Chapter 7

Quantitative Statistical Properties, a class of 1-d examples

Given a Dynamical System it is in general very hard to study its ergodic properties, especially if the goal is to have a *quantitative* understanding. To make clear what is meant by a *quantitative understanding* and which type of obstacles may prevent it, I devote this chapter to the study of a simple, but highly non-trivial, class of examples: one dimensional smooth expanding maps.

7.1 The problem

Recall from Examples 6.4.1 that a one dimensional smooth expanding map is a map $T \in \mathcal{C}^2(\mathbb{T}^1, \mathbb{T}^1)$ such that $|DT| \geq \lambda > 1$.

We know already that such maps have a unique absolutely continuous invariant measure (see sections 6.4.1, 6.5.1 Expanding maps).

We would like first to understand other invariant measures in order to have a clearer picture of which measurable Dynamical Systems can be associated to the topological Dynamical System $(\mathbb{T}^1, T)$. This is still at the qualitative level. In addition, we would like to have tools to actually compute such invariant measures with a given precision, and this is a first quantitative issue.

Next, we would like to study statistical properties more in depth. To this end we will restrict to the case $(\mathbb{T}^1, T, \mu)$, where μ is the measure absolutely continuous with respect to Lebesgue. The type of questions we would like to address are

If we make repeated finite time and precision measurements, what do we observe?

Remember that a measurement is represented by the evaluation of a function. The fact that the measurement has a finite precision correspond to the fact that the function has some uniform regularity (otherwise we could identify the point with an arbitrary precision). The fact that the measure is made for finite time means that we are able only to measure finite times averages. In other words we would like to understand the behavior of

$$\sum_{k=0}^{N-1} f \circ T^k$$

for large, but finite, N.

7.2 Invariant measures

Let $\mathcal{M}$ be the set of probability (Borel) measures on $\mathbb{T}^1$. We can then consider the new Dynamical System $(\mathcal{M}, T')$, where $T'\mu(f) = \mu \circ T$ for all $f \in \mathcal{C}^0(\mathbb{T}^1, \mathbb{R})$. The invariant measures are the fixed points of T', let us call them $\mathrm{Fix}(T')$. If $\mu \in \mathrm{Fix}(T')$ then for each $h \in L^\infty(\mathbb{T}^1, \mu)$, $h \geq 0$, $\mu(h) = 1$, we can consider the new probability measure defined by $\mu_h(f) = \mu(hf)$, for all $f \in \mathcal{C}^0(\mathbb{T}^1, \mathbb{R})$. Note that

$$|T'\mu_h(f)| = |\mu(hf \circ T)| \leq |h|_{L^\infty(\mu)}\mu(|f| \circ T) = |h|_{L^\infty(\mu)}\mu(|f|).$$

Hence $T'\mu_h$ is absolutely continuous with respect to μ and $\frac{dT'\mu_h}{d\mu} \in L^\infty(\mu)$. We can then define the operator $\mathcal{L}_\mu : L^\infty(\mathbb{T}^1, \mu) \to L^\infty(\mathbb{T}^1, \mu)$ by $\mathcal{L}_\mu h := \frac{dT'\mu_h}{d\mu}$.

Let $\{I_i\}$ be a partition in interval of $\mathbb{T}^1$ such that $T|_{I_i}$ is invertible, $T(I_i) = \mathbb{T}^1$ and $\cup_i I_i = \mathbb{T}^1$. Call S_i the inverse of the i-th branch of T. Then, setting $\rho_i := \frac{dT'\mu_{\mathbb{1}_{I_i}}}{d\mu}$

$$\begin{aligned} T'\mu_h(f) &= \sum_i \mu(h\mathbb{1}_{I_i} f \circ T) = \sum_i \mu(\mathbb{1}_{I_i}(h \circ S_i f) \circ T) \\ &= \mu\left(\left[\sum_i \rho_i h \circ S_i\right] f\right). \end{aligned}$$

Thus, setting $\rho = \sum_i \rho_i \circ T \mathbb{1}_{I_i}$ we have

$$\frac{dT'\mu_h}{d\mu} = \sum_i (\rho h) \circ S_i =: \mathcal{L}_\rho(h).$$

It follows that $\mathcal{L}_\rho(1) = 1$ and, for each $h \in L^\infty(\mu)$, $\mu(\mathcal{L}_\rho(h)) = T'\mu_h(1) = \mu(h)$.

Problem 7.1 *Compute ρ and $\mathcal{L}_\rho$, in the case in which μ is the unique invariant measure absolutely continuous with respect to Lebesgue.*

The relevant fact is that one has the following (partial) converse.

Lemma 7.2.1 *For $\rho \in \mathcal{C}^0$, $\rho \geq 0$, let $\mathcal{L}_\rho(h)(x) := \sum_{y\in T^{-1}x} \rho(y)h(y)$. If there exists $\lambda \in \mathbb{R}$, $h \in \mathcal{C}^0$, $h > 0$, such that $\mathcal{L}_\rho h = \lambda h$, then there exists a measure $\mu \in \mathcal{M}$ such that $\mu(\mathcal{L}_\rho f) = \lambda\mu(f)$ for all $f \in \mathcal{C}^0$ and there exists an invariant measure absolutely continuous with respect to μ.*

PROOF. By continuity there exists $\gamma > 0$ such that $h \geq \gamma > 0$. Thus

$$|\mathcal{L}_\rho^n f| \leq \gamma^{-1}|f|_\infty \mathcal{L}_\rho^n h = \lambda^n \gamma^{-1}|f|_\infty.$$

Hence, calling m the Lebesgue measure $\frac{1}{n}\sum_{k=0}^{n-1} \lambda^{-k}(\mathcal{L}'_\rho)^k m$ is a weakly compact sequence. Accordingly the same arguments used in Krylov-Bogoliubov Theorem 6.4.2 imply that there exists a measure μ such that $\lambda^{-1}\mathcal{L}'_\rho\mu = \mu$.

Next, define $\nu(f) := \mu(hf)$. Clearly ν is a measure absolutely continuous with respect to μ, in addition

$$\nu(f \circ T) = \lambda^{-1}(\mathcal{L}'_\rho\mu)(hf \circ T) = \lambda^{-1}\mu(f\mathcal{L}_\rho h) = \mu(fh) = \nu(f).$$

$\square$

7.3 Absolutely continuous invariant measure: revisited

We have already seen that there exists a unique invariant measure with respect to Lebesgue. Here we study this issue by a slightly different technique. Although the main idea is always to study the spectrum of the transfer operator, it is interesting to see how this can be achieved in many different ways, each way having its own advantages and disadvantages. Consider the transfer operator

$$\mathcal{L}h(x) := \sum_{y\in T^{-1}x} |D_yT|^{-1}h(y) \tag{7.3.1}$$

Problem 7.2 *Show that if $d\mu = hdm$, where m is the Lebesgue measure, then $\mu(f \circ T) = m(f\mathcal{L}h)$.*

Problem 7.3 *Show that, for each $n \in \mathbb{N}$,*

$$\mathcal{L}^n h(x) := \sum_{y\in T^{-n}x} |D_yT^n|^{-1}h(y)$$

Notice that, since DT cannot be zero, then its sign is constant. We limit ourselves, for simplicity, to the case $DT \geq \lambda$.

Problem 7.4 *Show that*

$$\begin{aligned}\frac{d}{dx}\mathcal{L}^n h(x) &= \sum_{y\in T^{-1}x}(D_yT)^{-2}h'(y) - D_y^2T(D_yT)^{-3}h(y)\\ &= \mathcal{L}((DT)^{-1}h') - \mathcal{L}(D^2T(DT)^{-2}h)\end{aligned}$$

7.3.1 A functional analytic setting

Let us consider first the Sobolev space $W^{1,1}$ and the space L^1.[1] Then, for each $h \in L^1(\mathbb{T}^1, m)$,

$$\int_{\mathbb{T}^1}|\mathcal{L}h|dm \le \int_{\mathbb{T}^1} 1\cdot\mathcal{L}|h|dm = \int_{\mathbb{T}^1} 1\circ T|h|dm = \int_{\mathbb{T}^1}|h|dm \tag{7.3.2}$$

that is $\mathcal{L}$ is a bounded operator on L^1 and its norm is bounded by one.

In addition, remembering Exercise 7.2,

$$\int_{\mathbb{T}^1}|\frac{d}{dx}\mathcal{L}h|dm \le \lambda^{-1}|h'|_{L^1} + D|h|_{L^1}, \tag{7.3.3}$$

where $D := \sup D^2T(DT)^{-2}$.

Problem 7.5 *Iterate the* (7.3.2), (7.3.3) *and prove, for all* $n\in\mathbb{N}$,

$$\begin{aligned}|\mathcal{L}^n h|_{L^1} &\le |h|_{L^1}\\ |\mathcal{L}^n h|_{W^{1,1}} &\le \lambda^{-n}|h|_{W^{1,1}} + B|h|_{L^1}\end{aligned}$$

where $B = 1 + (1-\lambda^{-1})^{-1}D$.

Since $W_{1,1}$ controls the L^∞ norm,[2] then we have that there exists $C > 0$ such that $|\mathcal{L}^n 1|_\infty < C$ for each $n \in \mathbb{N}$.

Using such a fact we can obtain similar inequalities in the Hilbert spaces L^2 and $W^{1,2}$. Indeed

$$\begin{aligned}\|\mathcal{L}^n h\|_{L^2}^2 = \int_{\mathbb{T}^1} h(\mathcal{L}^n h)\circ T^n \le \|h\|_{L^2}\left[\int_{\mathbb{T}^1}(\mathcal{L}^n h)^2\circ T^n\right]^{\frac{1}{2}} = \|h\|_{L^2}\\ \left[\int_{\mathbb{T}^1}(\mathcal{L}^n h)^2\mathcal{L}^n 1\right]^{\frac{1}{2}} \le C^{\frac{1}{2}}\|h\|_{L^2}\|\mathcal{L}^n h\|_{L^2}\end{aligned}$$

[1] For an open set $U \subset \mathbb{R}$, the spaces $W^{p,q}(U)$ are the completion of $\mathcal{C}^\infty(U,\mathbb{C})$ with respect to the norms $\left[|f|_{L^q}^q + |f'|_{L^q}^q + \cdots + |f^{(p)}|_{L^q}^q\right]^{\frac{1}{q}}$. Note that they are all Banach spaces by construction but the $W^{p,2}$ are also Hilbert spaces (**Exercise**: write the scalar product).

[2] If $f \in \mathcal{C}^\infty$, then the mean value theorem asserts $\int h = h(\xi)$ for some ξ. Then $h(x) = h(\xi) + \int_\xi^x h'(z)dz$. Thus $|h|_\infty \le |h|_{L^1} + |h'|_{L^1} = |h|_{W^{1,1}}$. The result extends then to all elements of $W^{1,1}$ by a standard approximation argument.

Which implies $\|\mathcal{L}^n h\|_{L^2} \le C^{\frac{1}{2}}\|h\|_{L^2}$ for each $n \in \mathbb{N}$. Hence,

$$\|\frac{d}{dx}\mathcal{L}^n h\|_{L^2} \le \lambda^{-n} C^{\frac{1}{2}}\|h'\|_{L^2} + D_n\|h\|_{L^2}.$$

Iterating as before we have, for all $n \in \mathbb{N}$,

$$\begin{aligned} &|\mathcal{L}^n h|_{L^2} \le C|h|_{L^2} \\ &|\mathcal{L}^n h|_{W^{1,2}} \le A\lambda^{-n}|h|_{W^{1,2}} + B|h|_{L^2}, \end{aligned} \tag{7.3.4}$$

for some appropriate constants A, B, C depending only on the map T.

To prove the existence of an invariant measure absolutely continuous with respect to Lebesgue we can try to mimic the Krylov-Bogolubov approach, but to do so we need a compactness result to substitute the weak compactness of the unit ball of the dual of a Banach space. This takes us in a very interesting detour in some fact of functional analysis.

7.3.2 Deeper in Functional analysis

Since we are on a circle it is a good idea to use Fourier series. For each function $h \in \mathcal{C}^\infty(\mathbb{T}, \mathbb{C})$ let h_k be its Fourier coefficients and define

$$(\mathbb{A}_n h)(x) = \sum_{|k|\le m} h_k e^{2\pi i k x} \tag{7.3.5}$$

Clearly, for all $m > 0$,

$$\begin{aligned} |h - \mathbb{A}_m|^2_{L^2} &= \sum_{|k|>m} |h_k|^2 = \sum_{|k|>m} |h_k|^2 |k|^{-2}|k|^2 \le m^{-2}\sum_{|k|>m} |(h')_k|^2 \\ &\le m^{-2}|h'|^2_{L^2} \le m^{-2}|h|^2_{W^{1,2}}. \end{aligned} \tag{7.3.6}$$

Using the above fact we can prove.

Lemma 7.3.1 *The unit ball of $W^{1,2}$ is (sequentially) compact in L^2.*

PROOF. Consider a sequence $\{h_m\} \subset W^{1,2}$, $|h_m|_{W^{1,2}} \le 1$. Since $\mathbb{A}_l$ are all finite rank operators, $\{\mathbb{A}_l h_m\}$ for l fixed are contained in a bounded finite dimensional (hence compact) set, thus there exists a converging subsequence for all l while (7.3.6) shows that the sequences for fixed m are all convergent. Using the usual diagonalization trick we can then extract a converging subsequence. □

Consider now $h_n := \frac{1}{n}\sum_{k=0}^{n-1}\mathcal{L}^k 1$. By the above lemma $\{h_n\}$ is relatively compact and thus we can extract a subsequence $\{h_{n_j}\}$ converging in L^2. Let h_* be the limit. Note that $\int h_n = 1$ for all $n \in \mathbb{N}$, thus $h_* \not\equiv 0$ and $\int h_* = 1$.

Problem 7.6 *Show that $\mathcal{L}h_* = h_*$, that is $d\mu := h_* dm$ is an invariant measure absolutely continuous with respect to Lebesgue and with L^2 density.*

Of course, at this point it is natural to ask if μ is the only measure with such a property or there exist others. To answer such a question we need some more facts.

7.3.3 Even deeper in Functional analysis

Since we have to do it, let us do in the following general setting.

Consider two Banach space $(\mathbb{B}, \|\cdot\|)$ and $(\mathbb{B}_0, |\cdot|)$ such that $\mathbb{B} \subset \mathbb{B}_0$ and

i. $|h| \leq \|h\|$ for all $h \in \mathbb{B}$,

ii. if $h \in \mathbb{B}$ and $|h| = 0$, then $h = 0$.

iii. There exists $C > 0 :$ for each $\varepsilon > 0$ there exists a finite rank operator $\mathbb{A}_\varepsilon \in L(\mathbb{B}, \mathbb{B})$ such that $\|\mathbb{A}_\varepsilon\| \leq C$ and $|h - \mathbb{A}_\varepsilon h| \leq \varepsilon \|h\|$ for all $h \in \mathbb{B}$.[3]

In addition consider a bounded operator $\mathcal{L} : \mathbb{B}_0 \to \mathbb{B}_0$, constants $A, B, C \in \mathbb{R}_+$, and $\lambda > 1$, such that

a. $|\mathcal{L}^n| \leq C$ for all $n \in \mathbb{N}$,

b. $\mathcal{L}(B) \subset B$

c. $\|\mathcal{L}^n h\| \leq A\lambda^{-n}\|h\| + B|h|$ for all $h \in \mathbb{B}$ and $n \in \mathbb{N}$.

In particular $\mathcal{L}$ can be seen as a bounded operator on $\mathbb{B}$.

Theorem 7.3.2 *The spectral radius of the operator $\mathcal{L} \in L(\mathbb{B}, \mathbb{B})$ is bounded by 1 while the essential spectral radius is bounded by λ^{-1}.*[4]

We can now prove our main result.

PROOF OF THEOREM 7.3.2. The first assertion is a trivial consequence of (c), (a) and (i).

---

[3] In fact, this last property can be weakened to: The unit ball $\{h \in \mathbb{B} : \|h\| \leq 1\}$ is relatively compact in $\mathbb{B}_0$. We use the present stronger condition since, on the one hand, it is true in all the applications we will be interested in and, on the other hand, drastically simplifies the argument. Note also that, if one uses the Fredholm alternative for compact operators rather than finite rank ones (Theorem D.0.1), then one can ask the $\mathbb{A}_\varepsilon$ to be compact instead than finite rank making easier their construction in concrete cases.

[4] The definition of *essential spectrum* varies a bit from book to book. Here we call essential spectrum the complement, in the spectrum, of the isolated eigenvalues with associated finite dimensional eigenspaces (which is also called the Fredholm spectrum).

The second part is much deeper. Let $\mathcal{L}_{n,\varepsilon} := \mathcal{L}^n \mathbb{A}_\varepsilon$, clearly such an operator is finite rank, in addition

$$\|\mathcal{L}^n h - \mathcal{L}_{n,\varepsilon} h\| \leq A\lambda^{-n}\|(\mathbb{1}-\mathbb{A}_\varepsilon)h\| + B|(\mathbb{1}-\mathbb{A}_\varepsilon)h| \leq A(1+C)\lambda^{-n}\|h\| + B\varepsilon\|h\|.$$

By choosing $\varepsilon = \lambda^{-n}$ we have that there exists $C_1 > 0$ such that

$$\|\mathcal{L}^n - \mathcal{L}_{n,\varepsilon}\| \leq C_1\lambda^{-n}.$$

For each $z \in \mathbb{C}$ we can now write

$$\mathbb{1} - z\mathcal{L} = (\mathbb{1} - z(\mathcal{L} - \mathcal{L}_{n,\varepsilon})) - z\mathcal{L}_{n,\varepsilon}.$$

Since

$$\|z(\mathcal{L} - \mathcal{L}_{n,\varepsilon})\| \leq |z|C_1\lambda^{-n} < \frac{1}{2},$$

provided that $|z| \leq \frac{1}{2C_1}\lambda^n$. Thus, given any z in the disk $D_n := \{|z| < \frac{1}{2C_1}\lambda^n\}$ the operator $B(z) := \mathbb{1} - z(\mathcal{L} - \mathcal{L}_{n,\varepsilon})$ is invertible.[5] Hence

$$\mathbb{1} - z\mathcal{L} = \left(\mathbb{1} - z\mathcal{L}_{n,\varepsilon}B(z)^{-1}\right)B(z) =: (\mathbb{1} - F(z))B(z).$$

By applying Fredholm analytic alternative (see Theorem D.0.1 for the statement and proof in a special case sufficient for the present purposes) to $F(z)$ we have that the operator is either never invertible or not invertible only in finitely many points in the disk D_n. Since for $|z| < 1$ we have $(\mathbb{1} - z\mathcal{L})^{-1} = \sum_{n=0}^{\infty} z^n\mathcal{L}^n$, the first alternative cannot hold hence the Theorem follows. □

7.3.4 The harvest

We are finally in the position to use all the above result to gain a deep understanding of the properties of the Dynamical Systems under consideration.

Problem 7.7 *Show that Theorem 7.3.2 implies that there exists $\sigma \in (0,1)$, $\{\theta_k\}_{k=1}^p$ and $L > 0$ such that*

$$\mathcal{L} = \sum_{k=1}^{p} e^{i\theta_k}\Pi_{\theta_k} + R$$

where Π_{θ_k} and R are operators on $W^{1,2}$ such that $\Pi_{\theta_k}\Pi_{\theta_j} = \delta_{jk}\Pi_{\theta_k}$ and $R\Pi_{\theta_k} = \Pi_{\theta_k}R = 0$. Moreover $|R^n| \leq L\sigma^n$.(Hint: Read section 6 of the Third Chapter of [Kat66] and recall that the operator is power bounded to exclude Jordan blocks.)

[5] Clearly $B(z)^{-1} = \sum_{n=0}^{\infty}[z(\mathcal{L} - \mathcal{L}_{n,\varepsilon})]^n$.

The above implies that

$$\Pi_\theta := \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} e^{-i\theta k}\mathcal{L}^k = \begin{cases} \Pi_{\theta_i} & \text{iff } \theta = \theta_j \\ 0 & \text{otherwise.} \end{cases} \tag{7.3.7}$$

Problem 7.8 *Using equations (7.3.4) show that, for each $h \in L^2$*

$$\|\Pi_\theta h\|_{W^{1,2}} \le C\|h\|_{L^2}.$$

(Hint: prove it first for $h \in W^{1,2}$ and then do a density argument).

Next, note that Exercise 7.6 implies that $h_* = \Pi_0 1 \neq 0$, that is one is in the spectrum on $\mathcal{L}$, this means that the spectral radius of $\mathcal{L}$ is one.

Accordingly, if $\Pi_\theta h = h$ we have $h \in W^{1,2} \subset \mathcal{C}^0$ and[6]

$$|h| = |\Pi_\theta h| \le \lim_{j\to\infty} \frac{1}{n_j}\sum_{k=0}^{n_j-1} \mathcal{L}^k|h| = \Pi_0|h| \le |h|_\infty h_*.$$

This means that all the eigenvectors of the peripheral spectrum are of the form $h = gh_*$ with $g \in \mathcal{C}^0$. Thus, if h_i is an $W^{1,2}$ orthonormal a base of the eigenspace associated to an eigenvalue θ, then the eigenprojector must have the form

$$\Pi_\theta h = \sum_i h_i \int \ell_i \cdot h,$$

with $\ell_i \in L^2$ and $\int \ell_i h_j = \delta_{ij}$. Hence $\Pi_\theta\mathcal{L} = e^{i\theta}\Pi_\theta$ implies

$$e^{i\theta}\sum_k h_k \int \ell_k \cdot h = \sum_k h_k \int \ell_k \cdot \mathcal{L}h = \sum_k h_k \int \ell_k \circ T \cdot h.$$

That is $e^{i\theta}\ell_k = \ell_k \circ T$. But then if we set $f_k := \bar{\ell}_k h_* \in L^2$, we have

$$\mathcal{L}f_k = e^{i\theta}\mathcal{L}(\bar{\ell}_k \circ T h_*) = e^{i\theta}\bar{\ell}_k\mathcal{L}h_* = e^{i\theta}\bar{\ell}_k h_* = e^{i\theta} f_k$$

By the above facts, this implies $\Pi_\theta f_k = f_k \in W^{1,2}$, that is $\ell_k \in \mathcal{C}^0$. But then for each $p \in \mathbb{N}$ we can set $h_p := \bar{\ell}_k^p h_*$ obtaining

$$\mathcal{L}h_p = e^{ip\theta}h_p.$$

Since the the peripheral spectrum consists of finitely many eigenvalues it follows that there must exist $p \in \mathbb{N}$ such that $p\theta = \theta \mod 2\pi$, that is the

[6] Remember that exercise 7.8 implies that the sequence in (7.3.7) converges in L^2, accordingly there exists a subsequence that converges almost everywhere with respect to Lebesgue.

spectrum on the unit circle must be the union of finitely many cyclic groups. In turn this implies that there exists $\bar p \in \mathbb{N}$ such that $\bar p\,\theta = 0 \mod 2\pi$, hence $\ell_k^{\bar p} = \ell_k^{\bar p} \circ T$. But this implies that if we define the sets $A_L := \{x \in \mathbb{T} \;:\; |\ell_k^{\bar p}| \leq L\}$, $L \in \mathbb{R}$, they are all invariant. So if χ_L is the characteristic function of the set A_L, then $\chi_L \circ T = \chi_L$ and $\mathcal{L}(\chi_L h_*) = \chi_L h_*$. We can thus produce a lot of eigenvalues of $\mathcal{L}$, but we know that such eigenvalues form a finite dimensional space. The only possibility is that only finitely many of the A_L are different. This is like saying that ℓ_k takes only finitely many values. But $\ell_k^{\bar p}$ is a continuous function, so it must be constant. Hence ℓ_k can assume only $\bar p$ different values, thus, again by continuity, must be constant. Finally this implies $\theta = 0$.

The conclusion is that one is the only eigenvalue on the unit circle and that the associated eigenprojector has rank one. So one is a simple eigenvalue and h_* is the only invariant density for the map.

7.3.5 conclusions

If we have any probability measure ν absolutely continuous with respect to Lebesgue and with density $h \in W^{1,2}$, then setting $d\mu = h_* dm$, for each $\varphi \in W^{1,2}$ we have

$$|\mu(\varphi \circ T^n) - \nu(\varphi \circ T^n)| = \left| \int \varphi \mathcal{L}^n (h - h_*) \right| \leq \|\varphi\|_{1,2} C \sigma^n \|h - h_*\|_{1,2}$$

where σ is the largest eigenvalue of modulus smaller than one (or λ^{-1} is no such eigenvalue exist).

Remark 7.3.3 *The above means that the evolution of the present* chaotic *system, if seen at the level of the absolutely continuous measures, becomes simply a dynamics with an uniformly attracting fixed point, the simplest dynamics of all!*

7.4 General transfer operators

In the previous sections we have been very successful in studying the measure absolutely continuous with respect to Lebesgue. We have seen in §7.2 (crf. Lemma 7.2.1) that to study other invariant measures one has to analyze more general transfer operators. Here we will restrict ourselves to studying

$$\mathcal{L}_g h := \mathcal{L}(e^g h)$$

where $\mathcal{L}$ is the usual transfer operator. This are called *transfer operators with weight* and g is sometime called the *potential*. We will consider first the case of $g : \mathbb{T}^1 \to \mathbb{C}$ and specialize to real potential later on.

For convenience, and also for didactical purposes, we will use the Banach spaces $\mathcal{C}^1$ and $\mathcal{C}^0$. Hence, form now on, we will assume $T \in \mathcal{C}^2(\mathbb{T}^1, \mathbb{T}^1)$ and $g \in \mathcal{C}^1(\mathbb{T}^1, \mathbb{C})$.

The first step is to compute the powers of $\mathcal{L}_g$ and study how they behave with respect to derivation.

Problem 7.9 *Show that, for each $n \in \mathbb{N}$, holds true*

$$\mathcal{L}_g^n h = \mathcal{L}^n \left[e^{g_n} h\right],$$

where $g_n = \sum_{k=0}^{n-1} g \circ T^k$.

Problem 7.10 *Show that for each $n \in \mathbb{N}$ and $h \in \mathcal{C}^1$ holds true*

$$\frac{d}{dx}\mathcal{L}_g^n h = \mathcal{L}_g^n \left[\frac{h'}{(T^n)'} - \frac{(T^n)''}{[(T^n)']^2}h + \frac{(g_n)'}{(T^n)'}h\right]$$

Note that $|\mathcal{L}_g^n h|_\infty \leq |h|_\infty \mathcal{L}_{\Re(g)}^n 1$. In addition,[7]

$$\begin{aligned}\left|\frac{(T^n)''(y)}{[(T^n)'(y)]^2}\right| &= \left|\frac{\frac{d}{dy}\prod_{k=0}^{n-1} T'(T^k y)}{[(T^n)'(y)]^2}\right| \\ &\leq \sum_{k=0}^{n-1}\left|\frac{T''(T^k y)}{(T^{n-k})'(T^k y)}\right| \leq \sum_{k=0}^{n-1} |T''|_\infty \lambda^{-n+k+1} \leq \frac{|T''|_\infty}{1-\lambda^{-1}}.\end{aligned}$$

Analogously,

$$\left|\frac{(g_n)'}{(T^n)'}\right| \leq \frac{|g'|_\infty}{1-\lambda^{-1}}.$$

The above inequalities imply

$$\left|\frac{d}{dx}\mathcal{L}_g^n h\right| \leq \lambda^{-n}\mathcal{L}_{\Re(g)}^n |h'| + B\mathcal{L}_{\Re(g)}^n |h|. \tag{7.4.8}$$

Which, taking the sup over x, yields

$$\left|\frac{d}{dx}\mathcal{L}_g^n h\right|_\infty \leq \lambda^{-n}|h'|_\infty \mathcal{L}_{\Re(g)}^n 1 + B_*|h|_\infty \mathcal{L}_{\Re(g)}^n 1,$$

Note that the above inequality implies that the spectral radius is bounded by $\rho = \lim_{n\to\infty} \|\mathcal{L}_{\Re(g)}^n 1\|_{\mathcal{C}^0}^{\frac{1}{n}}$ while the essential spectral radius is bounded by $\lambda^{-1}\rho$. The reader should notice that for positive potentials the above bounds are essentially sharp while for non positive, or complex, potential typically there will be cancellations that induce a smaller spectral radius. To control exactly such cancellations is, in general, a very hard problem.

[7] The quantity estimated here is usually called *distortion*. In fact, it measure how much the maps distorts intervals.

7.4.1 Real potential

In this section we will restrict to the case of $g \in \mathcal{C}^1(\mathbb{T}^1, \mathbb{R})$, i.e. real potentials.

If we define the cone $\mathcal{C}_a := \{h \in \mathcal{C}^1 \;:\; h > 0 \; |h'(x)| \leq ah(x)\}$, then equation (7.4.8), for $h > 0$, implies that, for each $\sigma \in (0, \lambda^{-1})$, $\mathcal{L}_g \mathcal{C}_a \subset \mathcal{C}_{\sigma_a}$ provided $a \geq B(\sigma - \lambda^{-1})^{-1}$.[8] We can then apply the theory of Appendix A to conclude the following.

Lemma 7.4.1 *For each real potential $g \in \mathcal{C}^1(\mathbb{T}^1, \mathbb{R})$, the transfer operator $\mathcal{L}_g$ has the Perron-Frobenius property, i.e. it has a simple strictly positive maximal eigenvalue and all the other eigenvalues are strictly smaller in modulus. In particular, the maximal eigenvalue of $\mathcal{L}_{\tau g}$, $\tau \in \mathbb{R}$, is analytic in τ.*[9]

7.4.2 Variational principle

....

7.5 Limit Theorems

Given $f \in \mathcal{C}^1$, $n \in \mathbb{N}$ and $a \in \mathbb{R}_+$ let

$$A_{a,n}(f) := \left\{ x \in \mathbb{T}^1 \;:\; \left| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) - \mu(f) \right| \geq a \right\}. \tag{7.5.9}$$

By the ergodic theorem $\lim_{n\to\infty} \mu(A_{a,n}(f)) = 0$. A natural question is:

Question 3 *How large is $m(A_{a,n})$?*

Note that we can write $\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) - \mu(f) = \frac{1}{n}\sum_{k=0}^{n-1} \hat f \circ T^k(x)$ where $\hat f := f - \mu(f)$. So we can reduce the question to the study of zero average function. A more refined question could be.

Question 4 *Does it exists a sequence $\{c_n\}$ such that*

$$\frac{1}{c_n} \sum_{k=0}^{n-1} \hat f \circ T^k(x)$$

converges in some sense to a non zero finite object?

---

[8] Note that this cone is almost the same than the one in Example 6.5.1, more precisely is its infinitesimal version.

[9] This follows from the fact that the maximal eigenvalue must always be simple and the results in Appendix C.4.

7.5.1 Large deviations. Upper bound

Note that it suffices to study the set

$$A_{a,n}^+(f) := \left\{ x \in \mathbb{T}^1 \; : \; \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) - \mu(f+a) \geq 0 \right\}.$$

since $A_{a,n}(f) = A_{a,n}^+(f) \cap A_{a,n}^+(-f)$. On the other hand, setting $\hat{f} := f - \mu(f)$, for each $\lambda \geq 0$ we have

$$\begin{aligned} m(A_{a,n}^+(f)) &= m(\{x \; : \; e^{\lambda \sum_{k=0}^{n-1} (\hat{f} \circ T^k(x) - a)} \geq 1\}) \leq e^{-n\lambda a} m(e^{\lambda \sum_{k=0}^{n-1} \hat{f} \circ T^k}) \\ &= e^{-n\lambda a} m(e^{\lambda \sum_{k=0}^{n-1} \hat{f} \circ T^k}). \end{aligned}$$

Accordingly,

$$m(A_{a,n}^+(f)) \leq e^{-n\lambda a} m(\mathcal{L}_\lambda^n 1) \tag{7.5.10}$$

where we have defined the operator $\mathcal{L}_\lambda g := \mathcal{L}(e^{\lambda \hat{f}} g)$, $\mathcal{L}$ being the Transfer operator of the map T.

By Lemma 7.4.1 $\mathcal{L}_\lambda$ has a maximal eigenvalue α_λ depending analytically on λ. Hence by the same argument used in Lemma 7.2.1 there exists $c \in \mathbb{R}$ such that

$$m(A_{a,n}^+(f)) \leq e^{-n(\lambda a - \ln \alpha_\lambda) + c}.$$

Since λ has been chosen arbitrarily we have obtained

$$m(A_{a,n}^+(f)) \leq e^{-n\tilde{I}(a) + c} \tag{7.5.11}$$

where $\tilde{I}(a) := \sup_{\lambda \in \mathbb{R}^+} \{\lambda a - \ln \alpha_\lambda\}$. The problem is then reduced to studying the function $I(a)$ which is commonly called *rate function*. Note that $\tilde{I}$ is not necessarily finite. Indeed if $a > \|\hat{f}\|_\infty$, then clearly $m(A_{a,n}^+(f)) = 0$.

To better understand the rate function it is helpful to make a little digression into convex analysis.

Recall that a function $f : \mathbb{R}^d \to \mathbb{R}^d$ is convex if for each $x, y \in \mathbb{R}^d$ and $t \in [0,1]$ we have $f(ty + (1-t)x) \leq tf(y) + (1-t)f(x)$ (if the inequality is everywhere strict, then the function is *stricly convex*.

Problem 7.11 *Show that if $f \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$, then f is convex iff $\frac{\partial^2 f}{\partial x^2}$ is a positive matrix.*[10] *Give a condition for strict convexity.*

Problem 7.12 *If a function $f : D \subset \mathbb{R}^d \to \mathbb{R}$, D convex,*[11] *is convex and bounded, then it is continuous.*

[10] A matrix $A \in GL(\mathbb{R}, d)$ is called *positive* if $A^T = A$ and $\langle v, Av \rangle \geq 0$ for each $v \in \mathbb{R}^d$.

[11] A set D is convex if, for all $x, y \in D$ and $t \in [0,1]$, olds true $ty + (1-t)x \in D$.

Given a function $f:\mathbb{R}^d\to\mathbb{R}$ let us define its *Legendre transform* as

$$f^*(x)=\sup_{y\in\mathbb{R}^d}\{\langle x,y\rangle - f(y)\} \tag{7.5.12}$$

Remark that f^* can take the value $+\infty$.

Problem 7.13 *Prove that f^* is convex.*

Problem 7.14 *Prove that $f^{**}\leq f$.*

Problem 7.15 *Prove that is $f\in\mathcal{C}^2(\mathbb{R}^d,\mathbb{R})$ is strictly convex, then the function $h(y):=\frac{\partial f}{\partial y}(y)$ is invertible and f^* is strictly convex. Moreover, calling g the inverse function of h, we have*

$$f^*(x)=\langle x,g(x)\rangle - f\circ g(x).$$

Problem 7.16 *Show that if $f\in\mathcal{C}^2$ is strictly convex, then $f^{**}=f$.*

Problem 7.17 *Show that, for each $x,y\in\mathbb{R}^d$, $\langle x,y\rangle\leq f^*(x)+f(y)$, (Young inequality).*

From the above discussion it follows that the rate function is defined very similarly to the Legendre transform of the logarithm of the maximal eigenvalue, which is commonly called *pressure of $\hat{f}$*. In fact, setting $I(a)=\max_{\lambda\in\mathbb{R}}(\lambda a-\ln\alpha_\lambda)$ we will se that, for $a\geq 0$, $I(a)=\tilde{I}(a)$. Unfortunately, to see that the rate function is exactly a Legendre transform takes some work. Let us start by studying the function α_λ.

Lemma 7.5.1 *There exists continuous functions $C_\lambda>0$ and $\rho_\lambda\in(0,1)$ such that, for $\lambda\leq 0$, $\mathcal{L}_\lambda=\alpha_\lambda\Pi_\lambda+Q_\lambda$, $\Pi_\lambda Q_\lambda=Q_\lambda\Pi_\lambda=0$, $\|Q_\lambda^n\|_{\mathcal{C}^1}\leq C_\lambda\rho_\lambda^n\alpha_\lambda^n$. Also $\Pi_\lambda(g)=h_\lambda\ell_\lambda(g)$, $\ell_\lambda(h_\lambda)=1$, $\ell_\lambda(h'_\lambda)=0$. In addition, $\mu_\lambda(\cdot):=\ell_\lambda(h_\lambda\,\cdot)$ is an invariant probability measure. Moreover everything is analytic in λ.*

Proof. As we have seen, there exists $h_\lambda\in\mathcal{C}^1$ and a measure ℓ_λ, both analytic in λ, such that the projection on the maximal eigenvalue of $\mathcal{L}_\lambda$ reads $\Pi_\lambda(h)=h_\lambda\ell_\lambda(h)$. Obviously

$$\mathcal{L}_\lambda h_\lambda=\alpha_\lambda h_\lambda, \tag{7.5.13}$$

and $\alpha_0=1$, $h_0=h$ and $\ell_0=m$. Notice that h_λ and ℓ_λ are not uniquely defined: by $\Pi_\lambda^2=\Pi_\lambda$ follows $\ell_\lambda(h_\lambda)=1$ but one normalization can be chosen freely.

Problem 7.18 *Show that the normalization of ℓ_λ,h_λ can be chosen so that $\ell_\lambda(h'_\lambda)=0$.*

□

Lemma 7.5.2 *The functions α_λ and $\ln\alpha_\lambda$ are convex. Moreover,*

$$\left|\frac{d}{d\lambda}\ln\alpha_\lambda\right| \leq |\hat{f}|_\infty.$$

PROOF. Note that

$$\frac{d^2}{d\lambda^2}\ln\alpha_\lambda = \frac{\alpha_\lambda''\alpha_\lambda - (\alpha_\lambda')^2}{\alpha_\lambda^2}, \tag{7.5.14}$$

thus the convexity of $\ln\alpha_\lambda$ implies the convexity of α_λ.

In view of the above fact we can differentiate (7.5.13) obtaining

$$\mathcal{L}_\lambda' h_\lambda + \mathcal{L}_\lambda h_\lambda' = \alpha_\lambda' h_\lambda + \alpha_\lambda h_\lambda'. \tag{7.5.15}$$

Applying ℓ_λ yields

$$\frac{d\alpha_\lambda}{d\lambda} = \alpha_\lambda \ell_\lambda(\hat{f}h_\lambda)) = \alpha_\lambda\mu_\lambda(\hat{f}). \tag{7.5.16}$$

Thus $\alpha_0' = 0$. Note that, as claimed,

$$\left|\frac{d}{d\lambda}\ln\alpha_\lambda\right| \leq |\mu_\lambda(\hat{f})| \leq |\hat{f}|_\infty.$$

Differentiating again yields

$$\frac{d^2\alpha_\lambda}{d\lambda^2} = \alpha_\lambda\mu_\lambda(\hat{f})^2 + \alpha_\lambda\ell_\lambda'(\hat{f}gh_\lambda) + \alpha_\lambda\ell_\lambda(\hat{f}h_\lambda'). \tag{7.5.17}$$

On the other hand, from (7.5.15) we have

$$(\mathbb{1}\alpha_\lambda - \mathcal{L}_\lambda)h_\lambda' = \mathcal{L}_\lambda(f_\lambda h_\lambda),$$

where $f_\lambda = \hat{f} - \mu_\lambda(\hat{f})$. Since, by construction, $\Pi_\lambda h_\lambda' = \Pi_\lambda(f_\lambda h_\lambda) = 0$, the above equation can be studied in the space $\mathbb{V}_\lambda = (\mathbb{1} - \Pi_\lambda)\mathcal{C}^1$ in which $\mathbb{1}\alpha_\lambda - \mathcal{L}_\lambda$ is invertible.

Setting $\hat{\mathcal{L}}_\lambda := \alpha_\lambda^{-1}\mathcal{L}_\lambda$, we have

$$h_\lambda' = (\mathbb{1} - \hat{\mathcal{L}}_\lambda)^{-1}\hat{\mathcal{L}}_\lambda(f_\lambda h_\lambda). \tag{7.5.18}$$

Doing similar considerations on the equation $\ell_\lambda(\mathcal{L}_\lambda) = \alpha_\lambda\ell_\lambda(g)$, we obtain

$$\begin{aligned}
\alpha_\lambda'' &= \alpha_\lambda\mu_\lambda(\hat{f})^2 + \alpha_\lambda\ell_\lambda(f_\lambda(\mathbb{1} - \hat{\mathcal{L}}_\lambda)^{-1}(\mathbb{1} + \hat{\mathcal{L}}_\lambda)(f_\lambda h_\lambda)) \\
&= \alpha_\lambda\mu_\lambda(\hat{f})^2 + \alpha_\lambda\sum_{n=1}^{\infty}\ell_\lambda(f_\lambda\hat{\mathcal{L}}_\lambda^n(\mathbb{1} + \hat{\mathcal{L}}_\lambda)(f_\lambda h_\lambda)) \\
&= \frac{(\alpha_\lambda')^2}{\alpha_\lambda} + \left[\mu_\lambda(f_\lambda^2) + 2\sum_{n=1}^{\infty}\ell_\lambda(f_\lambda\hat{\mathcal{L}}_\lambda^n(f_\lambda h_\lambda))\right]\alpha_\lambda.
\end{aligned} \tag{7.5.19}$$

Finally, notice that

$$\ell_\lambda(f_\lambda \hat{\mathcal{L}}_\lambda^n(f_\lambda h_\lambda)) = \ell_\lambda(\hat{\mathcal{L}}_\lambda^n(f_\lambda \circ T^n f_\lambda h_\lambda)) = \mu_\lambda(f_\lambda \circ T^n f_\lambda)$$

and

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\mu_\lambda\left(\left[\sum_{k=0}^{n-1} f_\lambda \circ T^k\right]^2\right) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k,j=0}^{n-1}\mu_\lambda(f_\lambda\circ T^k f_\lambda \circ T^j)\\ &= \mu_\lambda(f_\lambda^2) + \lim_{n\to\infty}\frac{2}{n}\sum_{k=1}^{n-1}(n-k)\mu_\lambda(f_\lambda\circ T^k f_\lambda)\\ &= \mu_\lambda(f_\lambda^2) + 2\sum_{k=1}^{\infty}\mu_\lambda(f_\lambda \circ T^k f_\lambda).\end{aligned} \tag{7.5.20}$$

The above two facts and equations (7.5.14), (7.5.19) yield

$$\frac{d^2}{d\lambda^2}\ln\alpha_\lambda = \lim_{n\to\infty}\frac{1}{n}\mu_\lambda\left(\left[\sum_{k=0}^{n-1} f_\lambda\circ T^k\right]^2\right) \geq 0. \tag{7.5.21}$$

□

Note that equation (7.5.16) implies $\alpha_0' = 0$, hence $\alpha_\lambda' \geq 0$ for $\lambda \geq 0$. Since the maximum of $\lambda a - \ln\alpha_\lambda$ is taken either at $\alpha_\lambda a = \alpha_\lambda'$ or at infinity (if $a > \sup_{\lambda>0}\frac{\alpha_\lambda'}{\alpha_\lambda}$), it follows that

$$\tilde{I}(a) = \sup_{\lambda\geq 0}(\lambda a - \ln\alpha_\lambda) = \sup_{\lambda}(\lambda a - \ln\alpha_\lambda) = I(a)$$

as announced. In fact, more can be said.

Lemma 7.5.3 *Either the rate function I is strictly convex, or there exists $\beta \in \mathbb{R}, \phi \in \mathcal{C}^0$ such that $f - \beta = \phi - \phi\circ T$.*

PROOF. By Problem 7.15 it suffices to prove that $\ln\alpha_\lambda$ is strictly convex. On the other hand equations (7.5.14) and (7.5.21) imply that if the second derivative of $\ln\alpha_\lambda$ is zero for some λ, then

$$\begin{aligned}\mu_\lambda\left(\left[\sum_{k=0}^{n-1} f_\lambda\circ T^k\right]^2\right) &= n\left[\mu_\lambda(\hat{f}^2) + 2\sum_{k=1}^{n-1}\frac{n-k}{n}\mu_\lambda(f_\lambda\circ T^k f_\lambda)\right]\\ &= -2n\sum_{k=n}^{\infty}\ell_\lambda(f_\lambda\hat{\mathcal{L}}_\lambda^k(f_\lambda h_\lambda)) - 2\sum_{k=1}^{n-1}k\ell_\lambda(f_\lambda\hat{\mathcal{L}}_\lambda^k(f_\lambda h_\lambda)) - \alpha_\lambda\mu_\lambda(\hat{f})^2\\ &\leq C(\lambda)\left[n\rho_\lambda^n + \sum_{k=0}^{\infty}k\rho_\lambda^k\right]\end{aligned}$$

Accordingly, the sequence $\sum_{k=0}^{n-1} f_\lambda \circ T^k$ is bounded in $L^2(\mathbb{T}^1, \mu_\lambda)$ and hence weakly compact. Let $\sum_{k=0}^{n_j-1} f_\lambda \circ T^k$ a weakly convergent subsequence,[12] that is there exists $\phi_\lambda \in L^2$ such that for each $\varphi \in L^2$ holds

$$\lim_{j\to\infty} \mu_\lambda(\varphi \sum_{k=0}^{n_j-1} f_\lambda \circ T^k) = \mu_\lambda(\varphi\phi_\lambda).$$

It follows that, for each $\varphi \in \mathcal{C}^1$,

$$\begin{aligned}\mu_\lambda(\varphi[f_\lambda - \phi_\lambda + \phi_\lambda \circ T]) &= \mu_\lambda(\varphi f_\lambda) + \lim_{j\to\infty} \sum_{k=0}^{n_j-1} \mu_\lambda(\varphi f_\lambda \circ T^{k+1} - \varphi f_\lambda \circ T^k)\\ &= \lim_{j\to\infty} \mu_\lambda(\varphi f_\lambda \circ T^{n_j}) = \lim_{j\to\infty} \ell_\lambda(f_\lambda \hat{\mathcal{L}}_\lambda^{n_j}(\varphi h_\lambda))\\ &= \mu_\lambda(\varphi)\mu_\lambda(f_\lambda) = 0.\end{aligned}$$

thus, since $\mathcal{C}^1$ is dense in L^2, it follows

$$f_\lambda = \phi_\lambda - \phi_\lambda \circ T\,, \qquad \mu_\lambda - \text{a.s.} \tag{7.5.22}$$

A function with the above property is called a *coboundary*, in this case an L^2 coboundary since we know only that $\phi_\lambda \in L^2(\mathbb{T}, \mu_\lambda)$. In fact, this it is not not enough to conclude the Lemma: we need to show, at least, that $\phi_\lambda \in \mathcal{C}^0$.

First of all notice that, since for each $\beta \in \mathbb{R}$ we have $f_\lambda = \phi_\lambda + \beta - (\phi_\lambda + \beta) \circ T$, we can assume without loss of generality $\mu_\lambda(\phi_\lambda) = 0$. But then

$$\hat{\mathcal{L}}_\lambda(f_\lambda\, h_\lambda) = \hat{\mathcal{L}}_\lambda(\phi_\lambda\, h_\lambda) - \phi_\lambda\, h_\lambda = -(\mathbb{1} - \hat{\mathcal{L}}_\lambda)\phi_\lambda\, h_\lambda.$$

Hence

$$\phi_\lambda = h_\lambda^{-1}(\mathbb{1} - \hat{\mathcal{L}}_\lambda)^{-1}\hat{\mathcal{L}}_\lambda(f_\lambda\, h_\lambda) \in \mathcal{C}^1.$$

□

Remark 7.5.4 *The above result is quite sharp. Indeed, it shows that if I is not strictly convex, then for each invariant measure ν holds $\nu(f) = \beta = \mu(f)$. So it suffices to find two invariant measures for which the average of f differs (for example the average on two periodic orbits) to infer that I is strictly convex.*

Problem 7.19 *Set $\sigma := \alpha''(0)$. Show that, for a small, $I(a) = \frac{a^2}{2\sigma} + \mathcal{O}(a^3)$. Show that if $a > |f|_\infty$, then $I(a) = +\infty$.*

[12] Such a subsequence always exists [LL01].

The above discussion allows to conclude

$$m(A^+_{a,n}(f)) \leq m(\mathcal{L}^n_{\lambda_-} h) \leq C e^{-\frac{a^2}{2\sigma^2} n + \mathcal{O}(a^3 n)}.$$

Since similar arguments hold for the set $A^+_{a,n}(-f)$, it follows that we have an exponentially small probability to observe a deviation from the average. Moreover, the expected size of a deviation is of order $n^{-\frac{1}{2}}$, to see if this is really the case we a lower bound.

7.5.2 Large deviations. Lower bound

Let $I = (\alpha, \beta)$, fix $c \in (0, \frac{\beta-\alpha}{2})$ and let us consider a $\lambda \in \mathbb{R}$ such that $\mu_\lambda(\hat{f}) \in (\alpha + c, \beta - c) = I_c$. Let $S_n = \sum_{k=0}^{n-1} \hat{f} \circ T^k$, then $\mu_\lambda(S_n) = n\mu_\lambda(\hat{f})$ and, by (7.5.20)

$$\mu_\lambda\left(\left[\sum_{k=0}^{n-1} \hat{f} \circ T^k - n\mu_\lambda(\hat{f})\right]^2\right) \leq C_\lambda n,$$

where C_λ depends continuously by λ. Thus, setting $A_{n,I} = \{x \in \mathbb{T}^1 \;:\; \frac{1}{n} S_n(x) \in I\}$,

$$\begin{aligned}\mu_\lambda(A^c_{n,I}) &\leq \mu_\lambda\left(\left\{\left|\sum_{k=0}^{n-1} f_\lambda \circ T^k\right| \geq cn\right\}\right)\\ &\leq c^{-2} n^{-2} \mu_\lambda\left(\left|\sum_{k=0}^{n-1} f_\lambda \circ T^k\right|^2\right) \leq C_\lambda c^{-2} n^{-1}.\end{aligned}$$

It follows that there exists $n_\lambda \in \mathbb{N}$ such that, for all $n \geq n_\lambda$, $\mu_\lambda(A_{n,I}) \geq \frac{1}{2}$. We can then write

$$\frac{1}{2} \leq \ell_\lambda(A_{n,I} h_\lambda) \leq C_\# e^{-(n+m)\ln \alpha_\lambda} \ell_\lambda\left(\mathcal{L}^{n+m}_\lambda(\mathbb{1}_{A_{n,I}})\right). \tag{7.5.23}$$

To conclude we must analyse a bit the characteristic function of $A_{n,I}$. First of all, notice that if $|T^k x - T^k y| \leq \varepsilon$ for each $k \leq n$, then $|T^k x - T^k y| \leq \lambda^{-n+k}\varepsilon$ for all $k \leq n$. Accordingly, for each $z \in [x, y]$

$$\begin{aligned}|D_x T^n - D_z T^n| &\leq |D_x T^n| \cdot (e^{\sum_{k=0}^{n-1} |\ln D_{T^k x} T - \ln D_{T^k z} T|} - 1)\\ &\leq |D_x T^n|(e^{C_\# \sum_{k=0}^{n-1} \lambda^{-k}\varepsilon} - 1) \leq C_\# |D_x T^n|.\end{aligned}$$

By a similar estimate follows $|D_x T^n - D_z T^n| \geq C_\# |D_x T^n|$ as well. Moreover,

$$|S_n(x) - S_n(y)| \leq \sum_{k=0}^{n-1} |f|_{\mathcal{C}^1} C_\# \lambda^{-k} \varepsilon \leq C_\# \varepsilon.$$

We can then write $A_{n,I} \supset \cup_l J_l \supset A_{n,I_c}$ where J_l are disjoint intervals such that $|T^n J_l| \leq \varepsilon$. Choosing ε small enough it follow that the oscillation of S_n on each J_l is smaller than c. Moreover

$$\begin{aligned}\|\mathcal{L}^n \mathbb{1}_{J_l}\|_{BV} &= \sup_{|\varphi|_\infty \leq 1} \int_{J_l} \varphi' \circ T^n \leq \sup_{|\varphi|_\infty \leq 1} \int_{J_l} \frac{d}{dx}\left[(DT^n)^{-1}\varphi \circ T^n\right] + B|J_l| \\ &\leq 2 \sup_{x \in J_l} |D_x T^n|^{-1} + B|J_l| \leq C_\# |J_l|.\end{aligned}$$

We can then continue our estimate started in (7.5.23),

$$\begin{aligned}\frac{1}{2} &\leq C_\# e^{-(n+m)\ln \alpha_\lambda + n\lambda\beta + mC_\#} \sum_l \ell_\lambda \left(\mathcal{L}^{n+m}(\mathbb{1}_{J_l})\right) \\ &= C_\# e^{-(n+m)\ln \alpha_\lambda + n\lambda\beta + mC_\#} \sum_l \ell_\lambda \left(m(J_l)(1 + \mathcal{O}(\rho^m))\right)) \\ &\leq C_\# e^{-n(\ln \alpha_\lambda - \lambda\beta)} m(A_{n,I}),\end{aligned}$$

where we have chosen m large but fixed. The above computations imply that, for each $L > 0$,

$$m(A_{n,I}) \geq C_L e^{-J_L(I)n}$$

where $J_L(I) = \max_{\{\lambda \leq L \,:\, \mu_\lambda(f) \in I_c\}} \lambda a - \ln \alpha_\lambda$. Note that, if f is not a coboundary and hence $\ln \alpha_\lambda$ is strictly convex, the maximum of $\lambda\beta - \ln \alpha_\lambda$ is attained at some finite value, hence, for L large enough, $J_L(I) = \sup_{\{\lambda \in \mathbb{R} \,:\, \mu_\lambda(f) \in I_c\}} \lambda\beta - \ln \alpha_\lambda$. This implies that

This must be fixed a bit

$$m(A^+_{a,n}) \geq C_\# e^{-J(a)n}$$

where $J(a) = \sup_{\{\lambda \,:\, \mu_\lambda(f) > a\}} \lambda a - \ln \alpha_\lambda$.

The surprising fact is that the upper and lower bound are essentially the same. To see this a little argument is needed.

7.5.3 Large deviations. Conclusions

In fact, it is possible to give a variational characterization of the rate function in the spirit of general Large deviation theory [Var84, Str84, DZ98].

Lemma 7.5.5 *Let $\mathcal{M}_T$ be the set of invariant probability measures invariant with respect to T. Then*

$$I(a) = -\sup_{\{\nu \in \mathcal{M}_T \,:\, \nu(f) \geq a\}} h_\nu(T) = J(a).$$

PROOF. By section 7.4.2 we have that, for each $\nu \in \mathcal{M}_T$, $\ln \alpha_\lambda = \sup_{\nu \in \mathcal{M}_T} \{h_\nu(T) + \lambda\nu(f)\} = h_{\mu_\lambda}(T) + \lambda\mu_\lambda(f)$. Thus for each $\nu \in \mathcal{M}_T$ such that $\nu(f) \geq a$, we can write

$$I(a) \leq \max_{\lambda \geq 0} \{\lambda(a - \nu(f)) - h_\nu(T)\} = -h_\nu(T).$$

On the other and

$$I(a) = \sup_{\lambda \geq 0}\{\lambda(a - \mu_\lambda(f)) - h_{\mu_\lambda}(T)\}.$$

If $a > \sup \mu_\lambda(f)$, then $I(a) = +\infty$, otherwise let λ_* be such that $\mu_{\lambda_*}(f) = a$,[13] then

$$I(a) \geq -h_{\mu_{\lambda_*}}(T) \geq - \sup_{\{\nu \in \mathcal{M}_T \,:\, \nu(f) \geq a\}} h_\nu(T).$$

Finally, since μ_λ and h_{μ_λ} depend smoothly from λ,

$$J(a) = \sup_{\{\lambda \,:\, \mu_\lambda(f) > a\}} \lambda a - \lambda \mu_\lambda(f) - h_{\mu_\lambda}(T) = I(a).$$

□

7.5.4 The Central Limit Theorem

We can now address the second question we have posed. From the above discussion is clear that we must chose $c_n = \sqrt{n}$.

Let $f \in BV$ and set $\hat{f} := f - \mu(f)$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \hat{f} \circ T^k(x) = 0 \quad m - \text{a.e.}$$

Let us set $\Psi_n := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \hat{f} \circ T^k$. We can consider Ψ_n a random variable with distribution $F_n(t) := \mu(\{x \;:\; \Psi_n(x) \leq t\})$. It is well known that, for each continuous function g holds[14]

$$\mu(g(\Psi_n)) = \int_{\mathbb{R}} g(t) dF_n(t)$$

where the integral is a Riemann-Stieltjes integral. It is thus clear that if we can control the distribution F_n, we have a very sharp understanding of the probability to have small deviations (of order $\sqrt{n}$) from the limit. From the

[13] Actually one must show that the sup is a max.

[14] If $g \in \mathcal{C}^1_0$, then

$$\int_{\mathbb{R}} g dF_n = -\int_{\mathbb{R}} F_n(t) g'(t) dt = -\int_{\mathbb{R}} dt \int_{\mathbb{T}^1} dx \chi_{\{z \,:\, \Psi_n(z) \leq t\}}(x) g'(t).$$

Applying Fubini yields

$$\int_{\mathbb{R}} g dF_n = -\int_{\mathbb{T}^1} dx \int_{\mathbb{R}} dt \chi_{\{z \,:\, \Psi_n(z) \leq t\}}(x) g'(t) = -\int_{\mathbb{T}^1} dx \int_{\Psi_n(x)}^{\infty} g'(t) dt = \int_{\mathbb{T}^1} dx g(\Psi_n(x)).$$

work in the previous section it follows that there exists $\delta > 0$ such that, for each $|\lambda| \leq \delta\sqrt{n}$,

$$\begin{aligned}\varphi_n(\lambda) := \mu(e^{i\lambda\Psi_n}) = \mu(\mathcal{L}^n_{i\lambda/\sqrt{n}}h) = (1 - \frac{\sigma^2\lambda^2}{2n} + \mathcal{O}(\lambda^3 n^{-\frac{3}{2}} + \rho^n)\|f\|_{BV})^n \\ = e^{-\frac{\sigma^2\lambda^2}{2}}(1 + \mathcal{O}(\lambda^3 n^{-\frac{1}{2}} + n\rho^n)\|f\|_{BV}).\end{aligned} \tag{7.5.24}$$

The above quantity is called *characteristic function* of the random variable and determines the distribution (at continuity points) via the formula

$$F_n(b) - F_n(a) = \lim_{\Lambda\to\infty} \frac{1}{2\pi}\int_{-\Lambda}^{\Lambda} \frac{e^{-ia\lambda} - e^{-ib\lambda}}{i\lambda}\varphi_n(\lambda)d\lambda,$$

as can be seen in any basic book of probability theory.[15]

Formula (7.5.24) means in particular that

$$\lim_{n\to\infty} m(e^{\lambda\Psi_n}) = e^{-\frac{\sigma^2\lambda^2}{2}} =: \varphi(\lambda).$$

What can we infer from the above facts? First of all a simple computation shows that

$$g(t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-it\lambda}\varphi(\lambda)d\lambda = \frac{1}{\sqrt{\pi}\sigma}e^{-\frac{t^2}{2\sigma^2}}$$

a random variable with such a density is called a Gaussian random variable with zero average and variance σ. Accordingly, formula (7.5.24) can be interpreted by saying that there exists a Gaussian random variable G such that

$$\frac{1}{n}\sum_{k=0}^{n-1}\hat{f}\circ T^k \sim \frac{1}{\sqrt{n}}G(1 + \mathcal{O}(n^{-\frac{1}{2}}))$$

in distribution. But what does this means concretely. Actual estimates are made difficult by the fact that the distribution under study not necessarily have a density, thus we are Fourier transforming function that behave quite badly at infinity. To overcome such a problem we can smoothen the quantities involved.

[15] In the case when there exists a density, that is an L^1 function f_n such that $F_n(b) - F_n(a) = \int_a^b f_n(t)dt$, then the formula above becomes simply

$$f_n(t) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-it\lambda}\varphi_n(\lambda)d\lambda,$$

and follows trivially by the inversion of the Fourier transform.

Let $j \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}_+)$ such that $\int_{\mathbb{R}} j(t)dt = 1$, $j(t) = j(-t)$, and $j(t) = 0$ for all $|t| > 1$, for each $\varepsilon > 0$ defined then $j_\varepsilon(t) := \varepsilon^{-1} j(\varepsilon^{-1} t)$ and

$$F_{n,\varepsilon}(t) := \int_{\mathbb{R}} j_\varepsilon(t-s) F_n(s) ds. \tag{7.5.25}$$

A simple computation shows that, for each $a, b \in \mathbb{R}$, holds

$$F_n(b+\varepsilon) - F_n(a-\varepsilon) \geq F_{n,\varepsilon}(b) - F_{n,\varepsilon}(a) \geq F_n(b-\varepsilon) - F_n(a+\varepsilon)$$

that is: if the measurements have a precision worst than 2ε, then $F_{n,\varepsilon}$ is as good as F_n to describe the resulting statistics. On the other hand calling $\varphi_{n,\varepsilon}$ the characteristic function associated to $F_{n,\varepsilon}$, holds $\varphi_{n,\varepsilon}(\lambda) = \varphi_n(\lambda)\hat{j}(\varepsilon\lambda)$, where $\hat{j}$ is the Fourier transform of j. Since now $F_{n,\varepsilon}$ is the law of a smooth random variable it has a density $f_{n,\varepsilon}$ and

$$f_{n,\varepsilon}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\lambda t} \varphi_n(\lambda) \hat{j}(\varepsilon\lambda) d\lambda$$

since j is smooth it follows that there exists $C > 0$ such that $|\hat{j}(\lambda)| \leq C(1 + \lambda^2)^{-2}$. We can finally use formula (7.5.24) to obtain a quantitative estimate

$$\begin{aligned}
f_{n,\varepsilon}(t) &= \frac{1}{2\pi} \int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}} e^{-i\lambda t} \varphi_n(\lambda) \hat{j}(\varepsilon\lambda) d\lambda + \mathcal{O}(\varepsilon^{-5} n^{-\frac{3}{2}}) \\
&= \frac{1}{2\pi} \int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}} e^{-i\lambda t} \varphi(\lambda) \hat{j}(\varepsilon\lambda) d\lambda + \mathcal{O}(\varepsilon^{-5} n^{-\frac{3}{2}} + n^{-\frac{1}{2}}) \\
&= g(t) + \mathcal{O}(\varepsilon + \varepsilon^{-5} n^{-\frac{3}{2}} + n^{-\frac{1}{2}}) = g(t) + \mathcal{O}(n^{-\frac{1}{2}})
\end{aligned}$$

provided we choose $n^{-\frac{1}{2}} \geq \varepsilon \geq n^{-5}$. Which, as announced, means that, if the precision of the instrument is compatible with the statistics, the typical fluctuations in measurements are of order $\frac{1}{\sqrt{n}}$ and Gaussian. This is well known by sperimentalists who routinely assume that the result of a measurement is distributed according to a Gaussian.[16]

7.6 Perturbation theory

To answer the questions posed at the beginning we need some perturbation theorems. Few such results are available (e.g., see [Kif88], [BY93] or [Bal00]

[16] Note however that our proof holds in a very special case that has little to do with a real experimental setting. To prove the analogous statement for a realistic experiment is a completely different ball game.

for a review), here we will follow mainly the theory developed in [KL99, GL06] adapted to the special cases at hand.

For simplicity let us work directly with the densities and in the case $d = 1$. Then $\mathcal{L}$ is the transfer operator for the densities. We will start by considering an abstract family of operators $\mathcal{L}_\varepsilon$ satisfying the following properties.

Condition 1 *Consider a family of operators $\mathcal{L}_\varepsilon$ with the following properties*

1. *A uniform Lasota-Yorke inequality:*
$$\|\mathcal{L}_\varepsilon^n h\|_{BV} \le A\lambda^{-n}\|h\|_{BV} + B|h|_{L^1}, \quad |\mathcal{L}_\varepsilon^n h|_{L^1} \le C|h|_{L^1} \ ;$$
2. $\int \mathcal{L}h(x)dx = \int h(x)dx$ *;*
3. *For $L : BV \to BV$ define the norm*
$$|||L||| := \sup_{\|h\|_{BV}\le 1} |Lf|_{L^1},$$
that is the norm of L as an operator from $BV \to L^1$. Then we require that there exists $D > 0$ such that
$$|||\mathcal{L} - \mathcal{L}_\varepsilon||| \le D\varepsilon.$$

Condition 1-(3) specifies in which sense the family $\mathcal{L}_\varepsilon$ can be considered an approximation of the unperturbed operator $\mathcal{L}$. Notice that the condition is rather weak, in particular the distance between $\mathcal{L}_\varepsilon$ and $\mathcal{L}$ as operators on BV can be always larger than 1. Such a notion of closeness is completely inadequate to apply standard perturbation theory, to get some perturbations results it is then necessary to drastically restrict the type of perturbations allowed, this is done by Conditions 1-(1,2) which state that all the approximating operators enjoys properties very similar to the limiting one.[17]

To state a precise result consider, for each operator L, the set
$$V_{\delta,r}(L) := \{z \in \mathbb{C} \mid |z| \le r \text{ or } \operatorname{dist}(z, \sigma(L)) \le \delta\}.$$
Since the complement of $V_{\delta,r}(L)$ belongs to the resolvent of L it follows that
$$H_{\delta,r}(L) := \sup\left\{\|(z-L)^{-1}\|_{BV} \mid z \in \mathbb{C}\setminus V_{\delta,r}(L)\right\} < \infty.$$
By $R(z)$ and $R_\varepsilon(z)$ we will mean respectively $(z - \mathcal{L})^{-1}$ and $(z - \mathcal{L}_\varepsilon)^{-1}$.

[17] Actually only Condition 1-(1) is needed in the following. Condition 1-(2) simply implies that the eigenvalue one is common to all the operators. If 1-(2) is not assumed, then the operator $\mathcal{L}_\varepsilon$ will always have one eigenvalue close to one, but the spectral radius could vary slightly, see [LMD03] for such a situation.

Theorem 7.6.1 ([KL99]) *Consider a family of operators $\mathcal{L}_\varepsilon : BV \to BV$ satisfying Conditions 1. Let $H_{\delta,r} := H_{\delta,r}(\mathcal{L})$; $V_{\delta,r} := V_{\delta,r}(\mathcal{L})$, $r > \lambda^{-1}$, $\delta > 0$, then, if $\varepsilon \le \varepsilon_1(\mathcal{L}, r, \delta)$, $\sigma(\mathcal{L}_\varepsilon) \subset V_{\delta,r}(\mathcal{L})$. In addition, if $\varepsilon \le \varepsilon_0(\mathcal{L}, r, \delta)$, there exists $a > 0$ such that, for each $z \notin V_{\delta,r}$, holds true*

$$|||R(z) - R_\varepsilon(z)||| \le C\varepsilon^a.$$

PROOF.[18] To start with we collect some trivial, but very useful algebraic identities.

For each operator $L : BV \to BV$ and $n \in \mathbb{Z}$ holds

$$\frac{1}{z}\sum_{i=0}^{n-1}(z^{-1}L)^i(z-L) + (z^{-1}L)^n = \mathbb{1} \tag{7.6.26}$$

$$R(z)(z-\mathcal{L}_\varepsilon) + \frac{1}{z}\sum_{i=0}^{n-1}(z^{-1}\mathcal{L})^i(\mathcal{L}_\varepsilon - \mathcal{L}) + R(z)(z^{-1}\mathcal{L})^n(\mathcal{L}_\varepsilon - \mathcal{L}) = \mathbb{1} \tag{7.6.27}$$

$$(z-\mathcal{L}_\varepsilon)\left[G_{n,\varepsilon} + (z^{-1}\mathcal{L}_\varepsilon)^n R(z)\right] = \mathbb{1} - (z^{-1}\mathcal{L}_\varepsilon)^n(\mathcal{L}_\varepsilon - \mathcal{L})R(z) \tag{7.6.28}$$

$$\left[G_{n,\varepsilon} + (z^{-1}\mathcal{L}_\varepsilon)^n R(z)\right](z-\mathcal{L}_\varepsilon) = \mathbb{1} - (z^{-1}\mathcal{L}_\varepsilon)^n R(z)(\mathcal{L}_\varepsilon - \mathcal{L}), \tag{7.6.29}$$

where we have set $G_{n,\varepsilon} := \frac{1}{z}\sum_{i=0}^{n-1}(z^{-1}\mathcal{L}_\varepsilon)^i$.

Let us start applying the above formulae. For each $h \in BV$ and $z \notin V_{r,\delta}$ holds

$$\begin{aligned}\|(z^{-1}\mathcal{L}_\varepsilon)^n(\mathcal{L}_\varepsilon - \mathcal{L})R(z)h\|_{BV} &\le (r\lambda)^{-n}A\|(\mathcal{L}_\varepsilon - \mathcal{L})R(z)h\|_{BV} + \frac{B}{r^n}|(\mathcal{L}_\varepsilon - \mathcal{L})R(z)h|_{L^1}\\ &\le [(r\lambda)^{-n}A2C_1 + Br^{-n}D\varepsilon]H_{r,\delta}\|h\|_{BV} < \|h\|_{BV}\end{aligned}$$

Thus $\|(z^{-1}\mathcal{L}_\varepsilon)^n(\mathcal{L}_\varepsilon - \mathcal{L})R(z)\|_{BV} < 1$ and the operator on the right hand side of (7.6.28) can be inverted by the usual Neumann series. Accordingly, $(z-\mathcal{L}_\varepsilon)$ has a well defined right inverse. Analogously,

$$\|(z^{-1}\mathcal{L}_\varepsilon)^n R(z)(\mathcal{L}_\varepsilon - \mathcal{L})h\|_{BV} \le (r\lambda)^{-n}A\|R(z)(\mathcal{L}_\varepsilon - \mathcal{L})h\|_{BV} + Br^{-n}|R(z)(\mathcal{L}_\varepsilon - \mathcal{L})h|_{L^1}.$$

This time to continue we need some informations on the L^1 norm of the resolvent. Let $g \in BV$, then equation (7.6.26) yields

$$\begin{aligned}|R(z)g|_{L^1} &\le \frac{1}{r}\sum_{i=0}^{n-1}|(z^{-1}\mathcal{L})^i g|_{L^1} + \|R(z)(z^{-1}\mathcal{L})^n g\|_{BV}\\ &\le \frac{1}{r^n(1-r)}|g|_{L^1} + H_{\delta,r}A(r\lambda)^{-n}\|g\|_{BV} + H_{\delta,r}Br^{-n}|g|_{L^1}\\ &\le r^{-n}(H_{\delta,r}B + (1-r)^{-1})|g|_{L^1} + H_{\delta,r}A(r\lambda)^{-n}\|g\|_{BV}\end{aligned}$$

[18] This proof is simpler than the one in [KL99], yet it gives worst bounds, although sufficient for the present purposes.

Substituting, we have

$$\|(z^{-1}\mathcal{L}_\varepsilon)^n R(z)(\mathcal{L}_\varepsilon - \mathcal{L})h\|_{BV} \le \{(r\lambda)^{-n} A H_{\delta,r} 2C_1[1 + Br^{-n}] + Br^{-2n}[H_{\delta,r}B + (1-r)^{-1}]D\varepsilon\}\|h\|_{BV} < 1,$$

again, provided ε is small enough and choosing n appropriately. Hence the operator on the right hand side of (7.6.29) can be inverted, thereby providing a left inverse for $(z - \mathcal{L}_\varepsilon)$. This implies that z does not belong to the spectrum of $\mathcal{L}_\varepsilon$.

To investigate the second statement note that (7.6.27) implies

$$R(z) - R_\varepsilon(z) = \frac{1}{z}\sum_{i=0}^{n-1}(z^{-1}\mathcal{L})^i(\mathcal{L}_\varepsilon - \mathcal{L})R_\varepsilon(z) - R(z)(z^{-1}\mathcal{L})^n(\mathcal{L}_\varepsilon - \mathcal{L})R_\varepsilon(z).$$

Accordingly, for each $\varphi \in BV$ holds

$$|R(z)\varphi - R_\varepsilon(z)\varphi|_{L^1} \le \{r^{-n}(1-r)^{-1}\varepsilon + H_{\delta,r}(\lambda r)^{-n}2AC_1 + H_{\delta,r}B\varepsilon\}\|R_\varepsilon(z)\varphi\|_{BV}.$$

□

7.6.1 Deterministic stability

The $\mathcal{L}_\varepsilon$ are Perron-Frobenius (Transfer) operators of maps T_ε which are $\mathcal{C}^1$–close to T, that is $d_{\mathcal{C}^1}(T_\varepsilon, T) = \varepsilon$ and such that $d_{\mathcal{C}^2}(T_\varepsilon, T) \le M$, for some fixed $M > 0$. In this case the uniform Lasota-Yorke inequality is trivial. On the other hand, for all $\varphi \in \mathcal{C}^1$ holds

$$\int(\mathcal{L}_\varepsilon f - \mathcal{L}f)\varphi = \int f(\varphi\circ T_\varepsilon - \varphi\circ T).$$

Now let $\Phi(x) := (D_xT)^{-1}\int_{Tx}^{T_\varepsilon x}\varphi(z)dz$, since

$$\Phi'(x) = -(D_xT)^{-1}D_x^2T\Phi(x) + D_xT_\varepsilon(D_xT)^{-1}\varphi(T_\varepsilon x) - \varphi(Tx)$$

follows

$$\int(\mathcal{L}_\varepsilon f - \mathcal{L}f)\varphi = \int f\Phi' + \int f(x)[(D_xT)^{-1}D_x^2T\Phi(x) + (1 - D_xT_\varepsilon(D_xT)^{-1})\varphi(T_\varepsilon x)].$$

Given that $|\Phi|_\infty \le \lambda^{-1}\varepsilon|\varphi|_\infty$ and $|1 - D_xT_\varepsilon(D_xT)^{-1}|_\infty \le \lambda^{-1}\varepsilon$, we have

$$\int(\mathcal{L}_\varepsilon f - \mathcal{L}f)\varphi \le \|f\|_{BV}\lambda^{-1}|\varphi|_\infty\varepsilon + |f|_{L^1}\lambda^{-1}(B+1)\varepsilon|\varphi|_\infty \le D\|f\|_{BV}\varepsilon|\varphi|_\infty.$$

By Lebesgue dominate convergence theorem we obtain the above inequality for each $\varphi \in L^\infty$, and taking the sup on such φ yields the wanted inequality.

$$|\mathcal{L}_\varepsilon f - \mathcal{L}f|_{L^1} \le D\|f\|_{BV}\varepsilon.$$

We have thus seen that all the requirements in Condition 1 are satisfied. See [Kel82] for a more general setting including piecewise smooth maps.

7.6.2 Stochastic stability

Next consider a set of maps $\{T_\omega\}$ depending on a parameter $\omega \in \Omega$. In addition assume that Ω is a probability space and consider a measure P on Ω. Consider the process $x_n = T_{\omega_n} \circ \cdots \circ T_{\omega_1} x_0$ where the ω are i.i.d. random variables distributed accordingly to P and let E_μ be the expectation of such process when x_0 is distributed according to μ. Then, calling $\mathcal{L}_\omega$ the transfer operator associated to T_ω, we have

$$E(f(x_{n+1}) \mid x_n) = \mathcal{L}_P f(x_n) := \int_\Omega \mathcal{L}_\omega f(x_n) P(d\omega).$$

Then if

$$|\mathcal{L}_\omega h|_{BV} \leq \lambda_\omega^{-1} |h|_{BV} + B_\omega |h|_{L^1}$$

integrating yields

$$|\mathcal{L}_P h|_{BV} \leq E(\lambda_\omega^{-1}) |h|_{BV} + E(B_\omega) |h|_{L^1}$$

And the operator $\mathcal{L}_P$ satisfy a Lasota-Yorke inequality provided that $E(\lambda^{-1}) < 1$ and $E(B) < \infty$.

In addition, if for some map T and associated transfer operator $\mathcal{L}$,

$$E(|\mathcal{L}_\omega h - \mathcal{L} h|) \leq \varepsilon |h|_{BV}$$

then we can apply perturbation theory and obtain stochastic stability.

7.6.3 Computability

If we want to compute the invariant measure and the rate of decay of correlations, we can use the operator P_t defined in (7.3.6) and define $\mathcal{L}_{t,m} = P_t \mathcal{L}^m$. By the estimates in Lemma **??** it follows

$$|\mathcal{L}_{t,m} h|_{BV} \leq 4^d \sigma^m |h|_{BV} + B|h|_{L^1}.$$

We can then chose the smallest m so that $4^d \sigma^m = \sigma_1 < 1$. Moreover, we also saw that

$$|\mathcal{L}_{t,m} h - \mathcal{L} h| \leq t^{-1} |h|_{BV}.$$

So we are again in the realm of our perturbation theory and we have that the finite dimensional operator $\mathcal{L}_{t,m}$ has spectrum close to the one of the transfer operator. We can then obtain all the info we want by diagonalizing a matrix.

7.6.4 Linear response

Linear response is a theory widely used by physicists. In essence it says the follow: consider a one parameter family of systems T_s and the associated (e.g.) invariant measures μ_s, then, for a given observable f one want to study the response of the system to a small change in s, and, not surprisingly, one expects $\mu_s(f) = \mu_0(f) + s\nu(f) + o(s)$. That is one expects differentiability in s. Yet differentiability is is not ensured by Theorem 7.6.1. Is it possible to ensure conditions under which linear response holds? The answer is yes (for example if holds if the maps are sufficiently smooth and the dependence on the parameter is also smooth in an appropriate sense). To prove it one need a sophistication of Theorem 7.6.1 that can be found in [GL06].

7.6.5 The hyperbolic case

One can wonder is the previous approach can be applied to uniformly hyperbolic systems and partially hyperbolic system. The answer is yes although the work in this direction is still in progress and the price to pay is the need to consider rather unusual functional spaces (space of anysotropic distributions). Just to give a vague idea let us look at a totally trivial example: toral automorphisms.

Then one can consider the norms:

$$\|f\|_{p,q} := \sum_{k\in\mathbb{Z}^{2d}\setminus\{0\}} |f_k| \frac{|k|^p}{1+|\langle v^s, k\rangle|^{p+q}} + |f_0|,$$

where f_k are the Fourier coefficients of f and v^s is the unit vector in the stable direction. Then

$$\begin{aligned} \|[\mathcal{L}f\|_{p,q} &\leq C_1\|f\|_{p,q}, \\ \|[\mathcal{L}^n f\|_{p,q} &\leq C_3\mu^n\|f\|_{p,q} + B\|f\|_{p-1,q+1}. \end{aligned} \tag{7.6.30}$$

we have thus the Lasota-Yorke inequality. Moreover on can easily check the relative compactness of $\{\|f\|_{p,q} \leq 1\}$ with respect to the topology induced by the norm $\|\cdot\|_{p-1,q+1}$, hence our previous theory applies almost verbatim.

To have a more precise idea of what can be done, see [GL06, BT07].

Hints to solving the Problems

7.18 Let ℓ_λ, h_λ be analytic. Let us define $z_\lambda = e^{-\int_0^\lambda \ell_\xi(h'_\xi)d\xi}$, define $\hat{h}_\lambda = z_\lambda h_\lambda$ and $\hat{\ell}_\lambda = z_\lambda^{-1}\ell_\lambda$ and check that they are normalized as required.

Notes

Large deviations are taken from Lai-Sang article and Keller book.

The stochastic stability is reasonably well understood (Cowienson) but what about the smooth dependence from a parameter (linear response)? Counterexamples in $d = 1$ but unknown in higher dimensions. The uniformly hyperbolic case is well understood but not much is know on how to apply the present ideas to the partially hyperbolic case and to the case of systems with discontinuities, although a concentrated effort is taking place to extend the theory in such directions.

Appendix A

Fixed Points Theorems (an idiosincratic selection)

In this appendinx I provide some standard and less standard Fixed poins theorems. These constitute a very partial introduction to the subject. The choice of the topics if motivated by the needs of the previous chapters.

A.1 Banach Fixed Point Theorem

Theorem A.1.1 (Fixed point contraction) *Given a Banach space $\mathcal{B}$, a bounded closed set $A \subset \mathcal{B}$ and a map $K : A \to \mathcal{B}$ if*

i) $K(A) \subset A$,

ii) there exists $\sigma \in (0,1)$ *such that* $\|K(v) - K(w)\| \leq \sigma\|v - w\|$ *for each* $v, w \in A$,

then there exists a unique $v_* \in A$ *such that* $Kv_* = v_*$.

Proof. Since A is bounded $\sup_{x,y\in A} \|x-y\| = L < \infty$, i.e. it has a finite diameter. Let $a_0 \in A$ and consider the sequence of points defined recursively by $a_{n+1} = K(a_n)$ and the sequence of sets $A_0 = A$ and $A_{n+1} = K(A_n) \subset A$. Let $d_n := \sup_{x,y\in A_n} \|x - y\|$ be the diameter of A_n. Then if $x, y \in A_n$, we have

$$\|K(y) - K(x)\| \leq \sigma\|x - y\| \leq \sigma d_n.$$

That is $d_{n+1} \leq \sigma d_n \leq \sigma^n L$. This means that, for each $n, m \in \mathbb{N}$, $a_n, a_0 \in A$ and $a_m, a_{n+m} \in A_m$, hence $\|a_{n+m} - a_m\| \leq \sigma^m L$. That is $\{a_n\} \subset A$ is a Cauchy sequence and, being $\mathcal{B}$ a Banach space, it must have an accumulation point $v_* \in \mathcal{B}$. Moreover since A is closed it must be $v_* \in A$. Clearly

$$\begin{aligned}\|Kv_* - v_*\| &= \lim_{n\to\infty} \|Kv_* - a_n\| = \lim_{n\to\infty} \|Kv_* - Ka_{n-1}\| \\ &\leq \lim_{n\to\infty} \sigma\|v_* - a_{n-1}\| = 0.\end{aligned}$$

Hence, v_* is a fixed point. Next, suppose there exist $u \in A$, such that $Ku = u$. Then

$$\|u - v_*\| = \|K(u - v_*)\| \le \sigma \|u - v * \|$$

implies $u = v_*$. □

Corollary A.1.2 *Given a Banach space $\mathcal{B}$ and a map $K : \mathcal{B} \to \mathcal{B}$ with the property that there exists $\sigma \in (0,1)$ such that $\|K(v) - K(w)\| \le \sigma \|v - w\|$ for each $v, w \in \mathcal{B}$, then there exists a unique $v_* \in \mathcal{B}$ such that $Kv_* = v_*$.*

PROOF. To prove the theorem, for each $L \in \mathbb{R}_+$ consider the sets $B_L := \{v \in \mathcal{B} : \|v\| \le L\}$. Then $\|K(v)\| \le \|K(v) - K(0)\| + \|K(0)\| \le \sigma\|v\| + \|K(0)\| \le \sigma L + \|K(0)\|$. Thus, for each $L \ge (1-\sigma)^{-1}\|K(0)\|$ we have that $K(B_L) \subset B_L$. The existence follows by applying Theorem A.1.1. The uniqueness follows by the same argument used at end of the proof of Theorem A.1.1. □

A.2 Hilbert metric and Birkhoff theorem

In this section we will see that the Banach fixed point theorem can produce unexpected results if used with respect to an appropriate metric: projective metric.

Projective metrics are widely used in geometry, not to mention the importance of their generalizations (e.g. Kobayashi metrics) for the study of complex manifolds [IK00]. It is quite surprising that they play a major rôle also in our situation, [Liv95].

Here we limit ourselves to a few word on the Hilbert metric, a quite important tool in hyperbolic geometry.

A.2.1 Projective metrics

Let $C \in \mathbb{R}^n$ be a strictly convex compact set. For each two point $x, y \in C$ consider the line $\ell = \{\lambda x + (1 - \lambda y) \mid \lambda \in \mathbb{R}\}$ passing through x and y. Let $\{u, v\} = \partial C \cap \ell$ and define[1]

$$\Theta(x, y) = \left| \ln \frac{\|x - u\|\|y - v\|}{\|x - v\|\|y - v\|} \right|$$

(the logarithm of the cross ratio). By remembering that the cross ratio is a projective invariant and looking at Figure A.1 it is easy to check that Θ is indeed a metric. Moreover the distance of an inner point from the boundary is always infinite. One can also check that if the convex set is a disc then the disc with the Hilbert metric is nothing else than the Poincaré disc.

[1]Remark that u, v can also be ∞.

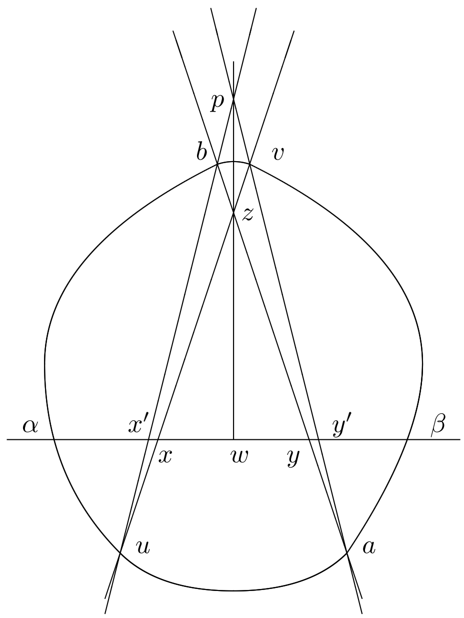

Figure A.1: Hilbert metric

The object that we will use in our subsequent discussion are not convex sets but rather convex cones, yet their projectivization is a convex set and one can define the Hilbert metric on it (whereby obtaining a semi-metric for the original cone). It turns out that there exists a more algebraic way of defining such a metric, which is easier to use in our context. Moreover, there exists a simple connection between vector spaces with a convex cone and vector lattices (in a vector lattice one can always consider the positive cone). This justifies the next digression in lattice theory.[2]

Consider a topological vector space $\mathbb{V}$ with a partial ordering "$\preceq$," that is a vector lattice.[3] We require the partial order to be "continuous," i.e. given $\{f_n\} \in \mathbb{V}$ $\lim_{n\to\infty} f_n = f$, if $f_n \succeq g$ for each n, then $f \succeq g$. We call such vector lattices "integrally closed." [4]

[2]For more details see [Bir57], and [Nus88] for an overview of the field.

[3]We are assuming the partial order to be well behaved with respect to the algebraic structure: for each $f, g \in \mathbb{V}$ $f \succeq g \iff f - g \succeq 0$; for each $f \in \mathbb{V}$, $\lambda \in \mathbb{R}^+ \setminus \{0\}$ $f \succeq 0 \Longrightarrow \lambda f \succeq 0$; for each $f \in \mathbb{V}$ $f \succeq 0$ and $f \preceq 0$ imply $f = 0$ (antisymmetry of the order relation).

[4]To be precise, in the literature "integrally closed" is used in a weaker sense. First, $\mathbb{V}$ does not need a topology. Second, it suffices that for $\{\alpha_n\} \in \mathbb{R}$, $\alpha_n \to \alpha$; $f, g \in \mathbb{V}$, if

We define the closed convex cone [5] $\mathcal{C} = \{f \in \mathbb{V} \mid f \neq 0,\ f \succeq 0\}$ (hereafter, the term "closed cone" $\mathcal{C}$ will mean that $\mathcal{C} \cup \{0\}$ is closed), and the equivalence relation "$\sim$": $f \sim g$ iff there exists $\lambda \in \mathbb{R}^+ \setminus \{0\}$ such that $f = \lambda g$. If we call $\widetilde{\mathcal{C}}$ the quotient of $\mathcal{C}$ with respect to $\sim$, then $\widetilde{\mathcal{C}}$ is a closed convex set. Conversely, 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. The reader must therefore be advised that "$\preceq$" will mean different things in different contexts.

It is then possible to define a projective metric Θ (Hilbert metric),[6] in $\mathcal{C}$, by the construction:

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

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

The relevance of the above metric in our coontex is due to the following Theorem by Garrett Birkhoff.

Theorem A.2.1 *Let $\mathbb{V}_1$, and $\mathbb{V}_2$ be two integrally closed vector lattices; $\mathcal{L} : \mathbb{V}_1 \to \mathbb{V}_2$ a linear map such that $\mathcal{L}(\mathcal{C}_1) \subset \mathcal{C}_2$, for two closed convex cones $\mathcal{C}_1 \subset \mathbb{V}_1$ and $\mathcal{C}_2 \subset \mathbb{V}_2$ with $\mathcal{C}_i \cap -\mathcal{C}_i = \emptyset$. Let Θ_i be the Hilbert metric corresponding to the cone $\mathcal{C}_i$. Setting $\Delta = \sup_{f,\, g \in T(\mathcal{C}_1)} \Theta_2(f,\, g)$ we have*

$$\Theta_2(\mathcal{L}f,\, \mathcal{L}g) \leq \tanh\left(\frac{\Delta}{4}\right) \Theta_1(f,\, g) \qquad \forall f,\, g \in \mathcal{C}_1$$

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

Proof. The proof is provided for the reader convenience.

Let $f,\, g \in \mathcal{C}_1$, on the one hand if $\alpha = 0$ or $\beta = \infty$, then the inequality is obviously satisfied. On the other hand, if $\alpha \neq 0$ and $\beta \neq \infty$, then

$$\Theta_1(f,\, g) = \ln \frac{\beta}{\alpha}$$

$\alpha_n f \succeq g$, then $\alpha f \succeq g$. Here we will ignore these and other subtleties: our task is limited to a brief account of the results relevant to the present context.

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

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

where $\alpha f \preceq g$ and $\beta f \succeq g$, since $\mathbb{V}_1$ is integrally closed. Notice that $\alpha \geq 0$, and $\beta \geq 0$ since $f \succeq 0$, $g \succeq 0$. If $\Delta = \infty$, then the result follows from $\alpha \mathcal{L} f \preceq \mathcal{L} g$ and $\beta \mathcal{L} f \succeq \mathcal{L} g$. If $\Delta < \infty$, then, by hypothesis,

$$\Theta_2\left(\mathcal{L}(g - \alpha f),\, \mathcal{L}(\beta f - g)\right) \leq \Delta$$

which means that there exist λ, $\mu \geq 0$ such that

$$\begin{aligned} \lambda \mathcal{L}(g - \alpha f) &\preceq \mathcal{L}(\beta f - g) \\ \mu \mathcal{L}(g - \alpha f) &\succeq \mathcal{L}(\beta f - g) \end{aligned}$$

with $\ln \frac{\mu}{\lambda} \leq \Delta$. The previous inequalities imply

$$\begin{aligned} \frac{\beta + \lambda\alpha}{1 + \lambda} \mathcal{L} f &\succeq \mathcal{L} g \\ \frac{\mu\alpha + \beta}{1 + \mu} \mathcal{L} f &\preceq \mathcal{L} g. \end{aligned}$$

Accordingly,

$$\begin{aligned} \Theta_2(\mathcal{L} f,\, \mathcal{L} g) &\leq \ln \frac{(\beta + \lambda\alpha)(1 + \mu)}{(1 + \lambda)(\mu\alpha + \beta)} = \ln \frac{e^{\Theta_1(f,\, g)} + \lambda}{e^{\Theta_1(f,\, g)} + \mu} - \ln \frac{1 + \lambda}{1 + \mu} \\ &= \int_0^{\Theta_1(f,\, g)} \frac{(\mu - \lambda) e^{\xi}}{(e^{\xi} + \lambda)(e^{\xi} + \mu)} d\xi \leq \Theta_1(f,\, g) \frac{1 - \frac{\lambda}{\mu}}{\left(1 + \sqrt{\frac{\lambda}{\mu}}\right)^2} \\ &\leq \tanh\left(\frac{\Delta}{4}\right) \Theta_1(f,\, g). \end{aligned}$$

□

Remark A.2.2 *If $\mathcal{L}(\mathcal{C}_1) \subset \mathcal{C}_2$, then it follows that $\Theta_2(\mathcal{L} f,\, \mathcal{L} g) \leq \Theta_1(f,\, g)$. However, a uniform rate of contraction depends on the diameter of the image being finite.*

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 $\mathcal{L}$ (see [Bir79] for details). The relevance of this theorem for the study of invariant measures and their ergodic properties is obvious.

It is natural to wonder about the strength of the Hilbert metric compared to other, more usual, metrics. While, in general, the answer depends on the cone, it is nevertheless possible to state an interesting result.

Lemma A.2.3 *Let $\|\cdot\|$ be a norm on the vector lattice $\mathbb{V}$, and suppose that, for each f, $g \in \mathbb{V}$,*

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

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

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

Proof. We know that $\Theta(f,\, g) = \ln \frac{\beta}{\alpha}$, where $\alpha f \preceq g$, $\beta f \succeq g$. This implies that $-g \preceq 0 \preceq \alpha f \preceq g$, i.e. $\|g\| \geq \alpha\|f\|$, or $\alpha \leq 1$. In the same manner it follows that $\beta \geq 1$. Hence,

$$\begin{aligned} g - f &\preceq (\beta - 1)f \preceq (\beta - \alpha)f \\ g - f &\succeq (\alpha - 1)f \succeq -(\beta - \alpha)f \end{aligned}$$

which implies

$$\|g - f\| \leq (\beta - \alpha)\|f\| \leq \frac{\beta - \alpha}{\alpha}\|f\| = \left(e^{\Theta(f,\, g)} - 1\right) \|f\|.$$

□

Many normed vector lattices satisfy the hypothesis of Lemma 1.3 (e.g. Banach lattices[7]); nevertheless, we will see that some important examples treated in this paper do not.

A.2.2 An application: Perron-Frobenius

Consider a matrix $L : \mathbb{R}^n \to \mathbb{R}^n$ of all strictly positive elements: $L_{ij} \geq \gamma > 0$. The Perron-Frobenius theorem states that there exists a unique eigenvector v^+ such that $v_i^+ > 0$, in addition the corresponding eigenvalue λ is simple, maximal and positive. There quite a few proofs of this theorem a possible one is based on Birkhoff theorem. Consider the cone $\mathcal{C}^+ = \{v \in \mathbb{R}^2 \mid v_i \geq 0\}$, then obviously $L\mathcal{C}^+ \subset \mathcal{C}^+$. Moreover an explicit computation (see

Problem A.1 *shows that*

$$\Theta(v, w) = \ln \sup_{ij} \frac{v_i w_j}{v_j w_i}. \tag{A.2.1}$$

[7]A Banach lattice $\mathbb{V}$ is a vector lattice equipped with a norm satisfying the property $\|\,|f|\,\| = \|f\|$ for each $f \in \mathbb{V}$, where $|f|$ is the least upper bound of f and $-f$. For this definition to make sense it is necessary to require that $\mathbb{V}$ is "directed," i.e. any two elements have an upper bound.

Then, setting $M = \max_{ij} L_{ij}$, it follows that

$$\Theta(Lv, Lw) \leq 2 \ln \frac{M}{\gamma} := \Delta < \infty.$$

We have then a contraction in the Hilbert metric and the result follows from usual fixed points theorems. Note that, since $\Theta(v, \lambda v) = 0$, for all $\lambda \in \mathbb{R}^+$, the fixed point $v_+ \in \mathbb{R}^n$ is only projective, that is $Lv_+ = \lambda v_+$ for some $\lambda \in \mathbb{R}$; in other words, we have an eigenvalue.

Remark that L^* satisfies the same conditions as L, thus there exists $w^+ \in \mathcal{C}^+$, $\mu \in \mathbb{R}^+$, such that $L^* w^+ = \mu w^+$. Next, define $\rho_1(v) = |\langle w^+, v\rangle|$ and $\rho_2(v) = \|v\|$. It is easy to check that they are two homogeneous forms of degree one adapted to the cone.

In addition, if $\rho_1(v) = \rho_2(v)$, then $\rho_1(L^n v) = \rho_1(L^n w)$. Hence, by Lemma A.2.3

$$\begin{aligned} \|L^n v - L^n w\| &\leq \left(e^{\Theta(L^n v, L^n w)} - 1\right) \min\{\|L^n v\|, \|L^n w\|\} \\ &\leq K \Lambda^n \min\{\|L^n v\|, \|L^n w\|\}, \end{aligned} \tag{A.2.2}$$

for some constant K depending only on v, w. The estimate A.2.2 means that all the vectors in the cone grow at the same rate. In fact, for all $v \in \text{int}\mathcal{C}$,

$$\|\lambda^{-n} L^n v - \lambda^{-n} L^n w\| \leq K \Lambda^n.$$

Hence, $\lim_{n\to\infty} \lambda^{-n} L^n v = v_+$.

Finally, consider $\mathbb{V}_1 = \{v \in \mathbb{V} \mid \langle w^+, v\rangle = 0\}$. Clearly $L\mathbb{V}_1 \subset \mathbb{V}_1$ and $\mathbb{V}_1 \oplus \text{span}\{v_+\} = \mathbb{V}$. Let $w \in \mathbb{V}_1$, clearly there exists $\alpha \in \mathbb{R}^+$ such that $\alpha v_+ + w \in \mathcal{C}$,[8] thus

$$\|L^n w\| \leq \|L^n(\alpha v_+ + w) - \alpha L^n v_+\| \leq L \Lambda^n \lambda^n.$$

This immediately implies that L restricted to the subspace $\mathbb{V}_1$ has spectral radius less that $\lambda\Lambda$. In other words, λ is the maximal eigenvalue, it is simple and any other eigenvalue must be smaller than $\lambda\Lambda$. We have thus obtained an estimate of the spectral gap between the first and the second eigenvalue.

Notes

For more details on Hilbert metrics see [Bir79], and [Nus88] for an overview of the field.

[8] this is a special case of the general fat that any vector can be written as the linear combination of two vectors belonging to the cone.

Appendix C

Perturbation Theory (a super-fast introduction)

The following is really super condensate (although self-consistent). If you want more details see [RS80, Kat66] in which you probably can find more than you are looking for.

C.1 Bounded operators

In the following we will consider only *separable* Banach spaces, i.e. Banach spaces that have a countable dense set.[1]

Given a Banch space $\mathcal{B}$ we can consider the set $L(\mathcal{B}, \mathcal{B})$ of the linear bounded operators from $\mathcal{B}$ to itself. We can then introduce the norm $\|B\| = \sup_{\|v\| \leq 1} \|Bv\|$.

Problem C.1 *Show that $(L(\mathcal{B}, \mathcal{B}), \|\cdot\|)$ is a Banach space. That is that $\|\cdot\|$ is really a norm and that the space is complete with respect to such a norm.*

Problem C.2 *Show that the $n \times n$ matrices form a Banach Algebra.*[2]

Problem C.3 *Show that $L(\mathcal{B}, \mathcal{B})$ form a Banach algebra.*[3]

To each $A \in L(\mathcal{B}, \mathcal{B})$ are associated two important subspaces: the range $R(A) = \{v \in \mathcal{B} \;:\; \exists\, w \in \mathcal{B} \text{ such that } v = Aw\}$ and the kernel $N(A) = \{v \in \mathcal{B} \;:\; Av = 0\}$.

[1] Recall that a Banach space is a complete normed vector space (in the following we will consider vector spaces on the field of complex numbers), that is a normed vector space in which all the Cauchy sequences have a limit in the space. Again, if you are uncomfortable with Banach spaces, in the following read $\mathbb{R}^d$ instead of $\mathcal{B}$ and matrices instead of operators, but be aware that we have to develop the theory without the use of the determinant that, in general, is not defined for operators on Banach spaces.

[2] A Banach Algebra $\mathcal{A}$ is a Banach space where it is defined the multiplications between element with the usual properties of an algebra and, in addition, for each $a, b \in \mathcal{A}$ holds $\|ab\| \leq \|a\| \cdot \|b\|$.

[3] The multiplication is given by the composition.

Problem C.4 *Prove, for each $A \in L(\mathcal{B}, \mathcal{B})$, that $N(A)$ is a closed linear subspaces of $\mathcal{B}$. Show that this is not necessarily the case for $R(A)$ if $\mathcal{B}$ is not finite dimensional.*

An very special, but very important, class of operators are the projectors.

Definition C.1.1 *An operator $\Pi \in L(\mathcal{B}, \mathcal{B})$ is called a* projector *iff $\Pi^2 = \Pi$.*

Note that if Π is a projector, so is $\mathbb{1} - \Pi$. We have the following interesting fact.

Lemma C.1.2 *If $\Pi \in L(\mathcal{B}, \mathcal{B})$ is a projector, then $N(\Pi) \oplus R(\Pi) = \mathcal{B}$.*

PROOF. If $v \in \mathcal{B}$, then $v = \Pi v + (\mathbb{1} - \Pi)v$. Notice that $R(\mathbb{1} - \Pi) = N(\Pi)$ and $R(\Pi) = N(\mathbb{1} - \Pi)$. Finally, if $v \in N(\Pi) \cap R(\Pi)$, then $v = 0$, which concludes the proof. □

Another, more general, very important class of operators are the compact ones.

Definition C.1.3 *An operator $K \in L(\mathcal{B}, \mathcal{B})$ is called* compact *iff for any bounded set B the closure of $K(B)$ is compact.*

Remark C.1.4 *Note that not all the linear operator on a Banach space are bounded. For example consider the derivative acting on $\mathcal{C}^1((0,1), \mathbb{R})$.*

C.2 Functional calculus

First of all recall that all the Riemannian theory of integration works verbatim for function $f \in \mathcal{C}^0(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ is a Banach space. We can thus talk of integrals of the type $\int_a^b f(t)dt$.[4] Next, we can talk of *analytic functions* for functions in $\mathcal{C}^0(\mathbb{C}, \mathcal{B})$: a function is analytic in an open region $U \subset \mathbb{C}$ iff at each point $z_0 \in U$ there exists a neighborhood $B \ni z_0$ and elements $\{a_n\} \subset \mathcal{B}$ such that

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad \forall z \in B. \tag{C.2.1}$$

Problem C.5 *Show that if $f \in \mathcal{C}^0(\mathbb{C}, \mathcal{B})$ is analytic in $U \subset \mathbb{C}$, then given any smooth closed curve γ, contained in a sufficiently small disk in U, holds*[5]

$$\int_\gamma f(z)dz = 0 \tag{C.2.2}$$

[4] This is special case of the so called Bochner integral [Yos95].

[5] Of course, by $\int_\gamma f(z)dz$ we mean that we have to consider any smooth parametrization $g : [a, b] \to \mathbb{C}$ of γ, $g(a) = g(b)$, and then $\int_\gamma f(z)dz := \int_a^b f \circ g(t)g'(t)dt$. Show that the definition does not depend on the parametrization and that one can use piecewise smooth parametrizations as well.

Then show that the same hold for any piecewise smooth closed curve with interior contained in U, provided U is simply connected.

Problem C.6 *Show that if $f \in \mathcal{C}^0(\mathbb{C}, \mathcal{B})$ is analytic in a simply connected $U \subset \mathbb{C}$, then given any smooth closed curve γ, with interior contained contained in U and having in its interior a point z, hods the formula*

$$f(z) = \frac{1}{2\pi i} \int_\gamma (\xi - z)^{-1} f(\xi) d\xi. \tag{C.2.3}$$

Problem C.7 *Show that if $f \in \mathcal{C}^0(\mathbb{C}, \mathcal{B})$ satisfies (C.2.3) for each smooth closed curve in a simply connected open set U, then f is analytic in U.*

C.3 Spectrum and resolvent

Given $A \in L(\mathcal{B}, \mathcal{B})$ we define the *resolvent*, called $\rho(A)$, as the set of the $z \in \mathbb{C}$ such that $(z\mathbb{1} - A)$ is invertible and the inverse belongs to $L(\mathcal{B}, \mathcal{B})$. The *spectrum* of A, called $\sigma(A)$ is the complement of $\rho(A)$ in $\mathbb{C}$.

Problem C.8 *Prove that, for each Banach space $\mathcal{B}$ and operator $A \in L(\mathcal{B}, \mathcal{B})$, if $z \in \rho(A)$, then there exists a neighborhood U of z such that $(z\mathbb{1} - A)^{-1}$ is analytic in U.*

From the above exercise follows that $\rho(A)$ is open, hence $\sigma(A)$ is closed.

Problem C.9 *Show that, for each $A \in L(\mathcal{B}, \mathcal{B})$, $\sigma(A) \neq \emptyset$.*

Problem C.10 *Show that if $\Pi \in L(\mathcal{B}, \mathcal{B})$ is a projector, then $\sigma(\Pi) = \{0, 1\}$.*

Up to now the theory for operators seems very similar to the one for matrices. Yet, the spectrum for matrices is always given by a finite number of points while the situation for operators can be very dfferenct.

Problem C.11 *Consider the operator $\mathcal{L} : \mathcal{C}^0([0,1], \mathbb{C}) \to \mathcal{C}^0([0,1], \mathbb{C})$ defined by*

$$(\mathcal{L}f)(x) = \frac{1}{2} f(x/2) + \frac{1}{2} f(x/2 + 1/2).$$

Show that $\sigma(\mathcal{L}) = \{z \in \mathbb{C} \; : \; |z| \leq 1\}$.

Problem C.12 *Show that, if $A \in L(\mathcal{B}, \mathcal{B})$ and p is any polynomial, then for each $n \in \mathbb{N}$ and smooth curve $\gamma \subset \mathbb{C}$, with $\sigma(A)$ in its interior,*

$$p(A) = \frac{1}{2\pi i} \int_\gamma p(z)(z\mathbb{1} - A)^{-1} dz.$$

Problem C.13 *Show that, for each $A \in L(\mathcal{B}, \mathcal{B})$ the limit*

$$r(A) = \lim_{n \to \infty} \|A^n\|^{\frac{1}{n}}$$

exists.

The above limit is called the *spectral radius* of A.

Lemma C.3.1 *For each $A \in L(\mathcal{B}, \mathcal{B})$ holds true* $\sup_{z \in \sigma(A)} |z| = r(A)$.

PROOF. Since we can write

$$(z\mathbb{1} - A)^{-1} = z^{-1}(\mathbb{1} - z^{-1}A)^{-1} = z^{-1} \sum_{n=0}^{\infty} z^{-n} A^n,$$

and since the series converges if it converges in norm, from the usual criteria for the convergence of a series follows $\sup_{z \in \sigma(A)} |z| \leq r(A)$. Suppose now that the inequality is strict, then there exists $0 < \eta < r(A)$ and a curve $\gamma \subset \{z \in \mathcal{C} \ : \ |z| \leq \eta\}$ which contains $\sigma(A)$ in its interior. Then applying Problem C.12 yields $\|A^n\| \leq C\eta^n$, which contradicts $\eta < r(A)$. □

Note that if $f(z) = \sum_{n=0}^{\infty} f_n z^n$ is an analytic function in all $\mathbb{C}$ (entire), then we can define

$$f(A) = \sum_{n=0}^{\infty} f_n A^n.$$

Problem C.14 *Show that, if $A \in L(\mathcal{B}, \mathcal{B})$ and f is an entire function, then for each smooth curve $\gamma \subset \mathbb{C}$, with $\sigma(A)$ in its interior,*

$$f(A) = \frac{1}{2\pi i} \int_{\gamma} f(z)(z\mathbb{1} - A)^{-1} dz.$$

In view of the above fact, the following definition is natural:

Definition C.3.2 *For each $A \in L(\mathcal{B}, \mathcal{B})$, f analytic in a region U containing $\sigma(A)$, then for each smooth curve $\gamma \subset U$, with $\sigma(A)$ in its interior, define*

$$f(A) = \frac{1}{2\pi i} \int_{\gamma} f(z)(z\mathbb{1} - A)^{-1} dz. \tag{C.3.4}$$

Problem C.15 *Show that the above definition does not depend on the curve γ.*

Problem C.16 *For each $A \in L(\mathcal{B}, \mathcal{B})$ and functions f, g analytic on a domain $D \supset \sigma(A)$, show that $f(A) + g(A) = (f + g)(A)$ and $f(A)g(A) = (f \cdot g)(A)$.*

Problem C.17 *In the hypotheses of the Definition C.3.2 show that* $f(\sigma(A)) = \sigma(f(A))$ *and* $[f(A), A] = 0$.

Problem C.18 *Consider* $f : \mathbb{C} \to \mathbb{C}$ *entire and* $A \in L(\mathcal{B}, \mathcal{B})$. *Suppose that* $\{z \in \mathbb{C} \ : \ f(z) = 0\} \cap \sigma(A) = \emptyset$. *Show that* $f(A)$ *is invertible and* $f(A)^{-1} = f^{-1}(A)$.

Problem C.19 *Let* $A \in L(\mathcal{B}, \mathcal{B})$. *Suppose there exists a semi-line* ℓ, *starting from the origin, such that* $\ell \cap \sigma(A) = \emptyset$. *Prove that it is possible to define an operator* $\ln A$ *such that* $e^{\ln A} = A$.

Remark C.3.3 *Note that not all the interesting functions can be constructed in such a way. In fact,* $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *is such that* $A^2 = -\mathbb{1}$, *thus it can be interpreted as a square rooth of* $-\mathbb{1}$ *but it cannot be obtained directly by a formula of the type* (C.3.4).

Problem C.20 *Suppose that* $A \in L(\mathcal{B}, \mathcal{B})$ *and* $\sigma(A) = B \cup C$, $B \cap C = \emptyset$, *suppose that the smooth closed curve* $\gamma \subset \rho(A)$ *contains* B, *but not* C, *in its interior, prove that*

$$P_B := \frac{1}{2\pi i} \int_\gamma (z\mathbb{1} - A)^{-1} dz$$

is a projector that does not depend on γ.

Note that by Problem C.17 easily follows that $P_B A = A P_B$. Hence, $AR(P_B) \subset R(P_B)$ and $AN(P_B) \subset N(P_B)$. Thus $\mathcal{B} = R(P_B) \oplus N(P_B)$ provides an invariant decomposition for A.

Problem C.21 *In the hypotheses of Problem C.20, prove that* $A = P_B A P_B + (\mathbb{1} - P_B) A (\mathbb{1} - P_B)$.

Problem C.22 *In the hypotheses of Problem C.20, prove that* $\sigma(P_B A P_B) = B \cup \{0\}$. *Moreover, if* $\dim(R(P_B)) = D < \infty$, *then the cardinality of* B *is* $\leq D$.

C.4 Perturbations

Let us consider $A, B \in L(\mathcal{B}, \mathcal{B})$ and the family of operators $A_\nu := A + \nu B$.

Lemma C.4.1 *For each* $\delta > 0$ *there exists* $\nu_\delta \in \mathbb{R}$ *such that, for all* $|\nu| \leq \nu_\delta$, $\rho(A_\nu) \supset \{z \in \mathbb{C} \ : \ d(z, \sigma(A)) > \delta\}$.

Proof. Let $d(z, \sigma(A)) > \delta$, then

$$(z\mathbb{1} - A_\nu) = (z\mathbb{1} - A)\left[\mathbb{1} - \nu(z\mathbb{1} - A)^{-1}B\right] \tag{C.4.5}$$

Now $\|(z\mathbb{1} - A)^{-1}B\|$ is a continuous function in z outside $\sigma(A)$, moreover it is bounded outside a ball of large enough radius, hence there exists $M_\delta > 0$ such that $\sum_{d(z,\sigma(A))>\delta} \|(z\mathbb{1} - A)^{-1}B\| \leq M_\delta$. Choosing $\nu_\delta = (2M_\delta)^{-1}$ yields the result. □

Suppose that $\bar{z} \in \mathbb{C}$ is an isolated point of $\sigma(A)$, that is there exists $\delta > 0$ such that $\{z \in \mathbb{C} \, : \, |z - \bar{z}| \leq \delta\} \cap (\sigma(A) \setminus \{\bar{z}\}) = \emptyset$, then the above Lemma shows that, for ν small enough, $\{z \in \mathbb{C} \, : \, |z - \bar{z}| \leq \delta\}$ still contains an isolated part of the spectrum of $\sigma(A_\nu)$, let us call it B_ν, clearly $B_0 = \{\bar{z}\}$.

Problem C.23 *Let P_{B_ν} be defined as in Problem C.20. Prove that, for ν small enough, it is an analytic function of ν.*

Problem C.24 *If P, Q are two projectors and $\|P - Q\| < 1$, then $\dim(R(P)) = \dim(R(Q))$.*

The above two exercises imply that the dimension of the eigenspace $R(P_{B_\nu})$ is constant.

Next, we consider the case in which B_0 consist of one point and $\dim(R(P_{B_0})) = 1$, it follows that also B_ν must consist of only one point, let us set $P_\nu := P_{B_\nu}$.

Lemma C.4.2 *If $\dim(R(P_0)) = 1$, then A_ν has a unique eigenvalue z_ν in a neighborhood of $\bar{z}$, $z_0 = \bar{z}$. In addition z_ν is an analytic function of ν.*

Proof. From the previous exercises it follows that P_ν is a rank one operator which depend analytically on ν. In addition, since P_ν is a rank one projector it must have the form $P_\nu w = v_\nu \ell_\nu(w)$, where $\ell_\nu \in \mathcal{B}'$.[6] Then $z_\nu P_\nu = P_\nu A_\nu P_\nu$. Next, setting $a(\nu) := \ell_0(P_\nu v_0) = \ell_\nu(v_0)\ell_0(v_\nu)$, we have that a is analytic and $a(0) = 1$. Thus $a \neq 0$ in a neighborhood of zero and $z_\nu = a(\nu)^{-1}\ell_0(P_\nu A_\nu P_\nu v_0)$ is analytic in such a neighborhood. □

Problem C.25 *If $\dim(R(P_0)) = 1$, then there exists $h_\nu \in \mathcal{B}$ and $\ell_\nu \in \mathcal{B}'$ such that $P_\nu f = h_\nu \ell_\nu(f)$ for each $f \in \mathcal{B}$. Prove that h_ν, ℓ_ν can be chosen to be analytic functions of ν.*

Hence in the case of $A \in L(\mathcal{B}, \mathcal{B})$ with an isolated simple[7] eigenvalue $\bar{z}$ we have that the corresponding eigenvalue z_ν of $A_\nu = A + \nu B$, $B \in L(\mathcal{B}, \mathcal{B})$, for ν small enough, depend smoothly from ν. In addition, using the notation

[6] By $\mathcal{B}'$, the dual space, we mean the set of bounded linear functionals on $\mathcal{B}$. Verify that is a Banach space with the norm $\|\ell\| = \sum_{w \in \mathcal{B}} \frac{|\ell(w)|}{\|w\|}$.

[7] That is with the associated eigenprojector of rank one.

of the previous Lemma, we can easily compute the derivative: differentiating $A_\nu v_\nu = z_\nu v_\nu$ with respect to ν and then setting $\nu = 0$, yields

$$Bv + Av_0' = z_0'v + \bar{z}v_0'.$$

But, for all $w \in \mathcal{B}$, $Pw = v\ell(w)$, with $\ell(Aw) = \bar{z}\ell(w)$ and $\ell(v) = 1$, thus applying ℓ to both sides of the above equation yields

$$z_0' = \ell(Bv).$$

Problem C.26 *Compute v_0'.*

Problem C.27 *What does it happen if the eigenspace associated to $\bar{z}$ is finite dimensional, but with dimension strictly larger than one?*

Hints to solving the Problems

C.1. The triangle inequality follows trivially from the triangle inequality of the norm of $\mathcal{B}$. To verify the completeness suppose that $\{B_n\}$ is a Cauchy sequence in $L(\mathcal{B}, \mathcal{B})$. Then, for each $v \in \mathcal{B}$, $\{B_n v\}$ is a Cauchy sequence in $\mathcal{B}$, hence it has a limit, call it $B(v)$. We have so defined a function from $\mathcal{B}$ to teself. Show that such a function is linear and bounded, hence it defines an element of $L(\mathcal{B}, \mathcal{B})$, which can easily be verified to be the limit of $\{B_n\}$.

C.2. Use the norm $\|A\| = \sup_{v \in \mathbb{R}^n} \frac{\|Av\|}{\|v\|}$.

C.3. Use the same norm as in Problem C.2.

C.4. The first part is trivial. For the second one can consider the vector space $\ell^2 = \{x \in \mathbb{R}^{\mathbb{N}} \ : \ \sum_{i=0}^\infty x_i^2 < \infty\}$. Equipped with the norm $\|x\| = \sqrt{\sum_{i=0}^\infty x_i^2}$ it is a Banach (actually Hilbert) space. Consider now the vectors $e_i \in \ell^2$ defined by $(e_i) = \delta_{ik}$ and the operator $(Ax)_k = \frac{1}{k} x_k$. Then $R(A) = \{x \in \ell^2 \ : \ \sum_{k=0}^\infty k^2 x_k^2 < \infty\}$, which is dense in ℓ^2 but strictly smaller.

C.5. Check that the same argument used in the well known case $\mathcal{B} = \mathbb{C}$ works also here.

C.6. Check that the same argument used in the well known case $\mathcal{B} = \mathbb{C}$ works also here.

C.7. Check that the same argument used in the well known case $\mathcal{B} = \mathbb{C}$ works also here.

C.8. Note that

$$(\zeta\mathbb{1} - A) = (z\mathbb{1} - A - (z-\zeta)\mathbb{1}) = (z\mathbb{1} - A)\left[\mathbb{1} - (z-\zeta)(z\mathbb{1} - A)^{-1}\right]$$

and that if $\|(z-\zeta)(z\mathbb{1}-A)^{-1}\| < 1$ then the inverse of $\mathbb{1} - (z-\zeta)(z\mathbb{1} - A)^{-1}$ is given by $\sum_{n=0}^{\infty}(z-\zeta)^n[(z\mathbb{1}-A)^{-1}]^n$ (the Von Neumann series–which really is just the geometric series).

C.9. If $\sigma(A) = \emptyset$, then $(z\mathbb{1} - A)^{-1}$ is an entire function, then the Von Neumann series shows that $(z\mathbb{1} - A)^{-1} = z^{-1}(\mathbb{1} - z^{-1}A)^{-1}$ goes to zero for large z, and then (C.2.3) shows that $(z\mathbb{1} - A)^{-1} = 0$ which is impossible.

C.10. Verify that $(z\mathbb{1} - \Pi)^{-1} = z^{-1}\left[\mathbb{1} - (z-1)^{-1}\Pi\right]$.

C.11. The idea is to look for eigenvalues by using Fourier series. Let $f = \sum_{k\in\mathbb{Z}} f_k e^{2\pi i k x}$ and consider the equation $\mathcal{L}f = zf$,

$$\sum_{k\in\mathbb{Z}} f_k \frac{1}{2}\left\{e^{\pi i k x} + e^{\pi i k x + \pi i k}\right\} = z\sum_{k\in\mathbb{Z}} f_k e^{2\pi i k x}.$$

Let us then restrict to the case in which $f_{2k+1} = 0$, then

$$\sum_{k\in\mathbb{Z}} f_{2k} e^{2\pi i k x} = z\sum_{k\in\mathbb{Z}} f_k e^{2\pi i k x}.$$

Thus we have a solution provided $f_{2k} = zf_k$, such conditions are satisfied by any sequence of the type

$$f_k = \begin{cases} z^j & \text{if } k = 2^j m, j \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

for $m \in \mathbb{N}$. It remains to verify that $\sum_{j=0}^{\infty} z^j e^{2\pi i 2^j x}$ belong to $\mathcal{C}^0$. This is the case if the series is uniformly convergent, which happens for $|z| < 1$. Thus all the points in $\{z \in \mathbb{C} \;:\; |z| < 1\}$ are point spectrum of infinite multiplicity. Since the spectrum is closed the statement of the Problem follows.

C.12. Let $p(z) = z^n$, then

$$\begin{aligned}\frac{1}{2\pi i}\int_\gamma z^n(z\mathbb{1} - A)^{-1}dz &= A^n + \frac{1}{2\pi i}\int_\gamma (z^n - A^n)(z\mathbb{1} - A)^{-1}dz \\ &= A^n + \sum_{k=0}^{n-1}\frac{1}{2\pi i}\int_\gamma z^k A^{n-k}dz = A^n.\end{aligned}$$

The statement for general polynomial follows trivially.

C.14. Approximate by polynomials.

C.17. For $z \notin f(\sigma(A))$ it is well defined

$$K(z) := \frac{1}{2\pi i} \int_\gamma (z - f(\zeta))^{-1} (\zeta \mathbb{1} - A)^{-1} \, d\zeta,$$

with γ containing $\sigma(A)$ in its interior. By direct computation, using definition C.3.2, one can verify that $(z\mathbb{1} - f(A))K(z) = \mathbb{1}$, thus $\sigma(f(A)) \subset f(\sigma(A))$. On the other hand if, if f is not constant, then for each $z \in \mathcal{C}$ $f(z) - f(\xi) = (z - \xi)g(\xi)$. Hence, applying Definition C.3.2 and Problem C.16 it follows $f(z)\mathbb{1} - f(A) = (z - A)g(A)$ which shows that if $z \in \sigma(A)$, then $f(z) \in \sigma(A)$ (otherwise $(z - A)\left[g(A)(f(z)\mathbb{1} - f(A))^{-1}\right] = \mathbb{1}$).

C.19. Since one can define the logarithm on $\mathbb{C} \setminus \ell$, one can use Definition C.3.2 to define $\ln A$. It suffices to prove that if $f : U \to \mathcal{C}$ and $g : V \to \mathcal{C}$, with $\sigma(A) \subset U$, $f(U) \subset V$, then $g(f(A)) = g \circ f(A)$. Whereby showing that the definition C.3.2 is a reasonable one. Indeed, rememebring Problems C.17, C.18,

$$\begin{aligned} g(f(A)) &= \frac{1}{2\pi i} \int_\gamma g(z)(z\mathbb{1} - f(A))^{-1} dz \\ &= \frac{1}{(2\pi i)^2} \int_{\gamma_1} \int_\gamma \frac{g(z)}{z - f(\xi)} (\xi \mathbb{1} - A)^{-1} dz d\xi \\ &= \frac{1}{2\pi i} \int_{\gamma_1} g(f(\xi))(\xi\mathbb{1} - A)^{-1} d\xi = f \circ g(A). \end{aligned}$$

From this imediately follows $e^{\ln A} = A$.

C.20. The non dependence on γ is obvious. A projector is characterized by the property $P^2 = P$. Thus

$$\begin{aligned} P_B^2 :=& \frac{1}{(2\pi i)^2} \int_{\gamma_1} \int_{\gamma_2} (z\mathbb{1} - A)^{-1} (\zeta\mathbb{1} - A)^{-1} dz d\zeta \\ =& \frac{1}{(2\pi i)^2} \int_{\gamma_1} dz \int_{\gamma_2} d\zeta (z - \zeta)^{-1} \left[(z\mathbb{1} - A)^{-1} - (\zeta \mathbb{1} - A)^{-1}\right]. \end{aligned}$$

If we have chosen γ_1 in the interior of γ_2, then $(z - \zeta)^{-1}(\zeta\mathbb{1} - A)^{-1}$ is analytic in the interior of γ_1, hence the corresponding integral gives zero. The other integral gives P_B, as announced.

C.21. Use the above decomposition and the fact that $(\mathbb{1} - P_B)$ is a projector.

C.22. The first part follows from the previous decomposition. Indeed, for z large (by Neumann series)

$$(z\mathbb{1} - A)^{-1} = (z\mathbb{1} - P_B A P_B)^{-1} + (z\mathbb{1} - (\mathbb{1} - P_B)A(\mathbb{1} - P_B))^{-1}.$$

Since the above functions are analytic in the respective resolvent sets it follows that $\sigma(A) \subset \sigma(P_B A P_B) \cup \sigma((\mathbb{1} - P_B)A(\mathbb{1} - P_B))$. Next, for $z \notin B$, define the operator

$$K(z) := \frac{1}{2\pi i}\int_\gamma (z-\xi)^{-1}(\xi\mathbb{1} - A)^{-1}\, d\xi,$$

where γ contains B, but no other part of the spectrum, in its interior. By direct computation (using Fubini and the standard facts about residues and integration of analytic functions) verify that

$$(z\mathbb{1} - P_B A P_B)K(z) = P_B.$$

This implies that, for $z \neq 0$, $(z\mathbb{1} - P_B A P_B)(K(z) + z^{-1}(\mathbb{1} - P_B)) = \mathbb{1}$, that is $(z\mathbb{1} - P_B A P_B)^{-1} = K(z) + z^{-1}(\mathbb{1} - P_B)$. Hence $\sigma(P_B A P_B) \subset B \cup \{0\}$. Since P_B has a kernel, zero must be in the spectrum. On the other hand the same argument applied to $\mathbb{1} - P_B$ yields $\sigma((\mathbb{1} - P_B)A)\mathbb{1} - P_B)) \subset C \cup \{0\}$, hence $\sigma(P_B A P_B) = B \cup \{0\}$.

The second property follows from the fact that $P_B A P_B$, when restricted to the space $R(P_B)$ is described by a $D \times D$ matrix A_B and the equation $\det(z\mathbb{1} - A_B) = 0$ is a polynomial of degree D in z and hence has exactly D solutions (counted with multiplicity).[8]

C.23. Use the representation in Problem C.20 and formula (C.4.5).

C.24. Note that $Q(\mathbb{1} + P - Q) = QP$, then $Q = (\mathbb{1} - (Q - P))^{-1}QP$, hence $\dim(R(P)) \geq \dim(R(Q))$, exchanging the role of P and Q the result follows.

C.25. Note that $\ell_\nu(h_\nu) = 1$ since P_ν is a projector, hence they are unique apart from a noralization factor. Then we can chose the normalization

[8] This is the real reason why spectral theory is done over the complex rather than the real. You should be well aquatinted with the fact that a polynomial p of degree D has D root over $\mathbb{C}$ but, in case you have forgotten, consider the following: first a polynomial of degree larger than zero must have at least a root, otherwise $\frac{1}{p(z)}$ would be an entire function and hence

$$\frac{1}{p(z)} = \lim_{r\to\infty}\frac{1}{2\pi}\int_0^{2\pi} d\theta \frac{1}{p(z + re^{i\theta})} = 0.$$

Let z_1 be a root. By the Taylor expansion in z_1 follows the decomposition $p(z) = (z - z_1)p_1(z)$ where p_1 has degree $D-1$. The result follows by induction.

$\ell_\nu(h_0) = 1$ for all ν small enough. Thus $P_\nu f = h_\nu$, that is h_ν is analytic. Hence, for each $g \in \mathcal{B}$ and ν small, $\ell_\nu(g)\ell_0(h_\nu) = \ell_0(P_\nu g)$, which implies ℓ_ν analytic for ν small.

C.27. Think hard.[9]

[9] A good idea is to start by considering concrete examples, for instance

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \mu \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad ; \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + \mu \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

Appendix D

Analytic Fredholm Theorem (fine rank)

Here we give a proof of the Analytic Fredholm alternative in a special case.

Theorem D.0.1 (Analytic Fredholm theorem–finite rank)[1] *Let D be an open connected subset of $\mathbb{C}$. Let $F : \mathbb{C} \to L(\mathbb{B}, \mathbb{B})$ be an analytic operator-valued function such that $F(z)$ is finite rank for each $z \in D$. Then, one of the following two alternatives holds true*

- $(\mathbb{1} - F(z))^{-1}$ *exists for no* $z \in D$
- $(\mathbb{1} - F(z))^{-1}$ *exists for all* $z \in D \backslash S$ *where S is a discrete subset of D (i.e. S has no limit points in D). In addition, if $z \in S$, then 1 is an eigenvalue for $F(z)$ and the associated eigenspace has finite multiplicity.*

Proof. First of all notice that, for each $z_0 \in D$ there exists $r > 0$ such that $D_{r(z_0)}(z_0) := \{z \in \mathbb{C} \;:\; |z - z_0| < r(z_0)\} \subset D$, and

$$\sup_{z \in D_{r(z_0)}(z_0)} \|F(z) - F(z_0)\| \leq \frac{1}{2}.$$

Clearly if we can prove the theorem in each such disk we are done.[2] Note that

$$\mathbb{1} - F(z) = \left(\mathbb{1} - F(z_0)(\mathbb{1} - [F(z) - F(z_0)])^{-1}\right)(\mathbb{1} - [F(z) - F(z_0)]).$$

[1] The present proof is patterned after the proof of the Analytic Fredholm alternative for compact operators (in Hilbert spaces) given in [RS80, Theorem VI.14]. There it is used the fact that compact operators in Hilbert spaces can always be approximated by finite rank ones. In fact the theorem holds also for compact operators in Banach spaces but the proof is a bit more involved.

[2] In fact, consider any connected compact set K contained in D. Let us suppose that for each $z_0 \in K$ we have a disk $D_{r(z_0)}(z_0)$ in the theorem holds. Since the disks $D_{r(z_0)/2}(z_0)$ form a covering for K we can extract a finite cover. If the first alternative holds in one such disk then, by connectness, it must hold on all K. Otherwise each $S \cap D_{r(z_0)/2}(z_0)$, and hence $K \cap S$, contains only finitely many points. The Theorem follows by the arbitrariness of K.

Thus the invertibility of $\mathbb{1}-F(z)$ in $D_r(z_0)$ depends on the invertibility of $\mathbb{1}-F(z_0)(\mathbb{1}-[F(z)-F(z_0)])^{-1}$. Let us set $F_0(z) := F(z_0)(\mathbb{1}-[F(z)-F(z_0)])^{-1}$.

Let us start by looking at the equation

$$(\mathbb{1} - F_0(z))h = 0. \tag{D.0.1}$$

Clearly if a solution exists, then $h \in \text{Range}(F_0(z)) = \text{Range}(F(z_0)) := \mathbb{V}_0$. Since $\mathbb{V}_0$ is finite dimensional there exists a basis $\{h_i\}_{i=1}^N$ such that $h = \sum_i \alpha_i h_i$. On the other hand there exists an analytic matrix $G(z)$ such that[3]

$$F_0(z)h = \sum_{ij} G(z)_{ij}\alpha_j h_i.$$

Thus (D.0.1) is equivalent to

$$(\mathbb{1} - G(z))\alpha = 0,$$

where $\alpha := (\alpha_i)$.

The above equation can be satisfied only if $\det(\mathbb{1} - G(z)) = 0$ but the determinant is analytic hence it is either always zero or zero only at isolated points.[4]

Suppose the determinant different from zero, and consider the equation

$$(\mathbb{1} - F_0(z))h = g.$$

Let us look for a solution of the type $h = \sum_i \alpha_i h_i + g$. Substituting yields

$$\alpha - G(z)\alpha = \beta$$

where $\beta := (\beta_i)$ with $F_0(z)g =: \sum_i \beta_i h_i$. Since the above equation admits a solution, we have $\text{Range}(\mathbb{1} - F_0(z)) = \mathbb{B}$, Thus we have an everywhere defined inverse, hence bounded by the open mapping theorem.

We are thus left with the analysis of the situation $z \in S$ in the second alternative. In such a case, there exists h such that $(\mathbb{1} - F(z))h = 0$, thus one is an eigenvalue. On the other hand, if we apply the above facts to the function $\Phi(\zeta) := \zeta^{-1}F(z)$ analytic in the domain $\{\zeta \neq 0\}$ we note that the first alternative cannot take place since for $|\zeta|$ large enough $\mathbb{1} - \Phi(\zeta)$ is obviously

[3] To see the analyticity notice that we can construct linear functionals $\{\ell_i\}$ on $\mathbb{V}_0$ such that $\ell_i(h_j) = \delta_{ij}$ and then extend them to all $\mathbb{B}$ by the Hahn-Banach theorem. Accordingly, $G(z)_{ij} := \ell_j(F_0(z)h_i)$, which is obviously analytic.

[4] The attentive reader has certainly noticed that this is the turning point of the theorem: the discreteness of S is reduced to the discreteness of the zeroes of an appropriate analytic function: a determinant. A moment thought will immediately explain the effort made by many mathematicians to extend the notion of determinant (that is to define an analytic function whose zeroes coincide with the spectrum of the operator) beyond the realm of matrices (the so called Fredholm determinants).

invertible. Hence, the spectrum of $F(z)$ is discrete and can accumulate only at zero. This means that there is a small neighborhood around one in which $F(z)$ has no other eigenvalues, we can thus surround one with a small circle γ and consider the projector

$$\begin{aligned} P :=& \frac{1}{2\pi i}\int_\gamma (\zeta - F(z))^{-1} d\zeta = \frac{1}{2\pi i}\int_\gamma \left[(\zeta - F(z))^{-1} - \zeta^{-1}\right] d\zeta \\ =& \frac{1}{2\pi i}\int_\gamma F(z)\zeta^{-1}(\zeta - F(z))^{-1} d\zeta. \end{aligned}$$

By standard functional calculus it follows that P is a projector and it clearly projects on the eigenspace of the eigenvector one. But the last formula shows that P must project on a subspace of the range of $F(z)$, hence it must be finite dimensional. $\square$

Bibliography

[AA68] V. I. Arnold and A. Avez. *Ergodic problems of classical mechanics.* Translated from the French by A. Avez. W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[Aar97] Jon Aaronson. *An introduction to infinite ergodic theory*, volume 50 of *Mathematical Surveys and Monographs.* American Mathematical Society, 1997.

[Arn83] V. I. Arnold. *Geometrical methods in the theory of ordinary differential equations*, volume 250 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science].* Springer-Verlag, New York, 1983. Translated from the Russian by Joseph Szücs, Translation edited by Mark Levi.

[Arn99] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 199? Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.

[Arn92] Vladimir I. Arnold. *Ordinary differential equations.* Springer Textbook. Springer-Verlag, Berlin, 1992. Translated from the third Russian edition by Roger Cooke.

[Bal00] 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.

[Bir57] Garrett Birkhoff. Extensions of Jentzsch's theorem. *Trans. Amer. Math. Soc.*, 85:219–227, 1957.

[Bir79] Garrett Birkhoff. *Lattice theory*, volume 25 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, R.I., third edition, 1979.

[BT07] Viviane Baladi and Masato Tsujii. Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms. *Ann. Inst. Fourier (Grenoble)*, 57(1):127–154, 2007.

[BY93] Viviane Baladi and Lai-Sang Young. On the spectra of randomly perturbed expanding maps. *Comm. Math. Phys.*, 156:355–385, 1993.

[CC95] Alessandra Celletti and Luigi Chierchia. A constructive theory of Lagrangian tori and computer-assisted applications. In *Dynamics reported*, volume 4 of *Dynam. Report. Expositions Dynam. Systems (N.S.)*, pages 60–129. Springer, Berlin, 1995.

[CH82] Shui Nee Chow and Jack K. Hale. *Methods of bifurcation theory*, volume 251 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science].* Springer-Verlag, New York, 1982.

[CL55] Earl A. Coddington and Norman Levinson. *Theory of ordinary differential equations.* McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.

[DS88] Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part I.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[DZ98] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.

[Euc78] Euclide. *Les éléments. I, II.* Éditions du Centre National de la Recherche Scientifique (CNRS), Paris, french edition, 1978. Translated from the Greek by Georges J. Kayas.

[Gal83] Giovanni Gallavotti. *The elements of mechanics*. Texts and Monographs in Physics. Springer-Verlag, New York, 1983. Translated from the Italian.

[GL06] Sébastien Gouëzel and Carlangelo Liverani. Banach spaces adapted to Anosov systems. *Ergodic Theory Dynam. Systems*, 26:189–217, 2006.

[Her83] Michael-R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, volume 103 of *Astérisque*. Société Mathématique de France, Paris, 1983. With an appendix by Albert Fathi, With an English summary.

[Her86] Michael-R. Herman. Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 2. *Astérisque*, (144):248, 1986. With a correction to: ıt On the curves invariant under diffeomorphisms of the annulus, Vol. 1 (French) [Astérisque No. 103-104, Soc. Math. France, Paris, 1983; MR0728564 (85m:58062)].

[HK95] Boris Hasselblatt and Anatole Katok. *Introduction to the modern theory of dynamical systems*. Cambridge university press, 1995.

[Hop39] Eberhard Hopf. Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung. *Ber. Verh. Sächs. Akad. Wiss. Leipzig*, 91:261–304, 1939.

[Hop40] Eberhard Hopf. Statistik der Lösungen geodätischer Probleme vom unstabilen Typus. II. *Math. Ann.*, 117:590–608, 1940.

[HS74] Morris W. Hirsch and Stephen Smale. *Differential equations, dynamical systems, and linear algebra*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Pure and Applied Mathematics, Vol. 60.

[IK00] Alexander V. Isaev and Steven G. Krantz. Invariant distances and metrics in complex analysis. *Notices Amer. Math. Soc.*, 47(5):546–553, 2000.

[Kat66] Tosio Kato. *Perturbation theory for linear operators*. Die Grundlehren der mathematischen Wissenschaften, Band 132. Springer-Verlag New York, Inc., New York, 1966.

[Kel82] Gerhard Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.*, 94(4):313–333, 1982.

[Kif88] Yuri Kifer. *Random perturbations of dynamical systems*, volume 16 of *Progress in Probability and Statistics*. Birkhäuser Boston Inc., Boston, MA, 1988.

[KL99] 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.

[Liv95] Carlangelo Liverani. Decay of correlations. *Ann. of Math. (2)*, 142(2):239–301, 1995.

[LL76] L. D. Landau and E. M. Lifshitz. *Course of theoretical physics. Vol. 1*. Pergamon Press, Oxford, third edition, 1976. Mechanics, Translated from the Russian by J. B. Skyes and J. S. Bell.

[LL01] Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[LMD03] Carlangelo Liverani and Véronique Maume-Deschamps. Lasota-Yorke maps with holes: conditionally invariant probability measures and invariant probability measures on the survivor set. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(3):385–412, 2003.

[Mos01] Jürgen Moser. *Stable and random motions in dynamical systems.* Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2001. With special emphasis on celestial mechanics, Reprint of the 1973 original, With a foreword by Philip J. Holmes.

[Nus88] Roger D. Nussbaum. Hilbert's projective metric and iterated nonlinear maps. *Mem. Amer. Math. Soc.*, 75(391):iv+137, 1988.

[NZ99] Igor Nikolaev and Evgeny Zhuzhoma. *Flows on 2-dimensional manifolds*, volume 1705 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1999. An overview.

[Poi87] Henri Poincaré. *Le Méthodes Nouvelles de la Mécanique Céleste, Tome I,II, III.* le grand clasiques Gauthier-Villards. Blanchard, Paris, 1987.

[PT93] Jacob Palis and Floris Takens. *Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations*, volume 35 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1993. Fractal dimensions and infinitely many attractors.

[Roh67] V. A. Rohlin. Lectures on the entropy theory of transformations with invariant measure. *Uspehi Mat. Nauk*, 22(5 (137)):3–56, 1967.

[Roy88] H. L. Royden. *Real analysis.* Macmillan Publishing Company, New York, third edition, 1988.

[RS80] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1980. Functional analysis.

[Sie45] Carl Ludwig Siegel. Note on the differential equations on the torus. *Ann. of Math. (2)*, 46:423–428, 1945.

[Str84] D. W. Stroock. *An introduction to the theory of large deviations.* Universitext. Springer-Verlag, New York, 1984.

[UvN47] S. M. Ulam and John von Neumann. On combination of stochastic and deterministic processes. (preliminary report at the summer meeting in new haven). *Bull.Amer. Math.Soc.*, 53, 11:1120–1120, 1947.

[Var84] S. R. S. Varadhan. *Large deviations and applications*, volume 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984.

[Yos95] Kōsaku Yosida. *Functional analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

[Zei86] Eberhard Zeidler. *Nonlinear functional analysis and its applications. I.* Springer-Verlag, New York, 1986. Fixed-point theorems, Translated from the German by Peter R. Wadsack.