Rigorous numerical investigation of the statistical properties of piecewise expanding maps–A feasibility study

Carlangelo Liverani*
Dipartimento di Matematica
Università di Roma (Tor Vergata)
Via della Ricerca Scientifica, 00133 Roma, Italy
liverani@mat.uniroma2.it

Rome, November 21, 2000

Abstract

I explore the concrete applicability of recent theoretical results to the rigorous computation of relevant statistical properties of a simple class of dynamical systems: piecewise expanding maps

1 Introduction

The aim of the present paper is to investigate the possibility of answering questions of the type:

- *Given a map of the interval is it possible to decide if it is expanding?*
- *Given a piecewise expanding map is it possible to decide if it is ergodic or mixing?*
- *Is it possible to determine with arbitrary precision its absolutely continuous invariant measure?*
- *If the map is mixing, is it possible to compute the exact rate of decay of correlations for a given function?*

Of course, the literature contains many papers in which some of these question are discussed either theoretically (especially, but not exclusively, as far as

*I would like to thank G.Keller and M.Dellnitz for motivating this work and G.Keller for many helpful discussions as well. In addition, it is a pleasure to acknowledge the support of the European Science Foundation Programme PRODYN. Finally, a warm thanks goes to the anonymous referees that help improve the paper with many remarks and suggestion. In particular, section 9 has been written to answer to an objection of a referee.

the invariant density is concerned) or numerically (e.g. [4], [5, 6, 7, 8], [9, 10], [15], [16], [21, 22], [24, 25], [26, 27, 28], [29, 30, 31, 32, 33, 34, 35], [36, 37], [40], [41, 42], [50], [51], [55], [58], [60], [69]). Nevertheless, on the one hand I find it hard to extract explicit estimates from the available papers,[1] not to mention that a few seem to contain wrong or confused results.[2] On the other hand many numerical computations are carried out without a real control on the errors. Errors that, at times, are estimated only by observing some stability of the result with respect to change of precision or (sometimes) of computer code or computer platform.

The focus of the present paper is on the possibility to answer the above questions, possibly via computer assisted strategies, but always with rigorous error bounds. To help the reader the paper is self-contained, apart from the results explicitly stated in section 3, at the beginning of section 8 and in section 9, and all the constants are always spelled out explicitly. This may make the reading a bit heavy at times, but I find it necessary to achieve the needed clarity and to make the present results readily usable.

The content of the paper is purely theoretical, it simply illustrates how to proceed rather than treating some specific example–from here the qualification: a feasibility study. In particular, once a problem is reduced to a finite computation it is not further analyzed. This means that some constants in the paper, although explicit, may be hard to compute if the required precision is too high. Of course, an analysis of the algorithmic complexity of the proposed methods would be required for a really complete study of the situation, but it exceeds the aims of the present paper.

Also, I would like to stress that, for the sake of completeness, I derive some known results and, due to my incomplete knowledge of the literature, some original results may not be properly acknowledged.

The proposed approach is based on recent theoretical results. In the last years the theoretical investigation of the statistical properties of hyperbolic dynamical systems has made substantial progresses. Here we will see how some new results can be systematically employed to tackle the above mentioned issues; whereby suppressing the need for any had hoc argument used by previous approaches.

The plan of the paper is as follows.

In section 2 some preliminary definitions and standard results are presented. This section can be skipped by the experienced reader.

Section 3 states explicitly a perturbation theorem from [44]. This can be skipped as well by readers familiar with the above mentioned work. Nevertheless, it should be noted that in [44] the various constants are defined in a much less explicit way (and I should also mention that there are some changes of notation).

Section 4 is devoted to showing how section 3 can be used to study the spectrum of the Transfer operator (when the expansion rate is larger than 2)

[1] With few notable exceptions (e.g., [46] where smooth circle maps are carefully investigated).

[2] In particular [16], [40], see [28], [42] for similar remarks.

and, in particular, to decide if the map is ergodic or mixing.

In section 5, I show how to compute, with error bounds, the invariant densities. The error is estimated in L^1 norm. It must be remarked that this section does not assume ergodicity or mixing of the map, thus these results can be used to study the invariant sets of a non-ergodic system.

Section 6 shows how to compute the asymptotic rate of decay of correlations for a given function. It gives optimal estimates only in L^1 norm while it provides correct results in bounded variation norm only for long times.

In section 7 the above bounds are employed to improve the investigation of the invariant density in the mixing case. Estimates in variation norm are also discussed.

Up to now we have just squeezed the Proposition stated in section 3. The shortcoming of the above approach is that we are able to control effectively only L^1 norms. In addition the proposed approach can be quite heavy from the computational point of view.

Section 8 shows that a technique based on projective metrics can be applied to easily gain some control over the BV norm.

In section 9 we address the first of the proposed questions and show how to apply the previous approach if the question is affirmative. In particular it is shown that, if some iterate of the map has expansion rate larger than 2, then it is possible, in principle, to study the Transfer operator associated to such an iterate with rigorous errors bounds.

Finally, section 10 is devoted to some concluding remarks and to showing that the present approach can be applied to a manifold of different systems and questions. As a matter of example I discuss two cases: Higher Dimensions and Higher Smoothness. The first discussion shows how the statistical properties of higher dimensional piecewise expanding maps can be investigated by using the ideas put forward in this paper. The second points out that if the map has further regularity properties ($C^{(3)}$ circle maps are explicitly considered, but $C^{(2+\varepsilon)}$ full branched Markov maps would suffice) then, by a simple change of norm, one can get much more accurate information on the invariant measure including an accurate knowledge in L^1 of its derivative.

In conclusion the answer to the above questions is mainly positive although the situation is not completely satisfactory.

The first problem, to establish if a map is expansive or not, is the less understood and, unless some trivial obstruction are present, can be investigated only by iterating the map. The positive part of the story is that it is possible to reduce such iteration to a finite computation that provides rigorous error bounds on the derivative (Proposition 9.1). Once a power of the map for which the derivative is strictly larger than two is known, one can reduce the study of the statistical properties of the map to the study of the iterated Transfer operator. I show that one can rigorously approximate the true Transfer operator with the Transfer operator of a piecewise linear approximation than can be used to rigorously compute the associated Ulam approximation (see the scheme at the end of the section 9).

Next, it is possible to determine the discrete spectrum of the transfer operator with arbitrary precision (Corollary 4.1).

It is thus possible to decide if the map is ergodic or mixing. Not so easy is the contrary: if one is not a simple eigenvalue in the computed precision it is possible both that one is an eigenvalue with higher multiplicity (and thus the system it is not ergodic) or that there are eigenvalues very close to one that have not been resolved yet at the current precision. It is possible to compute the invariant density with arbitrary precision in the L^1 norm (Lemma 5.1, Remark 5.1 and Lemma 7.1) and, less satisfactorily, in the BV norm (Lemma 5.1), but see also the end of section 7 for further comments. Finally, given a BV function it is possible to estimate accurately both the L^1 rate of decay (Lemma 6.1) and, (less accurately) the bounded variation decay (Lemma 6.2). Also a fast strategy to obtain reliable, but not optimal, estimates in bounded variation norm is provided (Lemma 8.3).

Of course, more refined questions could be asked, yet the results presented here show that, by working hard enough, we can often advance our understanding of a given system to a remarkable degree.

2 PRELIMINARIES AND DEFINITIONS

By a piecewise smooth expanding map I mean a piecewise $C^{(2)}$ map $T:[0,\,1]\to[0,\,1]$ such that[3]

$$|D_xT^n|\geq\lambda_n>1\quad m\text{–a.e.}$$

for $n\in\mathbb{N}$ large enough.

The idea that, through the years, has proven more fruitful in treating statistical properties of such systems is to investigate the spectral properties of the so called Transfer Operator (or Perron–Frobenius or Ruelle–Perron–Frobenius Operator).

To define the Transfer operator we start by noticing that the map T can be lifted to the algebra of $L^\infty([0,1],\mathbb{R},m)$ functions in the usual way:[4]

$$T_*\varphi:=\varphi\circ T\quad\forall\varphi\in L^\infty([0,1],\mathbb{R}).$$

and induces naturally a map on $L^1([0,1],\mathbb{R})$ defined by

$$\int_{[0,1]}f(x)T_*\varphi(x)dx=\int_{[0,1]}\mathcal{L}f(x)\varphi(x)dx\quad\forall\varphi\in L^\infty([0,1],\mathbb{R});\ f\in L^1([0,1],\mathbb{R}).$$

One can easily compute that

$$\mathcal{L}f(x)=\sum_{y\in T^{-1}x}|D_yT|^{-1}f(y).\tag{2.1}$$

[3]To be more precise, there exists a partition $\{I_i\}$, in intervals, such that T is $C^{(2)}$ on each $\bar{I}_i$. The restriction to a finite partition it is not really necessary, it is done just to simplify matters [66].

[4]By m I designate the Lebesgue measure. In the following I will omit it when this does not create confusion.

Note that, by definition, $\int_{[0,1]} \mathcal{L}f = \int_{[0,1]} f$. Accordingly

$$|\mathcal{L}f|_1 \leq |f|_1, \tag{2.2}$$

thus the operator has L^1 norm one.

Of course, the operator (2.1) has different spectral properties according to the function space on which it operates. The cumulative experience of more than a decade of work has shown that the space BV of functions of bounded variation is a very relevant space to consider. Clearly the formula (2.1) can be used to define $\mathcal{L}$ on BV.[5]

The first, instrumental, property of $\mathcal{L}$ is the so called *Lasota-Yorke inequality* [57]:

$$\|\mathcal{L}^n f\|_{\mathrm{BV}} \leq 2\lambda_n^{-1}\|f\|_{\mathrm{BV}} + B|f|_1, \tag{2.3}$$

where $\lambda_n := \inf_{x\in[0,1]} |D_x T^n|$.

The above inequality, together with (2.2), and $\int \mathcal{L}f = \int f$, imply that the spectral radius of $\mathcal{L}$, as an operator on BV, is one.

The fact that the unit BV ball is compact in $L^1([0,1],\mathbb{R})$ immediately implies that there exists at least one eigenvector of eigenvalue one: the density of an absolutely continuous invariant measure [57].

In addition, the two above facts imply that the operator $\mathcal{L}$ is quasi-compact and that the essential spectral radius $r_{ess}(\mathcal{L})$ is bounded (this was first proven by Keller [47] and later established in great generality by Hennion [38]) by

$$r_{ess}(\mathcal{L}) \leq \lim_{n\to\infty} \lambda_n^{-\frac{1}{n}} := \rho_\sigma. \tag{2.4}$$

In fact, [49], [61], [17] shows that the inequality can be an equality.[6]

The reader that wishes to have a more complete information concerning the statistical properties of expanding maps is referred to the very detailed book [3].

3 SPECTRAL STABILITY

Let $h \in \mathrm{BV}$ be such that $\mathcal{L}h = h$. If we define the measure $\mu(dx) = h(x)dx$, then μ is an invariant measure for the map T. As is well known the question of the ergodicity and mixing of the dynamical systems $([0,1], T, \mu)$ is then reduced to a question on the peripherical eigenvalues of $\mathcal{L}$. Namely the system is ergodic

[5]The space BV is the linear space of real functions on $[0,1]$ of bounded variation, that is such that the variation $\bigvee f := \sup_{\{x_1,\dots,x_n\}} \sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)| < \infty$. It is a Banach space when equipped with the norm $\|f\|_{\mathrm{BV}} := \bigvee f + |f|_1$.

[6]To have a simple idea of this type of arguments remark the following. For each $\nu < \lim_{n\to\infty}[\sup_x |D_x T^n|]^{-\frac{1}{n}} := \rho_\sigma^-$ and $f \in \mathrm{BV} : \ \mathcal{L}f = 0$, the series $\sum_{n=0}^{\infty} \nu^n f \circ T^n := \varphi_f$ is convergent in BV and $\mathcal{L}\varphi_f = \nu\varphi_f$. Since the nucleus of $\mathcal{L}$ is infinite dimensional (that is, zero is a point of essential spectrum) this shows that $\{z \in \mathbb{C} \mid |z| \leq \rho_\sigma^-\}$ is always contained in the essential spectrum of $\mathcal{L}$.

if 1 is a single eigenvalue and it is mixing if there are no other eigenvalues of modulus one.

The first step in studying the statistical properties of this type of dynamical systems is to determine the spectrum of the Transfer operator. As we already mentioned such an operator is quasi-compact and ρ_σ is an estimate of its essential spectral radius (see(2.4)).

The question that we will address in this, and the next, section is the following:
Given a precision δ is it possible to determine the peripherical spectrum of $\mathcal{L}$ with a precision smaller than δ?

By *peripherical* I mean with modulus strictly larger that ρ_σ.

The answer it known to be positive but it is unclear to me how to extract from the literature precise error bounds.

The basic idea to study the spectrum of $\mathcal{L}$ is to approximate it with a finite rank operator (which spectrum can be easily studied) and then rely on some spectral stability result.

There exist many spectral stability theorems that can be helpful (e.g. [15], [23]), yet the one that seems more efficient to me and that I will constantly refer to is an adaptation from [44]. Let us state it explicitly.

Suppose that two operators $\mathcal{L}_i$ on BV are given, both satisfying (2.2) and (2.3). Moreover, let us introduce the operator norm

$$|||\mathcal{L}||| := \sup_{\|f\|_{\mathrm{BV}}\leq 1} |\mathcal{L}f|_1 \; ; \tag{3.1}$$

and call

$$\varepsilon := |||\mathcal{L}_1 - \mathcal{L}_2|||.$$

In addition, for any operator $\mathcal{L}$, consider the set

$$V_{\delta,r}(\mathcal{L}) := \{z \in \mathbb{C} \mid |z| \leq r \text{ or } \operatorname{dist}(z, \sigma(\mathcal{L})) \leq \delta\}$$

where $\sigma(\mathcal{L})$ is the spectrum of $\mathcal{L}$.

Since the complement of $V_{\delta,r}(\mathcal{L})$ belongs to the resolvent of $\mathcal{L}$ it follows that

$$H_{\delta,r}(\mathcal{L}) := \sup\left\{\|(z-\mathcal{L})^{-1}\|_{\mathrm{BV}} \mid z \in \mathbb{C}\backslash V_{\delta,r}(\mathcal{L})\right\} < \infty.$$

I will always assume $r > \alpha$, for some fixed $\alpha \in (\rho_\sigma, 1)$. This implies that there exists a constant $A \in \mathbb{R}^+$ such that[7]

$$\|\mathcal{L}^n f\|_{\mathrm{BV}} \leq A\alpha^n \|f\|_{\mathrm{BV}} + B|f|_1. \tag{3.2}$$

From [44] we have the following.

[7]Let $\bar{n}$ be the first $n \in \mathbb{N}$ such that $\lambda_n^{-1} \leq \alpha^n$, then $A = \sup_{n\leq \bar{n}} \frac{2\lambda_n^{-1}}{\alpha^n}$. Clearly $2\lambda_n^{-1} = 2\lambda_{k\bar{n}+m}^{-1}$ for some $m < \bar{n}$, $k > 0$, thus $2\lambda_n^{-1} \leq 2\lambda_m^{-1}\lambda_{\bar{n}}^{-k} \leq A\alpha^{m+k\bar{n}} = A\alpha^n$.

Proposition 3.1 ([44]) *Consider two operators $\mathcal{L}_i : BV \to BV$ satisfying (3.2). Let $H^i_{\delta,r} := H_{\delta,r}(\mathcal{L}_i)$; $V^i_{\delta,r} := V_{\delta,r}(\mathcal{L}_i)$ and*

$$\begin{aligned} n_1 &:= \left[\tfrac{\ln 2A}{\ln r/\alpha}\right] \\ C &:= r^{-n_1};\ \ D := A(A+B+2) \\ n_2 &:= \left[\tfrac{\ln 8BDCH^1_{\delta,r}}{\ln r/\alpha}\right]. \end{aligned}$$

If

$$\varepsilon \le \frac{r^{n_1+n_2}}{8B(H^1_{\delta,r}B + (1-r)^{-1})} := \varepsilon_1(\mathcal{L}_1, r, \delta),$$

then, for each $z \in \mathbb{C}\setminus V^1_{\delta,r}$, we have

$$\|(z-\mathcal{L}_2)^{-1}f\|_{BV} \le \frac{4(A+B)}{1-r}r^{-n_1}\|f\|_{BV} + \frac{1}{2\varepsilon_1}|f|_1. \tag{3.3}$$

In addition, setting

$$\begin{aligned} \eta &:= \tfrac{\ln r/\alpha}{\ln \alpha^{-1}} \\ a &:= \tfrac{8[2A(A+B)+(1-r)^{-1}](A+B)^2 r^{-n_1}+1}{1-r} \\ b &:= 2[(4(A+B)^2(D+B)+B)(1-r)^{-1}r^{-n_1} + B], \end{aligned}$$

if

$$\begin{aligned} \varepsilon \le min &\left\{\varepsilon_1(\mathcal{L}_1,r,\delta),\ \left[\tfrac{r^{n_1}}{4B(H^1_{\delta,r}(D+B)+2A(A+B)+(1-r)^{-1})}\right]^{\frac{1}{\eta}}\right\} \\ &:= \varepsilon_0(\mathcal{L}_1,r,\delta), \end{aligned}$$

then, for each $z \in \mathbb{C}\setminus V^1_{\delta,r}$ holds

$$|||(z-\mathcal{L}_1)^{-1} - (z-\mathcal{L}_2)^{-1}||| \le \varepsilon^\eta \left[a\|(z-\mathcal{L}_1)^{-1}\|_{BV} + b\|(z-\mathcal{L}_1)^{-1}\|^2_{BV}\right]. \tag{3.4}$$

Remark 3.1 *Note that all the constants appearing in Proposition 3.1 depend only on A, B, r, α and $H^1_{\delta,r}$. In particular, ε_0 is essentially proportional to $(H^1_{\delta,r})^{\frac{1}{\eta}}$. Clearly $H^1_{\delta,r}$ is the only quantity that may be hard to estimate, we will see in the next section how this problem can be handled.*

The Proposition 3.1 implies immediately a spectral stability result.

Corollary 3.1 *If $\varepsilon \le \varepsilon_1(\mathcal{L}_1, r, \delta)$ then $\sigma(\mathcal{L}_2) \subset V_{\delta,r}(\mathcal{L}_1)$. In addition, if $\varepsilon \le \varepsilon_0(\mathcal{L}_1, r, \delta)$, then in each connected component of $V_{\delta,r}(\mathcal{L}_1)$ both $\sigma(\mathcal{L}_1)$ and $\sigma(\mathcal{L}_2)$ have the same multiplicity (that is: the associated spectral projections have the same rank).*

In fact, even stronger results are available on the eigenprojections, see [44] for details.

Remark 3.2 *The above Proposition holds in a quite general setting. In particular, the two norms involved need not be the BV and the L^1 norms. Any two Banach norms $\|\cdot\|$ and $|\cdot|$, $\|\cdot\| \geq |\cdot|$, such that the unit ball in the first is compact in the topology generated by the second will do. See [44] for full details.*

4 DETERMINATION OF THE SPECTRUM AND ERGODICITY

To determine the spectrum of $\mathcal{L}$ the basic idea is to perform an *Ulam approximation* [69].[8] Nevertheless, if one wants to perform it in reality a little bit of care is needed. Let us be more precise.

Consider a partition $\mathcal{P}_\varepsilon$ made of intervals of size at most ε. Let $\mathcal{F}_\varepsilon$ be the associate σ-algebra. The Ulam–approximation is defined as[9]

$$\mathcal{L}_\varepsilon f := \mathbb{E}(\mathcal{L}(\mathbb{E}(f|\mathcal{F}_\varepsilon))|\mathcal{F}_\varepsilon). \tag{4.1}$$

Clearly $\mathcal{L}_\varepsilon$ is a finite rank operator.[10] We will assume that $\lambda_1^{-1} \leq \alpha < r$ and $\lambda_1 > 2$, this can be always achieved, at least in principle, by eventually substituting $\mathcal{L}^k$ for $\mathcal{L}$ for some proper power k in the definition (4.1) of $\mathcal{L}_\varepsilon$.[11]

Lemma 4.1 *Setting* $\Gamma = \max\{2\lambda_1^{-1}+1, B\}(\frac{A}{1-\alpha}+B)$ *then*

$$|||\mathcal{L}^n - \mathcal{L}_\varepsilon^n||| \leq n\Gamma\varepsilon. \tag{4.2}$$

[8]Other types of approximations have also been proposed. See in particular [55, 26, 28] where piecewise linear and piecewise quadratic (rather than piecewise constant) function are used. Such improved schemes seem to yield faster convergence, yet they can be investigated exactly in the same manner as the simpler case treated here (see section 10.2 for an example).

[9]By $\mathbb{E}(f|\mathcal{F})$ we mean, as usual, the conditional expectation of f with respect to the σ-algebra $\mathcal{F}$ and the measure m. In our case if $x \in p \in \mathcal{P}_\varepsilon$ the expectation is simply given by

$$\mathbb{E}(f|\mathcal{F}_\varepsilon)(x) = \frac{1}{m(p)}\int_p f dm.$$

[10]In fact, if $x \in p \in \mathcal{P}_\varepsilon$,

$$\mathcal{L}_\varepsilon f(x) = \sum_{q\in\mathcal{P}_\varepsilon} \frac{m(T^{-1}p\cap q)}{m(p)m(q)}\int_q f dm,$$

which means that, if we define $F : \mathbb{R}^{\#\mathcal{P}_\varepsilon} \to \mathrm{BV}$ by $F(v)(x) := \frac{v_q}{m(q)} \quad \forall x \in q$, thus $[F^{-1}(f)]_q = \int_q f$, then setting

$$P_{pq} := \frac{m(T^{-1}p\cap q)}{m(q)},$$

the operator $\mathcal{L}_\varepsilon$ reduces to the matrix $P : \mathbb{R}^{\#\mathcal{P}_\varepsilon} \to \mathbb{R}^{\#\mathcal{P}_\varepsilon}$. More precisely, $F^{-1}\mathcal{L}_\varepsilon F = P$.

[11]This is done just to simplify matters. In alternative, estimate (4.3) can be improved and the condition $\lambda > 2$ relaxed, for details see [15] for the "no-turning points" case and [10] for a more general setup or [3] for a comprehensive discussion on perturbations of the Transfer operator. See section 9 for a more detailed discussion of how to take powers of the Transfer operator.

While[12]

$$\bigvee \mathcal{L}_\varepsilon f \leq 2\lambda_1^{-1} \bigvee f + B|f|_1 \tag{4.3}$$

PROOF This is rather standard, I just sketch the proof for completeness. Obviously $|\mathbb{E}(f \mid \mathcal{F})|_1 = |f|_1$, moreover $\bigvee \mathbb{E}(f \mid \mathcal{F}) = \bigvee f$. This last fact easily follows by noticing

$$\begin{aligned}\bigvee \mathbb{E}(f \mid \mathcal{F}) &= \sup_{\{x_1,\dots,x_n\}} \sum_{k=1}^{n-1} |\mathbb{E}(f \mid \mathcal{F})(x_{i+1}) - \mathbb{E}(f \mid \mathcal{F})(x_i)| \\ &= \sup_{\{x_1,\dots,x_n\}} \sum_{k=1}^{n-1} \left| \frac{1}{m(p(x_{i+1}))} \int_{p(x_{i+1})} f - \frac{1}{m(p(x_i))} \int_{p(x_i)} f \right|.\end{aligned}$$

A bit of thought shows that, given $f \in \mathrm{BV}$ and any interval $I \subset [0,1]$, it is always possible to change f in such a way that the new function $\tilde{f}$ has the following properties: *a)* $|f - \tilde{f}|_1 = 0$; *b)* $\bigvee f = \bigvee \tilde{f}$; *c)* $\exists x(I) \in I : \tilde{f}(x(I)) = m(I)^{-1} \int_I f$. Using these facts, and denoting $\mathcal{P}_\varepsilon = \{p_1, \dots, p_l\}$, yields

$$\bigvee \mathbb{E}(f \mid \mathcal{F}) \leq \sum_{k=1}^{l-1} \left| \tilde{f}(x(p_{i+1})) - \tilde{f}(x(p_i)) \right| \leq \bigvee \tilde{f} = \bigvee f.$$

In addition, for $f \in \mathrm{BV}$, holds

$$|\mathbb{E}(f \mid \mathcal{F}_\varepsilon) - f|_1 \leq \sup_{p \in \mathcal{P}_\varepsilon} \int_p \bigvee_p f \leq \varepsilon \bigvee f.$$

Thus, for all $f \in \mathrm{BV}$,

$$\begin{aligned}|(\mathcal{L}_\varepsilon - \mathcal{L})f|_1 \leq\ & |\mathbb{E}(\mathcal{L}\mathbb{E}(f \mid \mathcal{F}_\varepsilon) \mid \mathcal{F}_\varepsilon) - \mathcal{L}\mathbb{E}(f \mid \mathcal{F}_\varepsilon)|_1 + |\mathcal{L}\mathbb{E}(f \mid \mathcal{F}_\varepsilon) - \mathcal{L}f|_1 \\ \leq\ & \varepsilon \bigvee \mathcal{L}\mathbb{E}(f \mid \mathcal{F}_\varepsilon) + |\mathbb{E}(f \mid \mathcal{F}_\varepsilon) - f|_1 \\ \leq\ & \varepsilon 2\lambda_1^{-1} \bigvee f + \varepsilon B|f|_1 + \varepsilon \bigvee f \leq \varepsilon \max\{2\lambda_1^{-1} + 1, B\} \|f\|_{\mathrm{BV}}.\end{aligned}$$

Hence, for $f \in \mathrm{BV}$,

$$\begin{aligned}|(\mathcal{L}_\varepsilon^n - \mathcal{L}^n)f|_1 \leq\ & \sum_{i=1}^n |\mathcal{L}_\varepsilon^{n-i}(\mathcal{L}_\varepsilon - \mathcal{L})\mathcal{L}^{i-1}f|_1 = \sum_{i=1}^n |(\mathcal{L}_\varepsilon - \mathcal{L})\mathcal{L}^{i-1}f|_1 \\ \leq\ & \sum_{i=1}^n |||\mathcal{L}_\varepsilon - \mathcal{L}|||\ \|\mathcal{L}^{i-1}f\|_{\mathrm{BV}} \\ \leq\ & \sum_{i=1}^n |||\mathcal{L}_\varepsilon - \mathcal{L}|||\{A\alpha^{i-1} \bigvee f + B|f|_1\} \\ \leq\ & n(\tfrac{A}{1-\alpha} + B)|||\mathcal{L}_\varepsilon - \mathcal{L}|||\ \|f\|_{\mathrm{BV}}.\end{aligned}$$

□

A first naïve idea would be to apply directly Proposition 3.1 to this situation considering $\mathcal{L}_\varepsilon$ as a perturbation of $\mathcal{L}$. The problem with this approach is that the constants appearing in Proposition 3.1 depend on the norm of the resolvent

[12] By $\bigvee f$, as usual, we mean the variation of f, that is $\sup_{\{x_1,\dots,x_n\}} \sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)|$.

and I do not see any way to estimate directly such a norm for the operator $\mathcal{L}$. On the contrary all the constants can be explicitly computed for the operator $\mathcal{L}_\varepsilon$, yet it is not obvious that they can satisfy the needed condition, that is $\varepsilon \leq \Gamma^{-1}\varepsilon_0(\mathcal{L}_\varepsilon, r, \delta)$. It is interesting that Proposition 3.1 can be used also to answer affirmatively this problem.

Lemma 4.2 *There exists $\varepsilon_2 > 0$ such that, for each $\varepsilon \leq \varepsilon_2$*

$$\varepsilon \leq \Gamma^{-1}\varepsilon_0(\mathcal{L}_\varepsilon, r, \delta). \tag{4.4}$$

PROOF Just apply Proposition 3.1 to $\mathcal{L}$ and $\mathcal{L}_\varepsilon$ (with $\mathcal{L} = \mathcal{L}_1$ and $\mathcal{L}_\varepsilon = \mathcal{L}_2$). It follows that, if $\Gamma\varepsilon \leq \varepsilon_1(\mathcal{L}, r, \delta)$, there exists a constant $c_* > 0$ such that

$$\sup_{z\in\mathbb{C}\setminus V_{\delta/2,r}(\mathcal{L})} \|(z-\mathcal{L}_\varepsilon)^{-1}\|_{\mathrm{BV}} \leq c_*. \tag{4.5}$$

This means that $\sigma(\mathcal{L}_\varepsilon) \subset V_{\delta/2,r}(\mathcal{L})$. Hence, if $\lambda_\varepsilon \in \sigma(\mathcal{L}_\varepsilon)$, $|\lambda_\varepsilon| \geq r$, there exists $\lambda \in \sigma(\mathcal{L})$ such that $|\lambda_\varepsilon - \lambda| \leq \delta/2$ (See Corollary 3.1 or [44] for more details). Accordingly if $|z - \lambda_\varepsilon| \geq \delta$ we have

$$|z-\lambda| \geq |\lambda_\varepsilon - z| - |\lambda - \lambda_\varepsilon| \geq \frac{\delta}{2}.$$

The above estimate means that $V_{\delta,r}(\mathcal{L}_\varepsilon) \supset V_{\delta/2,r}(\mathcal{L})$, thus by (4.5)

$$H_{\delta,r}(\mathcal{L}_\varepsilon) \leq c_*\,.$$

Accordingly, $\varepsilon_0(\mathcal{L}_\varepsilon, r, \delta)$ is uniformly bounded from below. Define

$$\varepsilon_2 := \inf_\varepsilon \Gamma^{-1}\varepsilon_0(\mathcal{L}_\varepsilon, r, \delta) > 0$$

and this concludes the proof. □

The implication of Lemma 4.2 is that we can apply Proposition 3.1 with $\mathcal{L}_1 = \mathcal{L}_\varepsilon$ and $\mathcal{L}_2 = \mathcal{L}$; that is, we can regard $\mathcal{L}$ as a perturbation of the known operator $\mathcal{L}_\varepsilon$. As a result we can determine the spectrum of $\mathcal{L}$ since, for small ε, it is contained in a δ-neighborhood of the spectrum of $\mathcal{L}_\varepsilon$:

Corollary 4.1 *For each $\varepsilon > 0$, satisfying (4.4), we have $\sigma(\mathcal{L}) \subset V_{\delta,r}(\mathcal{L}_\varepsilon)$ with multiplicity.*

Remark 4.1 *It is important to remark that, since $\varepsilon_0(\mathcal{L}_\varepsilon, r, \delta)$ is* computable,[13] *to find ε that satisfies (4.4) it suffices to compute for smaller and smaller ε until (4.4) is verified.*[14]

---

[13] By *computable* I simply mean that it can be computed, eventually by using a computer, with rigorous error bounds. I will use the word in this loose sense rather then in one of its technical meanings.

[14] Just to have an idea of how small ε should be: if the maximal multiplicity of the peripherical spectrum is d_1, then ε_0 is proportional to $\delta^{\frac{d_1}{\eta}}$.

5 COMPUTING INVARIANT DENSITIES (PART I)

Here we will see how the results of the previous section suffice to gain considerable knowledge on the invariant densities. Suppose that the eigenvalue one has multiplicity d for the operator $\mathcal{L}_\varepsilon$, that no other eigenvalues are present in a $2\delta_0$-ball around one and the rest of the spectrum is contained in a disk of radius less than $1-2\delta_1$.[15] It is then natural to define the projection on the eigenspace of the eigenvalue one

$$\Pi_1^\varepsilon := \frac{1}{2\pi i}\int_{\{z\in\mathbb{C}\mid |z-1|=\delta_0\}} (z-\mathcal{L}_\varepsilon)^{-1}dz. \tag{5.1}$$

Lemma 5.1 *For each* $\varepsilon > 0$ *that satisfies (4.4) and* $\ln\alpha^{-1}/\ln\varepsilon^{-1} < \min\{\delta_0, \delta_1\}$, *we have*

$$\begin{aligned} &|||\Pi_1^0 - \Pi_1^\varepsilon||| \le K_\varepsilon \varepsilon \ln\varepsilon^{-1}, \\ &K_\varepsilon =: \left[\tfrac{b}{\ln\alpha^{-1}}\|\Pi_1^\varepsilon\|_{BV}^2 + \tfrac{(a+2b\tilde{H})}{\ln\varepsilon^{-1}}\|\Pi_1^\varepsilon\|_{BV} + \tfrac{b\tilde{H}^2\ln\alpha^{-1}}{(\ln\varepsilon^{-1})^2}\right], \end{aligned} \tag{5.2}$$

where

$$\tilde{H} := \sup_{|z|=1-\delta_1} \|(z-\mathcal{L}_\varepsilon(1-\Pi_1^\varepsilon))^{-1}\|_{BV}.$$

PROOF For each $\delta < 2\delta_0$, (3.4) implies

$$\begin{aligned} |||\Pi_1^0 - \Pi_1^\varepsilon||| &\le \tfrac{1}{2\pi}\int_{\{z\in\mathbb{C}\mid |z-1|=\delta\}} |||(z-\mathcal{L})^{-1} - (z-\mathcal{L}_\varepsilon)^{-1}|||\,|dz| \\ &\le \delta\varepsilon^\eta[aH_{\delta,1-\delta}(\mathcal{L}_\varepsilon) + bH_{\delta,1-\delta}(\mathcal{L}_\varepsilon)^2], \end{aligned}$$

where

$$\eta = \frac{\ln(1-\delta)\alpha^{-1}}{\ln\alpha^{-1}} \le 1 + \frac{\delta}{\ln\alpha}.$$

On the other hand the standard spectral theory of isolated non-essential point spectrum (see [43] page 181) implies that

$$(z-\mathcal{L}_\varepsilon)^{-1} = (z-1)^{-1}\Pi_1^\varepsilon + \sum_{i=1}^{d}(z-1)^{-i-1}D^i + R_1(z) \tag{5.3}$$

where

$$D = (1-\mathcal{L}_\varepsilon)\Pi_1^\varepsilon$$

and $R_1(z)$ is the resolvent of $\mathcal{L}_\varepsilon(1-\Pi_1^\varepsilon)$ which clearly has spectral radius strictly smaller than one. It must be noted that in the case at hand we have $D = 0$

[15] This last requirement can be relaxed. It just suffices to do the following argument considering the eigenprojections on all the eigenvalues of modulus one, I choose not to do this just for simplicity.

since $D \neq 0$ would imply a polynomial growth of $\mathcal{L}_\varepsilon^n$ incompatible with the a priory bound $|\mathcal{L}_\varepsilon^n f|_1 \leq |f|_1$. Accordingly,

$$H_{\delta,1-\delta}(\mathcal{L}_\varepsilon) \leq \|\Pi_1^\varepsilon\|_{\mathrm{BV}}\delta^{-1} + \tilde{H}.$$

By choosing $\delta = \ln \alpha^{-1}(\ln \varepsilon^{-1})^{-1} < \min\{\delta_0, \delta_1\}$ the lemma follows. □

The above Lemma allows to gain a very good idea of the vector space associated to the eigenvalue one. In particular, looking at the functions, in such a subspace, with compact support one obtains sets which mixes very slowly. Such quasi-invariant sets could be really invariant but, to establish it, one has to investigate the issue directly by making exact computations via the original map (see footnote 16 in section 6).

Remark 5.1 *In the case in which $d = 1$, that is when the map is mixing, Lemma 5.1 yields immediately an estimate of the L^1 distance between the eigenfunction h_ε of $\mathcal{L}_\varepsilon$ (that is, $\mathcal{L}_\varepsilon h_\varepsilon = h_\varepsilon$) and the eigenfunction h of $\mathcal{L}$. Yet, in section 8 we will see that in such a case the estimate of the constant K_ε can be substantially improved.*

6 Estimate of the rate of decay of correlations (part I)

By the results of the previous section is possible to decide if the map is ergodic or mixing by doing a sufficiently precise numerical computation.[16]

Let us suppose that we have found the map to be mixing. The next natural question is to determine the rate of decay of correlations. We will explore this issue in this section. We will see that what can be obtained by the perturbation Theorem 3.1, although very interesting, is not completely satisfactory. We will discuss later alternative approaches.

Let $\lambda_\varepsilon \in \sigma(\mathcal{L}_\varepsilon)\backslash\{1\}$, be the largest eigenvalue different from one, or set $\lambda_\varepsilon = r$ if there are no isolated eigenvalues of modulus larger than r. In this last case the rate of decay of correlations will be determined by the essential spectrum radius but the present techniques will not give exact constants in this regime, only (good) upper bounds, we will not discuss this particular case any further. Set $R = |\lambda_\varepsilon|$ and choose $\delta < 1 - R$. For each $R_1 \in (R + \delta, 1)$, let us define the eigenprojections

$$\Pi_{R_1}^\varepsilon := \frac{1}{2\pi i}\int_{\{z\in\mathbb{C}\;|\;|z|=R_1\}} (z - \mathcal{L}_\varepsilon)^{-1} dz. \tag{6.1}$$

Clearly $\mathcal{L}_\varepsilon f = h_\varepsilon \int f + \mathcal{L}_\varepsilon \Pi_{R_1}^\varepsilon f$.

[16] On the contrary, as already remarked in the introduction, it is not possible to decide that a map is not ergodic since, no matter the precision with which we know the eigenvalues, it is always possible that an eigenvalue we believe to lay on the unit circle is, in reality, of modulus slightly smaller than one.

Lemma 6.1 *Given $\varepsilon > 0$ satisfying (4.4), for each $f \in BV$, $\int f = 0$, we have*

$$\|\mathcal{L}^n f\|_{BV} \leq R_1^n \bar{C} \|f\|_{BV} \tag{6.2}$$

$$|\mathcal{L}^n f - \mathcal{L}_\varepsilon^n f|_1 \leq R_1^n \varepsilon^{\eta_1} C_{\varepsilon,\delta} \|f\|_{BV} \tag{6.3}$$

where

$$\begin{aligned}
\eta_1 &= \tfrac{\ln(|\lambda_\varepsilon|-\delta)\alpha^{-1}}{\ln \alpha^{-1}} \\
\bar{C} &= R_1 \tfrac{4(A+B)}{1+R_1} R_1^{-n_1} + \tfrac{1}{2\varepsilon_1(\mathcal{L}_\varepsilon, R_1, \delta)} \\
C_{\varepsilon,\delta} &= R_1 \left[a H_{\delta,R_1}(\mathcal{L}_\varepsilon) + b H_{\delta,R_1}(\mathcal{L}_\varepsilon)^2 \right].
\end{aligned}$$

PROOF Since $\int f = 0$ it follows $\Pi_{R_1}^\varepsilon f = \Pi_{R_1}^0 f = f$, thus

$$\mathcal{L}^n f = \frac{1}{2\pi i} \int_{\{|z|=R_1\}} z^n (z - \mathcal{L})^{-1} f dz.$$

Accordingly,

$$\|\mathcal{L}^n f\|_{\mathrm{BV}} \leq R_1^{n+1} \sup_{\{|z|=R_1\}} \|(z-\mathcal{L})^{-1}\|_{\mathrm{BV}} \|f\|_{\mathrm{BV}},$$

and the result follows by Proposition 3.1.

On the other hand

$$|\mathcal{L}^n f - \mathcal{L}_\varepsilon^n f|_1 \leq R_1^{n+1} |||(z-\mathcal{L})^{-1} - (z - \mathcal{L}_\varepsilon)^{-1}||| \, \|f\|_{\mathrm{BV}}.$$

From the above inequality and Proposition 3.1 the Lemma immediately follows.
□

Since $\mathcal{L}_\varepsilon$ has a spectral gap of size at least $1 - R_1$, it follows that there exists $C_{\varepsilon,\delta}^*$ such that, for each $f \in \mathrm{BV}$, $\int f = 0$, we have[17]

$$|\mathcal{L}_\varepsilon^n f|_1 \leq C_{\varepsilon,\delta}^* R_1^n |f|_1. \tag{6.4}$$

From the above Lemma it follows that if $f \in \mathrm{BV}$ and $g \in L^\infty([0,1], \mathbb{R})$, then

$$\left| \int f g \circ T^n - \int f \int g d\mu \right| \leq C_* |g|_\infty \|f\|_{\mathrm{BV}} R_1^n, \tag{6.5}$$

where $C_* \leq C_{\varepsilon,\delta}^* + C_{\varepsilon,\delta} \varepsilon^{\eta_1}$.

[17]Since $z - \mathcal{L}_\varepsilon$ is just a polynomial in z of degree given by the square of the number of elements of the partition, and since in a finite dimensional vector space all the norms are equivalent, one can compute explicitly

$$(C_{\varepsilon,\delta}^*)^{-1} = \inf_{\int f = 0} \frac{|(z - \mathcal{L}_\varepsilon) f|_1}{|f|_1} = \inf_{\int f = 0} \frac{|(z - \mathcal{L}_\varepsilon) \Pi_{R_1}^\varepsilon f|_1}{|f|_1} > 0.$$

Then

$$|(\mathcal{L}_\varepsilon^n \Pi_{R_1}^\varepsilon)^n|_1 \leq \frac{1}{2\pi} \int_{|z|=R_1} |z^n (z - \mathcal{L}_\varepsilon)^{-1} \Pi_{R_1}^\varepsilon|_1 |dz| \leq R_1^n C_{\varepsilon,\delta}^*.$$

Remark 6.1 *A reason for satisfaction is that the rate of decay in (6.5) is arbitrarily close to the correct one. The problem with (6.5) is clearly in $|g|_\infty$; it is in fact well known that $|g|_1$ should appear instead. Of course, for bounded functions, this can be adjusted by changing the constant C_*. Yet, the ratio between the L^1 and the L^∞ norm can be arbitrarily large and this shows that the estimate can be disastrous for certain functions, especially for short times.*

Note that, one would get the L^1 norm in (6.5) by using (6.2), yet the estimate of the constant $\bar{C}$ is far from optimal.

Remark 6.2 *The estimate of the constant $\bar{C}$ it is not a matter of mathematical finesse, it is a fundamental need for practical purposes. In fact, in any practical case one wishes to know the decay of correlation for finite time rather than the asymptotic behavior that can be completely non observable. Clearly, the finite time behavior is essentially determined by the size of $\bar{C}$.*

A first possibility to improve the estimate of the BV decay is to define $k := \inf\{j \in \mathbb{N} \mid 2A\alpha^j \leq R_1^j\}$. Then by the L^1 estimates and the Lasota-Yorke inequality follows

$$\begin{aligned} \bigvee \mathcal{L}^{kn} f &\leq \tfrac{1}{2} R_1^k \bigvee \mathcal{L}^{k(n-1)} f + B|\mathcal{L}^{k(n-1)} f|_1 \\ &\leq 2^{-n} R_1^{kn} \bigvee f + \textstyle\sum_{j=0}^{n-1} 2^{-j} R_1^{jk} |\mathcal{L}^{k(n-j)} f|_1 . \end{aligned} \tag{6.6}$$

This allows to estimate accurately the long time behavior and to obtain better bounds (although maybe not optimal) on the short time behavior.

Lemma 6.2 *Given $\varepsilon > 0$, satisfying (4.4), for each $f \in BV$, $\int f = 0$, holds*

$$\|\mathcal{L}^n f\|_{BV} \leq R_1^n \bar{C}^* \|f\|_{BV},$$

with

$$\bar{C}^* := \left\{ A C_{\varepsilon,\delta}^* + \left[2B(2A)^{\frac{\ln r_1^{-1}}{\ln R_1 \alpha^{-1}}} + 1 \right] \right\},$$

where $C_{\varepsilon,\delta}^*$ is defined in (6.4).

PROOF By induction from (6.6) it follows that

$$\bigvee \mathcal{L}^{kn} f \leq 2^{-n} R_1^{kn} \bigvee f + B R_1^{k(n-1)} C_{\varepsilon,\delta}^* \sum_{j=0}^{n-1} 2^{-j} |f|_1,$$

thus

$$\bigvee \mathcal{L}^n f \leq R_1^n \{A \bigvee f + 2B C_{\varepsilon,\delta}^* R_1^{-k} |f|_1\}.$$

Since, by definition, $k \leq \frac{\ln 2A}{\ln r_1 \alpha^{-1}}$ the lemma follows. □

7 Computing the invariant density (part II)

As already remarked, if the map is mixing, then it is possible to use the previous estimates to obtain better bounds on the difference between the invariant density h_ε of the Ulam approximation and the invariant density h of the original system.

Lemma 7.1 *Given $\varepsilon > 0$ satisfying (4.4), we have*

$$|h - h_\varepsilon|_1 \le \{2C^*_{\varepsilon,\delta} + \frac{\ln \varepsilon^{-1}}{\ln R_1^{-1}}\Gamma\}\varepsilon.$$

Proof By Lemma 4.1 and (6.4)

$$\begin{aligned} |h - h_\varepsilon|_1 &= |\mathcal{L}^n h - \mathcal{L}^n_\varepsilon h_\varepsilon|_1 \\ &\le |\mathcal{L}^n_\varepsilon (h - h_\varepsilon)|_1 + |(\mathcal{L}^n - \mathcal{L}^n_\varepsilon)h|_1 \\ &\le \{2C^*_{\varepsilon,\delta} R_1^n + n\Gamma\varepsilon\}\|h\|_{\mathrm{BV}}. \end{aligned}$$

To conclude it suffices to choose n appropriately. For example

$$n = \frac{\ln \varepsilon^{-1}}{\ln R_1^{-1}}$$

will do. □

The remaining issue is to estimate the invariant density in the BV norm. This can be done in various ways although I do not know of an optimal one (but see section 9.2 for a special situation in which this is possible). Just to give an idea of what can be done let us say few words.

Suppose that one knows a function $\bar{h}_\varepsilon$ such that[18]

$$\|\mathcal{L}\bar{h}_\varepsilon - \bar{h}_\varepsilon\|_{\mathrm{BV}} \le \varepsilon.$$

In this case, one easily obtains

$$\begin{aligned} \|h - \bar{h}_\varepsilon\|_{\mathrm{BV}} &\le \|\mathcal{L}^n h - \mathcal{L}^n \bar{h}_\varepsilon\|_{\mathrm{BV}} + \|\mathcal{L}^n \bar{h}_\varepsilon - \bar{h}_\varepsilon\|_{\mathrm{BV}} \\ &\le \bar{C}^* R_1^n \|h - \bar{h}_\varepsilon\|_{\mathrm{BV}} + \textstyle\sum_{j=1}^n \|\mathcal{L}\|_{\mathrm{BV}}^{n-j} \|\mathcal{L}\bar{h}_\varepsilon - \bar{h}_\varepsilon\|_{\mathrm{BV}} \\ &\le \bar{C}^* R_1^n \|h - \bar{h}_\varepsilon\|_{\mathrm{BV}} + \varepsilon \textstyle\sum_{j=0}^{n-1} \|\mathcal{L}\|_{\mathrm{BV}}^j . \end{aligned}$$

Thus, by choosing n_* such that $\bar{C}R_1^{n_*} \le \frac{1}{2}$, we have

$$\|h - \bar{h}_\varepsilon\|_{\mathrm{BV}} \le 2\varepsilon \sum_{j=0}^{n_*-1} \|\mathcal{L}\|_{\mathrm{BV}}^j.$$

Of course, these are only general remarks but they show that something can be done in the required direction although the theory it is not as satisfactory as for the L^1 estimates.

[18] For example a function $\bar{h}_\varepsilon$ with the property below could be obtained by iterating the function h_ε via the operator $\mathcal{L}$, procedure that, to be honest, can pose non trivial numerical problems.

8 ESTIMATE OF THE RATE OF DECAY OF CORRELATIONS (PART II)

In section 6 we have seen that the results on decay of correlations, although very strong, are not completely satisfactory (cf. Remarks 6.1, 6.2). Another reason for discontent with the results so far presented is that they require estimating the norm of the resolvent. I claimed that such a norm is computable, in fact the entries of the resolvent of $\mathcal{L}_\varepsilon$ are rational functions of z, so the maximum on a given domain can certainly be explicitly computed. Nevertheless, such a computation can be enormously time consuming to the point of being intractable if the matrix is large enough.

In view of the above comments it would be useful to have a quick way to obtain reasonable estimates. In the following I propose a method to obtain just this by using a projective metric technique that has proved rather successful in the theoretical study of convergence to equilibrium [53], [54], [56], [62], [59], [70].

We need to recall some results in lattice theory originally due to Garrett Birkhoff [2]. More details and the proofs of the following results can be found in [53].

Consider a topological vector space $\mathbb{V}$ with a closed convex cone[19] $\mathcal{C}$ (hereafter, the term "closed cone" $\mathcal{C}$ will mean that $\mathcal{C} \cup \{0\}$ is closed).

Given a closed convex cone $\mathcal{C} \subset \mathbb{V}$, enjoying the property $\mathcal{C} \cap -\mathcal{C} = \emptyset$, we can define an order relation by

$$f \preceq g \iff g - f \in \mathcal{C} \cup \{0\}.$$

Henceforth, each time that we specify a convex cone we will assume the corresponding order relation and vice versa.

It is then possible to define a projective metric Θ (Hilbert metric),[20] in $\mathcal{C}$, by the construction:

$$\begin{aligned}
\alpha(f,\, g) &= \sup\{\lambda \in \mathbb{R}^+ \mid \lambda f \preceq g\} \\
\beta(f,\, g) &= \inf\{\mu \in \mathbb{R}^+ \mid g \preceq \mu f\} \\
\Theta(f,\, g) &= \log\left[\tfrac{\beta(f,\, g)}{\alpha(f,\, g)}\right]
\end{aligned}$$

where we take $\alpha = 0$ and $\beta = \infty$ if the corresponding sets are empty.

The importance of the previous constructions is due, in our context, to the following theorem.

Theorem 8.1 ([2]) *Let $\mathbb{V}$ be a topological vector lattices and $\mathcal{C} \subset \mathbb{V}$ a closed cone; $T : \mathbb{V} \to \mathbb{V}$ a linear map such that $T(\mathcal{C}) \subset \mathcal{C}$. Let Θ be the Hilbert metric*

[19] Here, by "cone," we mean any set such that, if f belongs to the set, then λf belongs to it as well, for each $\lambda > 0$.

[20] In fact, we define a semi–metric, since $f = \lambda g \Rightarrow \Theta(f,\, g) = 0$. The metric that we describe corresponds to the conventional Hilbert metric on the projectivization of $\mathcal{C}$.

corresponding to the cone $\mathcal{C}$. Setting $\Delta = \sup_{f,\, g \in T(\mathcal{C}_1)} \Theta(f,\, g)$ *we have*

$$\Theta(Tf,\, Tg) \le \tanh\left(\frac{\Delta}{4}\right) \Theta(f,\, g) \qquad \forall f,\, g \in \mathcal{C}_1$$

($\tanh(\infty) \equiv 1$*).*

In particular, if an operator maps a convex cone strictly inside itself (in the sense that the diameter of the image is finite), then it is a contraction in the Hilbert metric. This implies the existence of a "positive" eigenfunction (provided the cone is complete with respect to the Hilbert metric), and, with some additional work, the existence of a gap in the spectrum of T (see [1] for details). The completeness of the Hilbert metric can be studied by using the following result.

Lemma 8.1 *Let* $\|\cdot\|$ *be a norm on the vector lattice* $\mathbb{V}$*, and suppose that, for each* $f,\, g \in \mathbb{V}$*,*

$$-f \preceq g \preceq f \Longrightarrow \|f\| \ge \|g\|.$$

Then, given $f,\, g \in \mathcal{C} \subset \mathbb{V}$ *for which* $\int f = \int g$*,*

$$\|f - g\| \le \left(e^{\Theta(f,\, g)} - 1\right) \|f\|.$$

See [53] for proofs and more details.

If the operator $\mathcal{L}_\varepsilon$ has a spectral gap, it follows that there must exist, computable, $a_* \in \mathbb{R}^+$ such that the cone[21]

$$\widetilde{\mathcal{C}}^{\varepsilon}_{a_*} = \left\{ f \in \mathrm{BV} \;\middle|\; \left| \mathbb{E}(f \mid \mathcal{F}_\varepsilon) - h_\varepsilon \int_{[0,1]} f \right|_1 \le a_* \int_{[0,1]} f \right\}, \tag{8.1}$$

is strictly invariant. More precisely, $\mathcal{L}^k_\varepsilon \widetilde{\mathcal{C}}^{\varepsilon}_{a_*} \subset \widetilde{\mathcal{C}}^{\varepsilon}_{\sigma_k a_*}$ where, clearly,[22]

$$\lim_{k \to \infty} {\sigma_k}^{\frac{1}{k}} = |\lambda_\varepsilon|.$$

accordingly the operators $\mathcal{L}^k_\varepsilon$ are all contractions in the Hilbert metric.

[21] If one wants an even quicker, but rougher, approach: consider cones of the type

$$\widehat{\mathcal{C}}^{\varepsilon}_{a_*} = \left\{ f \in \mathrm{BV} \;\middle|\; \left| \mathbb{E}(f \mid \mathcal{F}_\varepsilon) - \int_{[0,1]} f \right|_1 \le a_* \int_{[0,1]} f \right\},$$

which do not require any previous knowledge of the invariant measure.

[22] As in the previous section λ_ε is the second largest eigenvalue, here again we are not discussing explicitly the case in which there is no eigenvalue outside the essential spectral radius beside 1. Yet, the present technique allows to obtain good upper bounds also in this case, it just suffices to substitute to σ_{n_0} the maximum between σ_{n_0} and $(\rho_\sigma + \delta)^{n_0}$ for some small δ.

In fact, a simple computation shows

$$\Theta(f,\,h_\varepsilon)\le \ln\frac{|\mathbb{E}(f\mid \mathcal{F}_\varepsilon)-h_\varepsilon\int f|+a_*\int f}{a_*\int f-|\mathbb{E}(f\mid\mathcal{F}_\varepsilon)-h_\varepsilon\int f|}.$$

Thus, if $f\in\tilde{\mathcal{C}}^\varepsilon_{\sigma a_*}$,

$$\Theta(f,\,h_\varepsilon)\le\ln\frac{1+\sigma}{1-\sigma}.$$

Accordingly, the diameter $\Delta(\tilde{\mathcal{C}}^\varepsilon_{\sigma a_*})$, in $\tilde{\mathcal{C}}^\varepsilon_{a_*}$, is bounded by $2\ln(1+\sigma)(1-\sigma)^{-1}$ which yields precisely σ as the rate of contraction for the Hilbert metric.

The interest of the above fact is that it is possible to define a cone which is strictly invariant for the unperturbed operator:

$$\mathcal{C}_{ab}=\left\{f\in\mathrm{BV}\;\middle|\;\bigvee\left(f-h\int f\right)\le b\int f;\;\left|f-h\int f\right|_1\le a\int f\right\}\tag{8.2}$$

Lemma 8.2 *If* $[1+\frac{B}{a_*}]\sigma_{n_0}\ge A\alpha^{n_0}$, *and*

$$\varepsilon\le\min\left\{\frac{(1+Ba_*^{-1})(2a_*+4B)}{2a_*(1+B)n_1\Gamma}\sigma^2_{n_1},\,\frac{\sigma_{n_1}}{n_1\Gamma}\right\},$$

where a_* *is determined by (8.1), then, setting*

$$\begin{aligned}a&=2a_*\\ b&=\tfrac{(2a_*+4B)\sigma_{n_0}-2n_0\varepsilon\Gamma a_*}{n_0\varepsilon\Gamma}\\ c&=2+2Ba_*^{-1},\end{aligned}$$

for each $n\in[n_0,\,n_1]$, *we have*

$$\mathcal{L}^n\mathcal{C}_{a,b}\subset\mathcal{C}_{c\sigma_n a,\,c\sigma_n b}.$$

PROOF To verify the invariance of $\mathcal{C}_{ab}$ we start by computing the variation.

$$\begin{aligned}\bigvee\mathcal{L}^n(f-h\textstyle\int f)&\le A\alpha^n\bigvee(f-h\textstyle\int f)+B|f-h\textstyle\int f|_1\\ &\le\{A\alpha^n b+Ba\}\textstyle\int f.\end{aligned}$$

Moreover, remembering (4.2), the properties of (8.1) and since $\mathcal{L}^n_\varepsilon f=\mathcal{L}^n_\varepsilon\mathbb{E}(f\mid\mathcal{F}_\varepsilon)$, holds[23]

$$\begin{aligned}\left|\mathcal{L}^nf-h\textstyle\int f\right|_1&\le|\mathcal{L}^n_\varepsilon(f-h\textstyle\int f)|_1+n\varepsilon\Gamma\|f-h\textstyle\int f\|_{\mathrm{BV}}\\ &\le|\mathcal{L}^n_\varepsilon(f-h_\varepsilon\textstyle\int f)|_1+|\mathcal{L}^n_\varepsilon(h-h_\varepsilon\textstyle\int h)|_1\textstyle\int f+n\varepsilon\Gamma\|f-h\textstyle\int f\|_{\mathrm{BV}}\\ &\le\{\sigma_n\max\{2a_*,\,4B+a\}+n\varepsilon\Gamma(a+b)\}\textstyle\int f.\end{aligned}\tag{8.3}$$

[23] Note that, given $d>B$, the cone $\mathcal{C}^+_d=\{f\ge0;\ \bigvee f\le d\int f\}$ is invariant for $\mathcal{L}^n$, with n large enough, since if $f\in\mathcal{C}^+_d$, $\mathcal{L}^nf\ge0$ and

$$\bigvee\mathcal{L}^nf\le(A\alpha^nd+B)\int f\le d\int f$$

if $B(1-A\alpha^n)^{-1}\le d$. Hence, $\bigvee h\le B$. Similar estimates holds for h_ε. Thus if $\rho=$

The result then follows if

$$\begin{cases} \sigma_n \max\{2a_*,\, 4B + a\} + n\varepsilon\Gamma(a+b) \leq c\sigma_n a \\ A\alpha^n b + aB \leq c\sigma_n b \end{cases} \tag{8.4}$$

At this point one should optimize the choice of a, b, c. Since this is best done on a case by case basis, let us make some reasonable, but simple, choice.

First, let $a = 2a_*$. Then, the first of (8.4) is satisfied if

$$\{(c-1)\sigma_n - n\varepsilon\gamma\}a \geq 4\sigma_n B + n\varepsilon\Gamma b$$

which has solution only if

$$c \leq 1 + n\varepsilon\Gamma\sigma_n^{-1} + 2\frac{B}{a_*}.$$

Let $c = 2 + 2Ba_*^{-1}$ and n such that $n\varepsilon\Gamma\sigma_n^{-1} < 1$, then choose

$$b = \frac{(2a_* + 4B)\sigma_n - 2n\varepsilon\Gamma a_*}{n\varepsilon\Gamma}.$$

The second of (8.4) yields

$$2a_*(1+B)n\varepsilon\Gamma \leq ([2 + 2Ba_*^{-1}]\sigma_n - A\alpha^n)(2a_* + 4B)\sigma_n.$$

Now, let $n_0 = \inf\{n \mid [1 + Ba_*^{-1}]\sigma_n > A\alpha^n\}$. The, for $n \geq n_0$

$$2a_*(1+B)n\varepsilon\Gamma \leq (1 + Ba_*^{-1})\sigma_n^2(2a_* + 4B)$$

which can be solved only if

$$\varepsilon \leq \frac{(1 + Ba_*^{-1})(2a_* + 4B)\sigma_{n_0}^2}{2a_*(1+B)n_0\Gamma}. \tag{8.5}$$

□

Remark 8.1 *A quick computation shows that, if $c\sigma_n < 1$, the Hilbert metric associated to the cone is contracted by $c\sigma_n$. Since σ_n can be chosen at most proportional to $\sqrt{\varepsilon}$, it follows that the rate of decay is estimated by $\sigma_n^{\frac{1}{n}}(1 + \frac{const}{\ln \varepsilon^{-1}})$ instead of the correct $\sigma_n^{\frac{1}{n}}$. Nevertheless, the result is much easier to implement than the more precise approach outlined in the previous sections.*

$\max\{2Ba_*^{-1} - 1, 0\}$, then $(h + \rho h_\varepsilon)(1+\rho)^{-1} \in \widetilde{\mathcal{C}}^{\varepsilon}_{a_*}$ and this implies

$$\begin{aligned} |\mathcal{L}_\varepsilon^n(h - h_\varepsilon)|_1 &= (1+\rho)|\mathcal{L}_\varepsilon^n([h + \rho h_\varepsilon][1+\rho]^{-1} - h_\varepsilon)|_1 \\ &= (1+\rho)|\mathcal{L}_\varepsilon^n(\mathbb{E}([h + \rho h_\varepsilon][1+\rho]^{-1} \mid \mathcal{F}_\varepsilon) - h_\varepsilon \textstyle\int [h + \rho h_\varepsilon][1+\rho]^{-1})|_1 \\ &\leq \max\{a_*,\, 2B\}\sigma^n. \end{aligned}$$

Similar reasoning yields, for $f \in \mathcal{C}_{ab}$,

$$|\mathcal{L}_\varepsilon^n(f - h_\varepsilon \int f)|_1 \leq \max\{a_*,\, 2B + a\}\sigma^n.$$

The results of the present section are summarized by the following.

Lemma 8.3 *Give $\varepsilon > 0$ and $a_* > 0$ such that (8.1) satisfies the property $\mathcal{L}_\varepsilon^k \widetilde{\mathcal{C}}_{a_*}^\varepsilon \subset \widetilde{\mathcal{C}}_{\sigma_k a_*}^\varepsilon$, if n_0, n_1, a, b, c are chosen as in Lemma 8.2, then, for each $n_* \in [n_0, n_1]$, we have*

$$\|\mathcal{L}^n f\|_{BV} \leq \gamma e(a^{-1} + b^{-1})\nu^n \|f\|_{BV}$$

with $\nu := (c\sigma_{n_*})^{\frac{1}{n_*}}$, $\gamma := (1 + \bigvee h + a + b)$.

PROOF To gain information on the behavior of the BV norm we must use Lemma 8.1 to transfer our knowledge of the Hilbert metric to an equivalent norm that satisfies the needed conditions.

One can check directly that the norm

$$\|f\|_{\mathcal{C}} := |\textstyle\int f| + \max\left\{\frac{1}{a}|f - h\int f|;\ \frac{1}{b}\bigvee(f - h\int f)\right\}$$

satisfies the hypotheses of the Lemma 8.1 and that

$$\|f\|_{\mathrm{BV}} \leq (1 + \bigvee h + a + b)\|f\|_{\mathcal{C}} := \gamma\|f\|_{\mathcal{C}}.$$

Accordingly, for each $\beta > 0$,

$$\begin{aligned}\|\mathcal{L}^n f\|_{\mathrm{BV}} &\leq \gamma\|\mathcal{L}^n f\|_{\mathcal{C}} \leq \gamma\|\mathcal{L}^n(f + \beta h) - \mathcal{L}^n \beta h\|_{\mathcal{C}} \\ &\leq \gamma(\exp[\theta(\mathcal{L}^n(f + \beta h), \mathcal{L}^n \beta h] - 1)\beta.\end{aligned}$$

Since $f + \beta h \in \mathcal{C}_{ab}$, if $\beta \geq \max\{b^{-1}\bigvee f, a^{-1}|f|_1\}$, for such β and $n_* \in [n_0, n_1]$ holds

$$\|\mathcal{L}^n f\|_{\mathrm{BV}} \leq \gamma e(c\sigma_{n_*})^{\frac{n}{n_*}}\beta \leq \gamma e(a^{-1} + b^{-1})\|f\|_{\mathrm{BV}}\nu^n$$

with $\nu := (c\sigma_{n_*})^{\frac{1}{n_*}}$. □

9 WHEN IS A MAP PIECEWISE EXPANDING?

In this section we discuss briefly the problem of establishing if a piecewise smooth map is expanding or, more precisely, if there exists $n \in \mathbb{N}$ such that $\lambda_n > 2$. If such an n is known one can try to apply the previous ideas to the operator $\mathcal{L}^n$.[24] Clearly, two main problems arises in this task: a) find the n; b)

[24] One could object that we are really interested in $\mathcal{L}$ not in $\mathcal{L}^n$, yet consider the following. If $\mathcal{L}^n$ has one as simple eigenvalue then the map T^n is ergodic, but then T is ergodic. The same holds for mixing. The only ambiguity could come from the possibility that T is ergodic, but non mixing, and T^n it is not even ergodic. Since the spectral analysis of $\mathcal{L}^n$ provides us also with an approximate knowledge of the intervals belonging to the various ergodic components, it is always possible to check how T acts on such intervals and establish if T is ergodic or not. Finally, if h is the invariant density for $\mathcal{L}^n$, then it is the invariant density for $\mathcal{L}$ as well. Similarly the rate of decay for $\mathcal{L}$ can be deduced from the decay rate of $\mathcal{L}^n$.

get sufficient knowledge on $\mathcal{L}^n$ to obtain the wanted results. Let us discuss this two issues separately.

The situation with the first issue it is not very satisfactory. Indeed, if $\lambda_1 := \inf_y |D_yT| > 1$ then iterating the map a number of times $n \in \mathbb{N}$ such that $n > \frac{\ln 2}{\ln \lambda_1}$ we have $\inf_y |D_yT^n| > 2$. On the contrary it could be that $\lambda_1 < 1$ and yet $\inf_y |D_yT^n| > 2$ for some n. Obviously it would be interesting to have sharp *no-go* conditions that exclude such a possibility. Unfortunately I am not aware of any such result. Yet some obvious conditions do exists. First of all the presence of a periodic point x_p of period n with $|D_{x_p}T^n| \leq 1$ clearly excludes the possibility that the map be expanding. If this trivial test does not detect an impossibility, then I do not know of any way to verify the expansivity of the map beside iterating it numerically and seeing that there are no attractors.[25] If this is the case one can try to do the iterations in a way that keep the numerical errors under control and hope to be able to obtain the wanted result. Of course a failure does not mean necessarily that the map it is not expanding, it just may show a limitation of the method.

This takes us to the second issue: how can we iterate the map (and the transfer operator) numerically but with error bounds. This is a rather vast subject and I do not intend to address it extensively. I will make some simplifying hypothesis: let us assume that the map is piecewise $C^{(3)}$ and that each piece has negative Schwarzian derivative.[26]

For the reader that considers these hypotheses too restrictive lets us point out that one can generalize the following arguments to maps with the property that each branch has a $C^{(2)}$ extension and use the Koebe principle (see Theorem 9.1) and a control on the distortion, rather than negative Schwarzian derivative (see [20] for details on how to do such type of estimates).[27]

Since, in general, the iteration of the map cannot be done analytically (at least not for large iterates) the typical problem will be to control the errors in a numerical iteration of the map, clearly such errors can grow exponentially fast if the map is really expanding as we hope, and this can lead to the impossibility to control reliably the derivative of the higher iterates. Essentially, the only maps

[25]The problem being that if no attractors are visible it does not necessarily means that they are not present, for example they could have a very small basin of attraction of their rate of attraction could be too slow to show up in the numerical study.

[26]Let us remind that the Schwarzian derivative Sf is defined as

$$Sf := \frac{D^3f}{Df} - \frac{3}{2}\left(\frac{D^2f}{Df}\right)^2.$$

The importance of negative Schwarzian derivative lies mostly in the fact that

$$ST^n(x) = \sum_{i=0}^{n-1} ST(T^ix)|D_xT^i|^2$$

hence if a map has negative Schwarzian derivative the same holds for all its iterates. To have an idea of which type of condition we are talking about consider that any polynomial of degree ≥ 2 with real coefficients and all zeros of f' real has $Sf < 0$ (see [20] page 285 Exercise 1.7).

[27]In this case one has to iterate non only the end points of the dynamical partition but also the enlarged end points, apart from this the situation is quite similar.

that can be iterated analytically are piecewise linear maps. In such a case it suffices to iterate the end points of the monotonicity intervals to have a complete knowledge of the iterated map. Consequently, the determination of the iterates becomes a finite computation and thus, in principle, it can be performed with rigorous error bounds.[28] Of course, the number of intervals in the dynamical partition typically grows exponentially with the number of iterations, thus also the determination of iterates for piecewise linear maps may become numerically intractable if one needs to compute really large iterates.

In the following we will see that the general situation under our hypotheses it is not much worst than the piecewise linear case.

First of all the map must be monotone in each smoothness interval otherwise we would have $\min|DT^n| = 0$ for each n and the map would not be piecewise expanding. Let $\mathcal{Z}^{(n)}$ be the smoothness partition associated to T^n. The first basic result is the so called *Minimum Principle*.

Proposition 9.1 (Minimum Principle [20]) *Let I be a closed interval with end points a, b and $T : I \to \mathbb{R}$ a map with negative Schwarzian derivative. If $DT(x) \neq 0$ for all $x \in I$, then*

$$|D_xT| > \min\{|D_aT|, |D_bT|\}, \quad \forall x \in I.$$

The above result implies that it suffices to compute the derivatives at the end points of the intervals of $\mathcal{Z}^{(n)}$, if the derivative becomes larger than one at all such points, then the map is expanding.

Once the expansivity is established it remains to compute the proper power of the Transfer operator. The basic idea is to approximate it via the Transfer operators associated to piecewise linear maps by taking advantage of the following perturbative result.

Proposition 9.2 ([50]) *Let $\mathcal{L}_\varepsilon$ be the Transfer operators of maps T_ε which are close to T. The distance between two maps being defined by*

$$\begin{aligned} d(T, T_\varepsilon) := \inf\{\kappa > 0 \mid \exists\, A \subseteq I\, \exists\, \omega : I \to I \textit{ s.th. } m(A) > 1 - \kappa,\ \omega \textit{ is a diffeo-}\\ \textit{morphism, } T_{\varepsilon|_A} = T \circ \omega_{|_A}, \textit{ and } \forall x \in I\textit{: } |\omega(x) - x| < \kappa,\\ |1/\omega'(x) - 1| < \kappa\}. \end{aligned}$$

Then

$$|||\mathcal{L} - \mathcal{L}_\varepsilon||| \leq 12 d(T, T_\varepsilon).$$

To be more concrete, let $n_0 \in \mathbb{N}$ be the first integer for which the map has been established to have expansions coefficient strictly larger than 2. Let $\mathcal{Z}_* := \{[a_i, a_{i+1}]\}$ any partition finer than $\mathcal{Z}^{(n_0)}$, then define the map

$$T_*x := T^{n_0}a_i + \frac{T^{n_0}a_{i+1} - T^{n_0}a_i}{a_{i+1} - a_i}(x - a_i) \quad \forall x \in [a_i, a_{i+1}]. \tag{9.1}$$

[28] To be really on the safe side on can employ *interval arithmetic* that allows to do arithmetic operations with rigorous error bounds, see [52] for a survey.

We can then apply Proposition 9.2 to determine the difference between the true Transfer operator $\mathcal{L}^{n_0}$ and the operator $\mathcal{L}_*$ associated to the map T_*. In fact, we can define

$$\omega(x) := T^{-n_0}|_Z T_* x \quad \forall x \in Z \in \mathcal{Z}_*.$$

Clearly, $\omega(a_i) = a_i$ and $T^{n_0} = T_* \circ \omega$. Moreover,

$$|1 - \omega'(x)^{-1}| = \left|1 - \frac{D_{\sigma(x)}T^{n_0}}{D_x T_*}\right| \leq \sup_{Z\in\mathcal{Z}_*} \sup_{x,y\in Z} \left|1 - \frac{D_x T^{n_0}}{D_y T^{n_0}}\right| = D_*. \tag{9.2}$$

Hence, if $\delta = \sup_{Z\in\mathcal{Z}_*} |Z|$, and $D_* < 1$, we have

$$|x - \omega(x)| \leq \delta D_* < D_*.$$

Accordingly, $d(T^{n_0}, T_*) \leq D_*$.[29]

In conclusion, the difference between the piecewise linear approximation and the true Transfer operator is, not surprisingly, determined by the distortion of the map T^{n_0}. Such a distortion may not be easy to compute if n_0 is large. To accomplish this last task another fundamental results about maps of negative Schwarzian derivative can be helpful.

Let $I \subset J$ be two intervals. we say that J is an ε–scaled neighborhood of I if both components of $I\backslash J$ have length $\varepsilon|I|$.

Theorem 9.1 (Kobe Principle [20]) *Let $I \subset J \subset [0,1]$ be two intervals and let $T : J \to [0,1]$ be a $C^{(3)}$ diffeomorphism with negative Schwarzian derivative. If $T(J)$ contains a τ-scaled neighborhood of $T(I)$ then*

$$\left|\frac{D_x T}{D_y T} - 1\right| \leq \frac{1+2\tau}{\tau^2}, \quad \forall x, y \in I.$$

To apply the above result we need to extend the map in such a way to have scaled neighborhoods. Clearly, each branch T_Z, $Z \in \mathcal{Z}^{(1)}$, can be extended to a larger interval $\hat{Z}$ such that the extension is still strictly monotone, $C^{(3)}$ and with negative Schwarzian derivative.[30] Let $\rho > 0$ be the minimal length of the connected components of $Z\backslash\hat{Z}$. Let $\hat{\mathcal{Z}}^{(1)} := \{\hat{Z}\}_{Z\in\mathcal{Z}^{(1)}}$. We can then define a new map $\hat{T} : \cup_{\hat{Z}\in\hat{\mathcal{Z}}^{(1)}} \hat{Z} \to \mathbb{R}$.[31] A bit more care is needed to define its powers. Let us consider the partitions $\mathcal{Z}_*$ and $\mathcal{Z}^{(1)} = \{Z_i\}_{i=1}^N$, for each $Z \in \mathcal{Z}_*$ let $\sigma(Z) \in N^{n_0}$ be determined by the history of Z, that is $\sigma(Z)_i = k$ iff $T^i Z \in$

[29] If the reader is worried that ω it is not a diffeomorphism because the derivative is may be discontinuous at the points a_i, consider that one can smoothen ω^{-1} by convolving it with a smooth function of support ε. The resulting function will differ, in the sup norm, from the original one by less than $D_*\varepsilon$ and will be now a diffeomorphism, moreover its derivative will differ from one by less than D_* as well. Taking the limit $\varepsilon \to 0$ one obtains the above statement.

[30] For example one can use the Taylor formula to third order computed at the end points to make the extension.

[31] Note that if $\hat{Z}_1 \cap \hat{Z}_2 \neq \emptyset$, then we double the points so that $\hat{T}|_{\hat{Z}_1}$ and $\hat{T}|_{\hat{Z}_2}$ may differ on the overlap.

$Z_k \in \mathcal{Z}^{(1)}$, clearly there exists a set $\Sigma \subset N^{n_0}$ such that $\mathcal{Z}_* = \{Z_{i,\sigma}\}_{i \in N_\sigma, \sigma \in \Sigma}$ where N_σ is the number of intervals in which corresponding element of $\mathcal{Z}^{(n_0)}$ has been subdivided and $\sigma(Z_{i,\sigma}) = \sigma$. We can then define the new partition $\hat{\mathcal{Z}}_* := \{\hat{Z}_\sigma\}_{\sigma \in \Sigma}$ defined by

$$\hat{Z}_\sigma = \hat{T}|_{\hat{Z}_{\sigma_1}}^{-1} \circ \cdots \circ \hat{T}|_{\hat{Z}_{\sigma_{n_0}}}^{-1} \mathbb{R}.$$

It is now possible to clearly define the powers of $\hat{T}^i$ for $i \leq n_0$:

$$\hat{T}^i x := \hat{T}|_{\hat{Z}_{\sigma_i}} \circ \cdots \circ \hat{T}|_{\hat{Z}_{\sigma_1}}(x) \quad \forall x \in \hat{Z}_\sigma.$$

To apply the Koebe principle we will then consider the images of the intervals $Z_{i,\sigma} \subset \hat{Z}_\sigma$ under $\hat{T}^{n_0}$. Clearly there exists ρ_* such that the connected components of $T^{n_0} Z_\sigma \backslash \hat{T}^{n_0} \hat{Z}_\sigma$ are larger than ρ_*. We will then choose the partition $\mathcal{Z}_*$ so that $\sup_{Z \in \mathcal{Z}_*} |T^{n_0} Z| \leq \delta_*$. We can now apply Theorem 9.1 with

$$\tau = \rho_* \delta_*^{-1} \geq 5. \tag{9.3}$$

Relations (9.3), (9.2) and Proposition 9.2 allow to conclude that

$$|||\mathcal{L}^{n_0} - \mathcal{L}_*||| \leq 12 D_* = \bar{D} \delta_*, \tag{9.4}$$

where $\bar{D} = 12 \frac{\delta_* + 2\rho_*}{\rho_*^2} \leq 12(1 + 2\rho_*)\rho_*^{-2}$.

In conclusion, we have outlined the following scheme to tackle the problem:

1. *Compute the derivative at the end points of the dynamical partitions to find out the possible existence of an n_0 for which the expansion rate is larger than 2.*
2. *Consider an extension of the map and a refined partition such that the error in (9.4) is sufficiently small.*
3. *Consider the piecewise linear approximation T_* of the map T^{n_0}. Apply Lemma 4.1 to the operator $\mathcal{L}_*$ associated to T_*, then, using the estimate (9.4) we have $|||\mathcal{L}^{n_0} - \mathcal{L}_{*,\varepsilon}||| \leq \Gamma\varepsilon + \bar{D}\delta_*$.*[32]
4. *Apply the techniques of the previous section to obtain informations on the statistical properties of the map T.*
5. *If the δ_* and ε turns out to be too large to obtaining the needed informations, repeat the previous steps with smaller ε, δ_*.*

To sum it up, we could say that the linear approximation is used to control the non linearity while the Ulam scheme is used to control the lack of the Markov property.

[32]Note that both Γ and the matrix associated to an Ulam scheme can be computed analytically for a piecewise linear map.

10 CONCLUSIONS AND SUPPLEMENTARY CONSIDERATIONS

The conclusion is that it could be possible to write a computer code able to answer, in many cases, the above questions.[33] Yet there are some, possibly dynamically relevant, questions that we have not addressed and that I do not see how to tackle with the present techniques.

For example, one could be interested in approximating, in BV norm, the eigenprojectoins associated to the eigenvalues of modulus less than one.

What we have discussed allows to gain information, also in BV norm, only on the peripherical eigenprojections. The reason for this difference rests in the fact that peripherical eigenprojections are completely determined by measures with L^∞ density with respect to Lebesgue. This is typically false for the other eigenprojections. To see this consider the case of a simple eigenvalue λ, $|\lambda| < 1$. Let Π be the associated eigenprogection and suppose

$$\|\Pi f\|_{\mathrm{BV}} \leq c|f|_1. \tag{10.1}$$

This means that

$$\Pi f = h_1 \int gf$$

for some $h_1 \in \mathrm{BV}$ and $g \in L^\infty$. But then $g \circ T = \lambda g$ and therefore $\lambda^n \in \sigma(\mathcal{L})$ for each $n \in \mathbb{N}$.[34] Consequently, if such eigenvalues are not there, as happens in several examples, (10.1) must be false.

It is important to remark that the analysis carried out in this paper is not limited to the case of one dimensional maps and BV spaces. Given the fact that the two main tools (Proposition 3.1 and the projective metrics) are applicable in a rather general contest, it is possible to parallel the above arguments in many other cases. Just as an example let us mention two: expanding maps in higher dimensions and one dimensional Markov maps where the operator can be considered on spaces of functions with higher smoothness.

Let us say two words on each example.

10.1 Higher dimensions

The first possibility is to consider an higher dimensional expanding case. Namely a piecewise analytic expanding map $T : [0, 1]^d \to [0, 1]^d$.[35] This systems have been widely investigated and this it is not the place to review the history of the subject, it will suffice to mention that results analogous to the one dimensional case hold. One can see [62] for a general treatment of this type of maps along the lines mentioned in section 2. As for the study of finite dimensional

[33] This would be a very feeble attempt at what D.Ruelle (jokingly ?) calls *Galactic Mathematics* [65]. Of course, being an earthling, I just argue about feasibility, rather than actually do it.

[34] Since $g_n(x) = g(x)^n$ is an eigenfunction as well: $g_n \circ T = (g \circ T)^n = \lambda^n g_n$.

[35] The analyticity hypothesis is completely inessential, it is introduced here just for simplicity so that (10.2) holds without any extra technical hypothesis.

approximations some partial result are in [29, 35] where the invariant density is studied. Here we will use several estimates developed in [13] to show that the present case can be treated in perfect analogy with the theory developed so far.

If we consider a partition is cubes of size ε then [13] Lemma 10 shows that[36]

$$\|\mathbb{E}(f \mid \mathcal{F}_\varepsilon)\|_{\mathrm{BV}} \leq d\|f\|_{\mathrm{BV}}.$$

Moreover, from [13] Corollary 3

$$|f - \mathbb{E}(f \mid \mathcal{F}_\varepsilon)|_1 \leq d\varepsilon\|f\|_{\mathrm{BV}}.$$

Finally, the Lasota-Yorke inequality holds [13], [14], [62], [67].[37]

$$\|\mathcal{L}^n f\|_{\mathrm{BV}} \leq A\alpha^n\|f\|_{\mathrm{BV}} + B|f|_1. \tag{10.2}$$

If k is such that $d^2 A\alpha^k = \theta < 1$, then, setting

$$\mathcal{L}_\varepsilon f := \mathbb{E}(\mathcal{L}^k \mathbb{E}(f \mid \mathcal{F}_\varepsilon) \mid \mathcal{F}_\varepsilon)$$

we have

$$\|\mathcal{L}_\varepsilon f\|_{\mathrm{BV}} \leq \theta\|f\|_{\mathrm{BV}} + dB|f|_1.$$

so

$$\|\mathcal{L}_\varepsilon^n f\|_{\mathrm{BV}} \leq \theta^n\|f\|_{\mathrm{BV}} + \frac{dB}{1-\theta}|f|_1.$$

Thus all what we said about the one dimensional case can be applied to the present one almost verbatim.

Of course, if the dimension is very high the problem quickly becomes numerically intractable. To obtain useful and reliable informations in very high dimensional cases clearly would require a brand new set of ideas.

10.2 Higher smoothness

As far as the Markov case is concerned let us limit the comments to the easiest case: C^3 circle maps.[38] In this case it is natural to introduce the norms

$$\begin{aligned} |f| &= |f|_1 + |f'|_1 \\ \|f\| &= |f| + |f'|_1 + |f''|_1. \end{aligned} \tag{10.3}$$

In other word we consider the Sobolev spaces $W^{1,1}$ and $W^{2,1}$ rather then the spaces L^1 and BV.

[36] Let us remind that the variation of a function f it is defined, in general as

$$\bigvee f := \sup_{\varphi \in \mathcal{C}^{(1)}} \frac{\int f \operatorname{div}\varphi}{\|\varphi\|_\infty},$$

see [18] for a detailed discussion of such notion and its consequences.

[37] Here is the only place where we use the analyticity assumption. See [62] for a more general setting. The necessity of extra conditions for the non-analytic case is shown by a counterexample of Tsujii [68].

[38] The extension to full branched $C^{(3)}$ maps is trivial.

Lemma 10.1 *The following Lasota-Yorke type inequality holds*

$$\|\mathcal{L}^n f\| \le \lambda_n^{-2}\|f\| + B_1|f|.$$

PROOF The proof is by direct computation. First of all remember that $\mathcal{L}^n f(x) = \sum_{y\in T^{-n}x} g_n(y)f(y)$ where $g_n(y) := |D_yT^n|^{-1}$. Then note that, since $T \in C^{(3)}$ and the sign of DT^n is constant, $g_n \in C^{(2)}$.

$$\frac{d}{dx}\mathcal{L}^n f(x) = \sum_{y\in T^{-n}x} g_n'(y)g_n(y)f(y) + f'(y)g_n(y)^2,$$

since $\frac{dy}{dx} = g_n(y)$. Hence

$$\begin{aligned}\frac{d^2}{dx^2}\mathcal{L}^n f(x) = \sum_{y\in T^{-n}x} & g_n''(y)g_n(y)^2 f(y)\\ & + g_n'(y)^2 g_n(y)f(y) + 3g_n'(y)g_n(y)^2 f'(y) + g_n(y)^3 f''(y).\end{aligned}$$

In addition,

$$\begin{aligned}g_n'(y) &= \sum_{j=1}^n g_{n-j-1}(T^{j+1}y)\left[\frac{d}{dy}|D_{T^jy}T|^{-1}\right] g_{j-1}(y)\\ &= -\sum_{j=1}^n g_{n-j-1}(T^{j+1}y)\frac{D_{T^jy}|D_{T^jy}T|}{|D_{T^jy}T|^2} := \sum_{j+1}^n g_{n-j-1}(T^{j+1}y)Di(T^jy).\end{aligned}$$

Thus $Di \in C^{(1)}$ and $|g_n'|_\infty \le \sum_{j=1}^\infty \lambda_j^{-1}|Di|_\infty := A_1$. In addition,

$$g_n'' = \sum_{j=1}^n \frac{g_{n-j}'(T^jy)}{g_j(y)}Di(T^jy) + \frac{g_{n-j}(T^jy)Di'(T^jy)}{g_j(y)}$$

hence

$$|g_n g_n''|_\infty \le \sum_{j=1}^\infty \lambda_j^{-1}A_1|Di|_\infty + \sum_{j=1}^\infty \lambda_j^{-2}|Di'|_\infty := B_0.$$

Accordingly,

$$\begin{aligned}|\frac{d}{dx}\mathcal{L}^n f(x)| &\le \lambda_n^{-1}\mathcal{L}^n|f'|(x) + A_1\mathcal{L}^n|f|(x)\\ |\frac{d^2}{dx^2}\mathcal{L}^n f(x)| &\le (A_1^2 + B_0)\mathcal{L}^n|f|(x) + 3\lambda_n^{-1}A_1\mathcal{L}^n|f'|(x) + \lambda_n^{-2}\mathcal{L}|f''|(x).\end{aligned}$$

The lemma follows then by integrating and remembering that the Transfer operator is an L^1 contraction. □

Remark 10.1 *The inequality proved in Lemma (10.1) shows that the essential spectrum of $\mathcal{L}$ on $W^{2,1}$ is bounded by ρ_σ^2 rather than by ρ_σ which is the bound on BV (or on $W^{1,1}$). This is a particular instance of a general fact (see [63], [64]).*[39]

Of course, in this context a standard Ulam type approximation would be disastrous since it would not preserve the smoothness of the space. The problem can easily be overcame. Just consider a sufficiently smooth partition of unity $\{\varphi_i\}_{i=1}^n$ such that $\varphi_i \geq 0$ and with at most two overlap: $\varphi_i\varphi_j \not\equiv 0$ implies $|i-j| \leq 1$, moreover require $\varphi_1(0) = \varphi_n(1) = 1$. If we call Δ_i the support of φ_i, then $\{\Delta_i \cap \Delta_{i+1}\}$ is a collection of disjoint intervals. Assume that

$$|\Delta_i| \leq \varepsilon.$$

We can now define the following objects

$$\begin{aligned}
&z_i := \left[\int_{[0,1]} \varphi_i\right]^{-1}\\
&\phi_i(x) := z_i \int_0^x \varphi_i\\
&\mathbb{P}_\varepsilon f := f(0) + \sum_i \phi_i \int_{[0,1]} \varphi_i f'\\
&\qquad = (1-\phi_1)f(0) + \phi_n f(1) - \sum_i \phi_i \int_{[0,1]} \varphi_i' f\\
&\mathcal{L}_\varepsilon := \mathbb{P}_\varepsilon \mathcal{L} \mathbb{P}_\varepsilon.
\end{aligned} \tag{10.4}$$

The basic estimate that allows to rederive all the results of the paper in the present setting is the following.

Lemma 10.2 *For each $f \in W^{2,1}$ holds*

$$|\mathbb{P}_\varepsilon f - f| \leq 2\varepsilon \|f\|. \tag{10.5}$$

In addition,

$$\begin{aligned}
&|\mathbb{P}_\varepsilon f| \leq 2|f|\\
&\|\mathbb{P}_\varepsilon f\| \leq \|f\| + 2|f|.
\end{aligned} \tag{10.6}$$

Proof Start by noting that, by definition, $\mathbb{P}_\varepsilon f(0) = f(0)$ and $\mathbb{P}_\varepsilon f(1) = f(1)$. In addition, it is obvious that

$$(\mathbb{P}_\varepsilon f)' = \sum_i \varphi_i z_i \int \varphi_i f.$$

and hence

$$|(\mathbb{P}_\varepsilon f)'|_1 \leq |f'|_1.$$

[39] See (2.4) for a definition of ρ_σ.

Moreover,

$$|(\mathbb{P}_\varepsilon f)' - f'|_1 \le \int \left| \sum_i \varphi_i \left[z_i \int \varphi_i f' - f' \right] \right|_1 .$$

To continue notice that f' is continuous, thus in each Δ_i there must exists a x_i such that

$$z_i \int \varphi_i f' = f'(x_i). \tag{10.7}$$

Accordingly,

$$|(\mathbb{P}_\varepsilon f)' - f'|_1 \le \sum_i \int \varphi_i(x)|f'(x_i) - f'(x)|dx \le \sum_i \int \varphi_i \int_{\Delta_i} |f''| \le 2\varepsilon |f''|_1.$$

Using the above estimates easily follows

$$|(\mathbb{P}_\varepsilon f) - f|_1 \le \varepsilon |f''|_1.$$

To obtain the last needed estimate is convenient to define $\Delta_i^* = \Delta_i \cap \Delta_{i+1}$

$$\begin{aligned} |(\mathbb{P}_\varepsilon f)''|_1 &= \int \left| \sum_i \varphi_i' z_i \int \varphi_i f' \right|_1 = \int \left| \sum_i \varphi_i' f'(x_i) \right| \\ &\le \sum_{i=1}^{n-1} \int_{\Delta_i^*} \varphi_{i+1} |f'(x_{i+1}) - f'(x_i)| \\ &\le \sum_{i=1}^{n-1} \int_{x_i}^{x_{i+1}} |f''| \le |f''|_1, \end{aligned}$$

where $\{x_i\}$ are defined in (10.7). Finally

$$|\mathbb{P}_\varepsilon f|_1 \le |f|_1 + |\mathbb{P}_\varepsilon f - f|_1 \le |f|_1 + 2|f'|_1.$$

□

The above Lemma implies

$$\begin{aligned} \|\mathcal{L}_\varepsilon f\| &\le \|\mathcal{L}\mathbb{P}_\varepsilon f\| + 2|\mathcal{L}\mathbb{P}_\varepsilon f| \le \lambda_1^{-2} \|\mathbb{P}_\varepsilon f\| + 2\lambda_1^{-1} |\mathbb{P}_\varepsilon f| + 2B|\mathbb{P}_\varepsilon f|_1 + B_1 |\mathbb{P}_\varepsilon f| \\ &\le \lambda_1^{-2} \|f\| + \{2\lambda^{-2} + 4\lambda_1^{-1} + 4B + 2B_1\} |f| \end{aligned}$$

which yields the Lasota-Yorke inequality for $\lambda_1 > 1$.[40]

We are thus in a situation absolutely similar to the one treated in the previous sections, hence all the above considerations apply to the present case.

Note that, in this context, the proposed scheme will provide optimal estimates for the invariant measure and the decay of correlations in $W^{1,1}$, while the estimates in $W^{2,1}$ have the same problems the estimates in BV had in the preceding discussion.

Clearly, if the map has even higher regularity, then higher smoothness can be treated exactly as in the present discussion.

[40] As usual one can get better result using a power of $\mathcal{L}$ instead of $\mathcal{L}$ itself.

References

[1] G.Birkhoff, *Lattice Theory*, American Mathematical Society Colloquium Publications, Providence, Rhode Island, **25**, 3rd ed., (1967).

[2] G.Birkhoff, *Extension of Jentzsch's Theorem*, Transaction American Mathematical Society, **85**, (1957), 219-227.

[3] V.Baladi, *Positive Transfer Operators and Decay of Correlations*, Advanced Series in Nonlinear Dynamics, **16**, World Scientific, Singapore (2000).

[4] V.Baladi, M.Holschneider, *Approximation of nonessential spectrum of transfer operators*, Nonlinearity, **12**, (1999), no. 3, 525–538.

[5] M.L.Blank, *Stochastic properties of deterministic dynamical systems,* Sov. Sci. Rev. C. Math./Phys., **6**, (1987), 243-271.

[6] M.L.Blank, *Small perturbations of chaotic dynamical systems,* Usp. Mat. Nauk, **44**, (1989), 3-28.

[7] M.L.Blank, *Singular phenomena in chaotic dynamical systems* Dokl. Akad.Nauk. (Russia), **328**, (1993), 7-11.

[8] M.L.Blank, *Chaotic mappings and stochastic Markov chains*, Abstracts of congress IAMP-91, 6p.

[9] M.L.Blank, G.Keller, *Stochastic stability versus localization in chaotic dynamical systems,* Nonlinearity, **10**, (1997), 81-107.

[10] M.L.Blank, G.Keller, *Random perturbations of chaotic dynamical systems: Stability of the spectrum,* Preprint (1998).

[11] A.Boyarsky, P.Góra, *Absolutely continuous invariant measures for piecewise expanding C^2 transformations in $\mathbb{Z}^N$,* Israel J. Math., **67**, (1989), 272-286.

[12] A. Boyarsky, P. Góra, *Laws of Chaos. Invariant Measures and Dynamical Systems in One Dimension,* Birkhäuser, (1997).

[13] J.Buzzi, G.Keller, *Zeta functions and Transfer operators for multidimensional piecewise affine and expanding maps,* preprint.

[14] J.Buzzi, *A.c.i.m.'s for arbitrary expanding piecewise real-analytic mappings for the plane*, Ergod. Th.& Dynam. Sys., (to appear).

[15] V.Baladi, L.S.Young, *On the spectra of randomly perturbed expanding maps,* Comm. Math. Phys., **156**, (1993), 355-385 (erratum **166**, (1994), 219-220).

[16] C.Chiu, Q.Du, T.Y.Li, *Error estimates of the Markov finite approximation to the Frobenius-Perron operator,* Nonlinear Analysis, Theory, Methods & Appl., **19**, (1992), 291-308.

[17] P.Collet, S.Isola, *On the essential spectrum of the transfer operator for expanding Markov maps,* Comm. Math. Phys., **139**, (1991), no. 3, 551–557.

[18] L.C.Evans, R.F.Gariepy, *Measure theory and fine properties of functions,* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[19] M.Dellnitz, O.Junge, *On the approximation of complicated dynamical behavior*, SIAM J. Num. Anal., **36**, 491–515, (1999).

[20] W.de Melo, S.van Strien, *One-dimensional dynamics*, Springer Verlag, Berlin, Heidelberg, New York, (1993).

[21] J.Ding, Q.Du, T.-Y.Li, *The spectral analysis of Frobenius-Perron operator*, J. Math. Anal. Appl., **184**(2), (1994), 285–301.

[22] J.Ding, Q.Du, T.-Y.Li, *Higher order approximations of Frobenius-Perron operator*, Appl. Math. Compt., **53**, (1993) 157–171.

[23] J.Ding, L.J.Huang, *A generalization of a classic theorem in the perturbation theory for linear operators*, J. Math. Anal. Appl., **239**, (1999), no. 1, 118–123.

[24] J.Ding, *Computing invariant measures of piecewise convex transformations*, J. Statist. Phys., **83**, (1996), no. 3-4, 623–635.

[25] J.Ding, *Absolutely continuous invariant measure of a piecewise concave mapping of* $[0,1]$, Nonlinear Anal., **28**, (1997), no. 7, 1133–1140.

[26] J.Din, T.-Y.Li, *Markov finite approximations of Frobenius-Perron operator*, Nonlinear analysis, Theory, Methods & Applications, **17**, (1991), 759–772.

[27] J.Ding, T.-Y.Li, *Projection solutions of Frobenius-Perron operator equations*, Internat. J. Math. Math. Sci., **16**, (1993), no. 3, 465–484.

[28] J.Ding, T.-Y.Li, *A convergence analysis for Markov finite approximations to a class of Frobenius-Perron operator*, Nonlinear analysis, Theory, Methods & Applications, **31**, n. 5/6, (1998), 765–776.

[29] J.Ding, A.Zhou, *Finite approximation of Frobenius-Perron operators. A solution of Ulam's conjecture to multi-dimensional transformations,* Phys. D, **92**, n.1-2, (1996), 61-68.

[30] J.Ding, A.Zhou, *Computing fixed densities of a class of Frobenius-Perron operators,* Nonlinear Anal., **32**, (1998), no. 3, 409–416.

[31] J.Ding, A.Zhou *A convergence rate analysis for Markov finite approximations to a class of Frobenius-Perron operators*, Nonlinear Anal., **31**, (1998), no. 5-6, 765–777.

[32] J.Ding, A.Zhou, *The projection method for a class of Frobenius-Perron operators,* Appl. Math. Lett., **12**, (1999), no. 1, 71–74.

[33] J.Ding, A.Zhou, *A finite element method for the Frobenius-Perron operator equation*, Appl. Math. Comput., **102**, (1999), no. 2-3, 155–164.

[34] J.Ding, A.Zhou *Computing fixed points of quasi-compact Frobenius-Perron operators,* Nonlinear Anal., **39**, (2000), no. 2, Ser. A: Theory Methods, 231–239.

[35] J.Ding, A.Zhou, *Absolutely continuous invariant measures for piecewise* C^2 *and expanding mappings in higher dimensions*, Discrete Contin. Dynam. Systems, **6**, (2000), no. 2, 451–458.

[36] G.Froyland, *Computer-assisted bounds for the rate of decay of correlations,* Commun. Math. Phys., **189**, (1997), 237-257.

[37] G.Froyland, *Finite approximation of the Sinai-Bowen-Ruelle measures for Anosov systems in two dimensions,* Random Compt. Dynam., **3**, (1995), 251–264.

[38] H.Hennion, *Sur un théorème spectral et son application aux noyaux Lipchitziens,* Proc. Amer. Math. Soc., **118**, (1993), 627-634.

[39] F. Hofbauer, G.Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations,* Math. Zeitschrift, **180**, (1982), 119-140.

[40] F.Y.Hunt, W.Miller, *On the approximation of invariant measures,* J. Stat. Phys., **66**, (1992), 535-548.

[41] F.Y.Hunt, *A Monte Carlo approach to the approximation of invariant measures*, Random Comput. Dynam., **2**, (1994), 111-133.

[42] F.Y.Hunt, *Erratum: "A Monte Carlo approach to the approximation of invariant measures" [Random Comput. Dynam. 2 (1994), no. 1, 111–133]*, Random Comput. Dynam., **5** (1997), no. 4, 361–362.

[43] T.Kato, *Perturbation Theory for linear operators*, Springer Verlag, New York,(1966).

[44] G.Keller, C.Liverani, *Stability of the Spectral Gap for transfer operators*, Annali della Scuola Normale di Pisa, Classe di Scienze (4), Vol.**XXVIII**, 141-152, (1999).

[45] C.T.Ionescu Tulcea, G.Marinescu, *Thórie ergodique pour des classes d'opérations non complètement continues,* Ann. Math., **52**, (1950), 140-147.

[46] M.Keane, R.Murray, L.S.Young, *Computing invariant measures for expanding circle maps,* Nonlinearity, **11**, (1998), 27-46.

[47] G.Keller, *Ergodicité et mesures invariantes pour les transformations dilatantes par morceaux d'une région bornée du plan,* C. R. Acad. Sci. Paris, Série A, **289**, (1979), 625-627 (Kurzfassung der Dissertation).

[48] G.Keller, *Un théorème de la limite centrale pour une classe de transformations monotones par morceaux,* C. R. Acad. Sci. Paris, Série A, **291**, (1980), 155-158.

[49] G.Keller, *On the rate of convergence to equilibrium in one-dimensional systems,* Commun. Math. Phys., **96**, (1984), 181-193.

[50] G.Keller, *Stochastic stability in some chaotic dynamical systems,* Monatshefte Math., **94**, (1982), 313-333.

[51] Yu.Kifer *Random perturbations of dynamical systems* Boston, MA, Birkhauser, (1988).

[52] O.Lanford III *Computer assisted proofs in analysis*, Phys. A, 124, 465 (1984).

[53] C.Liverani, *Decay of Correlations*, Annals of Mathematics, **142**, pp. 239-301, (1995).

[54] C.Liverani, *Decay of Correlations in Piecewise Expanding maps*, Journal of Statistical Physics, **78**, 3/4, pp. 1111-1129, (1995).

[55] T.-Y.Li, *Finite approximations for the Frobenius-Perron operator: A solution to Ulam's conjecture,* J. Approx. Th., **17**, (1976), 177-186.

[56] C.Liverani, B.Saussol, S.Vaienti, *Conformal Measure and Decay of Correlations for Covering Weighted Systems*, Ergodic Theory and Dynamical Systems, **18**, 6, pp. 1399–1420 (1998).

[57] A.Lasota, J.A.Yorke, *On the existence of invariant measures for piecewise monotonic transformations,* Transact.Amer.Math.Soc. ,**186**, (1973), 481-488.

[58] W. Miller, *Stability and approximation of invariant measures for a class of nonexpanding transformations,* Nonlinear Analysis, Theory, Methods & Appl., **23**, (1994), 1013-1025.

[59] V. Maume-Deshamps *Propriétés de mélange pour des systèmes dynamiques Markoviens*, Thèse, Université de Bourgogne, Dijon (1998).

[60] R.Murray, *Approximation error for invariant density calculations*, Discrete Continuous Dyn. Sys., **4**, n.3, (1998), 535-557.

[61] M.Pollicott, *Meromorphic extensions of generalised zeta functions,* Invent. Math., **85**, (1986), no. 1, 147–164.

[62] B.Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps,* preprint.

[63] D.Ruelle, *The thermodynamics formalism for expanding maps*, Comm.Math.Phys., **125**, 239-262, (1989).

[64] D.Ruelle, *An extension of the theory of Fredholm determinants*, Int. Hautes Etudes Sci.Publ.Math., **72**, 175–193, (1990).

[65] D.Ruelle, *Conversations on mathematics with a visitor from outer space*, preprint.

[66] M.Rychlik, *Bounded variation and invariant measures,* Studia Math., **76**, (1983), 69-80.

[67] M.Tsujii, *Absolutely continuous invariant measures for piecewise real-analytic expanding maps on the plane,* Comm. Math. Phys.,**208**, (2000), no. 3, 605–622.

[68] M.Tsuji, *Piecewise Expanding maps on the plane with singular ergodic properties,* to appear in Ergodic Theory and Dynamical Systems.

[69] Ulam *Problems in Modern Mathematics*, New York, Interscience (1960).

[70] M.Viana, *Stochastic dynamics of deterministic systems*, Col. Bras. de Matemática, notes (1997).

[71] M.Walter, *Discrete approximation of invariant measures for multidimensional maps*, DIMACS Ser.Discrete Math.Theoret.Comput.Sci., **34**, (1997).