar{\mathfrak{K}}^{k,2}_{\ell_k,2}$-type term which satisfies the following extra bound:

$$\mathrm{Leb}\left[\mathring{\rho}_{\ell_k} - e^{\varepsilon \mathcal{K}_0} \tilde{\rho}_\varrho \right] = \mathcal{O}(\varepsilon^2 L_k^3 + \varrho);$$

(c) *$\mathcal{E}^*_{\ell_k}$ satisfies the bounds*

$$\begin{aligned} \|\mathcal{E}^*_{\ell_k}\|_{L^1} &\le C_\# e^{\Phi^+} \left[\|\Phi'\|_{\mathcal{C}^1} \varepsilon^2 L_k^2 + \varepsilon^2 L_k^3 + \varrho \right] \\ \|\mathcal{E}^*_{\ell_k}\|_{\mathrm{BV}} &\le C_\# e^{\Phi^+} [1 + |\sigma|] \end{aligned}$$

where $\Phi^+ = \max \mathrm{Re}(\Phi)$;

(d) *finally $\tilde{\rho}_\varrho \in \mathcal{C}^\infty(\mathbb{T}^1, \mathbb{R})$ is a positive function that is close to ρ_{ℓ_k} in the sense*

$$\|\tilde{\rho}_\varrho - \mathring{\rho}_{\ell_k}\|_{L^1} = \mathcal{O}(\varepsilon \min\{\delta_{\mathbb{C}}^{-1}, L^2\} + \varrho), \tag{9.20}$$

and such that, for any $r \in \mathbb{N}$, $\|\tilde{\rho}_\varrho\|_{W^{r,1}} \le C_\# \varrho^{-r+1} \delta_{\mathbb{C}}^{-r}$.

---

[64] The reader should think of Φ as a function whose real part is negative and has very large absolute value

Proof. Recall (see Remark 3.8) that, for a given $\ell = (\mathbb{G}, \rho)$ and for any n, we have $(x_n, \theta_n) = F_\varepsilon^n(\mathbb{G}(x))$; in other words, we consider x_n and θ_n to be random variables on ℓ. In particular, we have, for any smooth test function $g : \mathbb{T} \to \mathbb{R}$:

$$\sum_{\ell_{k+1} \in \mathfrak{L}^k_{\ell_k}} \text{Leb}\left(g \cdot \mathrm{v}_{k,\ell_k,\ell_{k+1}} e^{\Phi \circ G_{\ell_{k+1}}} \mathring{\rho}_{\ell_{k+1}}\right) = \sum_{\ell_{k+1} \in \mathfrak{L}^k_{\ell_k}} \mathrm{v}_{k,\ell_k,\ell_{k+1}} \mu_{\ell_{k+1}} \left(g(x_0) e^{\Phi(\theta_0)}\right).$$

Using (8.21) and the definition of $\mathrm{v}_{k,\ell_k,\ell_{k+1}}$ (see Lemma 8.9), we gather:

$$\begin{aligned} &\sum_{\ell_{k+1} \in \mathfrak{L}^k_{\ell_k}} \text{Leb}\left(g \cdot \mathrm{v}_{k,\ell_k,\ell_{k+1}} e^{\Phi \circ G_{\ell_{k+1}}} \mathring{\rho}_{\ell_{k+1}}\right) \\ &= \int_{\mathbb{T}} [g e^{\Phi}] \circ F_\varepsilon^{L_k} \circ \mathbb{G}_{\ell_k}\; e^{i\sigma\left[\varepsilon \mathfrak{C}^{k,1}_{\ell_k,0,\emptyset} + \sum_{j=0}^{L_k-1} \varpi^k_{\ell_k,j} \circ F_\varepsilon^j\right]} \mathring{\rho}_{\ell_k} \\ &\quad + \sum_{s=1}^{q-1} (i\sigma\varepsilon)^s \int_{\mathbb{T}} [g e^{\Phi}] \circ F_\varepsilon^{L_k} \circ \mathbb{G}_{\ell_k} \cdot e^{i\sigma\left[\varepsilon \mathfrak{C}^{k,1}_{\ell_k,0,\emptyset} + \sum_{j=0}^{L_k-1} \varpi^k_{\ell_k,j} \circ F_\varepsilon^j\right]} \mathfrak{K}^{k,2s}_{\ell_k,3s} \mathring{\rho}_{\ell_k}. \end{aligned} \tag{9.21}$$

In the following we will find convenient to use $\bar{x} = \Upsilon_{\ell_k,L_k}(x)$, rather than x, as our fundamental random variable.[65] This can be done using Lemma 9.1: indeed, the change of variable formula yields that the pushforward of the density is given by

$$\Upsilon_* \mathring{\rho}_{\ell_k} = (\Upsilon^{-1})' \cdot \mathring{\rho}_{\ell_k} \circ \Upsilon^{-1}. \tag{9.22}$$

For any smooth function φ of the random variables $\{(x_i, \theta_i)\}_{i=0}^{L_k-1}$, under μ_{ℓ_k}, we can write $\tilde{\varphi}(x) = \varphi(\{F_\varepsilon^i \circ \mathbb{G}_{\ell_k}(x)\})$, where x is distributed according to $\mathring{\rho}_{\ell_k}$. Then our change of variable corresponds to looking at the random variable $\bar{\varphi}(\bar{x}) = \varphi(\{F_\varepsilon^i \circ \mathbb{G}_{\ell_k} \circ \Upsilon^{-1}_{\ell_k,L_k}(\bar{x})\})$ under $\Upsilon_* \mathring{\rho}_{\ell_k}$. In particular,

$$\|\bar{\varphi}'\|_\infty \leq C_\# \|\varphi\|_{\mathcal{C}^1} \bar{\Lambda}^{-1}_{0,L_k}.$$

The above considerations would suffice to treat the small terms in (9.21), but, unfortunately, are not adequate to treat the main term since we only have an exponentially large bound on the derivative of $\Upsilon_* \rho_{\ell_k}$ (see the last of (9.4)) which would create serious problems in our subsequent arguments, unless they can be discarded by some a priori estimate. In order to deal with this problem, we first need to introduce some notation. Let

$$\rho_* = \frac{\mathring{\rho}_{\ell_k}}{[1 - G' s_L]} \circ \Upsilon^{-1}; \tag{9.23}$$

observe that Sub-Lemma 9.5 implies that $\|\rho_*\|_{\text{BV}} \leq C_\# \delta_{\mathbb{C}}^{-1}$. We can now state a more useful bound for (9.22) whose proof is, for convenience, postponed to the end to this section.

Lemma 9.8. *The following formula holds true*

$$\Upsilon_* \mathring{\rho}_{\ell_k} = \rho_* \exp[\varepsilon \mathcal{K}_0] + \overline{\mathfrak{R}}^{2,3}_{\ell_k} \mathring{\rho}_{\ell_k}.$$

Next, we proceed to eliminate the explicit dependence on x_j and θ_j: first observe that, by definition, for any smooth function $A(x, \theta)$. Observe that we have $\bar{\Delta}^*_{\ell_k,j}, \bar{\xi}_j = \overline{\mathfrak{R}}^{1,1}_{\ell_k}$; hence we can write

$$\begin{aligned} [A(x_j, \theta_j)] \circ \Upsilon^{-1} &= A(\bar{x}_j, \bar{\theta}^*_{\ell_k,j}) + \partial_x A(\bar{x}_j, \bar{\theta}^*_{\ell_k,j}) \bar{\xi}_j + \partial_\theta A(\bar{x}_j, \bar{\theta}^*_{\ell_k,j}) \bar{\Delta}^*_{\ell_k,j} \\ &\quad + \frac{1}{2}\left[\partial_{\theta\theta} A (\bar{\Delta}^*_{\ell_k,j})^2 + \partial_{\theta x} A\, \bar{\Delta}^*_{\ell_k,j} \bar{\xi}_j + \partial_{xx} A\, \bar{\xi}_j^2\right] + \overline{\mathfrak{R}}^{3,3}_{\ell_k}. \end{aligned}$$

[65] In the rest of the proof we will often suppress the subscripts ℓ_k, L_k in Υ_{ℓ_k,L_k}, and related quantities, when this does not create any confusions.

In particular, using (9.17) we gather

$$\sum_{j=0}^{L_k-1}\varpi^k_{\ell_k,j}(x_j,\theta_j)=\sum_{j=0}^{L_k-1}\varpi^k_{\ell_k,j}(\bar x_j,\bar\theta^*_{\ell_k,j})+\varepsilon\bar{\mathfrak{K}}^{k,2}_{\ell_k,3}+\varepsilon^2\bar{\mathfrak{K}}^{k,3}_{\ell_k,6}+\overline{\mathfrak{R}}^{3,4}_{\ell_k};$$

$$\begin{aligned}(i\sigma\varepsilon)^s\mathfrak{K}^{k,2s}_{\ell_k,3s}&=(i\sigma\varepsilon)^s\sum_{\bar\imath}\mathfrak{C}^{k,2s}_{\ell_k,3s,\bar\imath}\prod_{j=1}^{3s}\left[A_{j,\bar\imath}(\bar x_{i_j},\bar\theta^*_{\ell,i_j})+\varepsilon\bar{\mathfrak{K}}^{k,1}_{\ell_k,3}+\overline{\mathfrak{R}}^{2,2}_{\ell_k}\right]\\&=(i\sigma\varepsilon)^s\mathfrak{K}^{k,2s}_{\ell_i,3s,*}+(i\sigma\varepsilon)^s\varepsilon\bar{\mathfrak{K}}^{k,2s+1}_{\ell_k,3(s+1)}+\sigma^s\overline{\mathfrak{R}}^{s+2,2s+2}_{\ell_k}.\end{aligned}$$

The above will suffice to estimate the error terms. However, to deal with $\Phi(\theta_{L_k})$ we will need a more explicit formula. By definition (5.1b) we have $\theta_{L_k}=\bar\theta_{\ell_k,L_k}+\Delta_{L_k}$; by (8.9), and using (9.7) we conclude that

$$\theta_{L_k}=\bar\theta_{\ell_k,L_k}+\varepsilon\sum_{j=0}^{L_k-1}\Xi^*_{\ell,j,k}\hat\omega(\bar x_l\circ\Upsilon,\bar\theta^*_{\ell,l})+\mathfrak{R}^{2,2}_{\ell}\tag{9.24}$$

We can now collect all the above relations to write (9.21) in terms of the slowly varying dynamics

$$\begin{aligned}&\sum_{\ell_{k+1}\in\mathfrak{L}^k_{\ell_k}}\int_{\mathbb{T}}g\cdot\nu_{k,\ell_k,\ell_{k+1}}e^{\Phi\circ G_{\ell_{k+1}}}\mathring\rho_{\ell_{k+1}}=\\&=\int_{\mathbb{T}}g(x_{L_k})e^{i\sigma\left[\varepsilon\mathfrak{C}^{k,1}_{\ell_k,0,\emptyset}+\sum_{j=0}^{L_k-1}\varpi^k_{\ell_k,j}\circ F^j_\varepsilon\right]+\Phi(\bar\theta_{\ell_k,L_k})+\varepsilon\Phi'(\bar\theta^*_{\ell_k,L_k})\sum_{j=0}^{L_k-1}\Xi^*_{\ell,j,k}\hat\omega(\bar x_l,\bar\theta^*_{\ell,l})}\mathring\rho_{\ell_k}\\&\quad+\|\Phi'\|_{\mathcal{C}^1}e^{\Phi^+}\int_{\mathbb{T}}g(x_{L_k})\mathfrak{R}^{2,2}_{\ell_k}\mathring\rho_{\ell_k}\\&\quad+\sum_{s=1}^{q-1}i^s\sigma^s\varepsilon^s\int_{\mathbb{T}}[ge^\Phi]\circ F^{L_k}_\varepsilon\circ\mathbb{G}_{\ell_k}\cdot e^{i\sigma\left[\varepsilon\mathfrak{C}^{k,1}_{\ell_k,0,\emptyset}+\sum_{j=0}^{L_k-1}\varpi^k_{\ell_k,j}\circ F^j_\varepsilon\right]}\mathfrak{K}^{k,2s}_{\ell_k,3s}\mathring\rho_{\ell_k}\\&=\int_{\mathbb{T}}g\circ\bar F^{L_k}_\varepsilon\cdot e^{i\sigma\varepsilon\mathfrak{C}^{k,1}_{\ell_k,0,\emptyset}+\sum_{j=0}^{L_k-1}\Omega^{k,\Phi}_{\ell_k,j}\circ\bar F^j_\varepsilon}\Psi_{\ell_k,q}\,e^{\Phi(\bar\theta_{\ell_k,L_k})}\rho_*(\bar x)d\bar x\\&\quad+e^{\Phi^+}\int_{\mathbb{T}}g\circ\bar F^{L_k}_\varepsilon\cdot e^{\sum_{j=0}^{L_k-1}\Omega^{k,\Phi}_{\ell_k,j}\circ\bar F^j_\varepsilon}\left[\|\Phi'\|_{\mathcal{C}^1}\overline{\mathfrak{R}}^{2,2}_{\ell_k}+\overline{\mathfrak{R}}^{2,3}_{\ell_k}\right]\Upsilon_*\rho_{\ell_k}d\bar x,\end{aligned}\tag{9.25}$$

where we have used the fact that, by hypothesis, $\varepsilon^2L_k^3\leq\min\{|\sigma|\varepsilon^3L_k^4,\sigma^2\varepsilon^4L_k^4\}$ and by definition $\mathfrak{R}^{p,q}_\ell\circ\Upsilon^{-1}=\overline{\mathfrak{R}}^{p,q}_\ell$. Since the above equation holds for all g, we have

$$\begin{aligned}\sum_{\ell_{k+1}\in\mathfrak{L}^k_{\ell_k}}\nu_{k,\ell_k,\ell_{k+1}}e^{\Phi\circ G_{\ell_{k+1}}}\mathring\rho_{\ell_{k+1}}&=e^{i\sigma\varepsilon\mathfrak{C}^{k,1}_{\ell_k,0,\emptyset}}\mathcal{L}_{\ell_k,k,L-1}\cdots\mathcal{L}_{\ell_k,k,0}\Psi_{\ell_k,q}\,e^{\Phi(\bar\theta_{\ell_k,L_k})}\rho_*\\&\quad+e^{\Phi^+}\mathcal{L}_{\ell_k,k,L-1}\cdots\mathcal{L}_{\ell_k,k,0}\left[\|\Phi'\|_{\mathcal{C}^1}\overline{\mathfrak{R}}^{2,2}_{\ell_k}+\overline{\mathfrak{R}}^{2,3}_{\ell_k}\right]\Upsilon_*\rho_{\ell_k},\end{aligned}\tag{9.26}$$

where we used the fact that, by definition,

$$\int_{\mathbb{T}}g\circ\bar F^{L_k}_\varepsilon\cdot e^{\sum_{j=0}^{L_k-1}\Omega^{k,\Phi}_{\ell_k,j}\circ\bar F^j_\varepsilon}\varphi(x)dx=\int_{\mathbb{T}}g\mathcal{L}_{\ell_k,k,L-1}\mathcal{L}_{\ell_k,k,L-2}\cdots\mathcal{L}_{\ell_k,k,0}\varphi.$$

In the sequel we will need to deal with smooth density functions. We can obtain this by a mollification procedure; (see, e.g. [44, Lemma B.1]): for each $\varrho>0$ there exists a $\tilde\rho_\varrho$ such that

$$\|\tilde\rho_\varrho-\rho_*\|_{L^1}\leq C_\#\varrho\quad\text{and}\quad\|\tilde\rho_\varrho\|_{W^{r,1}}\leq C_\#\varrho^{-r+1}\delta_{\mathbb{C}}^{-r}.\tag{9.27}$$

Note that $\|\mathcal{L}_{\ell,k,j}\|_{L^1}\leq e^{c_\#\varepsilon\|\Phi'\|_\infty}$. Moreover, by iterating (A.2), we have, for each $\varphi\in W^{1,1}$,

$$\left\|\frac{d}{dx}\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\varphi\right\|_{L^1}\leq C_\#\left[\|\bar\Lambda_{0,L_k}\varphi'\|_{L^1}+(1+|\sigma|)\|\varphi\|_{L^1}\right].\tag{9.28}$$

We also have

$$
\begin{aligned}
&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\Psi_{\ell_k,q}e^{\Phi(\bar\theta_{\ell_k},L_k)}[\tilde\rho_\varrho-\rho_*]\|_{L^1}\leq C_\#\varrho\\
&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\Psi_{\ell_k,q}e^{\Phi(\bar\theta_{\ell_k},L_k)}\tilde\rho_\varrho\|_{W^{1,1}}\leq C_\#(1+|\sigma|)
\end{aligned}\tag{9.29}
$$

By the lower semicontinuity of the variation [21, Section 5.2.1, Theorem 1], since $\tilde\rho_\varrho\to\rho_*$ in L^1 as $\varrho\to 0$, we have

$$
\begin{aligned}
&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\Psi_{\ell_k,q}e^{\Phi(\bar\theta_{\ell_k},L_k)}[\tilde\rho_\varrho-\rho_*]\|_{\mathrm{BV}}\\
&\qquad=\lim_{\varrho'\to 0}\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\Psi_{\ell_k,q}e^{\Phi(\bar\theta_{\ell_k},L_k)}[\tilde\rho_\varrho-\tilde\rho_{\varrho'}]\|_{W^{1,1}}\leq C_\#(1+|\sigma|)
\end{aligned}
$$

By a similar argument, estimating the remainder terms of (9.26), follows part (c) of the proposition. Finally, to prove (d), recall that $|G's_L|\leq C_\#\varepsilon$. Then, by (9.23) and (9.27):

$$
\|\tilde\rho_\varrho-\mathring\rho_{\ell_k}\|_{L^1}\leq\|\mathring\rho_{\ell_k}-\mathring\rho_{\ell_k}\circ\Upsilon^{-1}\|_{L^1}+C_\#(\varrho+\varepsilon)\leq C_\#(\varepsilon\min\{L^2,\delta_{\mathbb C}^{-1}\}+\varrho)
$$

where in the last step we used the first bound in (9.4). □

We conclude with the missing proof.

Proof of Lemma 9.8. By Lemma 9.1, equations (9.22), (9.23), (9.3) we gather

$$
\begin{aligned}
\Upsilon_*\rho_{\ell_k}&=\rho_*\exp\left[\sum_{j=0}^{L-1}\log\left[\frac{\partial_x f(x_j,\theta_j)}{\partial_x f(\bar x_j,\bar\theta^*_{\ell_k,j})}+\varepsilon\frac{\partial_\theta f(x_j,\theta_j)}{\partial_x f(x_j,\theta_j)}u_j\right]+\mathcal{O}(\varepsilon^2L)\right]\\
&=\rho_*\exp\left[\varepsilon\sum_{j=0}^{L-1}\frac{\partial_{\theta x}f(\bar x_j,\bar\theta^*_{\ell_k,j})\overline{W}_j+\partial_{xx}f(\bar x_j,\bar\theta^*_{\ell_k,j})\overline{\mathfrak{W}}_j+\partial_\theta f(\bar x_j,\bar\theta^*_{\ell_k,j})u_j}{\partial_x f(\bar x_j,\bar\theta^*_{\ell_k,j})}\right]\\
&\quad+\overline{\mathfrak{R}}^{2,3}_{\ell_k}\mathring\rho_{\ell_k},
\end{aligned}\tag{9.30}
$$

where the second needed property of $\overline{\mathfrak{R}}^{2,3}_{\ell_k}$ follows immediately from (9.4). Note that the exponent in (9.30) is, at most, of size $C_\#\varepsilon L^2\leq C_\#\varepsilon^{6\delta_*}$ (recall (8.6)), and it is of correlation type. It is then natural to expand the exponential in Taylor series and to use Notation 9.6. We can then write

$$
\sum_{j=0}^{L-1}\frac{\partial_{\theta x}f(\bar x_j,\bar\theta^*_{\ell_k,j})}{\partial_x f(\bar x_j,\bar\theta^*_{\ell_k,j})}\overline{W}_j=\bar{\mathfrak{K}}^{k,2}_{\ell_k,2}.
$$

In order to show that also the second term in the exponent can be treated as a correlation term, let us set for convenience $\alpha(\bar x,\bar\theta)=\frac{\partial_{xx}f(\bar x,\bar\theta)}{\partial_x f(\bar x,\bar\theta)}$, then

$$
\begin{aligned}
\sum_{j=0}^{L-1}\frac{\partial_{xx}f}{\partial_x f}\circ\bar F^j_\varepsilon\cdot\overline{\mathfrak{W}}_j&=-\sum_{j,s=0}^{L-1}\alpha\circ\bar F^j_\varepsilon\cdot\hat\omega\circ\bar F^s_\varepsilon\sum_{l=\max\{j,s+1\}}^{L-1}\Xi^*_{s,l}\cdot\Lambda_{0,l-j}\partial_\theta f\circ\bar F^l_\varepsilon\\
&=-\sum_{j\geq s+1}^{L-1}A_{1,j,s}\circ\bar F^s_\varepsilon\cdot\alpha\circ\bar F^j_\varepsilon-\sum_{j<s+1}^{L-1}\hat\omega\circ\bar F^s_\varepsilon A_{2,j,s}\circ\bar F^j_\varepsilon
\end{aligned}
$$

where

$$
A_{1,j,s}=\Xi^*_{0,l-s}\hat\omega\sum_{l=j}^{L-1}\Lambda_{0,l-j}\partial_\theta f\circ\bar F^{l-s}_\varepsilon
$$

$$
A_{2,j,s}=\alpha\sum_{l=s+1}^{L-1}\Xi^*_{s-j,l-j}\Lambda_{0,l-j}\partial_\theta f\circ\bar F^{l-j}_\varepsilon
$$

(recall that we dropped the subscript ℓ from Ξ^*). A direct computation shows that $\sup_{\bar\theta} \|A_{i,j,s}(\cdot,\bar\theta)\|_{\mathcal{C}^1} \le C_\#$. In order to deal with the third term in the exponential, we need to define the auxiliary variables

$$\tilde u_{k+1} = \frac{\partial_x \omega(x_k,\theta_k) + \tilde u_k}{\partial_x f(x_k,\theta_k)}, \quad \varepsilon \tilde u_0 = G'_\ell(x).$$

By (9.15) it is immediate to observe that

$$u_{k+1} - \tilde u_{k+1} = \frac{u_k - \tilde u_k}{\partial_x f(x_k,\theta_k) + \varepsilon \partial_\theta f(x_k,\theta_k) u_k} + \mathcal{O}(\varepsilon)$$

from which it follows $u_k = \tilde u_k + \mathcal{O}(\varepsilon)$. We can thus replace u_k with $\tilde u_k$ in the third term in the exponential and computations similar to the previous ones yield that also the third term in the exponential can be interpreted as a correlation term. Recalling (9.23), the above discussion implies that we can write

$$\Upsilon_* \rho_{\ell_k} = \rho_* \exp\left[\varepsilon \bar{\mathfrak{K}}^{k,2}_{\ell_k,2}\right] + \overline{\mathfrak{R}}^{2,3}_{\ell_k} \mathring{\rho}_{\ell_k}. \tag{9.31}$$

Finally, we claim that the term $\bar{\mathfrak{K}}^{k,2}_{\ell_k,2}$ can be written as $\mathcal{K}_0$. In order to see this, it suffices to integrate the above relation to obtain

$$1 = \mathrm{Leb}(\mathring{\rho}_{\ell_k}) = \mathrm{Leb}\left[e^{\varepsilon \bar{\mathfrak{K}}^{k,2}_{\ell_k,2}} \rho_*\right] + \mathcal{O}(\varepsilon^2 L_k^3), \tag{9.32}$$

which implies our requirement by taking into account (9.27). □

10. One block estimate: the large σ regime

In the large σ regime it suffices to estimate the contribution of the last block. To this end we first need an estimate on the product of the transfer operators defined in (9.19). To ease notation, in this section we will omit the indices ℓ_R, ℓ_{R-1} and $R-1$, referring to the last block, as no confusion can arise: in particular $\mathcal{L}_j$ will stand for $\mathcal{L}_{\ell_{R-1},R-1,j}$ and L will stand for L_{R-1}. Also, the transfer operators are defined with respect to the purely imaginary potentials $i\sigma\varpi^k_{\ell,j}$, where $\varpi^k_{\ell,j}$ is defined in (8.19), i.e. we have $\Phi \equiv 0$ in (9.19).

Lemma 10.1. *There exists $C_3 > 0$ and $\tau_1 \in (0,1)$ such that, for all $n \in [C_3 \log \varepsilon^{-1}, L-1]$ and $j \in [0, L-1-n]$, any $g \in \mathcal{C}^1(\mathbb{T})$ we have*

$$\|\mathcal{L}_{j+n}\cdots\mathcal{L}_j g\|_{\mathcal{C}^1} \le \tau_1^n \|g\|_{\mathcal{C}^1}. \tag{10.1}$$

Proof. We begin with a preliminary estimate on $\|\mathcal{L}_{j+n}\cdots\mathcal{L}_j g\|_{\mathcal{C}^1}$; as already noticed, the potentials are purely imaginary, thus for any $0 \le n \le L-1$ and $0 \le i \le L-1$ we have[66]

$$\|\mathcal{L}_i^n\|_{\mathcal{C}^0\to\mathcal{C}^0} \le C_\#. \tag{10.2}$$

Observe that by the same token

$$\|\mathcal{L}_{i+n}\cdots\mathcal{L}_i\|_{\mathcal{C}^0\to\mathcal{C}^0} \le C_\#. \tag{10.3}$$

Using the Lasota-Yorke inequality (A.2) we gather

$$\begin{aligned}\|\mathcal{L}_{i+n}\cdots\mathcal{L}_i g\|_{\mathcal{C}^1} &\le C_\# \lambda^{-n}\|g\|_{\mathcal{C}^1} + C_\# \sum_{k=0}^{n} \lambda^{-k}(1+|\sigma|)\|\mathcal{L}_{i+n-k-1}\cdots\mathcal{L}_i g\|_{\mathcal{C}^0} \\ &\le C_\#(1+|\sigma|)\|g\|_{\mathcal{C}^1}\end{aligned} \tag{10.4}$$

We continue with an estimate of $\|\mathcal{L}_{L-1}\cdots\mathcal{L}_0\|_{\mathcal{C}^1\to\mathcal{C}^0}$. Since $\hat\omega$ is not a coboundary, ϖ_j is not a coboundary, and the potentials (ϖ_j) satisfy UUNI (see Corollary B.4).

[66] This follows since $|\mathcal{L}_i g| \le |\mathcal{L}_{i,\sigma=0} g|$, therefore $\|\mathcal{L}_i^n g\|_{\mathcal{C}^0} \le \|\mathcal{L}_i^n 1\|_{\mathcal{C}^0}\|g\|_{\mathcal{C}^0} \le C_\#\|g\|_{\mathcal{C}^0}$.

We can thus apply Theorem B.5, that implies that there exists $\tau \in (0,1)$ such that, for any $i \in \{0,\cdots,L-1\}$,

$$\|\mathcal{L}_i^n\|_{\mathcal{C}^1\to\mathcal{C}^1} \leq \begin{cases} C_\#(1+|\sigma|) & \text{for } n < n_\varepsilon = \lfloor C_3 \log \varepsilon^{-1}\rfloor \\ \tau^n & \text{for } n \geq n_\varepsilon. \end{cases} \tag{10.5}$$

Note that we can choose C_3 as large as needed. Also we have the following trivial estimate for the difference of operators with potentials Ω_i:

$$\|\mathcal{L}_{\theta,\Omega_1} - \mathcal{L}_{\theta,\Omega_2}\|_{\mathcal{C}^0} \leq C_\#\|\Omega_1-\Omega_2\|_{\mathcal{C}^0}\|\mathcal{L}_{\theta,\Omega_1}\|_{\mathcal{C}^0},$$

and, by (A.6), we have, for all $g\in\mathcal{C}^1$,

$$\begin{aligned}\|\mathcal{L}_{\theta_1,\Omega}g - \mathcal{L}_{\theta_2,\Omega}g\| &\leq C_\#|\theta_1-\theta_2| \sup_{\theta\in[\theta_1,\theta_2]}\left[\|\mathcal{L}_{\theta,0}|g'|\|_{\mathcal{C}^0} + \|\mathcal{L}_{\theta,0}(1+\|\Omega\|_{\mathcal{C}^1})|g|\|_{\mathcal{C}^0}\right] \\ &\leq C_\#|\theta_1-\theta_2|\left[\|g\|_{\mathcal{C}^1} + (1+\|\Omega\|_{\mathcal{C}^1})\|g\|_{\mathcal{C}^0}\right].\end{aligned}$$

Accordingly, using the explicit formula (9.19) we have, for each $0\leq i<k<L$,

$$\|(\mathcal{L}_k-\mathcal{L}_i)g\|_{\mathcal{C}^0} \leq C_\#\varepsilon(k-i)\left[\|g\|_{\mathcal{C}^1} + (1+|\sigma|)\|g\|_{\mathcal{C}^0}\right]. \tag{10.6}$$

Observe moreover that we can write

$$\mathcal{L}_{i+n}\cdots\mathcal{L}_i = \mathcal{L}_i^{n+1} + \sum_{k=1}^{n}\mathcal{L}_{i+n}\cdots\mathcal{L}_{i+k+1}(\mathcal{L}_{i+k}-\mathcal{L}_i)\mathcal{L}_i^k.$$

Thus, for $n\in[n_\varepsilon,3n_\varepsilon]$ and $i\in\{0,\cdots,L-n-1\}$, we can use (10.3), (10.6) and (10.5) to write

$$\begin{aligned}\|\mathcal{L}_{i+n}\cdots\mathcal{L}_i g\|_{\mathcal{C}^0} &\leq \sum_{k=0}^{n-1}\|\mathcal{L}_{i+n}\cdots\mathcal{L}_{i+k+2}(\mathcal{L}_{i+k+1}-\mathcal{L}_i)\mathcal{L}_i^{k+1}g\|_{\mathcal{C}^0} + \tau^{n+1}\|g\|_{\mathcal{C}^1} \\ &\leq \left(C_\#(1+|\sigma|)\varepsilon n_\varepsilon^2 + \tau^{n_\varepsilon}\right)\|g\|_{\mathcal{C}^1} \leq C_\#|\sigma|\varepsilon n_\varepsilon^2\|g\|_{\mathcal{C}^1},\end{aligned}$$

provided C_3 in the definition of n_ε has been chosen large enough and since $\sigma\geq\sigma_0$.

Note that, for $|\sigma|\varepsilon n_\varepsilon^3 < 1$, we can bootstrap the above estimate by writing, for $n\in[3n_\varepsilon,4n_\varepsilon]$,

$$\begin{aligned}\|\mathcal{L}_{i+n}\cdots\mathcal{L}_i g\|_{\mathcal{C}^0} &\leq \sum_{k=0}^{n_\varepsilon}\|\mathcal{L}_{i+n}\cdots\mathcal{L}_{i+k+2}(\mathcal{L}_{i+k+1}-\mathcal{L}_i)\mathcal{L}_i^{k+1}g\|_{\mathcal{C}^0} + 2\tau^{n_\varepsilon}\|g\|_{\mathcal{C}^1} \\ &\leq \sum_{k=0}^{n_\varepsilon} C_\#|\sigma|\varepsilon n_\varepsilon^2\|\mathcal{L}_{i+n-n_\varepsilon-1}\cdots\mathcal{L}_{i+k+2}(\mathcal{L}_{i+k+1}-\mathcal{L}_i)\mathcal{L}_i^{k+1}g\|_{\mathcal{C}^1} + 2\tau^{n_\varepsilon}\|g\|_{\mathcal{C}^1} \\ &\leq C_\#\sum_{k=0}^{n_\varepsilon}(\varepsilon|\sigma|n_\varepsilon^2\left[\lambda^{-n_\varepsilon}|\sigma| + C_\#\varepsilon|\sigma|n_\varepsilon^2\right] + \tau^{n_\varepsilon})\|g\|_{\mathcal{C}^1} \leq C_\#\varepsilon^2|\sigma|^2 n_\varepsilon^5\|g\|_{\mathcal{C}^1}\end{aligned}$$

where we have chosen, again, C_3 large enough and, in the last line, we have used the Lasota–Yorke inequality (A.2).

Finally, note that, by using the Lasota–Yorke inequality again, it follows, for all $n\in[3n_\varepsilon,4n_\varepsilon]$,

$$\begin{aligned}\|\mathcal{L}_{i+n}\cdots\mathcal{L}_i g\|_{\mathcal{C}^1} &\leq C_\#\lambda^{-n}\|g\|_{\mathcal{C}^1} + C_\#\sum_{k=0}^{n-1}\lambda^{-k}|\sigma|\|\mathcal{L}_{i+n-k-1}\cdots\mathcal{L}_i g\|_{\mathcal{C}^0} \\ &\leq C_\#\left[\lambda^{-n_\varepsilon}|\sigma| + \varepsilon^2|\sigma|^3 n_\varepsilon^6\right]\|g\|_{\mathcal{C}^1} \\ &\leq \tau_1^n\|g\|_{\mathcal{C}^1}\end{aligned}$$

for some $\tau_1\in(\tau,1)$, provided, again, C_3 has been chosen large enough and since $\varepsilon^2|\sigma|^3 n_\varepsilon^6 \leq \varepsilon^{1/4}$ for ε small enough. □

We are now able to provide the proof of the main result of this section.

Proof of Proposition 8.11. Observe that, by definition and by Proposition 9.7, with $\Phi \equiv 0$, we have[67]

$$\sum_{\ell_R \in \mathfrak{L}^{R-1}_{\ell_{R-1}}} \nu_{R-1,\ell_{R-1},\ell_R} = \operatorname{Leb} \sum_{\ell_R \in \mathfrak{L}^{R-1}_{\ell_{R-1}}} \nu_{R-1,\ell_{R-1},\ell_R} \mathring{\rho}_{\ell_R}$$
$$= \operatorname{Leb} \mathcal{L}_{L-1} \cdots \mathcal{L}_0 [\Psi_q \tilde{\rho}_\varrho] + \operatorname{Leb} \mathcal{E}^*.$$

Recall that $|\sigma| \le \varepsilon^{-\frac{1}{2}-2\delta_*}$ and $\delta_* \in (1/99, 1/32)$. By Proposition 9.7(c), we have

$$|\operatorname{Leb} \mathcal{E}^*| \le \|\mathcal{E}^*\|_{L^1} \le C_\# \{\varepsilon^{2-9\delta_*} + \varrho\}.$$

Next, note that for each $r < L/(3n_\varepsilon) - 2$, by (10.1), (10.3) and (10.4) we obtain

$$\|\mathcal{L}_{L-1} \cdots \mathcal{L}_{i_r} A_{r,i_r} \cdots \mathcal{L}_{i_1} A_{1,i_1} \cdots \mathcal{L}_0 \tilde{\rho}_{\ell,\varrho}\|_{\mathcal{C}^0} \le C_\#^r \tau_1^{L/r+1} \varrho^{-1} \delta_{\mathbb{C}}^{-2} \varepsilon^{-r/2-2r\delta_*}$$

since at least one string of operators must be longer than $L/(r+1) \ge 3n_\varepsilon$. This allows to estimate the contribution of $\Psi_{\ell_{R-1},q}$ by expanding is series the exponential. We thus obtain

$$\left| \sum_{\tilde{\ell} \in \mathfrak{L}^L_{\ell,\Omega}} \nu_{\tilde{\ell}} \rho_{\tilde{\ell}} \mathbb{1}_{[a_{\tilde{\ell}}, b_{\tilde{\ell}}]} \right| = \mathcal{O}(\varepsilon^{2-9\delta_*} + \varrho + \varepsilon^{-1} \tau_1^{L/(c_\# q)} \varrho^{-1} \delta_{\mathbb{C}}^{-2} \varepsilon^{-q/2-2q\delta_*}).$$

Thus, the proposition follows by choosing $\varrho = \varepsilon^{2-9\delta_*}$. □

11. One block estimate: the intermediate σ regime

The following two cases require a much more accurate description of the one-block contribution, which can only be obtained for small σ. It will be achieved thanks to the technical lemmata contained in this section.

The argument is similar to the one of the previous section, only a different idea is needed to compute the norms of the relevant operators: provided σ_0 is small enough, such norms can be computed via perturbation theory.

11.1. The Transfer operators product formula.

Our task here is to study the transfer operators defined in (9.19) and then their products in the perturbative regime.

Lemma 11.1. *Let σ_0 be chosen small enough. For any $\sigma \in \mathcal{J}_1 \cup \mathcal{J}_2$ and $k \in \{0, \cdots, R-1\}$, $j \in \{0, \cdots, L_k-1\}$ and Φ satisfying the hypotheses of Proposition 9.7 and, additionally, so that $\varepsilon\|\Phi'\|_{\mathcal{C}^1} \sup_{j \le L_k} \|\Xi^*_{\ell,j}\hat{\omega}\|_{\mathcal{C}^1} \le \sigma_0$ we have:*

(a) *$\mathcal{L}_{\ell_k,k,j}$ is of Perron–Frobenius type, i.e. we can write $\mathcal{L}_{\ell_k,k,j} = e^{\chi_{\ell_k,k,j}} \mathcal{P}_{\ell_k,k,j} + \mathcal{Q}_{\ell_k,k,j}$, where $e^{\chi_{\ell_k,k,j}}$ is the maximal eigenvalue of $\mathcal{L}_{\ell_k,k,j}$ (as an operator acting on $\mathcal{C}^1, W^{1,1}$ or BV), $\mathcal{P}_{\ell_k,k,j}, \mathcal{Q}_{\ell_k,k,j}$ are such that $\mathcal{P}^2_{\ell_k,k,j} = \mathcal{P}_{\ell_k,k,j}$, $\mathcal{P}_{\ell_k,k,j}\mathcal{Q}_{\ell_k,k,j} = \mathcal{Q}_{\ell_k,k,j}\mathcal{P}_{\ell_k,k,j} = 0$, the operators $\mathcal{P}_{\ell_k,k,j}$ are rank one, and there exists $\tau \in (0,1)$ so that*

$$\|\mathcal{Q}^n_{\ell_k,k,j}\|_{\mathcal{C}^1,W^{1,1},\mathrm{BV}} \le C_\# \tau^n |e^{n\chi_{\ell_k,k,j}}|\,; \quad \|\mathcal{P}_{\ell_k,k,j}\|_{\mathcal{C}^1,W^{1,1},\mathrm{BV}} \le C_\#. \tag{11.1}$$

(b) *If, additionally, Φ satisfies $\varepsilon\|\Phi'\|_{\mathcal{C}^0} \le C_\# \varepsilon \sigma^2 (R-1-k) L_*$, then there exists a twice differentiable function $\boldsymbol{\chi}(\sigma, T, s, \varphi, \theta)$, smooth in σ, with derivatives*

[67] Recall that we are suppressing the subscripts $\ell_R, \ell_{R-1}, R-1, R$, when this does not create confusion.

with respect to φ, θ, s uniformly bounded by $C_\#|\sigma|$, such that

$$\chi_{\ell_k,k,j} = \boldsymbol{\chi}(\sigma, t-\varepsilon S_{k-1}, \varepsilon j, \theta^*_{\ell_k}, \bar\theta^*_{\ell_k,j}) + \mathcal{O}\left(\sigma^2\varepsilon^2 L_* + \sigma^3\varepsilon(R-1-k)L_*\right) \tag{11.2a}$$

$$\begin{aligned}\boldsymbol{\chi}(\sigma,T,s,\varphi,\theta) &= -\frac{\sigma^2}{2}\widehat{\Xi}(T-s-\varepsilon,\bar\theta(s+\varepsilon,\varphi))^2\hat{\boldsymbol{\sigma}}^2(\theta)+\mathcal{O}(\sigma^3)\\ &\le -\frac{\sigma^2}{4}\widehat{\Xi}(T-s-\varepsilon,\bar\theta(s+\varepsilon,\varphi))^2\hat{\boldsymbol{\sigma}}^2(\theta)\end{aligned} \tag{11.2b}$$

where $\theta^*_{\ell_k}, \bar\theta^*_{\ell_k,j}$ are defined in (8.20), $\hat{\boldsymbol{\sigma}}^2 \in \mathcal{C}^1(\mathbb{T},\mathbb{R}_{\geq 0})$ is given by the Green–Kubo formula (2.18) and $\widehat{\Xi}$ is defined in (8.7).

Proof. We will use indifferently the notation introduced in Proposition 9.7 and the one used in Appendix A. Such notations are connected by the relation $\mathcal{L}_{\ell_k,k,j} = \mathcal{L}_{\bar\theta^*_{\ell_k,j},\Omega^{k,\Phi}_{\ell_k,j}}$, where $\Omega^{k,\Phi}_{\ell_k,j}$ is defined in (9.19). In order to apply the results of Appendix A, let us consider the transfer operator given by $\mathcal{L}_{\bar\theta^*_{\ell,j},\varsigma\Omega^{k,\Phi}_{\ell,j}}$, for $\varsigma \in [0,1]$. Since the operator, for $\varsigma = 0$, has 1 as a simple maximal eigenvalue and a spectral gap (in any of the above mentioned spaces), it follows that we can choose σ_0 such that, for any $\sigma \in [0,\sigma_0]$, the operator for $\varsigma \in [0,1]$ has still a simple maximal eigenvalue and a spectral gap (assuming ε to be sufficiently small). Observe that, since the resolvent is continuous in θ, σ_0 can be chosen uniformly in θ and, consequently, since we have a uniform control on all terms appearing in $\varpi^k_{\ell_k,j}$, σ_0 can be chosen to be uniform in k, ℓ_k, j and Φ as well. This proves item (a).

We now prove item (b); note that the definition of $\varpi^k_{\ell_k,j}$ in (8.19) implies

$$\begin{aligned}\varpi^k_{\ell_k,j} &= \widehat{\Xi}(t-\varepsilon S_k, \bar\theta^*_{\ell_k,L_k})\Xi^*_{\ell_k,j,L_k}\hat\omega\\ &= \exp\left[\int_0^{t-\varepsilon S_k}\bar\omega'(\bar\theta(s+\varepsilon L_k,\theta^*_{\ell_k}))ds + \sum_{l=j+1}^{L_k-1}\int_{l\varepsilon}^{(l+1)\varepsilon}\bar\omega'(\bar\theta(s,\theta^*_{\ell_k}))ds + \mathcal{O}(\varepsilon^2L_k)\right]\hat\omega\\ &= \exp\left[\int_{(j+1)\varepsilon}^{t-\varepsilon S_{k-1}}\bar\omega'(\bar\theta(s,\theta^*_{\ell_k}))ds+\mathcal{O}(\varepsilon^2 L_k)\right]\hat\omega\\ &= \widehat{\Xi}(t-\varepsilon[S_{k-1}+j+1],\bar\theta^*_{\ell_k,j+1})\hat\omega + \mathcal{O}(\varepsilon^2L_k)\hat\omega.\end{aligned} \tag{11.3}$$

It is then natural to introduce the potentials

$$\boldsymbol{\Omega}(T,s,\varphi,x,\theta) = \widehat{\Xi}(T-s-\varepsilon,\bar\theta(s+\varepsilon,\varphi))\hat\omega(x,\theta) \tag{11.4}$$

so that

$$\varpi^k_{\ell_k,j}(x,\theta) = \boldsymbol{\Omega}(t-\varepsilon S_{k-1},\varepsilon j,\theta^*_{\ell_k},x,\theta)+\mathcal{O}(\varepsilon^2L_*)\hat\omega(x,\theta); \tag{11.5}$$

in particular, by definition of $\Omega^{k,\Phi}_{\ell_k,j}$ we gather

$$i\sigma\Omega^{k,\Phi}_{\ell_k,j} = i\sigma\boldsymbol{\Omega}(t-\varepsilon S_{k-1},\varepsilon j,\theta^*_{\ell_k},x,\theta)+\mathcal{O}(\varepsilon^2L_*+\sigma^2\varepsilon(R-1-k)L_*)\hat\omega(x,\theta).$$

Let, $e^{\boldsymbol{\chi}(\sigma,T,s,\varphi,\theta)}$ be the maximal eigenvalue of the operator associated to the potential $i\sigma\boldsymbol{\Omega}(T,s,\varphi,\cdot,\theta)$ and dynamics $f(\cdot,\theta)$. Then, by (A.19), we have

$$\chi_{\ell_k,k,j} = \boldsymbol{\chi}(\sigma,t-\varepsilon S_{k-1},\varepsilon j,\theta^*_{\ell_k},\bar\theta^*_{\ell_k,j})+\mathcal{O}(\sigma\varepsilon^2L_*+\sigma^2\varepsilon(R-1-k)L_*)\int_0^1 m_\varrho(\hat\omega h_\varrho)d\varrho,$$

where $m_\varrho = m_{\bar\theta^*_{\ell_k,j},\Omega_\varrho}$, and $h_\varrho = h_{\bar\theta^*_{\ell_k,j},\Omega_\varrho}$ with

$$\Omega_\varrho = i\sigma[(1-\varrho)\boldsymbol{\Omega}(t-\varepsilon S_{k-1},\varepsilon j,\theta^*_{\ell_k},x,\theta)+\varrho\Omega^{k,\Phi}_{\ell_k,j}.]$$

Then Lemma A.7 implies

$$\begin{aligned}\chi_{\ell_k,k,j} &= \boldsymbol{\chi}(\sigma, t-\varepsilon S_{k-1}, \varepsilon j, \theta^*_{\ell_k}, \bar\theta^*_{\ell_k,j}) \\ &\quad + \mathcal{O}\left((\sigma\varepsilon^2 L_* + \sigma^2\varepsilon(R-1-k)L_*)m_{\bar\theta^*_{\ell_k,j},0}(\hat\omega h_{\bar\theta^*_{\ell_k,j},0})\right) \\ &\quad + \mathcal{O}\left(\sigma^2\varepsilon^2 L_* + \sigma^3\varepsilon(R-1-k)L_*\right).\end{aligned}$$

The above implies the first equation of (11.2) since $\hat\omega$ has zero average by construction. Next, we use (A.11a), (A.12a) with $\varsigma = 0$ and (A.13) to obtain:

$$\begin{aligned}e^{\boldsymbol{\chi}(\sigma,T,s,\varphi,\theta)} &= 1 - \frac{\sigma^2}{2}\mu_\theta\left(\boldsymbol{\Omega}(T,s,\varphi,\cdot,\theta)^2\right) \\ &\quad - \sigma^2\sum_{m=1}^{\infty}\mu_\theta\left(\boldsymbol{\Omega}(T,s,\varphi,\bar f^m_{\ell,\theta}(\cdot),\theta)\boldsymbol{\Omega}(T,s,\varphi,\cdot,\theta)\right) + \mathcal{O}(\sigma^3) \\ &= e^{-\frac{\sigma^2}{2}\left[\mu_\theta(\boldsymbol{\Omega}^2)+2\sum_{m=1}^{\infty}\mu_\theta(\boldsymbol{\Omega}\circ\bar f^m_{\ell,\theta}\cdot\boldsymbol{\Omega})\right]+\mathcal{O}(\sigma^3)}\end{aligned} \tag{11.6}$$

where we have used the the decay of correlations implied by item (a). The second equation of (11.2) follows immediately, provided σ_0 has been chosen small enough. □

Remark 11.2. *As we will use the results below for all blocks, not just the last one, we are interested in all the operators $\mathcal{L}_{\ell_k,k,j}$. Yet, since all our computations are uniform in k and ℓ_k, there is no harm in dropping, again, the subscripts k, ℓ_k when this does not create confusion. Thus from now to the end of the section, to ease notation, k, ℓ_k are fixed and implicit. For the same reason we will write L rather than L_k. Moreover, to further ease our notation let us set $\Omega_j = \Omega^{k,\Phi}_{\ell_k,j}$ and $\bar\theta^*_j = \bar\theta^*_{\ell_k,j}$. Also, $\chi_j = \chi_{\bar\theta^*_j,\Omega_j}$, $m_j = m_{\bar\theta^*_j,\Omega_j}$ and $h_j = h_{\bar\theta^*_j,\Omega_j}$.*

Remark that, since $\mathcal{P}_j$ is a one dimensional projector, it can be written as $\mathcal{P}_j = h_j \otimes m_j$, where we choose to normalize h_j and m_j according to Lemma A.6. Also, for future reference, we define

$$\overline{\mathfrak{E}}_i := m_{i+1}(h_i - h_{i+1}). \tag{11.7}$$

We are now ready to derive a formula for the products of transfer operators in the perturbative regime.

Lemma 11.3. *There exists $\varepsilon_0, C_1 > 0$ such that, for any $i, n \in \{0, \cdots, L\}$, $\varepsilon \le \varepsilon_0$, $|\sigma| \in [C_1\varepsilon^2 L_*, |\sigma_0|]$ and $\varepsilon\|\Phi'\|_{\mathcal{C}^0} \le C_\#\sigma^2$ we have, for any $g \in \mathrm{BV}$,*

$$\begin{aligned}\left\|\widehat{\mathcal{L}}_{i+n}\cdots\widehat{\mathcal{L}}_i g - \exp\left[\sum_{j=0}^{n-1}\overline{\mathfrak{E}}_{i+j}\right]\overline{h}_{i,n}\cdot\overline{m}_{i,n}g\right\|_{\mathrm{BV}} &\le C_\#\|g\|_{L^1}\varepsilon^2 n^2 \\ &\quad + C_\#\left[e^{-c_\# n} + n^2\varepsilon^2 + (\log|\sigma|)^2|\sigma|\varepsilon\right]\|g\|_{BV},\end{aligned}$$

where

$$\overline{h}_{i,n} = \sum_{k=0}^{n}\widehat{\mathcal{Q}}_{i+n}\cdots\widehat{\mathcal{Q}}_{i+k+1}h_{i+k} \qquad \overline{m}_{i,n} = \sum_{k=0}^{n}m_{i+k}\widehat{\mathcal{Q}}_{i+k-1}\cdots\widehat{\mathcal{Q}}_i,$$

with $\widehat{\mathcal{L}}_j = e^{-\chi_j}\mathcal{L}_j$, $\widehat{\mathcal{Q}}_j = e^{-\chi_j}\mathcal{Q}_j$.

Proof. Let us define

$$\mathbb{X}_{i,n} = \exp\left[\sum_{j=0}^{n}\chi_{i+j}\right], \tag{11.8}$$

and introduce the auxiliary operators[68]

$$\overleftarrow{\mathcal{L}}_j = e^{\chi_j} h_{j+1} \otimes m_j + \mathcal{Q}_j \qquad \overleftarrow{\mathcal{L}}_{i,n} = \overleftarrow{\mathcal{L}}_{i+n} \overleftarrow{\mathcal{L}}_{i+n-1} \cdots \overleftarrow{\mathcal{L}}_i. \tag{11.9}$$

Observe that, by construction:

$$\mathcal{L}_{i+n} \cdots \mathcal{L}_i - \overleftarrow{\mathcal{L}}_{i,n} = \sum_{k=0}^{n} e^{\chi_{i+k}} \mathcal{L}_{i+n} \cdots \mathcal{L}_{i+k+1}(h_{i+k} - h_{i+k+1}) \otimes m_{i+k} \overleftarrow{\mathcal{L}}_{i,k-1}, \tag{11.10}$$

and one can check, by induction, that

$$\begin{aligned}\overleftarrow{\mathcal{L}}_{i,n} &= h_{i+n+1} \otimes \left[\sum_{k=0}^{n} e^{\sum_{j=k}^{n} \chi_{i+j}} m_{i+k} \mathcal{Q}_{i+k-1} \cdots \mathcal{Q}_i\right] + \mathcal{Q}_{i+n} \cdots \mathcal{Q}_i \\ &= \mathbb{X}_{i,n} \left\{h_{i+n+1} \otimes \overline{m}_{i,n} + \widehat{\mathcal{Q}}_{i+n} \cdots \widehat{\mathcal{Q}}_i\right\}. \end{aligned} \tag{11.11}$$

In order to continue we need to compare adjacent operators; this can be done using perturbation theory.

Sub-lemma 11.4. *For any $i \in \{0, \cdots, L\}$ we have:*

$$|\chi_{i+1} - \chi_i| \leq C_\# \varepsilon |\sigma|.$$
$$\|h_{i+1} - h_i\|_{\mathcal{C}^1} \leq C_\# \varepsilon.$$

The same bounds hold for $m_{i+1} - m_i$ as a functional on $W^{2,1}$. Yet, we also have the bound, for any $g \in \mathrm{BV}$:

$$|m_{i+1}(g) - m_i(g)| \leq C_\# \varepsilon |\sigma| (\log|\sigma|)^2 \|g\|_{\mathrm{BV}}.$$

Finally, we have

$$|\overline{\mathfrak{E}}_i| \leq C_\# \varepsilon |\sigma|. \tag{11.12}$$

Remark. *The estimate (11.12) reported above suffices for the present level of precision. Yet, if one wanted to compute the first term of the Edgeworth expansion, then it would be necessary to introduce the function*

$$\mathfrak{E}_k(\sigma, s, \varphi) = m_{\bar\theta(s+\varepsilon,\varphi), \widetilde\Omega_k(\sigma, s+\varepsilon, \varphi, \cdot)} \left(h_{\bar\theta(s,\varphi),\widetilde\Omega_k(\sigma,s,\varphi,\cdot)} - h_{\bar\theta(s+\varepsilon,\varphi),\widetilde\Omega_k(\sigma,s+\varepsilon,\varphi,\cdot)}\right),$$

where

$$\begin{aligned}\widetilde\Omega_k(\sigma, s, \varphi, x, \theta) &= i\sigma \mathbf{\Omega}(t - \varepsilon S_{k-1}, s, \varphi, x, \theta) \\ &\quad + \varepsilon \Phi'(\bar\theta(L_k \varepsilon, \varphi)) \widehat{\Xi}(L_k\varepsilon - s, \bar\theta(s+\varepsilon, \varphi)) \hat\omega(x, \theta).\end{aligned}$$

One could then use Appendix A.3 to show that $\overline{\mathfrak{E}}_i = \mathfrak{E}_k(\sigma, \varepsilon i, \theta^*_{\ell_k}) + \mathcal{O}(\varepsilon^2 L_*)$ and $\|\mathfrak{E}_k(\sigma, \cdot, \cdot)\|_{\mathcal{C}^1} \leq C_\# \varepsilon |\sigma|$. So one can keep $\mathfrak{E}_k$ in the definition of $\Phi_{\bar r}$ in Section 13.

Proof of Sub-lemma 11.4. We will have to vary both the dynamics and the potentials. This makes convenient to use, at times, the heavier, but more precise, notation introduced in Appendix A. In this notation $\mathcal{L}_j = \mathcal{L}_{\ell_k,k,j} = \mathcal{L}_{\bar\theta^*_{\ell_k,j}, \Omega^{k,\Phi}_{\ell_k,j}}$. Note that $\|\Omega_j\|_{\mathcal{C}^2} \leq C_\#(|\sigma| + \varepsilon \|\Phi'\|_{\mathcal{C}^0})$. Also, recall that $m_{\bar\theta^*_j,0} = \mathrm{Leb}$ and hence $\mathrm{Leb}(h_{\bar\theta^*_j,0}) = 1$. Next, observe that, although m_i is a distribution, it is almost a measure: indeed using Lemma A.13 with $n = C_\# |\sigma|^{-1}$ implies

$$|m_i(g)| \leq C_\# \|g\|_{L^1} + C_\# \exp\left[-c_\# |\sigma|^{-1}\right] \|g\|_{\mathrm{BV}}. \tag{11.13}$$

[68] In this section we use the standard conventions that, given any sequence of operators $\{A_i\}$, $A_j A_{j-1} \cdots A_{i+1} A_i = \mathbb{1}$ if $j < i$.

In turn, this implies that $\widehat{\mathcal{Q}}_i$ satisfies a Lasota–Yorke inequality as well. In order to see this, recall equations (A.2) and (A.27) and note that, if σ_0 is small enough, then there exists $\lambda_1 \in (1, \lambda)$ such that

$$\|\widehat{\mathcal{Q}}_i g\|_{\mathrm{BV}} \leq \|h_i m_i(g)\|_{\mathrm{BV}} + \|\widehat{\mathcal{L}}_i g\|_{\mathrm{BV}} \leq \lambda_1^{-1}\|g\|_{\mathrm{BV}} + C_\#\|g\|_{L^1}. \tag{11.14}$$

By Lemma A.9 we have

$$\begin{aligned} &|\chi_{\bar\theta^*_{i-1},\Omega_i} - \chi_{\bar\theta^*_i,\Omega_i}| \leq C_\#|\sigma|\varepsilon \\ &\|h_{\bar\theta^*_{i-1},\Omega_i} - h_{\bar\theta^*_i,\Omega_i}\|_{\mathcal{C}^2} \leq C_\#\varepsilon \\ &|m_{\bar\theta^*_{i-1},\Omega_i}(g) - m_{\bar\theta^*_i,\Omega_i}(g)| \leq C_\#\varepsilon|\sigma|\|g\|_{W^{2,1}}. \end{aligned} \tag{11.15}$$

It turns out that the third of the above estimates is not very convenient owing to the higher derivative in the right hand side. However, Lemma A.15 implies

$$|m_{\bar\theta^*_{i-1},\Omega_i}(g) - m_{\bar\theta^*_i,\Omega_i}(g)| \leq C_\#(\log|\sigma|)^2|\sigma|\varepsilon\|g\|_{\mathrm{BV}}. \tag{11.16}$$

Next, equations (9.19) and (8.19) imply

$$\|\Omega_{i+1} - \Omega_i\|_{\mathcal{C}^1} \leq C_\#\varepsilon|\sigma|. \tag{11.17}$$

We can then use (A.19) and argue as in (A.20) to obtain,[69] for $i \leq L$,

$$\begin{aligned} &|\chi_{\bar\theta^*_{i-1},\Omega_{i-1}} - \chi_{\bar\theta^*_{i-1},\Omega_i}| \leq C_\#\varepsilon|\sigma| \\ &\|h_{\bar\theta^*_{i-1},\Omega_{i-1}} - h_{\bar\theta^*_{i-1},\Omega_i}\|_{\mathcal{C}^1} \leq C_\#|\sigma|\varepsilon \\ &|m_{\bar\theta^*_{i-1},\Omega_{i-1}}(g) - m_{\bar\theta^*_{i-1},\Omega_i}(g)| \leq C_\#|\sigma|\varepsilon\|g\|_{\mathrm{BV}}, \end{aligned} \tag{11.18}$$

Collecting the above facts, yields the first three inequalities of the Lemma.

Next, by the third of equations (11.15), using that $h_{\bar\theta^*_{i-1},\Omega_{i-1}} \in \mathcal{C}^2$, see Remark A.4, and (11.18), we can write

$$\begin{aligned} m_i(h_{i-1} - h_i) &= m_i(h_{i-1}) - 1 \\ &= \left[m_{\bar\theta^*_i,\Omega_i} - m_{\bar\theta^*_{i-1},\Omega_{i-1}}\right](h_{\bar\theta^*_{i-1},\Omega_{i-1}}) \\ &= \left[m_{\bar\theta^*_{i-1},\Omega_i} - m_{\bar\theta^*_{i-1},\Omega_{i-1}}\right](h_{\bar\theta^*_{i-1},\Omega_{i-1}}) + \mathcal{O}(\varepsilon|\sigma|) = \mathcal{O}(\varepsilon|\sigma|), \end{aligned}$$

which concludes the proof of the Sub-lemma. □

We also need a bound on products of $\widehat{\mathcal{Q}}_i$'s which is rather obvious but a bit lengthy to prove.

Sub-lemma 11.5. *There exists $n_* \in \mathbb{N}$ such that, for all $k \in \{0, \cdots, L - n_*\}$, we have*

$$\|\widehat{\mathcal{Q}}_{k+n_*} \cdots \widehat{\mathcal{Q}}_k\|_{\mathrm{BV}} \leq e^{-1}.$$

Proof. Note that, by Lemma 11.1, there exist $C_4 > 0$ such that $\sup_{k,j}\|\widehat{\mathcal{Q}}_k^j\|_{\mathrm{BV}} \leq C_4$. We are now going to prove, by induction, that there exists $C_Q \geq C_4$ such that for any $N_Q > 0$, there exists ε_0, σ_0 such that, for all $\varepsilon \leq \varepsilon_0$ and $|\sigma| \leq \sigma_0$, we have

$$\sup_k \sup_{j \leq N_Q} \|\widehat{\mathcal{Q}}_{k+j} \cdots \widehat{\mathcal{Q}}_k\|_{\mathrm{BV}} \leq C_Q. \tag{11.19}$$

The claim is trivially true for $N_Q = 0$. Suppose it is true for all $j \leq N_Q - 1$ for some σ_0, ε_0. Possibly by decreasing σ_0 assume that $\sigma_0^2 N_Q \leq 1$ and note that, since we assume $\varepsilon\|\Phi'\|_{\mathcal{C}^0} \leq C_\#\sigma^2$:

$$|\mathrm{Leb}(\widehat{\mathcal{L}}_i g)| \leq \|\widehat{\mathcal{L}}_i g\|_{L^1} \leq e^{-\chi_i}\mathrm{Leb}(e^{\mathrm{Re}(\Omega^{k,\Phi}_{\ell_{k,j}})}|g|) \leq e^{C_\#\sigma^2}\|g\|_{L^1}.$$

[69] The formula (A.20b) holds also with the BV norm on the left hand side due to the lower semicontinuity of the variation [21, Section 5.2.1, Theorem 1].

Together with Sub-Lemma 11.4, the above inequality implies

$$
\begin{aligned}
&\|\widehat{\mathcal{Q}}_{k+j}\cdots\widehat{\mathcal{Q}}_k g\|_{L^1}\\
&\qquad\leq \|\widehat{\mathcal{L}}_{k+j}\widehat{\mathcal{Q}}_{k+j-1}\cdots\widehat{\mathcal{Q}}_k g\|_{L^1}+C_\#|(m_{k+j}-m_{k+j-1})(\widehat{\mathcal{Q}}_{k+j-1}\cdots\widehat{\mathcal{Q}}_k g)|\\
&\qquad\leq e^{c_\#\sigma^2}\|\widehat{\mathcal{Q}}_{k+j-1}\cdots\widehat{\mathcal{Q}}_k g\|_{L^1}+C_\#(\log|\sigma|)^2|\sigma|\varepsilon\|\widehat{\mathcal{Q}}_{k+j-1}\cdots\widehat{\mathcal{Q}}_k g\|_{\mathrm{BV}}.
\end{aligned}
$$

Iterating the above argument, since $\sigma N_q\leq 1$, and by the inductive hypothesis:

$$
\begin{aligned}
\|\widehat{\mathcal{Q}}_{k+j}\cdots\widehat{\mathcal{Q}}_k g\|_{L^1}&\leq C_\#\|g\|_{L^1}+C_\#(\log|\sigma|)^2|\sigma|\varepsilon\sum_{l=0}^{j-1}\|\widehat{\mathcal{Q}}_{k+l}\cdots\widehat{\mathcal{Q}}_k g\|_{\mathrm{BV}}\\
&\leq C_\#\|g\|_{L^1}+C_\#(\log|\sigma|)^2|\sigma|\varepsilon jC_Q\|g\|_{\mathrm{BV}}.
\end{aligned}\tag{11.20}
$$

We can now use (11.14) to write

$$
\begin{aligned}
\|\widehat{\mathcal{Q}}_{k+j}\cdots\widehat{\mathcal{Q}}_k g\|_{\mathrm{BV}}&\leq \lambda_1^{-1}\|\widehat{\mathcal{Q}}_{k+j-1}\cdots\widehat{\mathcal{Q}}_k g\|_{\mathrm{BV}}+C_\#\|\widehat{\mathcal{Q}}_{k+j-1}\cdots\widehat{\mathcal{Q}}_k g\|_{L^1}\\
&\leq[(\lambda_1^{-1}+C_\#(\log|\sigma|)^2|\sigma|\varepsilon j)C_Q+C_\#]\|g\|_{\mathrm{BV}}
\end{aligned}
$$

from which the claim follows.

Next, by Sub-Lemma 11.4,

$$
\begin{aligned}
\|(\widehat{\mathcal{Q}}_{k+j}-\widehat{\mathcal{Q}}_k)g\|_{L^1}&\leq\|(\widehat{\mathcal{L}}_{k+j}-\widehat{\mathcal{L}}_k)g-(h_{k+j}\otimes m_{k+j}-h_j\otimes m_j)g\|_{L^1}\\
&\leq C_\#(\log|\sigma|)^2|\sigma|\varepsilon j\|g\|_{\mathrm{BV}}.
\end{aligned}\tag{11.21}
$$

Using Lemma 11.1 once again, there exists $m_*\in\mathbb{N}$, independent of ε and σ, such that, for all $k\in\{0,\cdots,L\}$, $\|\widehat{\mathcal{Q}}_k^{m_*}\|_{\mathrm{BV}}\leq(2eC_Q)^{-1}$ and $\lambda_1^{m_*}\geq 8eC_Q^2$. We will use the above claim with $N_Q=m_*$. Note that, in particular, this implies that $\sigma_0^2m_*\leq 1$.

Then, using equations (11.14), (11.19), (11.20) and (11.21):

$$
\begin{aligned}
&\|(\widehat{\mathcal{Q}}_{k+2m_*}\cdots\widehat{\mathcal{Q}}_k-\widehat{\mathcal{Q}}_{k+2m_*}\cdots\widehat{\mathcal{Q}}_{k+m_*+1}\widehat{\mathcal{Q}}_{k+m_*}^{m_*})g\|_{\mathrm{BV}}\\
&\leq(\lambda_1^{-m_*}+C_\#(\log|\sigma|)^2|\sigma|\varepsilon m_*C_Q)\|(\widehat{\mathcal{Q}}_{k+m_*}\cdots\widehat{\mathcal{Q}}_k-\widehat{\mathcal{Q}}_{k+m_*}^{m_*})g\|_{\mathrm{BV}}\\
&\quad+C_\#\|(\widehat{\mathcal{Q}}_{k+m_*}\cdots\widehat{\mathcal{Q}}_k-\widehat{\mathcal{Q}}_{k+m_*}^{m_*})g\|_{L^1}\\
&\leq 2C_Q(\lambda_1^{-m_*}+C_\#(\log|\sigma|)^2|\sigma|\varepsilon m_*C_Q)\|g\|_{\mathrm{BV}}\\
&\quad+C_\#\sum_{j=1}^{m_*}\|\widehat{\mathcal{Q}}_{k+m_*}^j(\widehat{\mathcal{Q}}_{k+m_*-j}-\widehat{\mathcal{Q}}_{k+m_*})\widehat{\mathcal{Q}}_{k+m_*-j-1}\cdots\widehat{\mathcal{Q}}_k g\|_{L^1}\\
&\leq\left[2\lambda_1^{-m_*}+C_\#(\log|\sigma|)^2|\sigma|\varepsilon m_*^2\right]C_Q^2\|g\|_{\mathrm{BV}}\leq\frac{1}{2e}\|g\|_{\mathrm{BV}}
\end{aligned}
$$

provided ε_0,σ_0 have been chosen small enough. Hence

$$
\begin{aligned}
\|\widehat{\mathcal{Q}}_{k+2m_*}\cdots\widehat{\mathcal{Q}}_k\|_{\mathrm{BV}}&\leq\|\widehat{\mathcal{Q}}_{k+2m_*}\cdots\widehat{\mathcal{Q}}_k-\widehat{\mathcal{Q}}_{k+2m_*}\cdots\widehat{\mathcal{Q}}_{k+m_*+1}\widehat{\mathcal{Q}}_{k+m_*}^{m_*})\|_{\mathrm{BV}}\\
&\quad+\|\widehat{\mathcal{Q}}_{k+2m_*}\cdots\widehat{\mathcal{Q}}_{k+m_*+1}\widehat{\mathcal{Q}}_{k+m_*}^{m_*}\|_{\mathrm{BV}}\leq\frac{1}{2e}+C_Q(2eC_Q)^{-1}\leq e^{-1}.
\end{aligned}
$$

We have thus proved our claim, with $n_*=2m_*$. □

We can now use Sub-Lemmata 11.4 and 11.5 to continue the argument that we left at (11.11): we immediately obtain

$$
\begin{aligned}
&\left\|\mathbb{X}_{i,n}^{-1}\overleftarrow{\mathcal{L}}_{i,n}-h_{i+n+1}\otimes\overline{m}_{i,n}\right\|_{\mathrm{BV}}\leq C_\#e^{-c_\#n}\\
&|\overline{m}_{i,n}(g)-m_i(g)|\leq\sum_{k=1}^{i+n}|(m_{i+k}-m_{i+k-1})(\widehat{\mathcal{Q}}_{i+k-1}\cdots\widehat{\mathcal{Q}}_i g)|\\
&\qquad\leq C_\#(\log|\sigma|)^2|\sigma|\varepsilon\|g\|_{\mathrm{BV}}.
\end{aligned}\tag{11.22}
$$

At this point we can write, using repeatedly (11.10):

$$
\begin{aligned}
\mathcal{L}_{i+n}\cdots\mathcal{L}_i - \overleftarrow{\mathcal{L}}_{i,n} &= \sum_{k=0}^{n} e^{\chi_{i+k}}\mathcal{L}_{i+n}\cdots\mathcal{L}_{i+k+1}(h_{i+k}-h_{i+k+1})\otimes m_{i+k}\overleftarrow{\mathcal{L}}_{i,k-1},\\
&= \sum_{k=0}^{n} e^{\chi_{i+k}}\overleftarrow{\mathcal{L}}_{i+k+1,n-k-1}(h_{i+k}-h_{i+k+1})\otimes m_{i+k}\overleftarrow{\mathcal{L}}_{i,k-1}\\
&\quad+\sum_{k=0}^{n}\sum_{j=0}^{n-k-1} e^{\chi_{i+k}+\chi_{i+k+j+1}}\mathcal{L}_{i+n}\cdots\mathcal{L}_{i+k+j+2}(h_{i+k+j+1}-h_{i+k+j+2})\\
&\quad\otimes m_{i+k+j+1}\overleftarrow{\mathcal{L}}_{i+k+1,j-1}(h_{i+k}-h_{i+k+1})\otimes m_{i+k}\overleftarrow{\mathcal{L}}_{i,k-1}.
\end{aligned}
\tag{11.23}
$$

Note that, by (11.22) and (11.13) we have

$$
\|\mathbb{X}_{i,j}^{-1}\overleftarrow{\mathcal{L}}_{i,j}\|_{\mathrm{BV}} \le C_\#.
$$

Then, the first line of (11.23), together with Sub-Lemma 11.4, suffices to write

$$
\left\|\mathcal{L}_{i+n}\cdots\mathcal{L}_i - \overleftarrow{\mathcal{L}}_{i,n}\right\|_{\mathrm{BV}} \le C_\#\varepsilon\sum_{k=0}^{n-1}\|\mathcal{L}_{i+n}\cdots\mathcal{L}_{i+k+1}\|_{\mathrm{BV}}\,|\mathbb{X}_{i,k-1}|\,.
$$

From the above it follows by induction:

$$
\|\mathcal{L}_{i+n}\cdots\mathcal{L}_i\|_{\mathrm{BV}} \le C_\#|\mathbb{X}_{i,n}|. \tag{11.24}
$$

Recall that Lemma A.14, implies that, for any $j\in\{0,\cdots,n\}$,

$$
|\overline{m}_{i,j}(g)| + |m_{i+j}(g)| \le C_\#\|g\|_{L^1} + C_\#|\sigma|^{100}\|g\|_{BV} = C_\#\|g\|_{BV_\sigma},
$$

where we have introduced the shorthand notation $\|g\|_{\mathrm{BV}_\sigma} = \|g\|_{L^1} + |\sigma|^{100}\|g\|_{\mathrm{BV}}$. Note that (11.11) and the definition of $\overline{m}_{i,k}$ in the statement of Lemma 11.3 imply

$$
\begin{aligned}
m_{i+k}\overleftarrow{\mathcal{L}}_{i,k-1} &= \mathbb{X}_{i,k-1}\left\{\overline{m}_{i,k-1} + m_{i+k}\widehat{\mathcal{Q}}_{i+k-1}\cdots\widehat{\mathcal{Q}}_i\right\}\\
&= \mathbb{X}_{i,k-1}\overline{m}_{i,k}.
\end{aligned}
$$

Given the above, we can now use the full force of (11.23) and Sub-Lemma 11.4, using (11.11):

$$
\begin{aligned}
\mathcal{L}_{i+n}\cdots\mathcal{L}_i g &= \overleftarrow{\mathcal{L}}_{i,n}g + \sum_{k=0}^{n}\mathbb{X}_{i,k}\overleftarrow{\mathcal{L}}_{i+k+1,n-k-1}(h_{i+k}-h_{i+k+1})\cdot\overline{m}_{i,k}(g)\\
&\quad+\mathbb{X}_{i,n}\|g\|_{\mathrm{BV}_\sigma}\mathcal{O}_{\mathcal{C}^1}(\varepsilon^2 n^2)\\
&= \overleftarrow{\mathcal{L}}_{i,n}g + \mathbb{X}_{i,n}(h_{i+n}-h_{i+n+1})\cdot\overline{m}_{i,n}(g)\\
&\quad+\mathbb{X}_{i,n}\sum_{k=0}^{n-1} h_{i+n+1}\cdot\overline{m}_{i+k+1,n-k-1}(h_{i+k}-h_{i+k+1})\cdot\overline{m}_{i,k}(g)\\
&\quad+\mathbb{X}_{i,n}\sum_{k=0}^{n-1}\widehat{\mathcal{Q}}_{i+n}\cdots\widehat{\mathcal{Q}}_{i+k+1}h_{i+k}\cdot\overline{m}_{i,k}(g)\\
&\quad+\mathbb{X}_{i,n}\|g\|_{\mathrm{BV}_\sigma}\mathcal{O}_{\mathcal{C}^1}(\varepsilon^2 n^2).
\end{aligned}
$$

Finally, by definition of $\overline{h}_{i,n}, \overline{m}_{i,n}$, (11.22), Sub-Lemmata 11.4, 11.5, since $L^2\varepsilon < 1$ and recalling (11.7) we have[70]

$$\begin{aligned}
\widehat{\mathcal{L}}_{i+n}\cdots\widehat{\mathcal{L}}_i g &= h_{i+n+1}\overline{m}_{i,n}(g) + (h_{i+n} - h_{i+n+1})\cdot\overline{m}_{i,n}(g)\\
&\quad + \sum_{k=0}^{n-1}\widehat{\mathcal{Q}}_{i+n}\cdots\widehat{\mathcal{Q}}_{i+k+1}h_{i+k}\cdot\overline{m}_{i,n}(g)\\
&\quad + \sum_{k=0}^{n-1} m_{i+k+1}(h_{i+k} - h_{i+k+1})\overline{h}_{i,n}\cdot\overline{m}_{i,n}(g)\\
&\quad + \|g\|_{\mathrm{BV}_\sigma}\mathcal{O}_{\mathcal{C}^1}(\varepsilon^2 n^2) + \|g\|_{\mathrm{BV}}\mathcal{O}_{\mathcal{C}^1}(e^{-c_\# n} + (\log|\sigma|)^2|\sigma|\varepsilon)\\
&= \exp\left[\sum_{k=0}^{n-1}\overline{\mathfrak{E}}_{i+k}\right]\overline{h}_{i,n}\cdot\overline{m}_{i,n}(g) + \|g\|_{L^1}\mathcal{O}_{\mathcal{C}^1}(\varepsilon^2 n^2)\\
&\quad + \|g\|_{\mathrm{BV}}\mathcal{O}_{\mathcal{C}^1}(e^{-c_\# n} + e^{c_\# n|\sigma|\varepsilon}n^2\varepsilon^2|\sigma|^2 + (\log|\sigma|)^2|\sigma|\varepsilon).
\end{aligned}$$

□

11.2. Main result for the intermediate regime.

Lemma 11.3 is the basic tool to conclude the proof of the Local Central Limit Theorem. In this subsection we see how to use the lemma to prove the results we are interested in for the (easier) intermediate regime. The case of the small regime will be dealt with in the next section.

Proof of Proposition 8.12. First of all recall (see Remark 8.10 that in this regime we are considering only families of long standard pairs. Let us apply Proposition 9.7 with $\Phi \equiv 0$: we have, choosing $\varrho = \varepsilon^2$ and recalling that $L_* = \mathcal{O}(\varepsilon^{-3\delta_*})$:

$$\begin{aligned}
\sum_{\ell_R\in\mathfrak{L}^{R-1}_{\ell_{R-1}}}\nu_{R-1,\ell_{R-1},\ell_R} &= \mathrm{Leb}\sum_{\ell_R\in\mathfrak{L}^{R-1}_{\ell_{R-1}}}\nu_{R-1,\ell_{R-1},\ell_R}\mathring{\rho}_{\ell_R}\\
&= e^{i\sigma\varepsilon\mathfrak{C}^{k,1}_{\ell_{R-1},0,\emptyset}}\mathrm{Leb}\,\mathcal{L}_{\ell_{R-1},R-1,L_{R-1}-1}\cdots\mathcal{L}_{\ell_{R-1},R-1,0}[\Psi_q\tilde{\rho}_{\ell_{R-1},\varrho}]\\
&\quad + \mathcal{O}(\varepsilon^{2-9\delta_*}).
\end{aligned}$$

Next, we analyze each of the terms separately. Lemma 11.3 and Sub-Lemma 11.4 imply

$$\begin{aligned}
&\left\|\left[\mathcal{L}_{\ell_{R-1},R-1,L_{R-1}-1}\cdots\mathcal{L}_{\ell_{R-1},R-1,0} - \mathbb{X}^*_{\ell_{R-1}}\overline{h}_{0,L_{R-1}-1}\otimes\overline{m}_{0,L_{R-1}-1}\right]\tilde{\rho}_{\ell_{R-1},\varrho}\right\|_{\mathrm{BV}} \le\\
&\qquad \le C_\#\left|\mathbb{X}_{0,L_{R-1}-1}\right|\varepsilon,
\end{aligned}$$

where

$$\mathbb{X}^*_{\ell_{R-1}} = \exp\left[\sum_{j=0}^{L_{R-1}}\chi_j + \overline{\mathfrak{E}}_j\right].$$

Notice that Lemma 11.1 and (11.12) allow to write:

$$\begin{aligned}
\log\mathbb{X}^*_{\ell_{R-1}} &= -\frac{\sigma^2}{2}\sum_{j=0}^{L_{R-1}}\widehat{\Xi}(t - \varepsilon[S_{R-2}+j+1], \bar{\theta}^*_{\ell_{R-1},j})^2\hat{\boldsymbol{\sigma}}^2(\bar{\theta}^*_{\ell_{R-1},j})\\
&\quad + \mathcal{O}(\sigma^3 + \sigma^2\varepsilon^2 L_{R-1} + \sigma\varepsilon L_{R-1}),
\end{aligned}\tag{11.25}$$

and the same estimate holds for $\mathbb{X}_{0,L_{R-1}}$. The above implies that for $|\sigma|\in[\varepsilon^{\delta_*},\sigma_0]$, given the choice $L = \varepsilon^{-3\delta_*}$, we have $\mathbb{X}^*_{\ell_{R-1}} = \mathcal{O}(e^{-c_\#\varepsilon^{-\delta_*}})$, and the same for

[70] Here we use repeatedly that $\widehat{\mathcal{Q}}_j h_{j-1} = \widehat{\mathcal{Q}}_j(h_{j-1} - h_j)$ and the similar relation for m_j.

$\mathbb{X}_{0,L_{R-1}}$. Also, by similar arguments, the correlation terms will give a smaller contribution since $9\delta_* < 1$. It follows that

$$\left|\sum_{\ell_R\in\mathfrak{L}^{R-1}_{\ell_{R-1}}} \nu_{R-1,\ell_{R-1},\ell_R}\right| \leq C_\# \varepsilon^{2-9\delta_*}. \qquad \square$$

12. One block estimate: the small σ regime

As already mentioned, in the small σ regime the contraction of a single block is not sufficient for our needs; we thus need to combine together several blocks. To this end, in this section, we provide a suitable description of the one block contribution. Given a complex standard pair $\ell = (\mathbb{G}_\ell, \rho_\ell)$, recall the notation $\mathring{\rho}_\ell = \rho_\ell \mathbb{1}_{[a_\ell,b_\ell]}$.

Proposition 12.1. *Let $\sigma \in \mathcal{J}_1$, $\varrho = \varepsilon^2$, $k \in \{0,\cdots,R-1\}$ and $\Phi \in \mathcal{C}^2(\mathbb{T},\mathbb{C})$ such that $\Phi^+ = \max \operatorname{Re}(\Phi) \leq C_\#$, $\|\Phi'\|_{\mathcal{C}^0} \leq C_\# \min\{\sigma^2(R-1-k)L_*, \varepsilon^{-1}L_*^{-1}\}$ and $\varepsilon\|\Phi'\|_{\mathcal{C}^1} \sup_{0\leq j\leq L_k} \|\Xi^*_{\ell,j}\hat{\omega}\|_{\mathcal{C}^1} \leq \sigma_0$. Then*[71] *we have the estimates:*

$$\sum_{\ell_{k+1}\in\mathfrak{L}^k_{\ell_k}} e^{\Phi\circ G_{\ell_{k+1}}} \nu_{k,\ell_k,\ell_{k+1}} \mathring{\rho}_{\ell_{k+1}} = \mathbb{X}^{**}_{\ell_k} \overline{h}_{0,L_k-1} \overline{m}_{0,L_k-1}\left[e^{\Phi(\bar{\theta}_{\ell_k,L_k})}\mathring{\rho}_{\ell_k}\right] + \mathcal{E}^{**}_{\ell_k,1} + \mathcal{E}^{**}_{\ell_k,2} \tag{12.1}$$

where $\overline{h}_{i,n}$ and $\overline{m}_{i,n}$ are defined in Lemma 11.3,

$$\begin{aligned} \mathbb{X}^{**}_{\ell_k} &= \exp\left[\varepsilon^{-1}\int_0^{\varepsilon L_k} \mathcal{G}_k(\theta^*_{\ell_k}, s, \sigma)ds\right] \\ \mathcal{G}_k(\theta, s, \sigma) &= \boldsymbol{\chi}(\sigma, t-\varepsilon S_{k-1}, s, \theta, \bar{\theta}(s,\theta)), \end{aligned} \tag{12.2}$$

$\boldsymbol{\chi}$ being defined just below (11.5). Finally

$$\begin{aligned} &\|\mathcal{E}^{**}_{\ell_k,1}\|_{L^1} \leq C_\# e^{\Phi^+}(\varepsilon^2 L_k^3 + |\sigma|\varepsilon L_k^2 + (R-k)|\sigma|^3\varepsilon L_k^2)\ ; \\ &\|\mathcal{E}^{**}_{\ell_k,2}\|_{L^1} \leq C_\# e^{\Phi^+}\varepsilon L_k^2\ ; \qquad |\mathrm{Leb}(\mathcal{E}^{**}_{\ell_k,2})| \leq e^{\Phi^+} C_\#(|\sigma|\varepsilon L_k^3 + \varepsilon^2 L_k^3) \\ &\|\mathcal{E}^{**}_{\ell_k,1}\|_{\mathrm{BV}} \leq C_\#\ ; \qquad \|\mathcal{E}^{**}_{\ell_k,2}\|_{\mathrm{BV}} \leq C_\#, \end{aligned}$$

Proof. First let us apply Proposition 9.7 to the left hand side of (12.1), obtaining:

$$\sum_{\ell_{k+1}\in\mathfrak{L}^k_{\ell_k}} e^{\Phi\circ G_{\ell_{k+1}}} \nu_{k,\ell_k,\ell_{k+1}} \mathring{\rho}_{\ell_{k+1}} = e^{i\sigma\varepsilon\mathfrak{C}^{k,1}_{\ell_k,0,\emptyset}} \mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}[\Psi_{\ell_k,q} e^{\Phi(\bar{\theta}_{\ell_k,L_k})}\tilde{\rho}_{\ell_k,\varrho}] + \mathcal{E}^*_{\ell_k}. \tag{12.3}$$

Then observe that, by definition

$$|e^{i\sigma\varepsilon\mathfrak{C}^{k,1}_{\ell_k,0,\emptyset}} - 1| \leq C_\#\varepsilon|\sigma|L_k.$$

Next, we rewrite the first term on the right hand side of (12.3)

$$\begin{aligned} &\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}[\Psi_{\ell_k,q} e^{\Phi(\bar{\theta}_{\ell_k,L_k})}\tilde{\rho}_{\ell_k,\varrho}] \\ &= \mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\Big\{[\Psi_{\ell_k,q} - e^{\varepsilon\mathcal{K}_0}]\tilde{\rho}_{\ell_k,\varrho} + [e^{\varepsilon\mathcal{K}_0}\tilde{\rho}_{\ell_k,\varrho} - \mathring{\rho}_{\ell_k}]\Big\} e^{\Phi(\bar{\theta}_{\ell_k,L_k})} \\ &\quad + \mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\mathring{\rho}_{\ell_k} e^{\Phi(\bar{\theta}_{\ell_k,L_k})}. \end{aligned} \tag{12.4}$$

[71] Recall (see Remark 8.10) that in this regime we are dealing with families of *long* standard pairs.

We can now apply Lemma 11.3 to each term separately:

$$\left\|\left[\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}-\mathbb{X}^*_{\ell_k}e^{-\overline{\mathfrak{E}}_{L_k-1}}\overline{h}_{0,L_k-1}\otimes\overline{m}_{0,L_k-1}\right]e^{\Phi}\mathring{\rho}_{\ell_k}\right\|_{\mathrm{BV}} \leq C_\# e^{\Phi^+}|\mathbb{X}_{0,L_k-1}|L_k^2\varepsilon^2$$

$$\mathbb{X}^*_{\ell_k}=\exp\left[\sum_{j=0}^{L_k}\chi_{\ell_k,k,j}+\overline{\mathfrak{E}}_j\right] = \exp\left[\varepsilon^{-1}\int_0^{\varepsilon L_k}\mathcal{G}_k(\theta^*_{\ell_k},s,\sigma)ds\right]+\mathcal{O}(|\sigma|\varepsilon L_k+|\sigma|^3\varepsilon(R-k)L_k^2), \tag{12.5}$$

where we used the fact that $\|e^{\Phi(\bar\theta_{\ell_k},L_k)}\|_{\mathrm{BV}}\leq C_\#\varepsilon\|\Phi'\|_{\mathcal{C}^0}$; in the second line, we have used the definition (11.8) and, in the last line, we have used Lemma 11.1-(b) and (11.12).

Next, we want to compute the correlation terms. They are sum of terms of the following type (possibly expanding in series the exponential), with $s\leq q$ which, recall, has been fixed $q=7$:

$$\begin{aligned}&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\prod_{i=1}^{s}A\circ\bar F_\varepsilon^{i_s}e^{\Phi(\bar\theta_{\ell_k},L_k)}\tilde\rho_{\ell,\varrho}\|_{\mathrm{BV}}=\\&=\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,i_s+1}A_{s,\bar\imath}\mathcal{L}_{\ell_k,k,i_s}\cdots A_{1,\bar\imath}\mathcal{L}_{\ell_k,k,i_1}\cdots\mathcal{L}_{\ell_k,k,0}e^{\Phi}\tilde\rho_{\ell,\varrho}\|_{\mathrm{BV}}\\&\leq(C_\#)^q,\end{aligned}$$

where we have used equations (11.24) and (11.2) (which implies $|\mathbb{X}_{i,n}|\leq e^{-c_\#\sigma^2 n}$). Thus, by Proposition 9.7-(b), Notation 8.8 and since, by hypothesis, $\|\mathcal{L}_{\ell_k,k,j}\|_{L^1}\leq e^{c_\# L_*^{-1}}$,

$$\begin{aligned}&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}[\Psi_{\ell_k,q}-e^{\varepsilon\mathcal{K}_0}]e^{\Phi}\tilde\rho_{\ell_k,\varrho}\|_{L^1}\leq C_\# e^{\Phi^+}\varepsilon|\sigma|L_k^2\\&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}[\Psi_{\ell_k,q}-e^{\varepsilon\mathcal{K}_0}]e^{\Phi}\tilde\rho_{\ell_k,\varrho}\|_{\mathrm{BV}}\leq C_\#.\end{aligned} \tag{12.6}$$

Finally, we compute the remaining term on the left hand side of (12.4). Since Proposition 9.7-(d) implies

$$\|e^{\varepsilon\mathcal{K}_0}\tilde\rho_{\ell_k,\varrho}-\mathring\rho_{\ell_k}\|_{L^1}\leq C_\#\varepsilon L_k^2, \tag{12.7}$$

a brute force estimate, as the one above, yields

$$\begin{aligned}&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\,e^{\Phi}[e^{\varepsilon\mathcal{K}_0}\tilde\rho_{\ell_k,\varrho}-\mathring\rho_{\ell_k}]\|_{L^1}\leq C_\# e^{\Phi^+}\varepsilon L_k^2\\&\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\,e^{\Phi}[e^{\varepsilon\mathcal{K}_0}\tilde\rho_{\ell_k,\varrho}-\mathring\rho_{\ell_k}]\|_{\mathrm{BV}}\leq C_\#.\end{aligned} \tag{12.8}$$

Unfortunately, inequalities (12.8) yields a mistake is too large for our needs. We must be a bit more careful and compute the term in more detail.[72] It happens that a more precise estimate of the average with respect to Lebesgue will suffice. Let us us call $\mathcal{L}_{\ell,k,j,0}$ the transfer operator $\mathcal{L}_{\ell,k,j}$ computed for $\sigma=0$. Then, by standard perturbation theory,

$$\|\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}-\mathcal{L}_{\ell_k,k,L_k-1,0}\cdots\mathcal{L}_{\ell_k,k,0,0}\|_{L^1}\leq C_\# L_k|\sigma|. \tag{12.9}$$

[72] Note that this could be done also for other terms, hence allowing for smaller errors.

Thus, since $\operatorname{Leb}\mathcal{L}_{\ell_k,k,j,0} = \operatorname{Leb}$, by equations (12.7), (12.9) and Proposition 9.7-(b) we have

$$
\begin{aligned}
&\operatorname{Leb}\left\{\mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\, e^{\Phi(\bar\theta_{\ell_k,L_k})}[e^{\varepsilon\mathcal{K}_0}\tilde\rho_{\ell_k,\varrho}-\mathring{\rho}_{\ell_k}]\right\}\\
&= \operatorname{Leb} e^{\Phi(\bar\theta_{\ell_k,L_k})}\left\{e^{\varepsilon\mathcal{K}_0}\tilde\rho_{\ell_k,\varrho}-\mathring{\rho}_{\ell_k}\right\}+e^{\Phi^+}\mathcal{O}(\sigma L_k^3\varepsilon)\\
&= e^{\Phi(\bar\theta^*_{\ell_k,L_k})}\operatorname{Leb}\left\{e^{\varepsilon\mathcal{K}_0}\tilde\rho_{\ell_k,\varrho}-\mathring{\rho}_{\ell_k}\right\}+e^{\Phi^+}\mathcal{O}(\sigma L_k^3\varepsilon+\varepsilon^2L_k^2)\\
&= e^{\Phi^+}\mathcal{O}(\sigma L_k^3\varepsilon+\varepsilon^2L_k^3).
\end{aligned}
\tag{12.10}
$$

The proposition then follows by collecting the previous inequalities, setting $\mathcal{E}^{**}_{\ell_k,2} = \mathcal{L}_{\ell_k,k,L_k-1}\cdots\mathcal{L}_{\ell_k,k,0}\, e^{\Phi(\bar\theta_{\ell_k,L_k})}[e^{\varepsilon\mathcal{K}_0}\tilde\rho_{\ell_k,\varrho}-\mathring{\rho}_{\ell_k}]$ and putting all the other error terms in $\mathcal{E}^{**}_{\ell_k,1}$. □

13. Combining many blocks: main result for the small regime

This section contains the proof of Proposition 8.13, which follows by iterating the one block estimates obtained in the previous section (i.e. Proposition 12.1); we also rely on the results detailed in Appendix A. The proof essentially follows from the next technical lemma. Before stating it let us fix and recall some notation.

Let $t \in \mathbb{R}_+$ be fixed. Given $(x,\theta)\in\mathbb{T}^2$, recall the definitions $(x_j,\theta_j) = F^j_\varepsilon(x,\theta)$, $\bar\theta^*_{\ell,k} = \bar\theta(\varepsilon k,\theta^*_\ell)$, $\theta^*_\ell = \operatorname{Re}(\mu_\ell(\theta))$, while $\bar\theta_{\ell,\varepsilon k}(x) = \bar\theta(\varepsilon k, G_\ell(x))$; finally recall that, as defined in Section 8.1, we defined $S_k = \sum_{j=0}^k L_j$ with $S_{-1}=0$ and $\varepsilon S_{R-1} = t$. Moreover, for convenience, let us define, for $0\le\bar r\le R-1$:

$$
\mathcal{S}_{\bar r,\ell_{\bar r}} = \sum_{\ell_{\bar r+1}\in\mathfrak{L}^{\bar r}_{\ell_{\bar r}}}\cdots\sum_{\ell_R\in\mathfrak{L}^{R-1}_{\ell_{R-1}}}\prod_{j=\bar r+1}^{R}\nu_{j-1,\ell_{j-1},\ell_j}.
\tag{13.1}
$$

Note that, by (8.23) and since, by hypotheses $q\ge 5$:

$$
\mu_{\ell_0}(e^{i\sigma\mathbb{A}}) = \mathcal{S}_{0,\ell_0}+\mathcal{O}(\sigma\varepsilon L_*).
\tag{13.2}
$$

Also we define, for $1\le\bar r\le R$,

$$
\Phi_{\bar r}(\theta) = \varepsilon^{-1}\int_0^{\varepsilon S_{R-1}-\varepsilon S_{\bar r-1}}\mathcal{G}_{\bar r}(\theta,s,\sigma)ds
\tag{13.3}
$$

where $\mathcal{G}_{\bar r}$ is defined in (12.2). We will also use the operators defined in (9.19), with the potentials $\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},j} = i\sigma\varpi^{\bar r}_{\ell_{\bar r},j}+\varepsilon\Phi'_{\bar r+1}(\bar\theta^*_{\ell_{\bar r},L_{\bar r}})\Xi^*_{\ell_{\bar r},j,L_{\bar r}}\hat\omega$, where $\varpi^{\bar r}_{\ell_{\bar r},j}$ is defined in (8.19) (but see (11.5) for a more convenient expression). As in the previous section we will use indifferently the notations $\mathcal{L}_{\ell_{\bar r},\bar r,j} = \mathcal{L}_{\ell_{\bar r},j} = \mathcal{L}_{\bar\theta^*_{\ell_{\bar r},j},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},j}}$ for such operators and similarly for all the corresponding related quantities. To simplify notations, let[73]

$$
\begin{aligned}
\overline{h}_{\ell_{\bar r}} &= \sum_{k=0}^{L_{\bar r}-1}\widehat{\mathcal{Q}}_{\ell_{\bar r},L_{\bar r}-1}\cdots\widehat{\mathcal{Q}}_{\ell_{\bar r},k+1}h_{\bar\theta^*_{\ell_{\bar r},k},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},k}}\\
\overline{m}_{\ell_{\bar r}} &= \sum_{k=0}^{L_{\bar r}-1}m_{\bar\theta^*_{\ell_{\bar r},k},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},k}}\widehat{\mathcal{Q}}_{\ell_{\bar r},k-1}\cdots\widehat{\mathcal{Q}}_{\ell_{\bar r},0},\\
\Gamma(\theta,\sigma) &= \operatorname{Leb} h_{\theta,i\sigma\hat\omega},
\end{aligned}
\tag{13.4}
$$

where $\widehat{\mathcal{Q}}_{\ell_{\bar r},k}$ are the operators introduced in Lemma 11.1 with the normalization specified in Lemma 11.3. Remark that, by (A.21b), (A.25) and Lemma A.6, $\Gamma\in\mathcal{C}^2$

[73] This is just a more convenient notation, limited to the present context, for the objects $\overline{h}_{0,L_k-1}$ and $\overline{m}_{0,L_k-1}$ defined in Lemma 11.3.

and

$$|\partial_\theta\Gamma| + |\partial_\theta^2\Gamma| \le C_\#|\sigma|. \tag{13.5}$$

Lemma 13.1. *There exists $\varepsilon_0 > 0$ such that, for all $\varepsilon \in (0,\varepsilon_0)$, $L \le C_\#\varepsilon^{-3\delta_*}$, $\sigma \in \mathcal{J}_1$, $q \ge 3$, $\bar r \in \{0,\cdots,R-1\}$, $R \le C_\#\varepsilon^{-1}L_*^{-1}$, and $\varepsilon\sigma^2 L(S_{R-1} - S_{\bar r-1}) \le C_\#$*

$$\mathcal{S}_{\bar r,\ell_{\bar r}} = \overline{m}_{\ell_{\bar r}}\left(e^{\Phi_{\bar r}\circ G_{\ell_{\bar r}}}\mathring{\rho}_{\ell_{\bar r}}\right)\Gamma(\bar\theta^*_{\ell_{\bar r},S_{R-1}-S_{\bar r-1}},\sigma) + \overline{\mathcal{E}}_{\bar r,\ell_{\bar r}}, \tag{13.6}$$

where $\overline{\mathcal{E}}_{\bar r,\ell_{\bar r}}$ is a remainder term satisfying the following bound:[74]

$$\begin{aligned}|\overline{\mathcal{E}}_{\bar r,\ell_{\bar r}}| \le C_\# \sum_{k=\bar r}^{R-1} e^{-(S_{R-1}-S_{k-1})\sigma^2\hat{\boldsymbol{\sigma}}_-^2} \sum_{\ell_{\bar r+1}\in\mathfrak{L}^{\bar r}_{\ell_{\bar r}}}\cdots\sum_{\ell_k\in\mathfrak{L}^{k-1}_{\ell_{k-1}}}\prod_{j=\bar r+1}^{k}|\nu_{j,\ell_j,\ell_{j+1}}|\\ \times\left(|\sigma|\varepsilon L_*^3 + \varepsilon^2 L_*^3 + \varepsilon L_*^2|\sigma|^3(R-\bar r)\right),\end{aligned}$$

with $\hat{\boldsymbol{\sigma}}_-^2 = \min_\theta \hat{\boldsymbol{\sigma}}^2(\theta)$.

Proof. Let $\Phi^+_{\bar r}(\theta) = \max_\theta \operatorname{Re}\Phi_{\bar r}(\theta)$; then (11.2b) implies that

$$\Phi^+_{\bar r} \le -c_\#\sigma^2\hat{\boldsymbol{\sigma}}_-^2(S_{R-1} - S_{\bar r-1}). \tag{13.7}$$

Next, we proceed to prove (13.6) by backward induction. The base step $\bar r = R-1$ follows from Proposition 12.1, with $\Phi = \Phi_R = 0$. Indeed,

$$\begin{aligned}\mathcal{S}_{R-1,\ell_{R-1}} &= \operatorname{Leb}\sum_{\ell_R\in\mathfrak{L}^{R-1}_{\ell_{R-1}}}\nu_{R-1,\ell_{R-1},\ell_R}\mathring{\rho}_{\ell_R}\\ &= \mathbb{X}^{**}_{\ell_{R-1}}\operatorname{Leb}\overline{h}_{\ell_{R-1}}\overline{m}_{\ell_{R-1}}\,\mathring{\rho}_{\ell_{R-1}} + \operatorname{Leb}[\mathcal{E}^{**}_{\ell_{R-1},1} + \mathcal{E}^{**}_{\ell_{R-1},2}]\\ &= \operatorname{Leb}\overline{h}_{\ell_{R-1}}\overline{m}_{\ell_{R-1}}\left(\mathbb{X}^{**}_{\ell_{R-1}}\mathring{\rho}_{\ell_{R-1}}\right) + \mathcal{O}(\varepsilon^2L^3_{R-1} + |\sigma|\varepsilon L^3_{R-1}).\end{aligned} \tag{13.8}$$

Next, using the orthogonality relations between eigenvector and the operators $\widehat{\mathcal{Q}}$:

$$\begin{aligned}\operatorname{Leb}\overline{h}_{\ell_{R-1}} &= \operatorname{Leb}h_{\bar\theta^*_{\ell_{R-1},L_{R-1}-1},\Omega^{R-1,0}_{\ell_{R-1},L_{R-1}-1}}\\ &+ \sum_{k=0}^{L_{R-1}-2}\left(\operatorname{Leb} - m_{\bar\theta^*_{\ell_{R-1},L_{R-1}-1},\Omega^{R-1,0}_{\ell_{R-1},L_{R-1}-1}}\right)\widehat{\mathcal{Q}}_{\ell_{R-1},L_{R-1}-1}\\ &\times\cdots\widehat{\mathcal{Q}}_{\ell_{R-1},k+1}\left(h_{\bar\theta^*_{\ell_{R-1},k},\Omega^{R-1,0}_{\ell_{R-1},k}} - h_{\bar\theta^*_{\ell_{R-1},k+1},\Omega^{R-1,0}_{\ell_{R-1},k+1}}\right),\end{aligned} \tag{13.9}$$

Recalling definitions (9.19) and (8.19) we see that $\Omega^{R-1,0}_{\ell_{R-1},L_{R-1}-1} = i\sigma\hat\omega$. Then, by equations (13.4),(13.9), Sub-Lemmata 11.4, 11.5 and (A.17b), we have

$$\operatorname{Leb}\overline{h}_{\ell_{R-1}} = \Gamma(\bar\theta^*_{\ell_{R-1},L_{R-1}-1},\sigma) + \mathcal{O}(|\sigma|\varepsilon). \tag{13.10}$$

In addition, by (12.2) we have

$$\mathbb{X}^{**}_{\ell_{R-1}} = \exp\left[\Phi_{R-1}(\theta^*_{\ell_{R-1}})\right] + \mathcal{O}(|\sigma|\varepsilon L_{R-1}).$$

On the other hand, since Φ_{R-1} is the integral of $\boldsymbol{\chi}(t - S_{R-2}, s, \theta^*_{\ell_{R-1}}, \bar\theta(s,\theta^*_{\ell_{R-1}}))$,[75] and since $\boldsymbol{\Omega}$ is zero-average with respect to the SRB measure (see (11.4)), we can use (A.22a) to obtain:

$$\|\partial_\theta\boldsymbol{\chi}(\sigma,T,s,\varphi,\theta)\| \le C_\#\sigma^2. \tag{13.11}$$

---

[74] We use the convention that the inner sums equal 1 when $k = \bar r$.

[75] Which is the logarithm of the maximal eigenvalue associated to the potential $i\sigma\boldsymbol{\Omega}(t - S_{R-2}, s, \theta^*_{\ell_{R-1}}, \cdot, \bar\theta(s,\theta^*_{\ell_{R-1}}))$ with respect to the dynamics $f_{\bar\theta(s,\theta^*_{\ell_{R-1}})}(\cdot)$, see (11.4) and related comments.

On the other hand, (A.19a) and Lemma A.7 similarly imply

$$\|\partial_\varphi \boldsymbol{\chi}(\sigma, T, s, \varphi, \theta)\| \leq C_\# \sigma^2. \tag{13.12}$$

The above equations yield, for any $x \in [a_{\ell_{R-1}}, b_{\ell_{R-1}}]$,

$$|\Phi_{R-1}(\theta^*_{\ell_{R-1}}) - \Phi_{R-1} \circ G_{\ell_{R-1}}(x)| \leq C_\# L_{R-1}\sigma^2\varepsilon,$$

hence

$$\mathbb{X}^{**}_{\ell_{R-1}} = \exp\left[\Phi_{R-1} \circ G_{\ell_{R-1}}\right] + \mathcal{O}(\sigma\varepsilon L_{R-1}). \tag{13.13}$$

Collecting equations (13.8), (13.10) and (13.13) proves the case $\bar{r} = R-1$.

Next, let us assume that (13.6) holds for $\bar{r}+1 \leq R-1$, then

$$\begin{aligned}\mathcal{S}_{\bar{r},\ell_{\bar{r}}} = \sum_{\ell_{\bar{r}+1}\in\mathfrak{L}^{\bar{r}}_{\ell_{\bar{r}}}} \nu_{\bar{r},\ell_{\bar{r}},\ell_{\bar{r}+1}} \mathcal{S}_{\bar{r}+1,\ell_{\bar{r}+1}} &= \sum_{\ell_{\bar{r}+1}\in\mathfrak{L}^{\bar{r}}_{\ell_{\bar{r}}}} \nu_{\bar{r},\ell_{\bar{r}},\ell_{\bar{r}+1}} \\ &\times \overline{m}_{\ell_{\bar{r}+1}}\left(\exp\left[\Phi_{\bar{r}+1}\circ G_{\ell_{\bar{r}+1}}\right]\mathring{\rho}_{\ell_{\bar{r}+1}}\right)\Gamma(\bar{\theta}^*_{\ell_{\bar{r}+1},S_{R-1}-S_{\bar{r}}},\sigma) \\ &+ \sum_{\ell_{\bar{r}+1}\in\mathfrak{L}^{\bar{r}}_{\ell_{\bar{r}}}} \nu_{\bar{r},\ell_{\bar{r}},\ell_{\bar{r}+1}} \overline{\mathcal{E}}_{\bar{r}+1,\ell_{\bar{r}+1}}.\end{aligned}$$

To continue, it is necessary to remove the dependence of $\overline{m}_{\ell_{\bar{r}+1}}$ and Γ on $\ell_{\bar{r}+1}$ in such a way that we can apply Proposition 12.1. This will be done in two steps: first notice that for (x,θ) in the support of $\mu_{\ell_{\bar{r}}}$,[76]

$$|\theta_{L_{\bar{r}}} - \bar{\theta}(\varepsilon L_{\bar{r}}, \theta^*_{\ell_{\bar{r}}})| \leq C_\# \varepsilon L_{\bar{r}}. \tag{13.14}$$

Accordingly, using (13.5),

$$|\Gamma(\bar{\theta}^*_{\ell_{\bar{r}+1},S_{R-1}-S_{\bar{r}}},\sigma) - \Gamma(\bar{\theta}^*_{\ell_{\bar{r}},S_{R-1}-S_{\bar{r}-1}},\sigma)| \leq C_\# \varepsilon|\sigma| L_{\bar{r}},$$

thus

$$\begin{aligned}\mathcal{S}_{\bar{r},\ell_{\bar{r}}} &= e^{\Phi^+_{\bar{r}+1}}\mathcal{O}(\varepsilon|\sigma|L_{\bar{r}}) + \sum_{\ell_{\bar{r}+1}\in\mathfrak{L}^{\bar{r}}_{\ell_{\bar{r}}}} \nu_{\bar{r},\ell_{\bar{r}},\ell_{\bar{r}+1}} \overline{\mathcal{E}}_{\bar{r}+1,\ell_{\bar{r}+1}} \\ &+ \sum_{\ell_{\bar{r}+1}\in\mathfrak{L}^{\bar{r}}_{\ell_{\bar{r}}}} \overline{m}_{\ell_{\bar{r}+1}}\left(\exp\left[\Phi_{\bar{r}+1}\circ G_{\ell_{\bar{r}+1}}\right]\nu_{\bar{r},\ell_{\bar{r}},\ell_{\bar{r}+1}}\mathring{\rho}_{\ell_{\bar{r}+1}}\right)\Gamma(\bar{\theta}^*_{\ell_{\bar{r}},S_{R-1}-S_{\bar{r}-1}},\sigma).\end{aligned} \tag{13.15}$$

Before continuing we need a bound on $\Phi_{\bar{r}}$.

Sub-lemma 13.2. *For any $j \in \{0, \cdots, R-1\}$ we have*

$$\begin{aligned}\|\Phi'_j\|_{\mathcal{C}^0} &\leq C_\# \sigma^2 (S_{R-1} - S_{j-1}) \\ \|\Phi'_j\|_{\mathcal{C}^1} &\leq C_\# |\sigma| (S_{R-1} - S_{j-1}).\end{aligned}$$

Proof. By equations (13.12) and (13.11) it follows

$$\begin{aligned}|\Phi'_j(\theta)| \leq C_\#\varepsilon^{-1}&\int_0^{\varepsilon S_{R-1}-\varepsilon S_{j-1}} |\partial_\theta \boldsymbol{\chi}(\sigma, t-s-\varepsilon S_{j-1}, \bar{\theta}(s,\theta), \bar{\theta}(s,\theta))|ds \\ &+ C_\#\varepsilon^{-1}\int_0^{\varepsilon S_{R-1}-\varepsilon S_{j-1}} |\partial_\varphi \boldsymbol{\chi}(\sigma, t-s-\varepsilon S_{j-1}, \bar{\theta}(s,\theta), \bar{\theta}(s,\theta))|ds \\ &C_\#\sigma^2(S_{R-1}-S_{j-1}).\end{aligned}$$

This proves the first inequality, the second is obtained similarly by using the above formulae, (A.19) and Lemma A.9 (in particular (A.22d)). □

[76] In fact, using large deviations, it is possible to have a better estimate with large probability. We will not push this possibility as it is not needed for the level of precision we are currently after.

Note that Lemma 11.1(b) implies that $\Phi^+_{\bar r} \le -C_\# \sigma^2 (R-\bar r) L_*$. Moreover, Sub-Lemma 13.2, together with our hypotheses on $\bar r$, implies that the hypotheses of Proposition 9.7, Lemma 11.3 and Lemma 12.1, are all satisfied for $\Phi = \Phi_{\bar r}$. We can therefore apply all such results to the present situation.

Observe, moreover, that Sub-Lemma 13.2 and the definition of $\Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},k}$ imply:

$$\begin{aligned}\|\Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},k} - \Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},k-1}\|_{\mathcal{C}^1} &\le C_\#(\varepsilon|\sigma| + \varepsilon^2\sigma^2 L_*(R-\bar r))\\ &\le C_\#\varepsilon|\sigma|,\end{aligned}$$

where we used the fact that by definition $(R-\bar r) < R = \varepsilon^{-1}L_*^{-1}$ and that since $\sigma \in \mathcal{J}_1$ we have $\sigma^2 < |\sigma|$. Similarly, using (11.3) we have

$$\begin{aligned}\|\Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},0} - \Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},L_{\bar r}-1}\|_{\mathcal{C}^1} &\le |\sigma| \left|\widehat\Xi(t-\varepsilon S_{\bar r}, \theta^*_{\ell_{\bar r+1}}) - \widehat\Xi(t-\varepsilon S_{\bar r}, \bar\theta^*_{\ell_{\bar r},L_{\bar r}})\right| \|\hat\omega\|_{\mathcal{C}^1}\\ &\quad + C_\#(\varepsilon^2|\sigma|L_* + \varepsilon^2|\sigma|L_*^2(R-\bar r))\\ &\le C_\#(\varepsilon|\sigma|L_* + \varepsilon^2|\sigma|L_*^2(R-\bar r))\\ &\le C_\#\varepsilon|\sigma|L_*\end{aligned}$$

We can now take care of $\overline{m}_{\ell_{\bar r+1}}$: observe that, by applying (A.20b) and recalling footnote 69 and Lemma A.15

$$\begin{aligned}\overline{m}_{\ell_{\bar r+1}}(\varphi) &= m_{\bar\theta^*_{\ell_{\bar r+1},0},\Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},0}}(\varphi)\\ &\quad+\sum_{k=1}^{L_{\bar r+1}-1}\left(m_{\bar\theta^*_{\ell_{\bar r+1},k},\Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},k}} - m_{\bar\theta^*_{\ell_{\bar r+1},k-1},\Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},k-1}}\right)\left(\widehat{\mathcal{Q}}_{\ell_{\bar r+1},k-1}\cdots\widehat{\mathcal{Q}}_{\ell_{\bar r+1},0}\varphi\right)\\ &= m_{\bar\theta^*_{\ell_{\bar r+1},0},\Omega^{\bar r+1,\Phi_{\bar r+2}}_{\ell_{\bar r+1},0}}(\varphi) + \mathcal{O}(\varepsilon|\sigma|(\log\sigma)^2)\|\varphi\|_{\mathrm{BV}};\end{aligned}$$

and applying once again (A.20b) and Lemma A.15 together with (13.14)

$$\overline{m}_{\ell_{\bar r+1}}(\varphi) = m_{\bar\theta^*_{\ell_{\bar r},L_{\bar r}-1},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},L_{\bar r}-1}}(\varphi) + \mathcal{O}\big(\varepsilon|\sigma|(\log\sigma)^2 L_*\big)\|\varphi\|_{\mathrm{BV}}. \tag{13.16}$$

We can now continue with the estimate we left at (13.15) and obtain

$$\begin{aligned}\mathcal{S}_{\bar r,\ell_{\bar r}} &= e^{\Phi^+_{\bar r+1}}\mathcal{O}(\varepsilon|\sigma|(\log\sigma)^2 L_{\bar r}) + \sum_{\ell_{\bar r+1}\in\mathfrak{L}^{\bar r}_{\ell_{\bar r}}} \mathsf{v}_{\bar r,\ell_{\bar r},\ell_{\bar r+1}}\overline{\mathcal{E}}_{\bar r+1,\ell_{\bar r+1}}\\ &\quad + m_{\bar\theta^*_{\ell_{\bar r},L_{\bar r}-1},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},L_{\bar r}-1}}\left(\sum_{\ell_{\bar r+1}\in\mathfrak{L}^{\bar r}_{\ell_{\bar r}}} e^{\Phi_{\bar r+1}\circ G_{\ell_{\bar r+1}}}\mathsf{v}_{\bar r,\ell_{\bar r},\ell_{\bar r+1}}\mathring{\rho}_{\ell_{\bar r+1}}\right)\\ &\quad\times\Gamma(\bar\theta^*_{\ell_{\bar r},S_{R-1}-S_{\bar r-1}},\sigma).\end{aligned}$$

Finally we can apply Proposition 12.1 with $\bar s = S_{R-1} - S_{\bar r-1}$:

$$\begin{aligned}\mathcal{S}_{\bar r,\ell_{\bar r}} =& e^{\Phi^+_{\bar r+1}}\mathcal{O}\left(|\sigma|(\log\sigma)^2\varepsilon L_{\bar r}\right) + \sum_{\ell_{\bar r+1}\in\mathfrak{L}^{\bar r}_{\ell_{\bar r}}} \mathsf{v}_{\bar r,\ell_{\bar r},\ell_{\bar r+1}}\overline{\mathcal{E}}_{\bar r+1,\ell_{\bar r+1}}\\ &+ e^{\varepsilon^{-1}\int_0^{\varepsilon L_{\bar r}}\mathcal{G}_{\bar r}(\theta^*_{\ell_{\bar r}},s,\sigma)ds}\Gamma(\bar\theta^*_{\ell_{\bar r},S_{R-1}-S_{\bar r-1}},\sigma)\\ &\times m_{\bar\theta^*_{\ell_{\bar r},L_{\bar r}-1},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},L_{\bar r}-1}}(\overline{h}_{\ell_{\bar r}})\overline{m}_{\ell_{\bar r}}\left(e^{\Phi_{\bar r+1}(\bar\theta_{\ell_{\bar r},L_{\bar r}})}\mathring{\rho}_{\ell_{\bar r}}\right)\\ &+\mathcal{O}\left(m_{\theta^*_{\ell_{\bar r}},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},L_{\bar r}-1}}(\mathcal{E}^{**}_{\ell_{\bar r},1})\right) + \mathcal{O}\left(m_{\theta^*_{\ell_{\bar r}},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},L_{\bar r}-1}}(\mathcal{E}^{**}_{\ell_{\bar r},2})\right),\end{aligned}$$

where $\mathcal{G}_{\bar r}$ is defined in (12.2). Observe that, by definition:

$$\begin{aligned}\Phi_{\bar r+1}(\bar\theta_{\ell_{\bar r},L_{\bar r}}) &= \varepsilon^{-1}\int_0^{\varepsilon(S_{R-1}-S_{\bar r})} \boldsymbol{\chi}(\sigma, t-s-\varepsilon S_{\bar r}),0,\bar\theta(s,\bar\theta_{\ell_{\bar r},L_{\bar r}}),\bar\theta(s,\bar\theta_{\ell_{\bar r},L_{\bar r}})ds\\ &= \varepsilon^{-1}\int_{\varepsilon L_{\bar r}}^{\varepsilon(S_{R-1}-S_{\bar r-1})} \boldsymbol{\chi}(\sigma, t-s'-\varepsilon S_{\bar r-1},0,\bar\theta(s',G_{\ell_{\bar r}}),\bar\theta(s',G_{\ell_{\bar r}}))ds'\end{aligned}$$

Hence, using 11.1(b), we can write

$$\begin{aligned}&\Phi_{\bar r+1}(\bar\theta_{\ell_{\bar r},L_{\bar r}}) + \varepsilon^{-1}\int_0^{\varepsilon L_{\bar r}} \mathcal{G}_{\bar r}(\theta^*_{\ell_{\bar r}},s,\sigma)ds\\ &= \varepsilon^{-1}\int_{\varepsilon L_{\bar r}}^{\varepsilon(S_{R-1}-S_{\bar r-1})} \boldsymbol{\chi}(\sigma,t-s-\varepsilon S_{\bar r-1},0,\bar\theta(s,G_{\ell_{\bar r}}),\bar\theta(s,G_{\ell_{\bar r}}))ds\\ &\quad+\varepsilon^{-1}\int_0^{\varepsilon L_{\bar r}} \boldsymbol{\chi}(\sigma,t-s-\varepsilon S_{\bar r-1},0,\bar\theta(s,G_{\ell_{\bar r}}),\bar\theta(s,G_{\ell_{\bar r}}))ds + \mathcal{O}(\varepsilon|\sigma|L_{\bar r})\\ &= \Phi_{\bar r}\circ G_{\ell_{\bar r}} + \mathcal{O}(\varepsilon|\sigma|L_*).\end{aligned}$$

Finally, recalling (13.4),

$$m_{\bar\theta^*_{\ell_{\bar r},L_{\bar r}-1},\Omega^{\bar r,\Phi_{\bar r+1}}_{\ell_{\bar r},L_{\bar r}-1}}(\overline{h}_{\ell_{\bar r}}) = 1$$

from which the lemma readily follows by using Lemma A.14 (observe[77] that $\sigma^{100} < \sigma\varepsilon$ since $\delta_* > 1/99$ and $|\sigma| < \varepsilon^{\delta_*}$) and the bounds on $\mathcal{E}^{**}_{\ell_{\bar r},i}$ provided in Proposition 12.1. □

We are now, finally, ready to prove the very last missing piece in our argument.

Proof of Proposition 8.13. The basic idea is to apply Lemma 13.1. Unfortunately, Lemma 13.1 holds only under the additional hypothesis $\varepsilon\sigma^2 L_*(S_{R-1} - S_{\bar r-1}) \le C_\#$. Note that if $|\sigma| \le \varepsilon^{2\delta_*}$, then

$$\varepsilon\sigma^2 L_*(S_{R-1} - S_{\bar r-1}) \le C_\#\varepsilon^{4\delta_*-3\delta_*} \le C_\#\varepsilon^{\delta_*}.$$

Yet, if $|\sigma| \in [\varepsilon^{2\delta_*},\varepsilon^{\delta_*}]$, we can apply Lemma 13.1 only for

$$(S_{R-1} - S_{\bar r-1}) \le C_\#\varepsilon^{-1+\delta_*}. \tag{13.17}$$

So, choose $\bar r_\varepsilon$ such that $S_{R-1} - S_{\bar r_\varepsilon-1} = C_\#\varepsilon^{-1+\delta_*}$. Then, for $|\sigma| \in [\varepsilon^{2\delta_*},\varepsilon^{\delta_*}]$, we can rewrite (8.23), with $q \ge 5$, and (13.1) as follows

$$\mu_{\ell_0}(e^{i\sigma\mathbb{A}}) = \sum_{\ell_1\in\mathfrak{L}^0_{\ell_0}}\cdots\sum_{\ell_{\bar r_\varepsilon}\in\mathfrak{L}^{\bar r_\varepsilon-1}_{\ell_{\bar r_\varepsilon-1}}}\prod_{j=1}^{\bar r_\varepsilon}\mathsf{v}_{j-1,\ell_{j-1},\ell_j}\mathcal{S}_{\bar r_\varepsilon,\ell_{\bar r_\varepsilon}} + \mathcal{O}(\varepsilon\sigma L_*).$$

Hence, by (8.22), we can bound

$$\left|\mu_{\ell_0}(e^{i\sigma\mathbb{A}})\right| \le \sup_{\ell_{\bar r_\varepsilon}}\left|\mathcal{S}_{\bar r_\varepsilon,\ell_{\bar r_\varepsilon}}\right| + C_\#|\sigma|\varepsilon L_*.$$

We can now apply Lemma 13.1 and (11.2) to write

$$\left|\mathcal{S}_{\bar r_\varepsilon,\ell_{\bar r_\varepsilon}}\right| \le C_\# e^{-c_\#(S_{R-1}-S_{\bar r_\varepsilon-1})\sigma^2} + |\overline{\mathcal{E}}_{\bar r_\varepsilon,\ell_{\bar r_\varepsilon}}| \le C_\# e^{-c_\#\varepsilon^{-1+3\delta_*}} + |\overline{\mathcal{E}}_{0,\ell_0}|.$$

Collecting the above facts yields

$$\begin{aligned}|\mu_{\ell_0}(e^{i\sigma\mathbb{A}})| &\le \overline{\mathcal{E}}_{0,\ell_0} + \mathcal{O}\left(\varepsilon\sigma L_* + e^{-c_\#\varepsilon^{-1+3\delta_*}}\right)\\ &\le \overline{\mathcal{E}}_{0,\ell_0} + \mathcal{O}\left(\varepsilon\sigma L_*\right).\end{aligned}$$

---

[77] Since the choice of the power 100 in Lemma A.14 is arbitrary (see Footnote 85), one could in principle work with values of δ_* smaller than 1/99, if needed.

In particular, since for $|\sigma| \in [\varepsilon^{2\delta_*}, \varepsilon^{\delta_*}]$:

$$\exp\left[-\frac{\sigma^2}{2\varepsilon}\boldsymbol{\sigma}_t^2(\theta_0)\right] \le C_\# e^{-c_\#\varepsilon^{-1+4\delta_*}} \le C_\#\varepsilon|\sigma|L_*,$$

we have:

$$\mu_{\ell_0}(e^{i\sigma\mathbb{A}}) = \exp\left[-\frac{\sigma^2}{2\varepsilon}\boldsymbol{\sigma}_t^2(\theta_0)\right] + \overline{\mathcal{E}}_{0,\ell_0} + \mathcal{O}\left(\varepsilon\sigma L_*\right). \tag{13.18}$$

Next, we consider the case $|\sigma| \le \varepsilon^{2\delta_*}$, hence Lemma 13.1 holds with $\bar{r} = 0$. Accordingly, since $q \ge 5$, we can apply (13.2) that implies:

$$\mu_{\ell_0}(e^{i\sigma\mathbb{A}}) = \overline{m}_{\ell_0}(e^{\Phi_0\circ G_{\ell_0}}\mathring{\rho}_{\ell_0})\Gamma(\bar{\theta}^*_{\ell_0,t\varepsilon^{-1}},\sigma) + \overline{\mathcal{E}}_{0,\ell_0} + \mathcal{O}(\sigma\varepsilon L_*).$$

Note that definition (13.3) and equations (13.11), (13.12), using 13.2, give

$$\|\Phi_0\circ G_{\ell_0} - \Phi_0(\theta^*_{\ell_0})\|_{\mathcal{C}^1} \le C_\#\sigma^2 \le C_\#\varepsilon^{2\delta_*}$$

and, by (11.2) and recalling definitions (8.7) and (2.20),

$$\begin{aligned}\varepsilon\Phi_0(\theta^*_{\ell_0}) &= -\frac{\sigma^2}{2}\int_0^t \widehat{\Xi}(t-\varepsilon-s,\bar{\theta}(s+\varepsilon,\theta^*_{\ell_0}))^2\hat{\boldsymbol{\sigma}}^2(\bar{\theta}(s,\theta^*_{\ell_0}))ds + \mathcal{O}(\sigma^3)\\ &= -\frac{\sigma^2}{2}\boldsymbol{\sigma}_t^2(\theta_0) + \mathcal{O}(\sigma^3+\sigma^2\varepsilon).\end{aligned}$$

In addition, by computations similar to (13.16) and using Lemma A.14, we have

$$\overline{m}_{\ell_0}\mathring{\rho}_{\ell_0} = m_{\theta^*_{\ell_0},\Omega^{0,\Phi_0}_{\ell_0,L_0-1}}(\mathring{\rho}_{\ell_0}) + \mathcal{O}(\varepsilon\sigma(\log\sigma)^2L_*) = \text{Leb}(\mathring{\rho}_{\ell_0}) + \mathcal{O}(\sigma\log|\sigma|^{-1}).$$

Also, recalling the definition (13.4) and using Lemma A.6 we have

$$\Gamma(\bar{\theta}^*_{\ell_0,t\varepsilon^{-1}},\sigma) = 1 + \mathcal{O}(\sigma).$$

Accordingly, recalling (13.18), for all $\sigma \in \mathcal{J}_1$ we have

$$\mu_{\ell_0}(e^{i\sigma\mathbb{A}}) = \exp\left[-\frac{\sigma^2}{2\varepsilon}\boldsymbol{\sigma}_t^2(\theta_0)\right]\left(1+\mathcal{O}(\sigma^3\varepsilon^{-1}+\sigma^2+\sigma\log|\sigma|^{-1})\right) + \overline{\mathcal{E}}_{0,\ell_0} + \mathcal{O}(\sigma\varepsilon L_*).$$

Note that

$$\begin{aligned}&\frac{1}{2\pi\varepsilon}\int_{\mathcal{J}_1}\exp\left[-\frac{\sigma^2}{2\varepsilon}\boldsymbol{\sigma}_t^2(\theta_0)\right](\sigma^3\varepsilon^{-1}+\sigma^2+\sigma\log|\sigma|^{-1})d\sigma\\ &=\frac{1}{2\pi}\int_{\mathbb{R}}\exp\left[-\frac{\eta^2}{2}\boldsymbol{\sigma}_t^2(\theta_0)\right](\eta^3+\varepsilon^{1/2}\eta^2+\eta\log|\eta\varepsilon^{1/2}|^{-1})d\eta = \mathcal{O}(\log\varepsilon^{-1}).\end{aligned}$$

To conclude the proof of the proposition it then suffices to estimate the integral of $\overline{\mathcal{E}}_{0,\ell_0}$. This is easily done by noting that, for all $p,q\in\mathbb{N}$,

$$\begin{aligned}\frac{1}{2\pi\varepsilon}\int_{\mathcal{J}_1}\sum_{k=0}^{R-1}e^{-c_\#\sigma^2(R-k)L}(R-k)^qL^q|\sigma|^pd\sigma &\le \frac{C_\#}{\varepsilon}\sum_{k=1}^R\int_{\mathbb{R}}e^{-c_\#\eta^2}\frac{|\eta|^p}{(kL)^{(p+1)/2-q}}d\eta\\ &\le \frac{C_\#}{\varepsilon L^{(p+1)/2-q}}\sum_{k=1}^R k^{-(p+1)/2+q}.\end{aligned}$$

Thus

$$\int_{\mathcal{J}_1}\sum_{k=0}^{R-1}\frac{|\sigma|^p(R-k)^qL^q}{2\pi\varepsilon\exp[c_\#\sigma^2(R-k)L]}d\sigma \le \begin{cases}C_\#\varepsilon^{-2-q+(p+1)/2}L^{-1} & \text{for } p<1+2q\\ C_\#\varepsilon^{-1}L^{-1}\log(\varepsilon^{-1}) & \text{for } p=1+2q\\ C_\#\varepsilon^{-1}L^{q-(p+1)/2} & \text{for } p>1+2q.\end{cases}$$

We can now apply the above estimates to compute the integrals of the various contributions to $\overline{\mathcal{E}}_{0,\ell_0}$ obtaining

$$\frac{1}{2\pi\varepsilon}\int_{\mathcal{J}_1}\sum_{k=0}^{R-1} e^{-c_\# \sigma^2(R-k)L}[\sigma\varepsilon L_*^3]d\sigma = \mathcal{O}(L_*^2\log\varepsilon^{-1}) \qquad (p,q)=(1,0)$$
$$\frac{1}{2\pi\varepsilon}\int_{\mathcal{J}_1}\sum_{k=0}^{R-1} e^{-c_\# \sigma^2(R-k)L}[\varepsilon^2 L_*^3]d\sigma = \mathcal{O}(\varepsilon^{1/2}L_*^2) \qquad (p,q)=(0,0)$$
$$\frac{1}{2\pi\varepsilon}\int_{\mathcal{J}_1}\sum_{k=0}^{R-1} e^{-c_\# \sigma^2(R-k)L}[\sigma^3\varepsilon L_*^2(R-\bar r)]d\sigma = \mathcal{O}(\log\varepsilon^{-1}) \qquad (p,q)=(3,1)$$

which prove the proposition. □

Appendix A. Spectral theory for transfer operators: a toolbox

In this appendix we collect some known and less known (or possibly unknown) results on transfer operators that are used in the main part of this paper. Let us fix $r \geq 3$ and let $f(\cdot,\theta) = f_\theta \in \mathcal{C}^r(\mathbb{T},\mathbb{T})$, with $\theta \in \mathbb{T}$, be a one parameter family of orientation preserving expanding maps (i.e., there exists $\lambda > 1$ such that $\inf_{x,\theta} f'_\theta(x) \geq \lambda$). Let $\Omega(\cdot,\theta) = \Omega_\theta \in \mathcal{C}^r(\mathbb{T},\mathbb{C})$ be a family of *potentials*. We further assume some regularity[78] in θ; more precisely we require that $f \in \mathcal{C}^r(\mathbb{T}^2,\mathbb{T})$ and that $\partial_x\Omega \in \mathcal{C}^{r-1}(\mathbb{T}^2,\mathbb{C})$. For any $\varsigma \in \mathbb{R}$ we can then consider the family of operators $\mathcal{L}_{\theta,\varsigma\Omega}$ defined as:

$$[\mathcal{L}_{\theta,\varsigma\Omega}\, g](x) = \sum_{y\in f_\theta^{-1}(x)} \frac{e^{\varsigma\Omega_\theta(y)}}{f'_\theta(y)} g(y). \tag{A.1}$$

It is well know that the spectrum of such operators depends drastically on the space on which they act. We will be interested in BV, $W^{s,1}$ and $\mathcal{C}^s$, for $s \leq r-1$.

Remark. *We will use $\mathcal{C}^s$ (and similarly for the other spaces) as a shorthand notation for $\mathcal{C}^s(\mathbb{T})$ (which in turn is a shorthand notation for $\mathcal{C}^s(\mathbb{T},\mathbb{R})$). When we need to consider functions defined on $\mathbb{T}^2$ we will write explicitly $\mathcal{C}^s(\mathbb{T}^2)$.*

A.1. General facts.

Let us start with a useful result for the case of real potentials.

Lemma A.1. *If Ω is real, then for any $\varsigma \in \mathbb{R}, \theta \in \mathbb{T}$, the operator $\mathcal{L}_{\theta,\varsigma\Omega} : \mathcal{C}^1 \to \mathcal{C}^1$ is of Perron–Frobenius type. That is, it has a simple maximal eigenvalue $e^{\chi_{\theta,\varsigma\Omega}}$ with left and right eigenvectors that we denote with $m_{\theta,\varsigma\Omega}$ and $h_{\theta,\varsigma\Omega}$ (respectively), normalized so that $m_{\theta,\varsigma\Omega}(h_{\theta,\varsigma\Omega}) = 1$.*

In addition, $m_{\theta,\varsigma\Omega}$ is a positive measure; $h_{\theta,\varsigma\Omega} > 0$ and

$$\|(\log h_{\theta,\varsigma\Omega})'\|_\infty \leq C_\#(|\varsigma|\|\Omega'_\theta\|_\infty + 1).$$

Also, the spectral gap is continuous in ς, θ and the leading eigenvalue and eigenprojector are analytic in ς and differentiable in θ.

Proof. The statement could be proven by reducing the system to symbolic dynamics and then using results on the induced transfer operator. Yet, a much more efficient and direct proof can be obtained by the Hilbert metric technique used, e.g., in [41, Section 2] or [42]. Namely, consider the cone $\mathcal{K}_a = \{g \in \mathcal{C}^1 \;:\; |g'(x)| \leq ag(x), \forall x \in \mathbb{T}\}$. Since

$$\frac{d}{dx}\mathcal{L}_{\theta,\varsigma\Omega} g = \mathcal{L}_{\theta,\varsigma\Omega}\left(\frac{g'}{f'_\theta} + \varsigma\frac{g\Omega'_\theta}{f'_\theta} - \frac{g f''_\theta}{(f'_\theta)^2}\right), \tag{A.2}$$

[78] The requirements on regularity are not optimal, but rather reflect our case of interest.

it follows that

$$\left|\frac{d}{dx}\mathcal{L}_{\theta,\varsigma\Omega}g\right| \leq \lambda^{-1}\{a + |\varsigma|\|\Omega_\theta'\|_\infty + \lambda C_\#\}\mathcal{L}_{\theta,\varsigma\Omega}g.$$

Hence, for any $\varrho \in (\lambda^{-1}, 1)$, $\mathcal{L}_{\theta,\varsigma\Omega}\mathcal{K}_a \subset \mathcal{K}_{\varrho a}$ provided

$$a \geq (\varrho\lambda - 1)^{-1}(|\varsigma|\|\Omega_\theta'\|_\infty + \lambda C_\#). \tag{A.3}$$

A simple computation shows that the diameter Δ, computed in the Hilbert metric Θ determined by $\mathcal{K}_a$, of the image is bounded by $2\frac{1+\varrho}{1-\varrho}e^a$.[79] From this fact and Birkhoff Theorem [41, Theorem 1.1] it follows that $\mathcal{L}_{\theta,\varsigma\Omega}$ contracts the Hilbert metric by a factor $\tanh\frac{\Delta}{4}$. Also, notice that if $f-g, f+g \in \mathcal{K}_a$, then $\|f\|_{L^\infty} \geq \|g\|_{L^\infty}$. Accordingly, [41, Lemma 1.3] implies that, if $\|f\|_{L^\infty} = \|g\|_{L^\infty}$, then

$$\|f-g\|_{L^\infty} \leq \left(e^{\Theta(f,g)} - 1\right)\|f\|_{L^\infty}.$$

Next, let $f, g \in \mathcal{K}_a$ with $\mathrm{Leb}(f) = \mathrm{Leb}(g) = 1$. Then $e^{-a} \leq f, g \leq e^a$ hence $e^{-a}\mathcal{L}^n_{\theta,\varsigma\Omega}1 \leq \mathcal{L}^n_{\theta,\varsigma\Omega}f, \mathcal{L}^n_{\theta,\varsigma\Omega}g \leq e^a\mathcal{L}^n_{\theta,\varsigma\Omega}1$. This means that, for any $n \in \mathbb{N}$, there exists $\alpha_n \in [e^{-a}, e^a]$ such that

$$\begin{aligned}\|\mathcal{L}^n_{\theta,\varsigma\Omega}f - \alpha_n\mathcal{L}^n_{\theta,\varsigma\Omega}g\|_{L^\infty} &\leq C_\#\Theta(\mathcal{L}^n_{\theta,\varsigma\Omega}f, \mathcal{L}^n_{\theta,\varsigma\Omega}g)\|\mathcal{L}^n_{\theta,\varsigma\Omega}1\|_{L^\infty}\\ &\leq C_\#\Delta\left[\tanh\frac{\Delta}{4}\right]^n\|\mathcal{L}^n_{\theta,\varsigma\Omega}1\|_{L^\infty}.\end{aligned} \tag{A.4}$$

Let $e^{\chi_{\theta,\varsigma\Omega}}$, be the maximal eigenvalue of $\mathcal{L}_{\theta,\varsigma\Omega}$ when acting on $\mathcal{C}^1$. The above displayed equations, together with (A.2), imply that $\mathcal{L}_{\theta,\varsigma\Omega}$, when acting on $\mathcal{C}^1$, has a simple maximal eigenvalue and a spectral gap of size at least $e^{\chi_{\theta,\varsigma\Omega}}\left(1 - \tanh\frac{\Delta}{4}\right)$.

Accordingly, there exists $h_{\theta,\varsigma\Omega} \in \mathcal{C}^1$ and a distribution $m_{\theta,\varsigma\Omega} \in (\mathcal{C}^1)'$ such that

$$\mathcal{L}_{\theta,\varsigma\Omega}(g) = e^{\chi_{\theta,\varsigma\Omega}}h_{\theta,\varsigma\Omega}\, m_{\theta,\varsigma\Omega}(g) + \mathcal{Q}_{\theta,\varsigma\Omega}(g) =: e^{\chi_{\theta,\varsigma\Omega}}\mathcal{P}_{\theta,\varsigma\Omega}(g) + \mathcal{Q}_{\theta,\varsigma\Omega}(g),$$

where, for any $n \in \mathbb{N}$, $\|\mathcal{Q}^n_{\theta,\varsigma\Omega}\|_{\mathcal{C}^1\to\mathcal{C}^1} \leq C_{\theta,\varsigma\Omega}e^{\chi_{\theta,\varsigma\Omega}n}\tau^n_{\theta,\varsigma\Omega}$, with $\tau_{\theta,\varsigma\Omega} \in (0, 1 - \tanh\frac{\Delta}{4}]$, $\mathcal{Q}_{\theta,\varsigma\Omega}\mathcal{P}_{\theta,\varsigma\Omega} = \mathcal{P}_{\theta,\varsigma\Omega}\mathcal{Q}_{\theta,\varsigma\Omega} = 0$ and $m_{\theta,\varsigma\Omega}(h_{\theta,\varsigma\Omega}) = 1$. Moreover, by standard perturbation theory all the above quantities are analytic in ς.

We now show that $m_{\theta,\varsigma\Omega}$ is not just an element of $(\mathcal{C}^1)'$, as follows automatically from the general theory, but indeed a measure (i.e. an element of $(\mathcal{C}^0)'$). We have seen that

$$e^a\mathrm{Leb}(h_{\theta,\varsigma\Omega}) \geq h_{\theta,\varsigma\Omega} = e^{-n\chi_{\theta,\varsigma\Omega}}\mathcal{L}^n_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega} \geq e^{-a}e^{-n\chi_{\theta,\varsigma\Omega}}\mathcal{L}^n_{\theta,\varsigma\Omega}1.$$

Thus, for any $n \in \mathbb{N}$ and $g \in \mathcal{C}^1$, $g \geq 0$,

$$0 \leq e^{-n\chi_{\theta,\varsigma\Omega}}\mathcal{L}^n_{\theta,\varsigma\Omega}g = h_{\theta,\varsigma\Omega}m_{\theta,\varsigma\Omega}(g) + C_{\theta,\varsigma\Omega}\tau^n_{\theta,\varsigma\Omega}\|g\|_{\mathcal{C}^1}$$

which shows that $m_{\theta,\varsigma\Omega}$ is a positive functional and hence a measure.

Finally, the perturbation theory in [29, Section 8] implies that $\chi_{\theta,\varsigma\Omega}$ and $\mathcal{P}_{\theta,\varsigma\Omega} = h_{\theta,\varsigma\Omega} \otimes m_{\theta,\varsigma\Omega}$ are differentiable in θ (the latter with respect to the $L(\mathcal{C}^2, \mathcal{C}^0)$ topology) and that $C_{\theta,\varsigma\Omega}, \tau_{\theta,\varsigma\Omega}$ can be chosen to be continuous in θ.[80] Indeed, a direct computation shows that, setting

$$\mathcal{D}_{\theta,\varsigma\Omega}(g) = -\left[\frac{\partial_\theta f_\theta}{f_\theta'}g\right]' + \varsigma\left[\partial_\theta\Omega_\theta - \frac{\partial_\theta f_\theta}{f_\theta'}\Omega_\theta'\right]g, \tag{A.5}$$

[79] It suffices to compute the distance of a function from the constant function and recall that, for $f, g \in \mathcal{K}_a$, $\Theta(f,g)$ is defined as $\log\frac{\mu}{\lambda}$ where μ is the inf of the α such that $\alpha f - g \in \mathcal{K}_a$ and λ is the sup of the β such that $f - \beta g \in \mathcal{K}_a$.

[80] The Banach spaces $\mathcal{B}^i$ in [29, Section 8] here are taken to be $\mathcal{C}^i$.

we have

$$\mathcal{L}_{\theta+s,\varsigma\Omega} = \mathcal{L}_{\theta,\varsigma\Omega} + \int_\theta^{\theta+s} \mathcal{L}_{\varphi,\varsigma\Omega}\mathcal{D}_{\varphi,\varsigma\Omega}d\varphi = \mathcal{L}_{\theta,\varsigma\Omega}(\mathbb{1} + s\mathcal{D}_{\theta,\varsigma\Omega}) + \frac{s^2}{2}R_{\theta,s,\varsigma\Omega}, \tag{A.6}$$

where $R_{\theta,s,\varsigma\Omega} = \frac{2}{s^2}\left[\mathcal{L}_{\theta+s,\varsigma\Omega} - \mathcal{L}_{\theta,\varsigma\Omega}(\mathbb{1} - s\mathcal{D}_{\theta,\varsigma\Omega})\right]$. Moreover, for any $0 < k \le r-2$ we have:[81]

$$\begin{aligned}
&\|\mathcal{L}^n_{\theta,\varsigma\Omega}\|_{\mathcal{C}^k} \le C_{\theta,\varsigma\Omega}e^{n\chi_{\theta,\varsigma\Omega}}\\
&\|\mathcal{L}_{\theta+s,\varsigma\Omega} - \mathcal{L}_{\theta,\varsigma\Omega}\|_{\mathcal{C}^{k+1}\to\mathcal{C}^k} \le |s| \sup_{\theta'\in[\theta,\theta+s]} \|\mathcal{L}_{\theta',\varsigma\Omega}\|_{\mathcal{C}^{k+1}}\|\mathcal{D}_{\theta',\varsigma\Omega}\|_{\mathcal{C}^{k+1}\to\mathcal{C}^k}\\
&\qquad\qquad\qquad\qquad\qquad\le C_{\theta,\varsigma\Omega}|s|(1 + \|\Omega_\theta\|_{\mathcal{C}^{k+1}} + \|\partial_\theta\Omega_\theta\|_{\mathcal{C}^k})\\
&\|R_{\theta,s,\varsigma\Omega}\|_{\mathcal{C}^{k+2}\to\mathcal{C}^k} \le C_{\theta,\varsigma\Omega}(1 + \|\Omega_\theta\|_{\mathcal{C}^{k+2}} + \|\partial_\theta\Omega_\theta\|_{\mathcal{C}^{k+1}} + \|\partial^2_\theta\Omega_\theta\|_{\mathcal{C}^k}).
\end{aligned} \tag{A.7}$$

Hence the hypotheses of [29, Theorem 8.1] are satisfied and the resolvent $\mathbb{1}z - \mathcal{L}_{\theta,\varsigma\Omega}$, viewed as an operator from $\mathcal{C}^2$ to $\mathcal{C}^0$, is differentiable in θ. This implies the same for all spectral data, since they can be recovered by integrating the resolvent over the complex plane. □

In the case of arbitrary complex potentials it is also possible to obtain information on the spectrum, as described in the following result.

Lemma A.2. *For $\varsigma \in \mathbb{R}, \theta \in \mathbb{T}$, let $e^{\tau_{\theta,\varsigma\Omega}}$ be the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$ as an element of $L(\mathcal{C}^0, \mathcal{C}^0)$. Then the spectral radius $e^{\chi_{\theta,\varsigma\Omega}}$ of $\mathcal{L}_{\theta,\varsigma\Omega}$ as an element of $L(\mathcal{C}^1, \mathcal{C}^1)$ is bounded by $e^{\tau_{\theta,\varsigma\Omega}}$. In addition, the essential spectral radius is bounded by $\lambda^{-1}e^{\tau_{\theta,\varsigma\Omega}}$. Finally, the spectrum outside the disk of radius $\lambda^{-1}e^{\tau_{\theta,\varsigma\Omega}}$ is the same when $\mathcal{L}_{\theta,\varsigma\Omega}$ acts on all $\mathcal{C}^k$, $k \in \{1, \cdots, r-1\}$.*

Proof. Note that the computation yielding (A.2) also holds for any power f^n_θ; this gives:

$$\frac{d}{dx}\mathcal{L}^n_{\theta,\varsigma\Omega}g = \mathcal{L}^n_{\theta,\varsigma\Omega}\left(\frac{g'}{(f^n_\theta)'} + \varsigma\frac{g\Omega'_{\theta,n}}{(f^n_\theta)'} - \frac{g(f^n_\theta)''}{[(f^n_\theta)']^2}\right) \tag{A.8}$$

where $\Omega_{\theta,n} = \sum_{k=0}^{n-1}\Omega_\theta \circ f^k_\theta$. Then a direct computation yields

$$\|\mathcal{L}^n_{\theta,\varsigma\Omega}g\|_{\mathcal{C}^1} \le \|\mathcal{L}^n_{\theta,\varsigma\Omega}\|_{\mathcal{C}^0\to\mathcal{C}^0}\left[\lambda^{-n}\|g\|_{\mathcal{C}^1} + C_\#(|\varsigma|\|\Omega'_\theta\|_{\mathcal{C}^0} + 1)\|g\|_{\mathcal{C}^0}\right]. \tag{A.9}$$

We conclude that the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$ as an element of $L(\mathcal{C}^1, \mathcal{C}^1)$ is bounded by $e^{\tau_{\theta,\varsigma\Omega}}$. In addition, it follows from the usual Hennion's argument [33] that the essential spectral radius is bounded by $\lambda^{-1}e^{\tau_{\theta,\varsigma\Omega}}$. To conclude note that, by differentiating (A.8) one see that the essential spectral radius on $\mathcal{C}^k$ is bounded by $\lambda^{-k}e^{\tau_{\theta,\varsigma\Omega}}$. On the other hand, an eigenvalue in $\mathcal{C}^k$ is also an eigenvalue in $\mathcal{C}^1$. To prove the contrary, define the smoothing operator $Q_\epsilon g(x) = \int \epsilon^{-1}q(\epsilon^{-1}(x - y))g(y)dy$, where q is a bump function: $q \in \mathcal{C}^\infty_0(\mathbb{R}, \mathbb{R}_{\ge 0})$ with $\int q = 1$. Define $\mathcal{L}_\epsilon = Q_\epsilon\mathcal{L}_{\theta,\varsigma\Omega}$. By the perturbation theory in [36] the spectrum of $\mathcal{L}_\epsilon$ and $\mathcal{L}_{\theta,\varsigma\Omega}$ are close on each $\mathcal{C}^k$. On the other hand $\mathcal{L}_\epsilon$ is a compact operator and its spectrum is the same on each $\mathcal{C}^k$ since each eigenvalue belongs to $\mathcal{C}^\infty$. □

Note that, in general, it could happen that the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$ on $\mathcal{C}^1$ is smaller than $\lambda^{-1}e^{\tau_{\theta,\varsigma\Omega}}$. In this case, the second part of the above lemma is of limited interest.

[81] To get the first inequality, use the spectral decomposition together with (A.2) and its obvious analog for higher derivatives.

Remark A.3. *If Ω is real, then the spectral radii of $\mathcal{L}_{\theta,\varsigma\Omega}$ on $\mathcal{C}^1$ and $\mathcal{C}^0$ coincide; in fact, for each $\tau < \tau_{\theta,\varsigma\Omega}$ and $\chi > \chi_{\theta,\varsigma\Omega}$, there exists $\bar{n} \in \mathbb{N}$ and $g \in \mathcal{C}^0$ such that, for all $n \geq \bar{n}$,*

$$e^{\tau n}\|g\|_{\mathcal{C}^0} \leq \|\mathcal{L}^n_{\theta,\varsigma\Omega} g\|_{\mathcal{C}^0} \leq \|\mathcal{L}^n_{\theta,\varsigma\Omega} 1\|_{\mathcal{C}^1}\|g\|_{\mathcal{C}^0} \leq e^{\chi n}\|g\|_{\mathcal{C}^0}.$$

The claim then follows by Lemma A.2.

It is worth stressing the fact that the functional $m_{\theta,\varsigma\Omega}$ is guaranteed to be a measure only provided that the potential Ω_θ is real: this is essentially due to the fact that, because of cancellations of complex phases, the spectral radius on $\mathcal{C}^1$ might be smaller than the spectral radius on $\mathcal{C}^0$ if the potential has a non-zero imaginary part.

Remark A.4. *Note that, by arguments similar to the one described in this subsection, $\mathcal{L}_{\theta,\varsigma\Omega}$ is a well defined operator also on $\mathcal{C}^{r-1}$ or $W^{r-1,1}$ and, on such spaces, it has essential spectrum bounded by λ^{-r+1}. In particular $h_{\theta,\varsigma\Omega} \in \mathcal{C}^{r-1}$ and for any $1 \leq s < r$, we have*

$$\|h_{\theta,\varsigma\Omega}\|_{\mathcal{C}^s} \leq C_\#(1 + |\varsigma|\|\Omega_\theta\|_{\mathcal{C}^s})^{s+1}. \tag{A.10}$$

A.2. **Perturbation Theory with respect to ς.**

In this and the following subsections we will consider only the case in which there is a unique maximal eigenvalue $e^{\chi_{\theta,\varsigma\Omega}}$ which is simple. Hence $m_{\theta,\varsigma\Omega}$ and $h_{\theta,\varsigma\Omega}$ are well defined, except for their normalization, which is not determined by the spectral projector $\mathcal{P}_{\theta,\varsigma\Omega} = h_{\theta,\varsigma\Omega} \otimes m_{\theta,\varsigma\Omega}$ associated to $e^{\chi_{\theta,\varsigma\Omega}}$. Note that $\chi_{\theta,0} = 0$; moreover, for $\varsigma = 0$ there exists a natural normalization for $m_{\theta,0}$ and $h_{\theta,0}$ so that $m_{\theta,0}$ is the Lebesgue measure and $h_{\theta,0}$ is the density of the invariant SRB probability measure μ_θ. There is, however, no natural normalization for $\varsigma \neq 0$; we thus proceed to define one that is particularly suitable to our purposes.

Remark A.5. *Recall that the spectral data is analytic– in ς in a neighborhood of zero. Standard perturbation theory implies that such neighborhood contains the ς such that, for all $\theta \in \mathbb{T}$ we have $\|\Omega_\theta\|_{\mathcal{C}^1}|\varsigma| \leq \sigma_1$ for some fixed $\sigma_1 \in (0, 1)$ small enough. From now on we will assume ς in this set unless otherwise specified; in this regime we are guaranteed that $\mathcal{L}_{\theta,\varsigma\Omega}$ is a Perron–Frobenius operator.*

Let us differentiate the relation $m_{\theta,\varsigma\Omega}\mathcal{L}_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega} = e^{\chi_{\theta,\varsigma\Omega}}$ with respect to ς and obtain

$$\partial_\varsigma \chi_{\theta,\varsigma\Omega} = \nu_{\theta,\varsigma\Omega}(\Omega_\theta) \tag{A.11a}$$

where $\nu_{\theta,\varsigma\Omega}(g) = m_{\theta,\varsigma\Omega}(g h_{\theta,\varsigma\Omega})$; observe that $\nu_{\theta,\varsigma\Omega}(1) = 1$. Let us introduce the renormalized operators $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega} = e^{-\chi_{\theta,\varsigma\Omega}}\mathcal{L}_{\theta,\varsigma\Omega}$. Notice that $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega} = \mathcal{P}_{\theta,\varsigma\Omega} + \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}$, where $\widehat{\mathcal{Q}}_{\theta,\varsigma\Omega} = e^{-\chi_{\theta,\varsigma\Omega}}\mathcal{Q}_{\theta,\varsigma\Omega}$. Then by (A.11a) and the definition of $\mathcal{L}_{\theta,\varsigma\Omega}$ we obtain

$$\partial_\varsigma \widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(g) = \widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega} g)$$

with $\Omega_{\theta,\varsigma\Omega} = \Omega_\theta - \nu_{\theta,\varsigma\Omega}(\Omega_\theta)$. Thus, differentiating the relations $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega} = h_{\theta,\varsigma\Omega}$ and $m_{\theta,\varsigma\Omega}(\widehat{\mathcal{L}}_{\theta,\varsigma\Omega} g) = m_{\theta,\varsigma\Omega}\, g$ yields

$$\partial_\varsigma h_{\theta,\varsigma\Omega} = \left[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}\right]^{-1}\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega}) - C(\theta,\varsigma)h_{\theta,\varsigma\Omega} \tag{A.11b}$$

$$\partial_\varsigma m_{\theta,\varsigma\Omega}(g) = m_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}\left[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}\right]^{-1}\tilde{g}) + C(\theta,\varsigma)m_{\theta,\varsigma\Omega}(g) \tag{A.11c}$$

where $\tilde{g} = (\mathbb{1} - \mathcal{P}_{\theta,\varsigma\Omega})g = g - h_{\theta,\varsigma\Omega}m_{\theta,\varsigma\Omega}(g)$ and $C(\theta,\varsigma)$ depends on the normalization of $h_{\theta,\varsigma\Omega}$ and $m_{\theta,\varsigma\Omega}$. Using the above expressions, and differentiating (A.11a), it is immediate to obtain

$$\partial^2_\varsigma \chi_{\theta,\varsigma\Omega} = m_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}\left[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}\right]^{-1}(\mathbb{1} + \widehat{\mathcal{L}}_{\theta,\varsigma\Omega})\Omega_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega}), \tag{A.11d}$$

which yields

$$\partial_\varsigma^2 \chi_{\theta,\varsigma\Omega} = \nu_{\theta,\varsigma\Omega}(\Omega^2_{\theta,\varsigma\Omega}) + 2\sum_{k=1}^{\infty} \nu_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega} \circ f_\theta^k \Omega_{\theta,\varsigma\Omega}) = \tag{A.12a}$$

$$= \lim_{n\to\infty} \frac{1}{n} \nu_{\theta,\varsigma\Omega}\left(\left[\sum_{k=0}^{n-1} \Omega_{\theta,\varsigma\Omega} \circ f_\theta^k\right]^2\right) \tag{A.12b}$$

where we used the identity $m_{\theta,\varsigma\Omega}(g_1 \widehat{\mathcal{L}}^k_{\theta,\varsigma\Omega} g_2) = m_{\theta,\varsigma\Omega}(g_1 \circ f^k g_2)$, which is obtained directly by definition of the Transfer operator $\mathcal{L}_{\theta,\varsigma\Omega}$. Observe that (A.12b) shows that $\chi_{\theta,\varsigma\Omega}$ is (for real potentials) a convex function of ς.

By further differentiation of (A.11d) it is simple to show that

$$\partial_\varsigma^3 \chi_{\theta,\varsigma\Omega} = \nu_{\theta,\varsigma\Omega}(\Omega^3_{\theta,\varsigma\Omega}) + 3\sum_{k=1}^{\infty} \nu_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega} \circ f_\theta^k \Omega^2_{\theta,\varsigma\Omega} + \Omega^2_{\theta,\varsigma\Omega} \circ f_\theta^k \Omega_{\theta,\varsigma\Omega}) +$$
$$+ 6\sum_{k=1}^{\infty} \sum_{j=k+1}^{\infty} \nu_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega} \circ f_\theta^j \Omega_{\theta,\varsigma\Omega} \circ f_\theta^k \Omega_{\theta,\varsigma\Omega}),$$

which implies the useful estimate

$$|\partial_\varsigma^3 \chi_{\theta,\varsigma\Omega}| \le C_\# \|\Omega_\theta\|^3_{\mathcal{C}^1}. \tag{A.13}$$

Next, for all $n \in \mathbb{N}$,

$$\nu_{\theta,\varsigma\Omega}(\phi g \circ f_\theta^n) = m_{\theta,\varsigma\Omega}(\widehat{\mathcal{L}}^n_{\theta,\varsigma\Omega}(\phi g \circ f_\theta^n h_{\theta,\varsigma\Omega})) = m_{\theta,\varsigma\Omega}(g \widehat{\mathcal{L}}^n_{\theta,\varsigma\Omega}(\phi h_{\theta,\varsigma\Omega})). \tag{A.14}$$

The above and the iteration of (A.2) imply, setting $g = 1$ and taking the limit for $n \to \infty$, that $\nu_{\theta,\varsigma\Omega}$ is a measure provided that $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}$ is power bounded as an operator on $\mathcal{C}^0$. In addition, taking $\phi = 1$ we see that, in general, it is an invariant distribution for f_θ.

Lemma A.6. *There exists a normalization for $h_{\theta,\varsigma\Omega}$ and $m_{\theta,\varsigma\Omega}$ so that $m_{\theta,0} = \mathrm{Leb}$ and the corresponding $C(\theta,\varsigma)$ is identically 0, that is:*

$$\partial_\varsigma h_{\theta,\varsigma\Omega} = [\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1} \widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega} h_{\theta,\varsigma\Omega}) \tag{A.15a}$$

$$\partial_\varsigma m_{\theta,\varsigma\Omega}\, g = m_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1}[g - h_{\theta,\varsigma\Omega} m_{\theta,\varsigma\Omega}\, g]), \tag{A.15b}$$

provided Ω is real or, for arbitrary potentials, if $\|\Omega_\theta\|_{\mathcal{C}^1}|\varsigma| \le \sigma_1$ (see Remark A.5).

Proof. Let us temporarily fix a normalization which defines $\bar{h}_{\theta,\varsigma\Omega}$ and $\bar{m}_{\theta,\varsigma\Omega}$ so that $\mathrm{Leb}(\bar{h}_{\theta,\varsigma\Omega}) = 1$ for any ς. Note that for real potentials this can always be done since $h_{\theta,\varsigma\Omega} > 0$ due to Lemma A.1. For arbitrary potentials it is possible only if $\mathrm{Leb}(\mathcal{P}_{\theta,\varsigma\Omega}(\phi)) \neq 0$ for some $\phi \in \mathcal{C}^0$. This is the case for small ς due to $\mathrm{Leb}(\mathcal{P}_{\theta,0}(\phi)) = \mathrm{Leb}(\phi)$ and the continuity of $\mathcal{P}_{\theta,\varsigma\Omega}$.

Using (A.11b) and differentiating this normalization condition with respect to ς we obtain

$$\bar{C}(\theta,\varsigma) = \mathrm{Leb}([\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1} \widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega} \bar{h}_{\theta,\varsigma\Omega})). \tag{A.16}$$

Define $\alpha(\theta,\varsigma) = \int_0^\varsigma \bar{C}(\theta,\varsigma_1) d\varsigma_1$ and choose a new normalization so that $h_{\theta,\varsigma\Omega} = e^{\alpha(\theta,\varsigma)} \bar{h}_{\theta,\varsigma\Omega}$ (and consequently $m_{\theta,\varsigma\Omega} = e^{-\alpha(\theta,\varsigma)} \bar{m}_{\theta,\varsigma\Omega}$). Then, an immediate computation shows that $h_{\theta,\varsigma\Omega}$ and $m_{\theta,\varsigma\Omega}$ satisfy equations (A.15). □

We now fix once and for all the normalization of $h_{\theta,\varsigma\Omega}$ and $m_{\theta,\varsigma\Omega}$ to be the one constructed in Lemma A.6 and refer to it as the *standard normalization*.

Lemma A.7. *For any $g \in W^{1,1}$ and under the assumptions described in Remark A.5, we have*

$$\|\partial_\varsigma h_{\theta,\varsigma\Omega}\|_{\mathcal{C}^1} \leq C_\# \|\Omega_\theta\|_{\mathcal{C}^1} \tag{A.17a}$$

$$|\partial_\varsigma m_{\theta,\varsigma\Omega}\, g| \leq C_\# \|\Omega_\theta\|_{\mathcal{C}^1} \|g\|_{W^{1,1}} \tag{A.17b}$$

and, moreover,

$$|m_{\theta,0}\, h_{\theta,\varsigma\Omega} - 1| \leq C_\# \varsigma^2 \|\Omega_\theta\|^2_{\mathcal{C}^1} \qquad |m_{\theta,\varsigma\Omega}(h_{\theta,0}) - 1| \leq C_\# \varsigma^2 \|\Omega_\theta\|^2_{\mathcal{C}^1}. \tag{A.18}$$

Proof. Since all the quantities are analytic in ς and ς belongs to a fixed compact set, we have uniform bounds on $\|h_{\theta,\varsigma\Omega}\|_{\mathcal{C}^1}$ and $\|m_{\theta,\varsigma\Omega}\|_{(W^{1,1})'}$. Thus, by (A.15a), taking the $\mathcal{C}^1$-norm, we obtain:

$$\|\partial_\varsigma h_{\theta,\varsigma\Omega}\|_{\mathcal{C}^1} \leq \|[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1} \widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega} h_{\theta,\varsigma\Omega})\|_{\mathcal{C}^1} \leq C_\# \|\Omega_\theta\|_{\mathcal{C}^1}.$$

which implies $\|h_{\theta,\varsigma\Omega}\|_{\mathcal{C}^1} \leq C_\#(1 + |\varsigma| \|\Omega_\theta\|_{\mathcal{C}^1})$. Similar computations yield the corresponding result for $m_{\theta,\varsigma\Omega}$.

Finally, in order to obtain equations (A.18), observe that $m_{\theta,0}\, h_{\theta,\varsigma\Omega} - 1 = m_{\theta,0}(\int_0^\varsigma d\varsigma_1 \partial_\varsigma h_{\theta,\varsigma_1\Omega})$; then, since $m_{\theta,0}\, \partial_\varsigma h_{\theta,0} = 0$ we can write

$$|m_{\theta,0}\, h_{\theta,\varsigma\Omega} - 1| \leq \int_0^\varsigma d\varsigma_1 \int_0^{\varsigma_1} d\varsigma_2 \left\| \partial_\varsigma \left([\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma_2\Omega}]^{-1} \widehat{\mathcal{L}}_{\theta,\varsigma_2\Omega}(\Omega_{\theta,\varsigma_2\Omega} h_{\theta,\varsigma_2\Omega}) \right) \right\|_{\mathcal{C}^0}$$

from which follows the first of (A.18); a similar computation yields the second estimate, which concludes the proof. □

We now deal with transfer operators weighted with two different families of potentials, which we denote by $\Omega^{(0)}$ and $\Omega^{(1)}$. If $\|\Omega^{(0)} - \Omega^{(1)}\|_{\mathcal{C}^1}$ is small enough, we can once again use perturbation theory to compare spectral data. Until the end of this subsection we assume θ to be fixed and we will drop it from our notation since it will not cause any confusion. Also, we assume that either both $\Omega^{(0)}$ and $\Omega^{(1)}$ are real, or $\|\Omega^{(0)}\|_{\mathcal{C}^1}$ and $\|\Omega^{(1)}\|_{\mathcal{C}^1}$ to be sufficiently small (i.e. smaller than σ_1) so that we can assume $\varsigma = 1$ and still be in the perturbative regime (see Remark A.5). For $\varrho \in [0, 1]$, let us define the convex interpolation $\Omega^{(\varrho)} = \Omega^{(0)} + \varrho(\Omega^{(1)} - \Omega^{(0)})$, and let $\delta\Omega = \partial_\varrho \Omega^{(\varrho)} = \Omega^{(1)} - \Omega^{(0)}$; consider the transfer operators $\mathcal{L}_\varrho = \mathcal{L}_{\theta,\Omega^{(\varrho)}}$; similarly let $h_\varrho = h_{\theta,\Omega^{(\varrho)}}$ and $m_\varrho = m_{\theta,\Omega^{(\varrho)}}$. Then, by arguments analogous to the ones leading to equations (A.11), we obtain

$$\partial_\varrho \chi_\varrho = m_\varrho(\delta\Omega h_\varrho) \tag{A.19a}$$

$$\partial_\varrho h_\varrho = [\mathbb{1} - \widehat{\mathcal{Q}}_\varrho]^{-1} \widehat{\mathcal{L}}_\varrho(\delta\widehat{\Omega}_\varrho h_\varrho) - C(\varrho) h_\varrho \tag{A.19b}$$

$$\partial_\varrho m_\varrho g = m_\varrho(\delta\widehat{\Omega}_\varrho [\mathbb{1} - \widehat{\mathcal{Q}}_\varrho]^{-1}(g - h_\varrho m_\varrho g)) + C(\varrho) m_\varrho g. \tag{A.19c}$$

where we defined $\delta\widehat{\Omega}_\varrho = \delta\Omega - m_\varrho(\delta\Omega h_\varrho)$ and the function $C(\varrho)$ depends on the normalization for h_ϱ and m_ϱ.

Lemma A.8. *For any $g \in W^{1,1}$ and under the assumptions described in Remark A.5, and choosing the standard normalization we have*

$$\|\partial_\varrho h_\varrho\|_{\mathcal{C}^k} \leq C_\# \|\delta\Omega\|_{\mathcal{C}^k} \tag{A.20a}$$

$$|\partial_\varrho m_\varrho g - m_\varrho(\delta\widehat{\Omega}_\varrho [\mathbb{1} - \widehat{\mathcal{Q}}_\varrho]^{-1}(g - h_\varrho m_\varrho g))| \leq C_\# \|\Omega^{(0)}\|_{\mathcal{C}^1} \|\delta\Omega\|_{\mathcal{C}^1} \|g\|_{W^{1,1}}. \tag{A.20b}$$

Proof. As in the proof of Lemma A.6, let us denote by $\bar{h}_\varrho$ the eigenvector normalized so that $\mathrm{Leb}(\bar{h}_\varrho) = 1$, let $\bar{\bar{C}}(\varrho)$ be the corresponding normalization in (A.19). Then a direct computation (differentiating the normalization condition and using (A.19b)) shows that

$$\bar{\bar{C}}(\varrho) = \mathrm{Leb}\left([\mathbb{1} - \widehat{\mathcal{Q}}_\varrho]^{-1} \widehat{\mathcal{L}}_\varrho \delta\widehat{\Omega}_\varrho \bar{h}_\varrho \right).$$

Moreover, let $\bar{C}_\varrho(\varsigma)$ be defined as in (A.16) with the choice $\Omega = \Omega^{(\varrho)}$; then

$$C(\varrho) = \bar{\bar{C}}(\varrho) - \int_0^1 \partial_\varrho \bar{C}_\varrho(\varsigma) d\varsigma.$$

Note that, setting $g_{\varrho,\varsigma} = \left[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega^{(\varrho)}}\right]^{-1} \widehat{\mathcal{L}}_{\theta,\varsigma\Omega^{(\varrho)}}(\widehat{\Omega}_{\varrho,\varsigma}\bar{h}_{\theta,\varsigma\Omega^{(\varrho)}})$, where $\widehat{\Omega}_{\varrho,\varsigma} = \Omega^{(\varrho)} - m_{\theta,\varsigma\Omega^{(\varrho)}}(\Omega^{(\varrho)} h_{\theta,\varsigma\Omega^{(\varrho)}})$,

$$\bar{C}_\varrho(\varsigma) = \mathrm{Leb}(g_\varrho) = m_{\theta,\varsigma\Omega^{(\varrho)}}(g_{\varrho,\varsigma}) - \int_0^\varsigma \partial_{\varsigma_1} m_{\theta,\varsigma_1\Omega^{(\varrho)}}(g_{\varrho,\varsigma}) d\varsigma_1.$$

Note that the first term of the rightmost hand side of the equation above is identically zero, hence, by (A.15b), we conclude that

$$\bar{C}_\varrho(\varsigma) = -\int_0^\varsigma m_{\theta,\varsigma_1\Omega^{(\varrho)}}(\widehat{\Omega}_{\varrho,\varsigma_1}\left[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma_1\Omega^{(\varrho)}}\right]^{-1} \widehat{g}_{\varrho,\varsigma}) d\varsigma_1$$

where $\widehat{g}_{\varrho,\varsigma} = g_{\varrho,\varsigma} - m_{\theta,\varsigma_1\Omega^{(\varrho)}}(g_{\varrho,\varsigma}) h_{\theta,\varsigma_1\Omega^{(\varrho)}}$; this implies $|\bar{C}_\varrho(\varsigma)| \leq C_\# \|\Omega^{(0)}\|^2_{\mathcal{C}^1}$ and, since (A.19) implies that each derivative with respect to ϱ of the eigenvectors or operators yields an extra factor $\|\delta\Omega\|_{\mathcal{C}^1}$, $|\partial_\varrho \bar{C}_\varrho(\varsigma)| \leq C_\# \|\Omega^{(0)}\|^2_{\mathcal{C}^1} \|\delta\Omega\|_{\mathcal{C}^1}$. By similar arguments we obtain $|\bar{\bar{C}}_\varrho(\varsigma)| \leq C_\# \|\Omega^{(0)}\|_{\mathcal{C}^1} \|\delta\Omega\|_{\mathcal{C}^1}$, which then implies equations (A.20). □

A.3. **Perturbation Theory with respect to θ.**

Recalling the notation and computations at the end of the proof of Lemma A.1 and by argument analogous to the ones leading to equations (A.11), but differentiating with respect to θ, we gather:

$$\partial_\theta \chi_{\theta,\varsigma\Omega} = m_{\theta,\varsigma\Omega}\left(\mathcal{D}_{\theta,\varsigma\Omega} h_{\theta,\varsigma\Omega}\right) \tag{A.21a}$$

$$\partial_\theta h_{\theta,\varsigma\Omega} = [\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1} \widehat{\mathcal{L}}_{\theta,\varsigma\Omega} \widetilde{\mathcal{D}}_{\theta,\varsigma\Omega} h_{\theta,\varsigma\Omega} - D(\theta,\varsigma) h_{\theta,\varsigma\Omega} \tag{A.21b}$$

$$\partial_\theta m_{\theta,\varsigma\Omega}\, g = m_{\theta,\varsigma\Omega}(\mathcal{D}_{\theta,\varsigma\Omega}[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1} \tilde{g}) + D(\theta,\varsigma) m_{\theta,\varsigma\Omega}\, g \tag{A.21c}$$

where $\widetilde{\mathcal{D}}_{\theta,\varsigma\Omega} g = \mathcal{D}_{\theta,\varsigma\Omega} g - h_{\theta,\varsigma\Omega} m_{\theta,\varsigma\Omega}(\mathcal{D}_{\theta,\varsigma\Omega} g)$, and recall that $\tilde{g} = g - h_{\theta,\varsigma\Omega} m_{\theta,\varsigma\Omega}(g)$ and $\mathcal{D}_{\theta,\varsigma\Omega}$ is defined in (A.5). Once again $D(\theta,\varsigma)$ is a function which depends on the normalization for $h_{\theta,\varsigma\Omega}$ and $m_{\theta,\varsigma\Omega}$. Note that we cannot, in general, assume that $D = 0$; since $m_{\theta,0} = \mathrm{Leb}$, it is however true that $D(\theta, 0) = 0$ for any θ. Similarly, since $\chi_{\theta,0} = 0$, we have $\partial_\theta \chi_{\theta,0} = 0$.

Lemma A.9. *There exists $\sigma_0 \in (0, \sigma_1)$ such that, if*

$$\|\partial_\theta^2 \Omega\|_{\mathcal{C}^1(\mathbb{T}^2)} + \|\partial_x \Omega\|_{\mathcal{C}^2(\mathbb{T}^2)} + \|\Omega\|_{\mathcal{C}^2(\mathbb{T}^2)} \leq \sigma_0$$

we have, for any $1 \leq k < r - 1$ and using the standard normalization:

$$|\partial_\theta \chi_{\theta,\Omega} - \partial_\theta \mathrm{Leb}(\Omega_\theta h_{\theta,0})| \leq C_\# \|\Omega\|^2_{\mathcal{C}^1(\mathbb{T}^2)}. \tag{A.22a}$$

$$\|\partial_\theta h_{\theta,\Omega}\|_{\mathcal{C}^k} \leq C_\#(1 + \|\Omega_\theta\|_{\mathcal{C}^{k+1}} + \|\partial_\theta \Omega_\theta\|_{\mathcal{C}^k}) \tag{A.22b}$$

$$|\partial_\theta m_{\theta,\Omega}\, g| \leq C_\# \|\Omega\|_{\mathcal{C}^2(\mathbb{T}^2)} \|g\|_{W^{2,1}} \tag{A.22c}$$

$$\left|\partial_\theta^2 \chi_{\theta,\Omega}\right| \leq C_\#(\|\partial_\theta^2 \Omega\|_{\mathcal{C}^1(\mathbb{T}^2)} + \|\partial_x \Omega\|_{\mathcal{C}^2(\mathbb{T}^2)} + \|\Omega\|_{\mathcal{C}^2(\mathbb{T}^2)}). \tag{A.22d}$$

Additionally, $\partial_\varsigma h_{\theta,\varsigma\Omega}$, $\partial_\varsigma m_{\theta,\varsigma\Omega}$ and $\partial_\varsigma^2 \chi_{\theta,\varsigma\Omega}$ are differentiable in θ.

Proof. Plugging (A.15b), (A.15a) and (A.5) in (A.21a), we have:

$$\begin{aligned}\partial_\theta\chi_{\theta,\Omega} &= \text{Leb}(\mathcal{D}_{\theta,\Omega}h_{\theta,\Omega}) + \int_0^1 m_{\theta,\varsigma\Omega}\left(\Omega_{\theta,\varsigma\Omega}[\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1}\widetilde{\mathcal{D}}_{\theta,\Omega}h_{\theta,\Omega}\right)d\varsigma \\ &= \text{Leb}\left(\partial_\theta\Omega_\theta\cdot h_{\theta,0} - \frac{\partial_\theta f_\theta}{f'_\theta}\Omega'_\theta h_{\theta,0}\right) - \text{Leb}\left(\Omega_{\theta,0}[\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,0}]^{-1}\left[\frac{\partial_\theta f_\theta}{f'_\theta}h_{\theta,0}\right]'\right) \\ &\quad + \mathcal{O}(\|\Omega\|^2_{\mathcal{C}^1(\mathbb{T}^2)}),\end{aligned} \tag{A.23}$$

where the term having a derivative in (A.5) disappears by integration by parts against Lebesgue. Next, note that (A.21b) implies

$$\begin{aligned}\partial_\theta\text{Leb}(\Omega_\theta h_{\theta,0}) &= \text{Leb}(\partial_\theta\Omega_\theta h_{\theta,0}) + \text{Leb}(\Omega_\theta[\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,0}]^{-1}\widehat{\mathcal{Q}}_{\theta,0}\widetilde{\mathcal{D}}_{\theta,0}h_{\theta,0}) \\ &= \text{Leb}(\partial_\theta\Omega_\theta h_{\theta,0}) - \text{Leb}\left(\Omega_\theta[\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,0}]^{-1}\left[\frac{\partial_\theta f_\theta}{f'_\theta}h_{\theta,0}\right]'\right) \\ &\quad + \text{Leb}\left(\Omega_\theta\left[\frac{\partial_\theta f_\theta}{f'_\theta}h_{\theta,0}\right]'\right) + \mathcal{O}(\|\Omega\|^2_{\mathcal{C}^1(\mathbb{T}^2)}).\end{aligned}$$

In addition,

$$\begin{aligned}\text{Leb}\left([\Omega_{\theta,0}-\Omega_\theta][\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,0}]^{-1}\left[\frac{\partial_\theta f_\theta}{f'_\theta}h_{\theta,0}\right]'\right) &= \text{Leb}(\Omega_\theta h_{\theta,0}) \\ &\quad\times\text{Leb}\left([\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,0}]^{-1}\left[\frac{\partial_\theta f_\theta}{f'_\theta}h_{\theta,0}\right]'\right) \\ &= \text{Leb}(\Omega_\theta h_{\theta,0})\text{Leb}\left(\left[\frac{\partial_\theta f_\theta}{f'_\theta}h_{\theta,0}\right]'\right) = 0.\end{aligned}$$

where the last term disappears again by integration by part against Lebesgue. Combining the above expressions with (A.23) yields (A.22a). Next, recall that, by the construction of the standard normalization given in Lemma A.6 we have set $h_{\theta,\varsigma\Omega} = e^{\alpha(\theta,\varsigma)}\bar{h}_{\theta,\varsigma\Omega}$, where $\bar{h}_{\theta,\varsigma\Omega}$ is normalized so that $\text{Leb}(\bar{h}_{\theta,\varsigma\Omega}) = 1$ and $\alpha(\theta,\varsigma) = \int_0^\varsigma \bar{C}(\theta,\varsigma_1)d\varsigma_1$, $\bar{C}$ being given by (A.16). Observe that, differentiating the normalization condition for $\bar{h}_{\theta,\varsigma\Omega}$ with respect to θ, we obtain, using equations (A.21):

$$\bar{D}(\theta,\varsigma) = \text{Leb}([\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1}\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}\widetilde{\mathcal{D}}_{\theta,\varsigma\Omega}\bar{h}_{\theta,\varsigma\Omega}).$$

Then, by definition of $h_{\theta,\varsigma\Omega}$ we get:

$$D(\theta,\varsigma) = \bar{D}(\theta,\varsigma) - \partial_\theta\alpha(\theta,\varsigma).$$

Thus, by the definition of $\alpha(\theta,\varsigma)$ and using the hypothesis on $\|\Omega\|_{\mathcal{C}^0}$, a direct computation, which is left to the reader, yields

$$|\partial_\theta\alpha(\theta,\varsigma)| \le C_\#\|\Omega\|_{\mathcal{C}^1(\mathbb{T}^2)}. \tag{A.24}$$

The proof of (A.22b) immediately follows from (A.21b) using the definition (A.5). In order to prove (A.22c), one has to examine $\bar{D}$ in more detail. By (A.5) we have

$$\begin{aligned}\bar{D}(\theta,1) = -\text{Leb}\left([\mathbb{1}-\widehat{\mathcal{Q}}_{\theta,\Omega}]^{-1}\widehat{\mathcal{L}}_{\theta,\Omega}\left\{\left[\frac{\partial_\theta f_\theta}{f'_\theta}\bar{h}_{\theta,\Omega}\right]' - h_{\theta,\Omega}m_{\theta,\Omega}\left(\left[\frac{\partial_\theta f_\theta}{f'_\theta}\bar{h}_{\theta,\Omega}\right]'\right)\right\}\right) \\ + \mathcal{O}(\|\Omega\|_{\mathcal{C}^1(\mathbb{T}^2)}).\end{aligned}$$

Let $\widehat{g} = [\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\Omega}]^{-1}\widehat{\mathcal{L}}_{\theta,\Omega}\left\{\left[\frac{\partial_\theta f_\theta}{f'_\theta}\bar{h}_{\theta,\Omega}\right]' - h_{\theta,\Omega}m_{\theta,\Omega}\left(\left[\frac{\partial_\theta f_\theta}{f'_\theta}\bar{h}_{\theta,\Omega}\right]'\right)\right\}$ and observe that $m_{\theta,\Omega}(\widehat{g}) = 0$ and $\|\widehat{g}\|_{\mathrm{BV}} \leq C_\#$; then, (A.17b) implies

$$|\mathrm{Leb}(\widehat{g})| \leq |m_{\theta,\Omega}(\widehat{g})| + \left|\int_1^0 \partial_\varsigma m_{\theta,\varsigma\Omega}(\widehat{g})d\varsigma\right| \leq C_\#\|\Omega_\theta\|_{\mathcal{C}^1}\|\widehat{g}\|_{W^{1,1}}.$$

It follows that

$$|D(\theta,1)| \leq C_\#\|\Omega\|_{\mathcal{C}^1(\mathbb{T}^2)}. \tag{A.25}$$

We thus obtain (A.22c) by (A.25) and applying to (A.21c) similar arguments. In order to prove (A.22d), observe that differentiating (A.21a) yields

$$\partial_\theta^2\chi_{\theta,\varsigma\Omega} = (\partial_\theta m_{\theta,\varsigma\Omega})(\mathcal{D}_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega}) + m_{\theta,\varsigma\Omega}((\partial_\theta\mathcal{D}_{\theta,\varsigma\Omega})h_{\theta,\varsigma\Omega}) + m_{\theta,\varsigma\Omega}(\mathcal{D}_{\theta,\varsigma\Omega}(\partial_\theta h_{\theta,\varsigma\Omega})).$$

Substituting (A.21c), (A.5) and (A.21b) in the above expression we get

$$\partial_\theta^2\chi_{\theta,\Omega} = m_{\theta,\Omega}(A' + B)$$

where A, B are two functions that (using (A.10) and our assumptions on Ω) satisfy

$$\|A'\|_{W^{1,1}} \leq \|A\|_{\mathcal{C}^2} \leq C_\#$$
$$\|B\|_{\mathcal{C}^1} \leq C_\#(\|\partial_\theta^2\Omega\|_{\mathcal{C}^1(\mathbb{T}^2)} + \|\partial_x\Omega\|_{\mathcal{C}^2(\mathbb{T}^2)} + \|\Omega\|_{\mathcal{C}^2(\mathbb{T}^2)})$$

To conclude the proof, we use (A.15b) as in the proof of (A.22a) which yields the result since $\mathrm{Leb}(A') = 0$ by integration by parts. Finally, the last statement follows from the above considerations and the formulae (A.15) and (A.11d). □

We conclude the subsection with a non-perturbative result

Lemma A.10. *Assume that Ω is real and $\mathrm{Leb}(\Omega h_{0,\theta}) = 0$, then*

$$\|\partial_\theta\chi_{\theta,\Omega}\| \leq C_\# \min\{\|\Omega\|_{\mathcal{C}^1} + 1, \|\Omega\|_{\mathcal{C}^1}^2\}.$$

Proof. If $\|\Omega\|_{\mathcal{C}^0} \leq \sigma_0$, the estimate follows from (A.22a). In the non-perturbative regime, i.e. for potentials of larger norm, recall that $m_{\theta,\Omega}$ is a positive measure (see Lemma A.1), and therefore (A.21a) with (A.5) imply

$$|\partial_\theta\chi_{\theta,\Omega}| \leq C_\# m_{\theta,\Omega}(\|\Omega\|_{\mathcal{C}^1}h_{\theta,\Omega} + |h'_{\theta,\Omega}|).$$

The claim then follows by recalling the normalization $1 = m_{\theta,\Omega}(h_{\theta,\Omega})$ and that, using once again Lemma A.1:

$$|h'_{\theta,\Omega}(x)| \leq C_\#(\|\Omega\|_{\mathcal{C}^1} + 1)h_{\theta,\Omega}(x)$$ □

A.4. **Results for functions of bounded variation.**

In certain parts of the paper it is convenient to consider transfer operators acting on the Sobolev Space $W^{1,1}$ or on the space of function of bounded variations BV. Since $W^{1,1} \subset \mathrm{BV}$, with the same norm, we will limit our discussion to the second, more general, case.

For functions of bounded variation, the Lasota–Yorke inequality reads as follows: for any $\psi \in \mathcal{C}^1$ and any $n \in \mathbb{N}$, setting $\Omega_{\theta,n} = \sum_{k=0}^{n-1}\Omega_\theta \circ f_\theta^k$,

$$\begin{aligned}\left|\int \psi'\mathcal{L}_{\theta,\varsigma\Omega}^n g\right| &= \left|\int g e^{\varsigma\Omega_{\theta,n}}(\psi') \circ f_\theta^n\right| \\ &\leq \left|\int g\left[\frac{e^{\varsigma\Omega_{\theta,n}}\psi \circ f_\theta^n}{(f_\theta^n)'}\right]'\right| + C_\#\int |\psi|\mathcal{L}_{\theta,0}^n\left[e^{\varsigma\mathrm{Re}(\Omega_{\theta,n})}(1 + |\varsigma|\|\Omega'_\theta\|_{\mathcal{C}^0})|g|\right].\end{aligned} \tag{A.26}$$

Thus, setting

$$\tilde{\chi}_n = \log \|\mathcal{L}^n_{\theta,\varsigma\mathrm{Re}(\Omega_{\theta,n})}\|_{L^1} = \log \|e^{\varsigma\mathrm{Re}(\Omega_{\theta,n})}\|_{L^1}$$
$$\tau_n = \log \left\| \frac{e^{\varsigma\mathrm{Re}(\Omega_{\theta,n})}}{(f^n_\theta)'} \right\|_{L^\infty}$$

we have

$$\|\mathcal{L}^n_{\theta,\varsigma\Omega} g\|_{\mathrm{BV}} \le e^{\tau_n}\|g\|_{\mathrm{BV}} + C_\# e^{\tilde{\chi}_n}(1 + |\varsigma|\|\Omega_\theta\|_{\mathcal{C}^1})\|g\|_{L^1}. \tag{A.27}$$

By the usual Hennion argument [33], the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$ is bounded by $e^{\tilde{\chi}_n/n}$ and the essential spectral radius by $e^{\tau_n/n}$. Note that (A.26) also implies[82]

$$\|\mathcal{L}^n_{\theta,\varsigma\Omega} g\|_{\mathrm{BV}} \le \left[e^{\tau_n} + C_\#(1 + |\varsigma|\|\Omega_\theta\|_{\mathcal{C}^1}) \int e^{\varsigma\mathrm{Re}(\Omega_{\theta,n})}\right] \|g\|_{\mathrm{BV}}. \tag{A.28}$$

In addition, calling $\mathcal{H}_n$ the set of inverse branches of f^n_θ, we have, by standard distortion estimates,

$$\begin{aligned} \int_{\mathbb{T}^1} e^{\varsigma\mathrm{Re}(\Omega_{\theta,n})} &= \int_{\mathbb{T}^1} \mathcal{L}^n_{\theta,0} e^{\varsigma\mathrm{Re}(\Omega_{\theta,n})} \\ &= \sum_{h\in\mathcal{H}_n} \int_{\mathbb{T}^1} e^{\varsigma\mathrm{Re}(\Omega_{\theta,n})\circ h(x)} h'(x)dx \ge C_\# e^{\tau_n}. \end{aligned} \tag{A.29}$$

Thus our bound on the spectral radius is larger or equal than our bound on the essential spectral radius. Nevertheless, these are just estimates: the real values could be much smaller.

Remark A.11. *For real potentials and $\varsigma \ge 0$, more can be said. If χ denotes the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega_\theta}$ as an operator on $\mathcal{C}^1$ and e^χ its maximal eigenvalue, then*[83]

$$e^{-c_\#|\varsigma|}e^{n\chi} \le \int_{\mathbb{T}^1} \mathcal{L}^n_{\theta,\varsigma\Omega_\theta} 1 = \int_{\mathbb{T}^1} e^{\varsigma\Omega_{\theta,n}} = \int_{\mathbb{T}^1} \mathcal{L}^n_{\theta,\varsigma\Omega_\theta} 1 \le e^{c_\#|\varsigma|}e^{n\chi}.$$

This, together with (A.28) and (A.29), implies that the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$ on BV coincides with the spectral radius on $\mathcal{C}^1$ and, moreover, $e^{-\chi}\mathcal{L}_{\theta,\varsigma\Omega_\theta}$ is power bounded on BV. This does not, however, imply that $\mathcal{L}_{\theta,\varsigma\Omega}$ is a Perron–Frobenius operator also when acting on BV: in fact, for large ς, the essential spectral radius could a priori coincide with the spectral radius. Nevertheless, we can find a simple condition that prevents this pathological behavior. Let $\operatorname{Osc}\Omega_\theta := \sup\Omega_\theta - \inf\Omega_\theta$; then

$$e^{\tau_n} \le \log\left[\lambda^{-n}e^{\varsigma\sup\Omega_{\theta,n}}\right] \le \log\left[\lambda^{-n}e^{n\varsigma\operatorname{Osc}\Omega_\theta}\left\|e^{\varsigma\Omega_{\theta,n}}\right\|_{L^1}\right].$$

The above implies that if $|\varsigma|\operatorname{Osc}\Omega_\theta < \log\lambda$, then the essential spectral radius is strictly smaller than the spectral radius.

Remark A.12. *No such general bounds are available for arbitrary complex potentials and we must then rely on perturbation theory. If the potential is purely imaginary, then (A.27) implies that the essential spectral radius is smaller than λ^{-1}. Since the point spectrum is independent on the space on which the operators act,*[84] *it follows that the spectrum outside the disk $\{|z| \le \lambda^{-1}\}$ on BV coincides with the spectrum on $\mathcal{C}^1$.*

---

[82] Recall that, in one dimension, $\|g\|_{L^\infty} \le \|g\|_{\mathrm{BV}}$.

[83] Since $m_{\theta,\varsigma\Omega}\mathcal{L}^n_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega} = e^{n\chi_{\theta,\varsigma\Omega}}$ and $m_{\theta,\varsigma\Omega}$ is a measure (see Lemma A.1), then there exists $x_* \in \mathbb{T}$ such that $\mathcal{L}^n_{\theta,\varsigma\Omega}h_{\theta,\varsigma\Omega}(x_*) = e^{n\chi_{\theta,\varsigma\Omega}}$, then the claim follows recalling (A.3).

[84] This can be proven as in Lemma A.2.

We conclude this brief discussion with a number of estimates on the left eigenvector; observe first that by definition and by the analytic dependence of all objects on ς, we have, for $\|\Omega_\theta\|_{\mathcal{C}^1} < \sigma_1$ that $|m_{\theta,\Omega} g| \leq C_\# \|g\|_{\mathrm{BV}}$, and using (A.15b) we thus conclude that, similarly to (A.17b):

$$|\partial_\varsigma m_{\theta,\varsigma\Omega}(g)| \leq C_\# \|\Omega_\theta\|_{\mathcal{C}^1} \|g\|_{\mathrm{BV}}, \tag{A.30}$$

in particular:

$$|m_{\theta,\varsigma\Omega}(g) - \mathrm{Leb}(g)| < C_\# \|\Omega_\theta\|_{\mathcal{C}^1} \|g\|_{\mathrm{BV}}. \tag{A.31}$$

We now proceed to obtain a refinement of the above estimates.

Lemma A.13. *There exists $\sigma_2, C_5 > 0$ such that, provided $\|\Omega_\theta\|_{\mathcal{C}^1} < \sigma_2$ and Ω is either real or it satisfies the estimate $\|\Omega_\theta\|_{\mathcal{C}^1} \exp\big[-C_5 \|\Omega_\theta\|_{\mathcal{C}^1}^{-1}\big] \leq \|\Omega_\theta\|_{\mathcal{C}^0}$, for any $g \in \mathrm{BV}$ and $n \in \mathbb{N}$, we have*

$$|m_{\theta,\Omega}\, g| \leq C_\# e^{c_\# n \|\Omega_\theta\|_{\mathcal{C}^0}} \|g\|_{L^1} + C_\# e^{-c_\# n} \|g\|_{\mathrm{BV}}. \tag{A.32}$$

Proof. First of all we choose $\sigma_2 \leq \sigma_1$ so that $\mathcal{L}_{\theta,\Omega}$ is of Perron–Frobenius type and let $h_{\theta,\Omega} \otimes m_{\theta,\Omega}$ be the eigenprojector associated to its maximal eigenvalue $e^{\chi_{\theta,\Omega}}$. In particular we have, for any $n > 0$

$$\widehat{\mathcal{L}}^n_{\theta,\Omega} g = h_{\theta,\Omega} m_{\theta,\Omega}(g) + \widehat{\mathcal{Q}}^n_{\theta,\Omega}(g)$$

and therefore there exists $\tau \in (0,1)$ such that:

$$\begin{aligned} |m_{\theta,\Omega}\, g| &\leq \left| \frac{\mathrm{Leb}\, \widehat{\mathcal{L}}^n_{\theta,\Omega} g}{\mathrm{Leb}\, h_{\theta,\Omega}} \right| + C_\# \tau^n \|g\|_{\mathrm{BV}} \\ &\leq C_\# \mathrm{Leb} \left| e^{\Omega_{\theta,n} - n\chi_{\theta,\Omega}} g \right| + C_\# \tau^n \|g\|_{\mathrm{BV}}. \end{aligned} \tag{A.33}$$

Observe that if Ω is real, then Lemma A.1 states that $m_{\theta,\Omega}$ is a measure. Therefore (A.11a) implies that

$$|\chi_{\theta,\Omega}| < C_\# \|\Omega_\theta\|_{\mathcal{C}^0}. \tag{A.34}$$

We claim that the same estimate holds also for complex potentials satisfying the assumption given in the statement. In fact, a priori $|\chi_{\theta,\Omega}| < C_\# \|\Omega_\theta\|_{\mathcal{C}^1}$, thus (A.33) implies

$$|m_{\theta,\Omega}\, g| \leq C_\# e^{n \|\Omega_\theta\|_{\mathcal{C}^1}} \|g\|_{L^1} + C_\# \tau^n \|g\|_{\mathrm{BV}}.$$

We can then choose $n = \|\Omega_\theta\|_{\mathcal{C}^1}^{-1}$ and apply the resulting bound to (A.11a), obtaining the better estimate:

$$|\chi_{\theta,\Omega}| \leq C_\# \|\Omega_\theta\|_{\mathcal{C}^0} + C_\# e^{-c_\# \|\Omega_\theta\|_{\mathcal{C}^1}^{-1}} \|\Omega_\theta\|_{\mathcal{C}^1} \leq C_\# \|\Omega_\theta\|_{\mathcal{C}^0}, \tag{A.35}$$

where we have used the hypotheses of the lemma choosing $C_5 = |\log \tau|$. We can thus use again (A.33) and obtain (A.32), concluding the proof of the lemma. □

Lemma A.14. *Under the assumptions of Lemma A.13, for any $g \in \mathrm{BV}$:*

$$|\partial_\varsigma m_{\theta,\varsigma\Omega}\, g| \leq C_\# \|\Omega_{\theta,\varsigma\Omega}\|_{\mathcal{C}^0} \log \|\Omega_\theta\|_{\mathcal{C}^1}^{-1} \|g\|_{L^1} + C_\# \|\Omega_\theta\|_{\mathcal{C}^1}^{100} \|g\|_{\mathrm{BV}} \tag{A.36}$$

in particular:[85]

$$|m_{\theta,\Omega}\, g - \mathrm{Leb}\, g| \leq C_\# \|\Omega_{\theta,\Omega}\|_{\mathcal{C}^0} \log \|\Omega_\theta\|_{\mathcal{C}^1}^{-1} \|g\|_{L^1} + C_\# \|\Omega_\theta\|_{\mathcal{C}^1}^{100} \|g\|_{\mathrm{BV}}. \tag{A.37}$$

[85] Our choice of the power 100 is clearly arbitrary: one could substitute it with any sufficiently large number at the expense of increasing the constants.

Proof. Again we choose $\sigma_2 \le \sigma_1$ so that $\mathcal{L}_{\theta,\Omega}$ is of Perron–Frobenius type and let $h_{\theta,\Omega} \otimes m_{\theta,\Omega}$ denote the eigenprojector associated to its maximal eigenvalue $e^{\chi_{\theta,\Omega}}$. Combining (A.15b) and (A.32) with the choice $n = \|\Omega_\theta\|_{\mathcal{C}^0}^{-1}$ we obtain:

$$
\begin{aligned}
|\partial_\varsigma m_{\theta,\varsigma\Omega}| \le{}& C_\# \|(\Omega_{\theta,\varsigma\Omega}[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1}\tilde g)\|_{L^1} \\
&+ C_\# \|\Omega_\theta\|_{\mathcal{C}^1} \exp\big[-c_\# \|\Omega_\theta\|_{\mathcal{C}^0}^{-1}\big] \|g\|_{\mathrm{BV}}
\end{aligned}
\tag{A.38}
$$

where recall that $\tilde g = g - h_{\theta,\varsigma\Omega} m_{\theta,\varsigma\Omega}\, g$. On the other hand, for any $K \in \mathbb{N}$

$$
\begin{aligned}
\|\Omega_{\theta,\varsigma\Omega}[\mathbb{1} - \widehat{\mathcal{Q}}_{\theta,\varsigma\Omega}]^{-1}\tilde g\|_{L^1} &\le \sum_{k=0}^{\infty} \|\Omega_{\theta,\varsigma\Omega} \widehat{\mathcal{L}}^k_{\theta,\varsigma\Omega}\tilde g)\|_{L^1} \\
&\le C_\# \|\Omega_{\theta,\varsigma\Omega}\|_{\mathcal{C}^0} \left[K\|g\|_{L^1} + K|m_{\theta,\varsigma\Omega}\, g| + \tau^K \|\tilde g\|_{\mathrm{BV}}\right]
\end{aligned}
\tag{A.39}
$$

since $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}$ is power bounded in L^1 and where $\tau \in (0,1)$ is determined by the spectral gap in BV. Note that in the considered range of ς we can assume τ to be independent on ς. We can now choose $K = C \log \|\Omega_\theta\|_{\mathcal{C}^1}^{-1}$. By Lemma A.13, with the choice $n = \|\Omega_\theta\|_{\mathcal{C}^0}^{-1}$, we can substitute (A.39) in (A.38) to obtain (A.36), provided that C has been chosen large enough. Integrating (A.36) with respect to ς from 0 to 1 yields (A.37). □

Lemma A.15. *Under the hypotheses of Lemma A.13, for any $\theta, \theta' \in \mathbb{T}$ sufficiently close, we have*

$$
|m_{\theta,\Omega}(g) - m_{\theta',\Omega}(g)| \le C_\# [\log \|\Omega\|_{\mathcal{C}^0}]^2 |\theta - \theta'| \|\Omega\|_{\mathcal{C}^1} \|g\|_{\mathrm{BV}}.
$$

Proof. We choose σ_2 so that the operators $\mathcal{L}_{\theta,\Omega}$ are of Perron–Frobenius type for all θ (see Remark A.12). Accordingly, there exists $\tau \in (0,1)$ such that[86] for any $N \in \mathbb{N}$

$$
|m_{\theta,\Omega}(g) - m_{\theta',\Omega}(g)| \le \left| \frac{\mathrm{Leb}(\widehat{\mathcal{L}}^N_{\theta,\Omega} g)}{\mathrm{Leb}(h_{\theta,\Omega})} - \frac{\mathrm{Leb}(\widehat{\mathcal{L}}^N_{\theta',\Omega} g)}{\mathrm{Leb}(h_{\theta',\Omega})} \right| + C_\# \tau^N \|g\|_{\mathrm{BV}}.
$$

In addition, by (A.21b), (A.25) and (A.17b) we have

$$
\begin{aligned}
|\mathrm{Leb}(h_{\theta,\Omega}) - \mathrm{Leb}(h_{\theta',\Omega})| \le{}& \int_\theta^{\theta'} d\varphi \left|(\mathrm{Leb} - m_{\varphi,\Omega})([\mathbb{1} - \widehat{\mathcal{Q}}_{\varphi,\Omega}]^{-1} \widehat{\mathcal{L}}_{\varphi,\Omega} \widetilde{\mathcal{D}}_{\varphi,\Omega} h_{\varphi,\Omega})\right| \\
&+ \mathcal{O}(|\theta - \theta'| \|\Omega\|_{\mathcal{C}^1}) = \mathcal{O}(|\theta - \theta'| \|\Omega\|_{\mathcal{C}^1}),
\end{aligned}
$$

where recall $\widetilde{\mathcal{D}}_{\theta,\Omega} g = \widetilde{\mathcal{D}}_{\theta,\Omega} g - h_{\theta,\Omega} m_{\theta,\Omega}(\mathcal{D}_{\theta,\Omega} g)$ and $\mathcal{D}_{\theta,\Omega}$ is defined in (A.5). Also note that

$$
\begin{aligned}
\frac{\mathrm{Leb}(\widehat{\mathcal{L}}^N_{\theta',\Omega} g)}{\mathrm{Leb}(h_{\theta,\Omega})} - \frac{\mathrm{Leb}(\widehat{\mathcal{L}}^N_{\theta',\Omega} g)}{\mathrm{Leb}(h_{\theta',\Omega})} &= \frac{\mathrm{Leb}(\widehat{\mathcal{L}}^N_{\theta',\Omega} g)}{\mathrm{Leb}(h_{\theta',\Omega})} \left[\frac{\mathrm{Leb}(h_{\theta',\Omega})}{\mathrm{Leb}(h_{\theta,\Omega})} - 1\right] \\
&= m_{\theta',\Omega}(g) \mathcal{O}(|\theta - \theta'| \|\Omega\|_{\mathcal{C}^1}) + \mathcal{O}(\tau^N \|g\|_{\mathrm{BV}}) \\
&= \mathcal{O}(|\theta - \theta'| \|\Omega\|_{\mathcal{C}^1}) \left(\|g\|_{L^1} + C_\# \exp\big[-c_\# \|\Omega_\theta\|_{\mathcal{C}^0}^{-1}\big] \|g\|_{\mathrm{BV}}\right) + \mathcal{O}(\tau^N \|g\|_{\mathrm{BV}})
\end{aligned}
$$

[86] By perturbation theory the spectral gap varies continuously in θ, hence by compactness there exist an uniform spectral gap.

where we have used Lemma A.13 with the choice $n = \|\Omega_\theta\|_{\mathcal{C}^0}^{-1}$. We can then choose $N = C \log \|\Omega\|_{\mathcal{C}^0}^{-1}$ for C large enough and continue our estimate to write[87]

$$\begin{aligned}|m_{\theta,\Omega}(g) - m_{\theta',\Omega}(g)| &\leq C_\# \sum_{k=1}^{N} \int_\theta^{\theta'} \left|\mathrm{Leb}(\widehat{\mathcal{L}}_{\theta,\Omega}^{N-k}\widehat{\mathcal{L}}_{\varphi,\Omega}\mathcal{D}_{\varphi,\Omega}\widehat{\mathcal{L}}_{\theta,\Omega}^{k-1}g)d\varphi\right| \\ &\quad + C_\#[\log \|\Omega\|_{\mathcal{C}^0}^{-1}]|\theta-\theta'|\|\Omega\|_{\mathcal{C}^1}\|g\|_{L^1} + C_\#\|\Omega\|_{\mathcal{C}^0}^{100}\|g\|_{\mathrm{BV}} \\ &\leq \sum_{k=1}^{N}\int_\theta^{\theta'}\left|\mathrm{Leb}\left(\left[e^{\sum_{j=0}^{N-k}\Omega_\theta\circ \bar f_{\ell,\theta}^j\circ\bar f_{\ell,\varphi}+\Omega_\varphi}-1\right]\mathcal{D}_{\varphi,\Omega}\widehat{\mathcal{L}}_{\theta,\Omega}^{k-1}g\right)d\varphi\right| \\ &\quad + C_\#[\log \|\Omega\|_{\mathcal{C}^0}^{-1}]|\theta-\theta'|\|\Omega\|_{\mathcal{C}^1}\|g\|_{L^1} + \|\Omega\|_{\mathcal{C}^0}^{100}\|g\|_{\mathrm{BV}},\end{aligned}$$

which, integrating by parts, yields the lemma.k □

A.5. **Generic conditions.**

Here we discuss some conditions that prevent non generic behavior of the transfer operator. They are arranged by (apparent) increasing strength. Yet, we will see at the end of the section that, although in general they are all different, in the particularly simple case we are considering, they are in fact all equivalent to the weaker condition: the potential should not be cohomologous to a constant. As the latter condition holds generically, all the conditions stated below also hold generically.

Lemma A.16. *Let θ, ς be values for which $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}$ has a spectral gap, $m_{\theta,\varsigma\Omega}$ is a measure and $|h_{\theta,\varsigma\Omega}| > 0$. If $\partial_\varsigma^2\chi_{\theta,\varsigma\Omega}$ is zero, then Ω_θ is* cohomologous to a constant, *i.e. there exists $\beta \in \mathbb{R}$ and $\phi \in \mathcal{C}^1$ such that*

$$\Omega_\theta = \beta + \phi - \phi \circ f_\theta.$$

Proof. Note that if the second derivative is zero for some θ and ς then, by the computation implicit in (A.12b), it follows that the sequence $\sum_{k=0}^{n-1}\Omega_{\theta,\varsigma\Omega}\circ f_\theta^k$ is uniformly bounded in $L^2(\mathbb{T}, m_{\theta,\varsigma\Omega})$ and hence weakly compact.[88] Let $\sum_{k=0}^{n_j-1}\Omega_{\theta,\varsigma\Omega}\circ f_\theta^k$ be a weakly convergent subsequence and let $\phi \in L^2$ be its limit. Hence, for any $\varphi \in L^2$ holds

$$\lim_{j\to\infty}\nu_{\theta,\varsigma\Omega}\left(\varphi\sum_{k=0}^{n_j-1}\Omega_{\theta,\varsigma\Omega}\circ f_\theta^k\right) = \nu_{\theta,\varsigma\Omega}(\varphi\phi).$$

[87] Remember (A.6), from which $\partial_\theta\widehat{\mathcal{L}}_{\theta,\Omega} = \widehat{\mathcal{L}}_{\theta,\Omega}\mathcal{D}_{\theta,\Omega} - \widehat{\mathcal{L}}_{\theta,\Omega}\partial_\theta\log\chi_{\theta,\Omega}$, and (A.22a). Also, in the second line, we use (A.11a) to exchange $\chi_{\theta,\Omega}$ with 1.

[88] Indeed, recalling (A.12a),

$$\begin{aligned}\nu_{\theta,\varsigma\Omega}\left(\sum_{k=0}^{n-1}\Omega_{\theta,\varsigma\Omega}\circ f_\theta^k\right)^2 &= n\left\{\nu_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}^2) + 2\sum_{j=1}^{n-1}\nu_{\theta,\varsigma\Omega}\left(\Omega_{\theta,\varsigma\Omega}\circ f_\theta^j\Omega_{\theta,\varsigma\Omega}\right)\right\} \\ &\quad - 2\sum_{j=1}^{n-1} j\nu_{\theta,\varsigma\Omega}\left(\Omega_{\theta,\varsigma\Omega}\circ f_\theta^j\Omega_{\theta,\varsigma\Omega}\right) \\ &= -2\sum_{j=n}^{\infty}\nu_{\theta,\varsigma\Omega}\left(\Omega_{\theta,\varsigma\Omega}\circ f_\theta^j\Omega_{\theta,\varsigma\Omega}\right) - 2\sum_{j=1}^{n-1} j\nu_{\theta,\varsigma\Omega}\left(\Omega_{\theta,\varsigma\Omega}\circ f_\theta^j\Omega_{\theta,\varsigma\Omega}\right)\end{aligned}$$

which is bounded by (A.14) and the spectral gap of $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}$.

It follows that, for any $\varphi \in \mathcal{C}^1$,

$$\begin{aligned}\nu_{\theta,\varsigma\Omega}(\varphi[\Omega_{\theta,\varsigma\Omega} - \phi + \phi\circ f_\theta]) = \\ &= \nu_{\theta,\varsigma\Omega}(\varphi\Omega_{\theta,\varsigma\Omega}) + \lim_{j\to\infty}\sum_{k=0}^{n_j-1}\nu_{\theta,\varsigma\Omega}(\varphi[\Omega_{\theta,\varsigma\Omega}\circ f_\theta - \Omega_{\theta,\varsigma\Omega}]\circ f_\theta^k)\\ &= \lim_{j\to\infty}\nu_{\theta,\varsigma\Omega}(\varphi\Omega_{\theta,\varsigma\Omega}\circ f_\theta^{n_j}) = \lim_{j\to\infty} m_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}\widehat{\mathcal{L}}^{n_j}_{\theta,\varsigma\Omega}(\varphi h_{\theta,\varsigma\Omega}))\\ &= \nu_{\theta,\varsigma\Omega}(\varphi)\nu_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}) = 0.\end{aligned}$$

Since $\mathcal{C}^1$ is dense in L^2, it follows that

$$\Omega_{\theta,\varsigma\Omega} = \phi - \phi\circ f_\theta\,, \qquad m_{\theta,\varsigma\Omega} - \text{a.s.} \tag{A.40}$$

A function with the above property is called a *coboundary*, in this case an L^2 coboundary. In fact, more is true: $\phi \in \mathcal{C}^1$. Indeed, recalling that $\nu_{\theta,\varsigma\Omega}(\phi) = 0$,

$$\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}\, h_{\theta,\varsigma\Omega}) = \widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\phi\, h_{\theta,\varsigma\Omega}) - \phi\, h_{\theta,\varsigma\Omega} = -(\mathbb{1} - \widehat{\mathcal{L}}_{\theta,\varsigma\Omega})(\phi\, h_{\theta,\varsigma\Omega}).$$

Note that the above equation has a unique $L^2(\mathbb{T}, m_{\theta,\varsigma\Omega})$ solution.[89] Hence

$$\phi = -h^{-1}_{\theta,\varsigma\Omega}(\mathbb{1} - \widehat{\mathcal{L}}_{\theta,\varsigma\Omega})^{-1}\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}(\Omega_{\theta,\varsigma\Omega}\, h_{\theta,\varsigma\Omega}) \in \mathcal{C}^1.$$

The proof then follows recalling that $\Omega_{\theta,\varsigma\Omega} = \Omega_\theta - \nu_{\theta,\varsigma\Omega}(\Omega_\theta)$ and setting $\beta = \nu_{\theta,\varsigma\Omega}(\Omega_\theta)$. □

Remark A.17. *Note that the above lemma applies in particular to the case of real potentials (since $m_{\theta,\varsigma\Omega}$ is a measure and $h_{\theta,\varsigma\Omega} > 0$ by Lemma A.1) and for $\varsigma = 0$.*

Following [32] we introduce

Definition A.18. *A real function $A \in \mathcal{C}^1$ is called* aperiodic*, with respect to the dynamics f, if there is no* BV *function β and $\nu_0, \nu_1 \in \mathbb{R}$ such that $A + \beta\circ f - \beta$ is constant on each domain of invertibility of f, and has range in $2\pi\nu_1\mathbb{Z} + \nu_0$.*

Also in the following we will need the, seemingly stronger, condition.

Definition A.19. *A real function $A \in \mathcal{C}^1$ is called* c-constant*, with respect to the dynamics f, if there is a* BV *function β such that $A+\beta\circ f-\beta$ is constant on each domain of invertibility of f.*

We conclude with the announced proof that all the above properties are equivalent in our case of interest.

Lemma A.20. *If $f \in \mathcal{C}^2(\mathbb{T}, \mathbb{T})$ and expanding, then any c-constant zero average function $A \in \mathcal{C}^1(\mathbb{T}, \mathbb{R})$ is necessarily a coboundary.*

Proof. By definition there exists $\beta \in$ BV such that $\alpha = A + \beta\circ f - \beta$ where α is constant on the invertibility domains of f. If we apply the normalized transfer operator $\widehat{\mathcal{L}}$ to the previous relation we have $\beta = (\mathbb{1} - \widehat{\mathcal{L}})^{-1}\widehat{\mathcal{L}}(\alpha - A)$. Note that, by hypothesis, $\widehat{\mathcal{L}}^k(\alpha - A)$ is $\mathcal{C}^1$ except for at most one point (the common image of the boundary points of invertibility domains), for each $k > 0$. Thus β has at most one (jump) discontinuity, which we assume without loss of generality to be at $x = 0$. Since A is smooth on $\mathbb{T}$, calling $\mathcal{P}$ the partition of invertibility domains, we have:

$$0 = \int_0^1 A' dx = \sum_{p\in\mathcal{P}}\int_p (\beta' - (\beta\circ f)')dx = (1-d)(\beta(1^-) - \beta(0^+))$$

[89] Assume otherwise that the equation $\widehat{\mathcal{L}}_{\theta,\varsigma\Omega}\psi = \psi$ has more than one solution in L^2. But for any such solution let $\{\psi_\varepsilon\} \subset \mathcal{C}^1$ be a sequence that converges to ψ in L^2, then it converges in L^1, moreover $m_{\theta,\varsigma\Omega}(|\widehat{\mathcal{L}}^n_{\theta,\varsigma\Omega}(\psi - \psi_\varepsilon)|) \le m_{\theta,\varsigma\Omega}\,|\psi - \psi_\varepsilon|$. Thus, $\psi = \widehat{\mathcal{L}}^n_{\theta,\varsigma\Omega}\psi_\varepsilon + o(1) = h_{\theta,\varsigma\Omega}m_{\theta,\varsigma\Omega}\,\psi + o(1)$. Hence $\psi = h_{\theta,\varsigma\Omega}m_{\theta,\varsigma\Omega}\,\psi$.

where d is the number if invertibility domains of f, i.e. its topological degree. We thus conclude that β is in fact continuous on $\mathbb{T}$ and therefore α has to be constant on $\mathbb{T}$ (hence identically zero). Consequently, β must be smooth and correspondingly A is then a $\mathcal{C}^1$-coboundary. □

A.6. **Non perturbative results.**
In this section we collect some results that hold when ς is large, i.e. well outside the perturbative regime. Such results hold under the generic conditions discussed in the previous section. Even though we have proven that all the conditions are equivalent, we will state the next lemmata under the conditions that are most natural in the proof (and for which the lemma might naturally hold in greater generality).

Lemma A.21. *If $i\Omega_\theta$ is real, of zero average with respect to $\nu_{\theta,0}$ and is aperiodic, then, for all $\varsigma \neq 0$, the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$ when acting on both $\mathcal{C}^1$ and BV is strictly less than 1 and varies continuously with ς unless it is smaller than λ^{-1}.*

Proof. We start by noticing that, for $\varsigma = 0$, the maximal eigenvalue is 1 and all other eigenvalues have modulus strictly smaller than 1. As pointed out in Remark A.12 the relevant spectrum on $\mathcal{C}^1$ and BV is the same. Hence, for small ς we can apply perturbation theory and the first and last of (A.11) imply that $\chi_{\theta,\varsigma\Omega} = -\frac{\Sigma(\theta)}{2}\varsigma^2 + \mathcal{O}(\varsigma^3)$ for some $\Sigma(\theta) > 0$. Note that $\Sigma(\theta)$ is continuous in θ,[90] hence, by Lemma A.16 and Lemma A.20, $\inf_\theta \Sigma(\theta) > 0$. This yields the results for small ς. On the other hand, suppose by contradiction that for $\varsigma \in \mathbb{R} \setminus \{0\}$, $\mathcal{L}_{\theta,\varsigma\Omega} h_* = e^{i\vartheta} h_*$ for some $h_* \in \mathcal{C}^1(\mathbb{T}, \mathbb{C})$ and $\vartheta \in [0, 2\pi)$. Then $|h_*| \leq \mathcal{L}_{\theta,0}|h_*|$, but $\int [\mathcal{L}_{\theta,0}|h_*| - |h_*|] = 0$ implies $|h_*| = \mathcal{L}_{\theta,0}|h_*|$, so $|h_*| = h_0$, the maximal eigenvector of $\mathcal{L}_{\theta,0}$. Accordingly, $h_* = e^{i\beta} h_0$ where β is some real-valued function. Note that we can choose β so that it is smooth a part, at most, a jump of $2\pi n$, for some $n \in \mathbb{N}$, at a fixed point of f_θ. Next, notice that

$$h_0 = e^{-i\beta - i\vartheta} \mathcal{L}_{\theta,\varsigma\Omega} h_* = \mathcal{L}_{\theta,0}\left(e^{i(-i\varsigma\Omega_\theta + \beta - \beta\circ f_\theta - \vartheta)} h_0\right).$$

If we set $\alpha = -i\varsigma\Omega_\theta + \beta - \beta \circ f_\theta - \vartheta$ and we take the real part of the above we get

$$0 = \mathcal{L}_{\theta,0}\left([1 - \cos\alpha] h_0\right).$$

Since the function to which the operator is applied is non negative the range of α must be a subset of $2\pi\mathbb{Z}$ and can have discontinuities only at the preimages of the discontinuity of β.

Finally, the continuity of the maximal eigenvalue follows from standard perturbation theory [34] unless the essential spectral radius coincides with the spectral radius. □

The above theorem implies that the spectral radius is smaller than 1 but does not provide any uniform bound. Since we will need a uniform bound, more information is necessary. This, as already noticed in [7, 28], can be gained by using Dolgopyat's technique [18].

Lemma A.22. *If $i\Omega_\theta$ is real, of zero average with respect to $\nu_{\theta,0}$ and is a non c-constant function with respect to f_θ, then for each $\varsigma_0 > 0$ there exists $\tau \in [0, 1)$, such that, for any $\varsigma \notin [-\varsigma_0, \varsigma_0]$ and $\theta \in \mathbb{T}^1$, the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$, when acting on $\mathcal{C}^1$, is less than τ.*

Proof. By Lemma B.2 and Theorem B.5 of Appendix B, there exists $\varsigma_1, A > 0$ and $\tau \in (0, 1)$ such that for all $\varsigma \notin [-\varsigma_1, \varsigma_1]$, $n \geq A \log|\varsigma|$ and $\theta \in \mathbb{T}^1$,

$$\|\mathcal{L}^n_{\theta,\varsigma\Omega}\|_{1,\varsigma} \leq \tau^n \tag{A.41}$$

[90] This follows from (A.12b) and the perturbation theory in [36].

where $\|f\|_{1,\varsigma} = |f|_\infty + |\varsigma|^{-1}|f'|_\infty$. Next, by Lemma A.21, the spectral radius of $\mathcal{L}_{\theta,\varsigma\Omega}$, for $|\varsigma| \in [\varsigma_0, \varsigma_1]$ is uniformly smaller than one, hence (A.41) is valid also in such a range perhaps modifying A and τ accordingly. Also note that (A.2) implies, for all $n \in \mathbb{N}$, $\|\mathcal{L}^n_{\theta,\varsigma\Omega}\|_{1,\varsigma} \leq C_\# \|\Omega\|_{\mathcal{C}^1}$. In particular, by expanding via the Newman series, for any $z \in \mathbb{C}$, $\tau < |z| \leq 1$:

$$\|(\mathbb{1}z - \mathcal{L}_{\theta,\varsigma\Omega})^{-1}\|_{1,\varsigma} \leq C_\# \left\{ \frac{\|\Omega\|_{\mathcal{C}^1} A \log|\varsigma|}{|z|^{A\log|\varsigma|}} + \frac{(\tau|z|^{-1})^{A\log|\varsigma|}}{1-\tau|z|^{-1}} \right\}. \tag{A.42}$$

Hence the spectral radius is bounded by τ while (A.2) implies that the essential spectral radius is bounded by λ^{-1}. □

Appendix B. Dolgopyat's theory

In this appendix we prove a bound for the transfer operator for large ς. The proof is after the work of Dolgopyat on the decay of correlation in Anosov flows [18]. Unfortunately, we need uniform results in θ, so we cannot use directly the results in [47, 8, 2]. Although the results below can be obtained by carefully tracing the dependence on the parameters in published proofs, e.g., in [2, 8], this is a non trivial endeavor. Therefore we believe the reader will appreciate the following presentation that collects a variety of results and benefits from several simplifications allowed by the fact that we treat smooth maps (even though the arguments can be easily upgraded to cover all the results in the above mentioned papers).

B.1. **Setting.**

Let $f \in \mathcal{C}^r(\mathbb{T}^2, \mathbb{T}^1)$ and $\omega \in \mathcal{C}^{r-1}(\mathbb{T}^2, \mathbb{R})$, $r \geq 2$. We will consider the one parameter family of *dynamics* $f_\theta(x) = f(x, \theta)$, of *potentials* $\Omega_\theta(x) = \omega(x, \theta)$ and the associated *Transfer Operators*

$$\mathcal{L}_{\theta,i\varsigma\Omega_\theta} g(x) = \sum_{y \in f_\theta^{-1}(x)} \frac{e^{i\varsigma\Omega_\theta(y)} g(y)}{f'_\theta(y)}.$$

Also, we assume that there exists $\lambda > 1$ such that $\inf_{x,\theta} f'_\theta(x) \geq \lambda$ (uniform *expansivity*). It is convenient to fix a partition $\mathcal{P}_\theta = \{I_i\}$ of $\mathbb{T}^1$, such that each I_i is a maximal invertibility domain for f_θ. We adopt the convention that the leftmost point of the interval I_1 is always zero, which we assume to be a fixed point for every f_θ.[91] Note however that all the following is independent of such a choice of the partition.

Remark B.1. *Since the maps f_θ are all topologically conjugate (by structural stability of smooth expanding maps), there is a natural isomorphism between $\mathcal{P}_0$ and $\mathcal{P}_\theta$, $\theta \in \mathbb{T}^1$. From now on we will implicitly identify elements of the partitions (and their corresponding inverse branches) for different θ via this isomorphism and will therefore drop the subscript θ when this does not any create confusion.*

At last we require that the Ω_θ satisfies a condition (in general, although not in the present context, see Appendix A.5) stronger than aperiodicity; namely we assume it is not *c-constant* (see Definition A.19).

Let $\mathcal{H}_n$ be the collection of the inverse branches of f^n_θ as defined by the partition $\mathcal{P}$. Note that an element of $\mathcal{H}_n$ can be written as $h_1 \circ \cdots \circ h_n$ where $h_i \in \mathcal{H}_1$, thus $\mathcal{H}_n$ is isomorphic to $\mathcal{H}_1^n$. It is then natural to define $\mathcal{H}_\infty = \mathcal{H}_1^{\mathbb{N}}$.

---

[91] Note that this latter assumption does not imply a loss of generality only if the lines of fixed points of f_θ are homotopic to $\{(0, \theta)\}_{\theta \in \mathbb{T}}$. If not, one can simply consider a finite open cover of the torus (in the θ variable), and make the following argument for each element of the covering.

B.2. **Uniform uniform non integrability (UUNI).**
The first goal of this section is to prove the following fact.

Lemma B.2. *In the hypotheses specified in Subsection B.1 there exist $C_6 > 0$ and $n_0 \in \mathbb{N}$ such that, for each $n \geq n_0$ and $\theta \in \mathbb{T}^1$,*

$$\sup_{h_\theta,\kappa_\theta \in \mathcal{H}_n} \inf_{x\in\mathbb{T}^1} \left| \frac{d}{dx}(\Omega_{n,\theta}\circ h_\theta)(x) - \frac{d}{dx}(\Omega_{n,\theta}\circ \kappa_\theta)(x)\right| \geq C_6, \tag{B.1}$$

where, $\Omega_{n,\theta} := \sum_{k=0}^{n-1}\Omega_\theta \circ f_\theta^k$.

Remark B.3. *Condition* (B.1)*, when referred to a single map, is commonly called* uniform non integrability *(UNI for short) and has been originally introduced by Chernov in* [10]*, a remarkable paper which constituted the first breakthrough in the quantitative study of decay of correlations for flows. The difference here is due to the fact that we have a family of dynamics, rather that only one, and we require a further level of uniformity. The relation between UNI and not being cohomologous to a piecewise constant function was first showed in* [2, Proposition 7.4]*. The above Lemma constitutes a not very surprising extension of the aforementioned proposition.*

Proof of Lemma B.2. Suppose the lemma to be false, then given $a \in \mathbb{N}$ large enough to be chosen later, there exist sequences $\{n_j, \theta_j\}$, $\lambda^{-n_j} < \frac{1-\lambda^{-1}}{2j^a}$, such that for each $h, \kappa \in \mathcal{H}_{\theta_j,n_j}$ there exists $x_{j,h,\kappa} \in \mathbb{T}^1$ such that

$$\left| \frac{d}{dx}(\Omega_{n_j,\theta_j}\circ h)(x_{j,h,\kappa}) - \frac{d}{dx}(\Omega_{n_j,\theta_j}\circ \kappa)(x_{j,h,\kappa})\right| \leq \frac{1}{j^a}.$$

Start by noting that if $h \in \mathcal{H}_n$, then $h = h_1 \circ \cdots \circ h_n$ with $h_i \in \mathcal{H}_1$, and

$$\Omega_{n,\theta}\circ h = \sum_{k=0}^{n} \Omega_\theta \circ h_{k+1}\circ\cdots\circ h_n.$$

Note that all the $\Omega_{n,\theta}\circ h$ belong to $\mathcal{C}^{r-1}(\mathbb{T}^1\setminus\{0\},\mathbb{R})$. For further use, given $h \in \mathcal{H}_\infty$, let us define, for each $n \in \mathbb{N}$,

$$\Xi_{\theta,n,h}(x) = \sum_{k=0}^{n} \Omega_\theta \circ h_k \circ\cdots\circ h_n(x) - \sum_{k=0}^{n} \Omega_\theta \circ h_k \circ\cdots\circ h_n(x_0), \tag{B.2}$$

for some fixed $x_0 \in \mathbb{T}^1$. We remark that, by usual distortion arguments, for each $h \in \mathcal{H}_\infty$, we have $\|\Xi_{\theta,n,h}\|_{\mathcal{C}^1} \leq C_\#$. For each $h \in \mathcal{H}_n$ and $k \leq n$ let $\bar h_k := f_\theta^k \circ h$. Next, for each $j \in \mathbb{N}$ and $p_j \leq n_j$, let $\ell \in \mathcal{H}_{\theta_j,p_j}$. Then, for each $h, \kappa \in \mathcal{H}_{\theta_j,n_j-p_j}$, letting $x_{j,h\circ\ell,\kappa\circ\ell} = z$,

$$\begin{aligned}
\frac{1}{j^a} &\geq \left| \frac{d}{dx}(\Omega_{n_j,\theta_j}\circ h\circ\ell)(z) - \frac{d}{dx}(\Omega_{n_j,\theta_j}\circ \kappa\circ\ell)(z)\right| \\
&= \left| \sum_{k=0}^{n_j-p_j} \Omega'_{\theta_j}\circ \bar h_k\circ\ell(z)\cdot(\bar h_k\circ\ell)'(z) - \Omega'_{\theta_j}\circ\bar\kappa_k\circ\ell(z)\cdot(\bar\kappa_k\circ\ell)'(z)\right| \\
&\geq \left| \frac{d}{dx}(\Omega_{n_j-p_j,\theta_j}\circ h)(\ell(z)) - \frac{d}{dx}(\Omega_{n_j-p_j,\theta_j}\circ \kappa)(\ell(z))\right| \Lambda^{-p_j},
\end{aligned}$$

where $\Lambda = \sup_{\theta,x}|f'_\theta(x)|$. Accordingly, setting $\mathcal{P}_{\theta,n} = \{h(\mathbb{T}^1)\}_{h\in\mathcal{H}_n}$, for each $I \in \mathcal{P}_{\theta_j,p_j}$ we have

$$\sup_{x\in I}\left| \frac{d}{dx}(\Omega_{n_j-p_j,\theta_j}\circ h)(x) - \frac{d}{dx}(\Omega_{n_j-p_j,\theta_j}\circ \kappa)(x)\right| \leq \frac{\Lambda^{p_j}}{j^a} + C_\#\lambda^{-p_j}.$$

Thus, setting $\bar n_j = n_j - p_j$, choosing $p_j = C_\# \log j$ and provided that a has been chosen large enough, we have that for each $h, \kappa \in \mathcal{H}_{\theta_j, \bar n_j}$,

$$\left\| \frac{d}{dx}(\Omega_{\bar n_j,\theta_j} \circ h) - \frac{d}{dx}(\Omega_{\bar n_j,\theta_j} \circ \kappa) \right\|_\infty \leq \frac{C_\#}{j}.$$

Next, for $h_{\theta,1}, \cdots, h_{\theta,m} \in \mathcal{H}_1$, let $C_m = \sup_\theta \sup_{\{h_{\theta,i}\}} \|\partial_\theta[h_{\theta,1} \circ \cdots \circ h_{\theta,m}]\|_\infty$. Then

$$\partial_\theta[h_{\theta,1} \circ \cdots \circ h_{\theta,m}] = [\partial_\theta h_{\theta,1}] \circ h_{\theta,2} \cdots \circ h_{\theta,m} + h'_{\theta,1} \circ h_{\theta,2} \cdots \circ h_{\theta,m} \cdot \{\partial_\theta[h_{\theta,2} \cdots \circ h_{\theta,m}]\}$$

implies $C_m \leq C_\# + \lambda^{-1} C_{m-1}$, that is $C_m \leq C_\#$. This implies that

$$\|\partial_\theta \partial_x \left[\Omega_{n,\theta} \circ h_\theta\right]\|_\infty \leq C_\#.$$

We can then consider a subsequence $\{j_k\}$ such that $\{\theta_{j_k}\}$ converges, let $\bar\theta$ be its limit. Also, without loss of generality, we can assume that $j_k \geq 2C_\# k$ and$|\theta_{j_k} - \bar\theta| \leq k^{-1}$. Thus, for k large enough and for each $h, \kappa \in \mathcal{H}_{\bar n_{j_k}\bar\theta}$, we have

$$\left\| \frac{d}{dx}(\Omega_{\bar n_{j_k},\bar\theta} \circ h)(x) - \frac{d}{dx}(\Omega_{\bar n_{j_k},\bar\theta} \circ \kappa) \right\|_\infty \leq \frac{1}{k}. \tag{B.3}$$

We are now done with the preliminary considerations and we can conclude the argument. Let $\overline{\Xi}_{k,h} = \Xi_{\bar\theta, n_{j_k}, h}$. Since

$$\begin{aligned}|\Omega_\theta \circ h_{k+1} \circ \cdots \circ h_n(x) - \Omega_\theta \circ h_{k+1} \circ \cdots \circ h_n(x_0)| &\leq \|\frac{d}{dx}\Omega_\theta \circ h_{k+1} \circ \cdots \circ h_n\|_\infty \\ &\leq C_\# \lambda^{-n+k},\end{aligned}$$

it follows that the limit $\overline{\Xi}_h = \lim_{k\to\infty} \overline{\Xi}_{k,h}$ exists in the uniform topology. Note that, since the derivative of $\overline{\Xi}_{k,h}$ are uniformly bounded, $\overline{\Xi}$ is Lipschitz in $\mathbb{T}^1 \setminus \{0\}$. In addition, since $\overline{\Xi}_h(x_0) = 0$, equation(B.3) implies, for each $h, \kappa \in \mathcal{H}_{\bar\theta,\infty}$,

$$\left\| \frac{d}{dx}\left[\overline{\Xi}_{k,h} - \overline{\Xi}_{k,\kappa}\right] \right\|_\infty \leq \frac{2}{k}.$$

It follows that $\overline{\Xi}_h = \Phi$ is independent of h. Finally, choose $h \in \mathcal{H}_1$ and $\bar h \in \mathcal{H}_\infty$ such that $\bar h = h \circ h \circ \cdots$, then, if $x \in h^{-1}(\mathbb{T}^1)$,

$$\Xi_{\theta,n,\bar h} \circ f_\theta(x) - \Xi_{\theta,n,\bar h}(x) = \Omega_\theta - \Omega_\theta(h^n(x)).$$

Since f_θ has exactly one fixed point x_I in each $I \in \mathcal{P}_\theta$, $\lim\limits_{n\to\infty} h^n(x) = x_I$, where $I = h(\mathbb{T}^1)$. From the above considerations it follows

$$\Phi \circ f_{\bar\theta} - \Phi = \Omega_{\bar\theta} + \Psi$$

where Ψ is constant on the elements of $\mathcal{P}_{\bar\theta}$ and $\Phi \in \mathrm{BV}$. That is, $\Omega_{\bar\theta}$ is c-constant, contrary to the hypothesis. □

It is now easy to obtain the result we are really interested in.

Corollary B.4. *In the hypotheses specified in Subsection B.1, there exists $n_1 \in \mathbb{N}$ and $h, \kappa \in \mathcal{H}_{n_1}$ such that, for each $n \geq n_1$, $\theta \in \mathbb{T}^1$ and $\ell \in \mathcal{H}_{n-n_1}$*

$$\inf_{x\in\mathbb{T}^1} \left| \frac{d}{dx}(\Omega_{n,\theta} \circ \ell \circ h)(x) - \frac{d}{dx}(\Omega_{n,\theta} \circ \ell \circ \kappa)(x) \right| \geq \frac{C_6}{2}, \tag{B.4}$$

Proof. Let $n_1 \geq n_0$. Then, for $h \in \mathcal{H}_{n_1}$ and $\ell \in \mathcal{H}_{n-n_1}$

$$\Omega_{n,\theta} \circ \ell \circ h = \Omega_{n_1,\theta} \circ h + \Omega_{n-n_1,\theta} \circ \ell \circ h.$$

Thus, by Lemma B.2, we can choose h, κ so that

$$\inf_{x\in\mathbb{T}^1}\left|\frac{d}{dx}\left(\Omega_{n,\theta}\circ\ell\circ h-\Omega_{n,\theta}\circ\ell\circ\kappa\right)(x)\right|\geq\frac{3}{4}C_6-C_\#(|h'|_\infty+|\kappa'|_\infty)$$
$$\geq\frac{3}{4}C_6-C_\#\lambda^{-n_1}.$$

The result follows by choosing n_1 large enough. □

B.3. **Dolgopyat inequality.**
In order to investigate the operator $\mathcal{L}_{\theta,i\varsigma\Omega_\theta}$ for large ς it is convenient to use slightly different norms and operators. The reason is that on the one hand, the main estimate is better done in a ς dependent norm and, on the other hand, it is convenient to have operators that are contractions. Let $\rho := h_{\theta,0}[\int h_{\theta,0}]^{-1}$ be the invariant density of the operator $\mathcal{L}_{\theta,0}$ and define

$$\check{\mathcal{L}}_{\theta,i\varsigma\Omega_\theta}(g)=\rho^{-1}\mathcal{L}_{\theta,i\varsigma\Omega_\theta}(\rho g)\ ,\qquad\qquad \|g\|_{1,\varsigma}=\|g\|_{\mathcal{C}^0}+\frac{\|g'\|_{\mathcal{C}^0}}{|\varsigma|},$$

Then we have[92]

$$\text{(B.5)}\qquad \|\check{\mathcal{L}}_{\theta,i\varsigma\Omega_\theta}(g)\|_{\mathcal{C}^0}\leq\|g\|_{\mathcal{C}^0},\qquad\qquad \|\check{\mathcal{L}}_{\theta,i\varsigma\Omega_\theta}(g)\|_{L^1_\rho}\leq\|g\|_{L^1_\rho},$$

as announced. Moreover, by (A.2), it follows that, for $\varsigma\geq\varsigma_0$, with $\varsigma_0>0$ large enough, and $n\in\mathbb{N}$,

$$\text{(B.6)}\qquad \|\check{\mathcal{L}}^n_{\theta,i\varsigma\Omega_\theta}(g)\|_{1,\varsigma}\leq C_\#\lambda^{-n}\|g\|_{1,\varsigma}+B_0\|g\|_{\mathcal{C}^0},$$

for a fixed constant B_0. Fix $\bar\lambda\in(\lambda,1)$ and choose $n_2\in\mathbb{N}$ such that $C_\#\lambda^{-n_2}\leq\bar\lambda^{-n_2}$. Also, for future use, we chose n_2 so that $\bar\lambda^{-n_2}\leq\frac12$. Iterating the above inequalities by steps of length $n_3\in\mathbb{N}$, $n_3\geq n_2$, we have

$$\text{(B.7)}\qquad \|\check{\mathcal{L}}^{kn_3}_{\theta,i\varsigma\Omega_\theta}(g)\|_{1,\varsigma}\leq\bar\lambda^{-kn_3}\|g\|_{1,\varsigma}+B_0\sum_{j=0}^{k-1}\bar\lambda^{-(k-j-1)n_3}\|\check{\mathcal{L}}^{jn_3}_{\theta,i\varsigma\Omega_\theta}(g)\|_{\mathcal{C}^0}.$$

Theorem B.5. *If condition (B.4) is satisfied, then there exists $A,B>0$, and $\gamma<1$ such that for all $|\varsigma|\geq B$ and $n\geq A\log|\varsigma|$, we have*

$$\sup_{\theta\in\mathbb{T}^1}\|\check{\mathcal{L}}^n_{\theta,i\varsigma\Omega_\theta}\|_{1,\varsigma}\leq\gamma^n.$$

Remark B.6. *In fact, Theorem B.5, for fixed θ, is a special case of* [8, Theorem 1.1]. *To be precise,* [8, Theorem 1.1] *is stated for a single map and with strictly positive roof functions (a role here played by Ω_θ). The latter can easily be arranged by multiplying the transfer operator by $e^{i2\|\Omega_\theta\|}$, which does not change the norm. In addition, a careful look at the proof should show that A,B,γ depend on the map and potential only via n_1, C_6 of (B.4) and $\|f'_\theta\|_\infty$, $\|(f'_\theta)^{-1}\|_\infty$, $\|f''_\theta(f'_\theta)^{-1}\|_\infty$, $\|\Omega_\theta\|_{\mathcal{C}^2}$ which, in the present case, are all uniformly bounded. Nevertheless, we think the reader may appreciate the following simpler, self-contained, proof rather than being referred to the guts of* [8].

Proof of Theorem B.5. For each $g\in\mathcal{C}^1$ set $g_k=\check{\mathcal{L}}^{kn_3}_{\theta,i\varsigma\Omega_\theta}g$, with $n_3\geq n_1$ from Corollary B.4 and $n_3\geq n_2$ as in equation (B.7). The basic idea, going back to Dolgopyat [19], is to construct iteratively functions $u_k\in\mathcal{C}^1(\mathbb{T}^1,\mathbb{R}_{\geq0})$ such that $|g_k(x)|\leq u_k(x)$ for all $k\in\mathbb{N}$ and $x\in\mathbb{T}^1$ and on which one has good bounds. More precisely:

92 The first follows trivially from $|\check{\mathcal{L}}_{\theta,i\varsigma\Omega_\theta}g|\leq\|g\|_{L^\infty}\check{\mathcal{L}}_{\theta,0}1$ and $\mathcal{L}_{\theta,0}\rho=\rho$. The second from the standard $\|\mathcal{L}_{\theta,i\varsigma\Omega_\theta}(g)\|_{L^1}\leq\|g\|_{L^1}$.

Lemma B.7. *There exists constants $K, \beta, B_1, \varsigma_0 > 1$, $\tau > 0$, $n_3 \geq \max\{n_1, n_2\}$ and, for all $g \in \mathcal{C}^1$ with $\|g'\|_{\mathcal{C}^0} \leq \beta|\varsigma|^{-1}\|g\|_{\mathcal{C}^0}$, functions $\{\Gamma_{g,k}\}_{k\in\mathbb{N}} \in \mathcal{C}^1(\mathbb{T}^1, [4/5, 1])$ such that, for all $|\varsigma| \geq \varsigma_0$ and $k \in \mathbb{N}$,*

$$\|\Gamma'_{g,k}\|_{L^\infty} \leq B_1|\varsigma|, \tag{B.8}$$

and, setting $u_0 = \|g\|_\infty + \beta^{-1}|\varsigma|^{-1}\|g'\|_\infty$ and $u_{k+1} = \check{\mathcal{L}}^{n_3}_{\theta,0}(\Gamma_{g,k}u_k)$, we have, for any $x \in \mathbb{T}^1$,

$$\max\{|u_k'(x)|, |g_k'(x)|\} \leq \beta|\varsigma|u_k(x)\ , \qquad |g_k(x)| \leq u_k(x)$$

and, for any $I = [a_1, a_2] \subset \mathbb{T}^1$ so that $|a_2 - a_1| \geq 4K|\varsigma|^{-1}$:

$$\int_I \check{\mathcal{L}}^{n_3}_{\theta,0}\Gamma^2_{g,k} \leq e^{-3\tau}|I|.$$

Let us postpone the proof of Lemma B.7 and see how it implies the wanted result. First of all, note that if $\|g_k'\|_{\mathcal{C}^0} \geq \beta|\varsigma|^{-1}\|g_k\|_{\mathcal{C}^0}$, then equation (B.7) implies $\|g_{k+1}\|_{1,\varsigma} \leq \gamma\|g_k\|_{1,\varsigma}$, provided β has been chosen large enough. We can thus assume $\|g'\|_{\mathcal{C}^0} \leq \beta|\varsigma|^{-1}\|g\|_{\mathcal{C}^0}$ without loss of generality.

Next, note that, for any $j_0 \in \mathbb{N}$, by equation (A.2) and choosing $\bar\lambda$ as in equation (B.7), we can write

$$\left|\frac{d}{dx}\check{\mathcal{L}}^{j_0n_3}_{\theta,0}(u_k)\right| \leq \bar\lambda^{-j_0n_3}\check{\mathcal{L}}^{j_0n_3}_{\theta,0}(|u_k'|) + B\check{\mathcal{L}}^{j_0n_3}_{\theta,0}(u_k) \leq (\bar\lambda^{-j_0n_3}\beta|\varsigma| + B)\check{\mathcal{L}}^{j_0n_3}_{\theta,0}(u_k).$$

By eventually increasing ς_0, we can choose j_0 so that, for all $\varsigma \geq \varsigma_0$,

$$\left|\frac{d}{dx}\rho\left[\check{\mathcal{L}}^{j_0n_3}_{\theta,0}(u_k)\right]^2\right| \leq \frac{\tau|\varsigma|}{4K}\rho\left[\check{\mathcal{L}}^{j_0n_3}_{\theta,0}(u_k)\right]^2. \tag{B.9}$$

Thus, given any partition $\{p_m\}$ of $\mathbb{T}^1$ in intervals of size between $3K|\varsigma|^{-1}$ and $4K|\varsigma|^{-1}$ we have

$$\begin{aligned}\int_{\mathbb{T}^1} u^2_{k+j_0}\rho &\leq \int_{\mathbb{T}^1}\left\{\check{\mathcal{L}}^{n_3}_{\theta,0}[\Gamma_{k+j_0-1}(\check{\mathcal{L}}^{(j_0-1)n_3}_{\theta,0}u_k)]\right\}^2\rho \leq \int_{\mathbb{T}^1}\rho\check{\mathcal{L}}^{n_3}_{\theta,0}\Gamma^2_{k+j_0-1}\cdot\check{\mathcal{L}}^{j_0n_3}_{\theta,0}u_k^2\\ &\leq \sum_m\int_{p_m}\check{\mathcal{L}}^{n_3}_{\theta,0}\Gamma^2_{k+j_0-1}\frac{e^\tau}{|p_m|}\int_{p_m}\rho\check{\mathcal{L}}^{j_0n_3}_{\theta,0}u_k^2\\ &\leq e^{-2\tau}\int_{\mathbb{T}^1}\rho\check{\mathcal{L}}^{j_0n_3}_{\theta,0}u_k^2 = e^{-2\tau}\int_{\mathbb{T}^1}u_k^2\,\rho,\end{aligned}$$

where in the second inequality of the first line we have used Schwarz inequality with respect to the sum implicit in $\check{\mathcal{L}}^{n_3}_{\theta,0}$ and $\check{\mathcal{L}}^{n_3(j_0-1)}_{\theta,0}$; the second line follows from (B.9); the first inequality of the third line follows from the last assertion of Lemma B.7, while the last inequality follows from the well known contraction of $\check{\mathcal{L}}_{\theta,0}$ in L^1_ρ.

Finally, iterating the above equation, we obtain

$$\|u_{kj_0}\|_{L^2_\rho} \leq e^{-k\tau}\|u_0\|_{L^2_\rho}.$$

Accordingly, there exists $A > 0$ such that, for all $n \geq \frac{A}{2}\log|\varsigma|$ we have $\bar\lambda \leq |\varsigma|^{-2}$ and

$$\|g_n\|_{L^2_\rho} \leq \|u_n\|_{L^2_\rho} \leq |\varsigma|^{-4}\|u_0\|_{L^2_\rho}.$$

The above equation together with (B.7) and the fact that, for all $\tilde g \in \mathcal{C}^1$,[93]

$$\|\tilde g\|_\infty \leq \|\tilde g\|^{\frac12}_{L^2_\rho}\left[\|\tilde g\|_{L^2_\rho} + 2\|\tilde g'\rho^{-1}\|_\infty\right]^{\frac12}$$

yields $\|\check{\mathcal{L}}^n_{\theta,i\varsigma\Omega_\theta}g\|_{1,\varsigma} \leq |\varsigma|^{-1}\|g\|_{1,\varsigma}$ for all $n \in [A\log|\varsigma|, 2A\log|\varsigma|] \cap \mathbb{N}$. The latter readily implies Theorem B.5. □

[93] Indeed, $|\tilde g(x)|^2 \leq \|\tilde g\|^2_{L^2_\rho} + 2\int_{\mathbb{T}^1}|\tilde g|\,|\tilde g'|$.

Proof of Lemma B.7. Since $u_0 = \|g\|_\infty + \beta^{-1}|\varsigma|^{-1}\|g'\|_\infty$, trivially, $\max\{|u_0'|, |g_0'|\} \le \beta|\varsigma|u_0$ and $|g_0(x)| \le u_0(x)$ for all $x \in \mathbb{T}^1$. Suppose, by induction, that $\max\{|u_k'|, |g_k'|\} \le \beta|\varsigma|u_k$, and $|g_k(x)| \le u_k(x)$ for all $x \in \mathbb{T}^1$, then (A.2) implies

$$\begin{aligned}|g_{k+1}'(x)| &\le \bar\lambda^{-n_3}(\check{\mathcal{L}}_{\theta,0}^{n_3}|g_k'|)(x) + B|\varsigma|(\check{\mathcal{L}}_{\theta,0}^{n_3}|g_k|)(x)\\ &\le \beta|\varsigma|\left[\bar\lambda^{-n_3} + B\beta^{-1}\right](\check{\mathcal{L}}_{\theta,0}^{n_3}u_k)(x) \le \beta|\varsigma|\frac{5}{4}\left[\bar\lambda^{-n_3} + B\beta^{-1}\right]u_{k+1}\end{aligned} \tag{B.10}$$

where we have assumed the existence of the wanted $\Gamma_{g,k}$ that remains to be constructed. By choosing β large enough it follows

$$|g_{k+1}'(x)| \le \beta|\varsigma|u_{k+1}.$$

The proof of the analogous inequality for u_k being similar, but it uses B.8.

Next, let $h_*, \kappa_* \in \mathcal{H}_{n_3}$ be two branches satisfying (B.4), whose existence follows by Corollary B.4, and let us define the set $\widehat{\mathcal{H}} = \mathcal{H}_{n_3} \setminus \{h_*, \kappa_*\}$. Then,

$$|g_{k+1}(x)| \le \sum_{h\in\widehat{\mathcal{H}}} \frac{(u_k\rho)\circ h(x)}{\rho(x)(f_\theta^{n_3})'\circ h(x)} + \left|\sum_{h\in\{h_*,\kappa_*\}} \frac{e^{i\varsigma\Omega_{n_3,\theta}\circ h(x)}(\rho g_k)\circ h(x)}{\rho(x)\,(f_\theta^{n_3})'\circ h(x)}\right|. \tag{B.11}$$

To conclude we need a sharp estimate for the second term in (B.11), where a cancellation may take place. To this end it is helpful to introduce a partition of unity. This can be obtained by a function $\phi \in \mathcal{C}^2(\mathbb{R}_{\ge 0}, [0,1])$ such that $\phi(x) = 1$ for $|x| \le \frac12$ and $\phi(x) = 0$ for $|x| \ge 1$ and $1 = \sum_{n\in\mathbb{N}} \phi(x-n)$, for all $x \in \mathbb{R}$. We then define $L_\varsigma = \lfloor K^{-1}|\varsigma|\rfloor$, $\psi_m(x) = \phi(L_\varsigma x - m)$. Note that, by construction, $\sum_{m=0}^{L_\varsigma-1}\psi_m = 1$ (here we are interpreting the ψ_m as functions on $\mathbb{T}^1$). Let $I_m = \operatorname{supp}\psi_m$ and let x_m be its middle point. Note that $\frac{K}{|\varsigma|} \le |I_m| \le \frac{2K}{|\varsigma|}$.

To continue, for each $m \in \{1, \cdots, L_\varsigma - 1\}$, we must consider two different cases. First suppose that there exists $\bar h_m \in \{h_*, \kappa_*\}$ such that $g_k(\bar h_m(x_m)) \le \frac12 u_k(\bar h_m(x_m))$. Note that, for $z \in h(I_m)$,

$$e^{-2K\lambda^{-n_3}\beta}u_k(z) \le u_k(h(x_m)) \le e^{2K\lambda^{-n_3}\beta}u_k(z)$$

hence

$$|g_k'(z)| \le \beta|\varsigma|u_k(z) \le \frac32\beta|\varsigma|u_k(\bar h_m(x_m))$$

provided $K\lambda^{-n_3}\beta \le \frac18$. Which implies, for all $x \in I_m$,

$$|g_k \circ \bar h_m(x)| \le \frac45|u_k \circ \bar h_m(x)|, \tag{B.12}$$

provided $K\lambda^{-n_3}\beta \le \frac{1}{90}$.

Second, suppose that, for each $h \in \{h_*, \kappa_*\}$, $|g_k(h(x_m))| \ge \frac12 u_k(h(x_m))$. Then

$$|g_k'(z)| \le \beta|\varsigma|u_k(z) \le 2\beta|\varsigma|u_k(h(x_m)) \le 4\beta|\varsigma|\,|g_k(h(x_m))|.$$

The above implies $\frac12|g_k(h(x_m))| \le |g_k(z)| \le 2|g_k(h(x_m))|$ provided $K\lambda^{-n_3}\beta \le \frac18$.

Thus, setting $A_h = \frac{(\rho g_k)\circ h(x)}{\rho(x)\,(f_\theta^{n_3})'\circ h(x)}$, we have

$$\begin{aligned}|A_h'(x)| &\le C_\#(h'(x)^2\beta|\varsigma| + h'(x))|u_k(h(x))| \le C_\#(h'(x)^2\beta|\varsigma| + h'(x))|g_k(h(x))|\\ &\le C_\#(h'(x)\beta|\varsigma| + 1)|A_h(x)|.\end{aligned}$$

Defining $A_h = e^{i\theta_h}B_h$, with θ_h, B_h real and $B_h \ge 0$, we have

$$|A_h'| \ge \frac{1}{\sqrt2}\left(|\theta_h' B_h| + |B_h'|\right).$$

The above implies that, given β, we can chose n_3 and ς_0 large enough so that

$$|\theta_h'(x)| \le C_\#(h'(x)\beta|\varsigma| + 1) \le \frac{C_6|\varsigma|}{32\pi},$$
$$|B_h'(x)| \le \frac{C_6|\varsigma|}{32\pi} B_h(x).$$

Hence, setting $\Theta := \Omega_{n_3,\theta} \circ h_* - \Omega_{n_3,\theta} \circ \kappa_* + |\varsigma|^{-1}(\theta_{h_*} - \theta_{\kappa_*})$,

$$C_\# \ge \left|\frac{d}{dx}\Theta\right| \ge \frac{C_6}{4}.$$

In turns, this implies that the phase Θ has at least one full oscillation in I_m provided $K \ge \frac{8\pi}{C_6}$. Also, $\inf_{I_m} B_h \ge \frac{1}{2}\sup_{I_m}|B_h|$, provided $K \le \frac{10\pi}{C_6}$. Next, suppose that $\|B_{h_*}\|_\infty \ge \|B_{\kappa_*}\|_\infty$, (hence $4|B_{h_*}(x)| \ge |B_{\kappa_*}(x)|$), and set $\bar{h}_m = \kappa_*$, the other case being treated exactly in the same way (interchanging the role of h_* and κ_*, hence setting $\bar{h}_m = \kappa_*$). Given the above notation, the last term of (B.11) reads

$$|B_{h_*} - e^{i\varsigma\Theta}B_{\kappa_*}| = \left[B_{h_*}^2 + B_{\kappa_*}^2 - 2B_{h_*}B_{\kappa_*}\cos\varsigma\Theta\right]^{\frac{1}{2}}.$$

It follows that there exists a constant $C_7 > 0$ and intervals $J \subset \overset{\circ}{I}_m$, $\frac{4\pi}{C_6|\varsigma|} \ge |J| \ge \frac{C_7}{|\varsigma|}$ on which $\cos\varsigma\Theta \ge 0$. Then, on each such interval J,

$$|B_{h_*} - e^{i\varsigma\Theta}B_{\kappa_*}| \le \left[B_{h_*}^2 + B_{\kappa_*}^2\right]^{\frac{1}{2}} \le B_{h_*} + \frac{4}{5}B_{\kappa_*}.$$

We can then define $\Xi_m \in \mathcal{C}^\infty(I_m, [\frac{4}{5}, 1])$ such that $\Xi_m(x) = 1$ outside the intervals J, $\Xi_m(x) = \frac{4}{5}$ on the mid third of each J and $\|\Xi_m\|_{\mathcal{C}^1} \le C_\#|\varsigma|$. It follows

$$\text{(B.13)} \qquad \begin{aligned} |B_{h_*} - e^{i\varsigma\Theta}B_{\kappa_*}| &\le \left|\frac{(g_k\rho)\circ h_*}{(\rho f_\theta^{n_3})'\circ h_*}\right| + \Xi_m\left|\frac{(g_k\rho)\circ \kappa_*}{(\rho f_\theta^{n_3})'\circ \kappa_*}\right| \\ &\le \frac{(u_k\rho)\circ h_*}{(\rho f_\theta^{n_3})'\circ h_*} + \Xi_m\frac{(u_k\rho)\circ \bar{h}_m}{(\rho f_\theta^{n_3})'\circ \bar{h}_m}. \end{aligned}$$

We can finally define the function $\Gamma_{g,k} \in \mathcal{C}^1(\mathbb{T}^1, [0,1])$ as

$$\Gamma_{g,k}(x) = \sum_{m=0}^{L_\varsigma - 1} \psi_m \circ f^{n_3}(x)\Gamma_{k,m}(x).$$

where

$$\Gamma_{k,m}(x) = \begin{cases} 1 & \text{if } x \in h(I_m), h \ne \bar{h}_m, \\ \Xi_m \circ f^{n_3}(x) & \text{if } x \in \bar{h}_m(I_m). \end{cases}$$

Note that with the above definition, condition (B.8) is satisfied. Also, by equations (B.12) and (B.13), it follows $|g_{k+1}| \le u_{k+1}$.

Finally, we must check the last claim of the Lemma. Note that it suffices to consider intervals I of size between $4K|\varsigma|^{-1}$ and $8K|\varsigma|^{-1}$.

$$\begin{aligned} \int_I \check{\mathcal{L}}_{\theta,0}^{n_3}\Gamma_{g,k}^2 &\le \sum_m \int_I \psi_m \cdot \check{\mathcal{L}}_{\theta,0}^{n_3}(\Gamma_{k,m}^2) \\ &\le \sum_m \left[\sum_{h \in \mathcal{H}_{n_3}/\{\bar{h}_m\}} \int_I \psi_m \frac{\rho\circ h \cdot h'}{\rho} + \int_I \psi_m \Xi_m^2 \frac{\rho \circ \bar{h}_m \cdot \bar{h}_m'}{\rho}\right]. \end{aligned}$$

Note that there exists at least one m_* such that $I_{m_*} \subset I$. Moreover, at least $\frac{C_7}{3K}$ of I_{m_*} (hence at least $\frac{C_7}{24K}$ of I) is covered by intervals J on which $\Xi_{m_*} = 4/5$ and

$\psi_{m_*} = 1$. Let J_* be the union of such intervals. Since $\frac{\rho(x)}{\rho(y)} \le e^{C_\#|x-y|}$, for each $\eta \in (0,1)$,

$$\begin{aligned}\int_I \psi_{m_*}\Xi_{m_*}\frac{\rho\circ\bar h_{m_*}\cdot\bar h'_{m_*}}{\rho} &\le \int_{I\setminus J_*}\psi_{m_*}\frac{\rho\circ\bar h_{m_*}\cdot\bar h'_{m_*}}{\rho} + \frac{16}{25}\int_{J_*}\frac{\rho\circ\bar h_{m_*}\cdot\bar h'_{m_*}}{\rho}\\ &\le (1-\eta)\int_{I\setminus J_*}\psi_{m_*}\frac{\rho\circ\bar h_{m_*}\cdot\bar h'_{m_*}}{\rho} + \left(\eta\frac{|I|}{|J_*|}+\frac{16}{25}\right)\int_{J_*}\psi_{m_*}\frac{\rho\circ\bar h_{m_*}\cdot\bar h'_{m_*}}{\rho}.\end{aligned}$$

Thus, choosing $\eta = \frac{9C_7}{25(C_7+24K)}$ we have

$$\int_I \psi_{m_*}\Xi_{m_*}\frac{\rho\circ\bar h_{m_*}\cdot\bar h'_{m_*}}{\rho} \le (1-\eta)\int_I\psi_{m_*}\frac{\rho\circ\bar h_{m_*}\cdot\bar h'_{m_*}}{\rho}.$$

Also note that there exists $M>0$ such that, for all $h\in\mathcal H_{n_3}$ and $m\in\{1,\cdots,L_\varsigma-1\}$,

$$\int_I \psi_m\frac{\rho\circ hh'}{\rho} \le M\int_I\psi_{m_*}\frac{\rho\circ\bar h_{m_*}\bar h'_{m_*}}{\rho}.$$

Moreover, note that I can intersect at most 9 intervals I_m. By an argument similar to the above it then follows that there exists $\tau>0$ such that

$$\begin{aligned}\int_I \check{\mathcal L}^{n_3}_{\theta,0}\Gamma^2_{g,k} &\le e^{-3\tau}\sum_m\sum_{h\in\mathcal H_{n_3}}\int_I\psi_m\frac{\rho\circ h\cdot h'}{\rho}\\ &= e^{-3\tau}\int_I\check{\mathcal L}^{n_3}_{\theta,0}1 = e^{-3\tau}|I|.\end{aligned}$$

□

Appendix C. A tedious computation

Here we perform explicitly the computations that lead to (8.25). These are simple but tedious computations that, in subsequent occasions, will be left to the reader. We provide this appendix so that the reader can see precisely how such computations are done and be able to reproduce them when equally detailed proofs are not provided.

Let us recall the starting point (see (8.12)):

$$\begin{aligned}\mathbb M_k &= \widehat\Xi(t-\varepsilon S_k,\bar\theta_{L_k})\left\{\sum_{j=0}^{L_k-1}\Xi_{j,L_k}\left[\hat\omega(x_j,\theta_j)-\frac{\varepsilon}{2}\bar\omega'(\bar\theta_j)\bar\omega(\bar\theta_j)\right]\right\}+\varepsilon\,\mathbb C_k\\ \mathbb C_k &= \frac12\widehat\Xi(t-\varepsilon S_k,\bar\theta_{L_k})P(t-\varepsilon S_k,\bar\theta_{L_k})\left[\sum_{j=0}^{L_k-1}\Xi_{j,L_k}\hat\omega(x_j,\theta_j)\right]^2.\end{aligned}$$

Recall, as already observed in Section 2, that $\bar\omega\in\mathcal C^{3-\alpha}$ for any $\alpha>0$. Let us compute term by term.

$$\begin{aligned}&\widehat\Xi(t-\varepsilon S_k,\bar\theta_{L_k})\sum_{j=0}^{L_k-1}\Xi_{j,L_k}\hat\omega(x_j,\theta_j) = \widehat\Xi(t-\varepsilon S_k,\bar\theta^*_{\ell,L_k})\sum_{j=0}^{L_k-1}\Xi^*_{\ell,j,L_k}\hat\omega(x_j,\theta_j)\\ &+\partial_\theta\widehat\Xi(t-\varepsilon S_k,\bar\theta^*_{\ell,L_k})\partial_\theta\bar\theta(\varepsilon L_k,\theta^*_\ell)\sum_{j=0}^{L_k-1}\Xi^*_{\ell,j,L_k}\hat\omega(x_j,\theta_j)(\theta_0-\theta^*_\ell)\\ &+\varepsilon\widehat\Xi(t-\varepsilon S_k,\bar\theta^*_{\ell,L_k})\sum_{j=0}^{L_k-1}\sum_{l=0}^{L_k-1}\frac{\Xi^*_{\ell,j,L_k}}{\left(1+\varepsilon\bar\omega'(\bar\theta^*_{\ell,l})\right)}\bar\omega''(\bar\theta^*_{\ell,l})\partial_\theta\bar\theta(\varepsilon l,\theta^*_\ell)\hat\omega(x_j,\theta_j)(\theta_0-\theta^*_\ell)\\ &+\mathcal O(\varepsilon^2\delta_{\mathbb C}^2L_k+\varepsilon^{3-\delta_*}\delta_{\mathbb C}^{2-\delta_*}L_k^2),\end{aligned}$$

where we have used that $|\theta_0-\theta^*_\ell|\le C_\#\varepsilon\delta_{\mathbb C}$, by $\partial_\theta\widehat\Xi$ we mean the derivative with respect to the second variable and we have used the definition (8.8) of $\Xi_{i,j}$. Now note that the term on the second line of the previous equation is of type (recall Notation 8.8)

$$\varepsilon \mathfrak{K}_{\ell,2}^{k,1} = \varepsilon \sum_{i_1,i_2} \mathfrak{C}_{\ell,2,(i_1,i_2)}^{k,1} A_{1,i_1} A_{2,i_2}$$

where we have $\mathfrak{C}_{\ell,2,(i_1,i_2)}^{k,1} = 0$, if $i_1 \neq 0$, and $A_{1,0}(x,\theta) = \varepsilon^{-1} \cdot (\theta - \theta_\ell^*)$ while $A_{2,i_2} = \hat{\omega}(x,\theta)$. Note that, provided C^* has been chosen large enough, $\|A_{j,i_j}\|_{\mathcal{C}^1} \leq C^*$, as required. In fact, the terms has the extra property $\|\mathfrak{K}_{\ell,2}^{k,1}\|_{\mathcal{C}^0} \leq C^* \delta_{\mathbb{C}} L_k$, but we will not use this in the following. Similar arguments show that the term on the third line is of type $\varepsilon^2 \mathfrak{K}_{\ell,3}^{k,2}$: we can thus subsume both terms as a $\varepsilon \mathfrak{K}_{\ell,3}^{k,1}$ term.

Next,

$$\begin{aligned}
&- \frac{\varepsilon}{2} \widehat{\Xi}(t - \varepsilon S_k, \bar{\theta}_{L_k}) \Xi_{j,L_k} \sum_{j=0}^{L_k-1} \bar{\omega}'(\bar{\theta}_j) \bar{\omega}(\bar{\theta}_j) \\
&= - \frac{\varepsilon}{2} \widehat{\Xi}(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) \Xi_{\ell,j,L_k}^* \sum_{j=0}^{L_k-1} \bar{\omega}'(\bar{\theta}_{\ell,j}^*) \bar{\omega}(\bar{\theta}_{\ell,j}^*) \\
&- \frac{\varepsilon}{2} \partial_\theta \widehat{\Xi}(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) \partial_\theta \bar{\theta}(\varepsilon L_k, \theta_\ell^*) \Xi_{\ell,j,L_k}^* \sum_{j=0}^{L_k-1} \bar{\omega}'(\bar{\theta}_{\ell,j}^*) \bar{\omega}(\bar{\theta}_{\ell,j}^*)(\theta_0 - \theta_\ell^*) \\
&- \frac{\varepsilon}{2} \widehat{\Xi}(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) \Xi_{\ell,j,L_k}^* \sum_{j=0}^{L_k-1} \left[\bar{\omega}''(\bar{\theta}_{\ell,j}^*) \bar{\omega}(\bar{\theta}_{\ell,j}^*) + \bar{\omega}'(\bar{\theta}_{\ell,j}^*)^2\right] \partial_\theta \bar{\theta}(\varepsilon j, \theta_\ell^*)(\theta_0 - \theta_\ell^*) \\
&+ \mathcal{O}(L_k^2 \varepsilon^3 \delta_{\mathbb{C}} + \varepsilon^{3-\delta_*} \delta_{\mathbb{C}}^{2-\delta_*} L_k).
\end{aligned}$$

The terms in the second and third line are of type $\varepsilon^2 \mathfrak{K}_{\ell,1}^{k,1}$, which is a bound smaller that the one for the correlation terms already obtained. Finally, for the last term we have

$$\begin{aligned}
\mathbb{C}_k &= \frac{1}{2} \widehat{\widehat{\Xi}}(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) P(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) \left[\sum_{j=0}^{L_k-1} \Xi_{\ell,j,L_k}^* \hat{\omega}(x_j, \theta_j)\right]^2 \\
&+ \frac{1}{2} \partial_\theta \widehat{\Xi}(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) \partial_\theta \bar{\theta}(\varepsilon L_k, \theta_\ell^*) P(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) \left[\sum_{j=0}^{L_k-1} \Xi_{\ell,j,L_k}^* \hat{\omega}(x_j, \theta_j)\right]^2 (\theta_0 - \theta_\ell^*) \\
&+ \frac{1}{2} \partial_\theta \widehat{\Xi}(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) P(t - \varepsilon S_k, \bar{\theta}_{\ell,L_k}^*) \partial_\theta \bar{\theta}(\varepsilon L_k, \theta_\ell^*) \left[\sum_{j=0}^{L_k-1} \Xi_{\ell,j,L_k}^* \hat{\omega}(x_j, \theta_j)\right]^2 (\theta_0 - \theta_\ell^*) \\
&+ \mathcal{O}(L_k^3 \varepsilon^2 \delta_{\mathbb{C}} + \varepsilon^2 L_k^2 \delta_{\mathbb{C}}^2).
\end{aligned}$$

Note that the first two lines can be interpreted as a $\mathfrak{K}_{\ell,3}^{k,2}$ term; also, any previous correlation term can be interpreted as a term of this type. Collecting the above facts, and recalling the constraints on L_* and $\delta_{\mathbb{C}}$, we obtain (8.25).

References

[1] D. V. Anosov. Averaging in systems of ordinary differential equations with rapidly oscillating solutions. *Izv. Akad. Nauk SSSR Ser. Mat.*, 24:721–742, 1960.

[2] Artur Avila, Sébastien Gouëzel, and Jean-Christophe Yoccoz. Exponential mixing for the Teichmüller flow. *Publ. Math. Inst. Hautes Études Sci.*, (104):143–211, 2006.

[3] V. I. Bakhtin. Cramér asymptotics in the averaging method for systems with fast hyperbolic motions. *Tr. Mat. Inst. Steklova*, 244(Din. Sist. i Smezhnye Vopr. Geom.):65–86, 2004.

[4] Victor I. Bakhtin. Cramér's asymptotics in systems with fast and slow motions. *Stoch. Stoch. Rep.*, 75(5):319–341, 2003.

[5] Viviane Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.
[6] Viviane Baladi and Masato Tsujii. Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms. *Ann. Inst. Fourier (Grenoble)*, 57(1):127–154, 2007.
[7] Viviane Baladi and Brigitte Vallée. Euclidean algorithms are Gaussian. *J. Number Theory*, 110(2):331–386, 2005.
[8] Viviane Baladi and Brigitte Vallée. Exponential decay of correlations for surface semi-flows without finite Markov partitions. *Proc. Amer. Math. Soc.*, 133(3):865–874 (electronic), 2005.
[9] Michael Blank, Gerhard Keller, and Carlangelo Liverani. Ruelle-Perron-Frobenius spectrum for Anosov maps. *Nonlinearity*, 15(6):1905–1973, 2002.
[10] Nikolai I. Chernov. Markov approximations and decay of correlations for Anosov flows. *Ann. of Math. (2)*, 147(2):269–324, 1998.
[11] D. Collier and I. D. Morris. Approximating the maximum ergodic average via periodic orbits. *Ergodic Theory Dynam. Systems*, 28(4):1081–1090, 2008.
[12] Jacopo De Simoi and Carlangelo Liverani. Fast-slow partially hyperbolic systems beyond averaging. statistical properties. *Inventiones*. Online first. Preprint arXiv:1408.5454.
[13] Jacopo De Simoi and Carlangelo Liverani. The martingale approach after Varadhan and Dolgopyat. In *Hyperbolic dynamics, fluctuations and large deviations*, volume 89 of *Proc. Sympos. Pure Math.*, pages 311–339. Amer. Math. Soc., Providence, RI, 2015.
[14] Jacopo De Simoi, Carlangelo Liverani, Christophe Poquet, and Denis Volk. Fast–Slow Partially Hyperbolic Systems Versus Freidlin–Wentzell Random Systems. *J. Stat. Phys.*, 166(3-4):650–679, 2017.
[15] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
[16] Mark F. Demers and Carlangelo Liverani. Stability of statistical properties in two-dimensional piecewise hyperbolic maps. *Trans. Amer. Math. Soc.*, 360(9):4777–4814, 2008.
[17] Mark F. Demers and Hong-Kun Zhang. Spectral analysis of the transfer operator for the Lorentz gas. *J. Mod. Dyn.*, 5(4):665–709, 2011.
[18] Dmitry Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math. (2)*, 147(2):357–390, 1998.
[19] Dmitry Dolgopyat. Prevalence of rapid mixing. II. Topological prevalence. *Ergodic Theory Dynam. Systems*, 20(4):1045–1059, 2000.
[20] Dmitry Dolgopyat. Averaging and invariant measures. *Mosc. Math. J.*, 5(3):537–576, 742, 2005.
[21] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
[22] Frédéric Faure. Semiclassical origin of the spectral gap for transfer operators of a partially expanding map. *Nonlinearity*, 24(5):1473–1498, 2011.
[23] Frédéric Faure and Johannes Sjöstrand. Upper bound on the density of Ruelle resonances for Anosov flows. *Comm. Math. Phys.*, 308(2):325–364, 2011.
[24] Frédéric Faure and Masato Tsujii. Band structure of the Ruelle spectrum of contact Anosov flows. *C. R. Math. Acad. Sci. Paris*, 351(9-10):385–391, 2013.
[25] William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.
[26] P. Giulietti, C. Liverani, and M. Pollicott. Anosov flows and dynamical zeta functions. *Ann. of Math. (2)*, 178(2):687–773, 2013.
[27] Georg A. Gottwald and Ian Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 469(2156):20130201, 16, 2013.
[28] Sébastien Gouëzel. Local limit theorem for nonuniformly partially hyperbolic skew-products and Farey sequences. *Duke Math. J.*, 147(2):193–284, 2009.
[29] Sébastien Gouëzel and Carlangelo Liverani. Banach spaces adapted to Anosov systems. *Ergodic Theory and Dynamical Systems*, 26(1):189–217, 2006.
[30] Sébastien Gouëzel and Carlangelo Liverani. Compact locally maximal hyperbolic sets for smooth maps: fine statistical properties. *J. Differential Geom.*, 79(3):433–477, 2008.
[31] Y. Guivarc'h and J. Hardy. Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(1):73–98, 1988.
[32] Y. Guivarc'h and J. Hardy. Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(1):73–98, 1988.

[33] Hubert Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118(2):627–634, 1993.
[34] Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.
[35] Gerhard Keller. *Equilibrium states in ergodic theory*, volume 42 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1998.
[36] Gerhard Keller and Carlangelo Liverani. Stability of the spectrum for transfer operators. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 28(1):141–152, 1999.
[37] Yuri Kifer. Averaging principle for fully coupled dynamical systems and large deviations. *Ergodic Theory Dynam. Systems*, 24(3):847–871, 2004.
[38] Yuri Kifer. Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging. *Mem. Amer. Math. Soc.*, 201(944):viii+129, 2009.
[39] Hidetoshi Komiya. Elementary proof for Sion's minimax theorem. *Kodai Math. J.*, 11(1):5–7, 1988.
[40] Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997.
[41] Carlangelo Liverani. Decay of correlations. *Ann. of Math. (2)*, 142(2):239–301, 1995.
[42] Carlangelo Liverani. Decay of correlations for piecewise expanding maps. *J. Statist. Phys.*, 78(3-4):1111–1129, 1995.
[43] Carlangelo Liverani. On contact Anosov flows. *Ann. of Math. (2)*, 159(3):1275–1312, 2004.
[44] Carlangelo Liverani. Multidimensional expanding maps with singularities: a pedestrian approach. *Ergodic Theory Dynam. Systems*, 33(1):168–182, 2013.
[45] I. Melbourne and A. M. Stuart. A note on diffusion limits of chaotic skew-product flows. *Nonlinearity*, 24(4):1361–1367, 2011.
[46] K. R. Parthasarathy. On the category of ergodic measures. *Illinois J. Math.*, 5:648–656, 1961.
[47] Mark Pollicott. On the mixing of Axiom A attracting flows and a conjecture of Ruelle. *Ergodic Theory Dynam. Systems*, 19(2):535–548, 1999.
[48] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.
[49] Maurice Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958.
[50] Masato Tsujii. Exponential mixing for generic volume-preserving anosov flows in dimension three. Preprint arXiv:1601.00063.
[51] Masato Tsujii. Quasi-compactness of transfer operators for contact Anosov flows. *Nonlinearity*, 23(7):1495–1545, 2010.
[52] Lai-Sang Young. Large deviations in dynamical systems. *Trans. Amer. Math. Soc.*, 318(2):525–543, 1990.

Jacopo De Simoi, Department of Mathematics, University of Toronto, 40 St George St. Toronto, ON, Canada M5S 2E4
E-mail address: jacopods@math.utoronto.ca
URL: http://www.math.utoronto.ca/jacopods

Carlangelo Liverani, Dipartimento di Matematica, II Università di Roma (Tor Vergata), Via della Ricerca Scientifica, 00133 Roma, Italy.
E-mail address: liverani@mat.uniroma2.it
URL: http://www.mat.uniroma2.it/~liverani