Spectral methods in dynamics (part 2)

Vaughn Climenhaga

February 8, 2013

This is a continuation of the last post, which were notes from the first in a series of talks at the Houston dynamics seminar on spectral methods for transfer operators as tools to establish statistical properties of dynamical systems. This week we apply the methods introduced last week to the example of a piecewise expanding interval map.

1   Piecewise expanding interval maps

We consider maps $T\colon X \to X$ of the form shown in Figure 1, where $X = [0, 1]$ is the unit interval. The map T is assumed to be C^2 on each of finitely many intervals whose union is X – these are called the *basic intervals* for T. Moreover, we assume that $\lambda > 1$ is such that $|T'(x)| \geq \lambda$ for every $x \in X$.
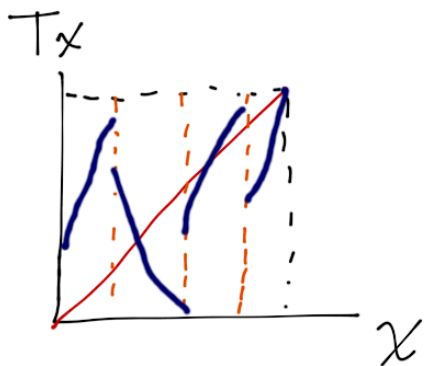

Figure 1: A piecewise expanding interval map.

Our goal is to show that the transfer operator for such maps has a spectral gap when it acts on suitable Banach spaces. Existence of a spectral gap can

be interpreted as the statement that apart from functions which are densities of absolutely continuous invariant measures (and hence are fixed by $\mathcal{P}_T$), the transfer operator acts as a contraction on a certain space of functions; the mechanism driving this contractive property is the fact that T expands distances on the phase space $[0, 1]$. We note that the action of $\mathcal{P}_T$ on L^1 satisfies

$$
\begin{aligned}
\|\mathcal{P}_T\varphi\|_1 &= \sup\left\{ \int (\mathcal{P}_T\varphi) \cdot \psi \, dx \,\big|\, \psi \in L^\infty, \|\psi\|_\infty \le 1\right\} \\
&= \sup\left\{ \int \varphi \cdot (\psi \circ T) \, dx \,\big|\, \psi \in L^\infty, \|\psi\|_\infty \le 1\right\} \\
&\le \|\varphi\|_1.
\end{aligned}
\tag{1}
$$

In fact, (1) holds for any measurable transformation T that is *non-singular* – that is, T does not map a set of positive Lebesgue measure into a set of zero measure. Non-singular maps are precisely those maps for which every $\psi \in L^\infty$ has $\|\psi \circ T\|_\infty \le \|\psi\|_\infty$. In other words, non-singularity of T implies that the Koopman operator does not expand distances in L^∞, which in turn implies that the transfer operator does not expand distances in L^1. However, (1) is not enough to deduce any information on decay of correlations for T, because the contraction is not strict.

In fact, (1) does not even let us deduce the existence of an absolutely continuous invariant measure. How might we hope to find such a measure? Recall the proof of the Krylov–Bogolyubov theorem, which establishes the existence of an invariant measure for a continuous map on a compact metric space (though there is no mention of absolute continuity): one starts with a measure μ that is not necessarily invariant, and then considers the sequence of Cesàro averages $\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} \mu \circ T^{-k}$. Any limit point of this sequence is an invariant measure, and compactness of the space of measures shows that such limit points exist.

In our setting we want an *absolutely continuous* invariant measure, which means we should play the same game on the set of density functions: starting with the constant function $\mathbf{1}$, representing Lebesgue measure, we may consider the sequence

$$
\varphi_n = \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{P}_T^k \mathbf{1}.
\tag{2}
$$

If $\varphi_{n_j} \to \varphi \in L^1$, then $d\mu = \varphi \, dx$ defines an invariant measure μ, which is an acip. (Note that $\int \varphi_n \, dx = 1$ and $\varphi_n \ge 0$ for all n.) But how do we obtain a

2

convergent subsequence? Thanks to (1) we know that every φ_n is contained in the unit ball in L^1 – but this ball is not compact.

The solution is to consider an auxiliary Banach space $\mathcal{B} \subset L^1$ such that the unit ball of $\mathcal{B}$ is relatively compact in L^1. If $\mathcal{B}$ can be chosen such that the sequence φ_n is uniformly bounded in the $\mathcal{B}$-norm, then relative compactness implies the existence of a subsequence that converges (in L^1) to some $\varphi \in L^1$, which is the desired density. (Indeed, it is often the case that $\varphi \in \mathcal{B}$.)

For the doubling map, which we studied last time, the appropriate Banach space to use was the space of Lipschitz functions, whose unit ball embeds compactly into L^1 by the Arzelà–Ascoli theorem. However, this choice does not fare so well for general piecewise expanding interval maps.

Say that the map T is *full-branched* if $T(J_i) = [0, 1]$ for each basic interval J_i. If T is not full-branched (such as the map in Figure 1), then one can choose points x_1, x_2 that are arbitrarily close together but have different numbers of pre-images, and so in particular the quantities $\sum_{y \in T^{-1}(x_j)} |T'(y)|^{-1}$ for $j = 1, 2$ do not approach each other as $x_1 \to x_2$. This means that $\mathcal{P}_T \mathbf{1}$ has a discontinuity at the endpoints of a non-full branch of T, and so the space of continuous functions is not $\mathcal{P}_T$-invariant.

In a future talk we will see how this problem can be remedied if the map T has a Markov structure, but for the time being we deal with the situation by replacing the space of Lipschitz functions with a different space, which *is* invariant under the action of $\mathcal{P}_T$.

2 Functions of bounded variation

To this end, we recall that the total variation of a function $\varphi \colon [0, 1] \to \mathbb{C}$ is

$$|\varphi|_{BV} = \sup \left\{ \sum_{k=1}^n |\varphi(x_k) - \varphi(x_{k-1})| \,\Big|\, 0 = x_0 < x_1 < \cdots < x_n = 1 \right\}. \quad (3)$$

A function φ has bounded variation if $|\varphi|_{BV} < \infty$, and we denote by BV the vector space of such functions. A useful example to keep in mind is the following: Given any $\alpha \geq 0$, the function $\varphi_\alpha(x) = x^\alpha \sin(1/x)$ is defined on $(0, 1]$ and can be extended to $[0, 1]$ by $\varphi_\alpha(0) = 0$. It has bounded variation if and only if $\alpha > 1$.

Remark 1 *A bounded variation function is continuous except perhaps on a countable set of jump discontinuities, and differentiable Lebesgue-a.e. (Think*

of the examples just mentioned – the function φ_α is continuous at 0 as long as $\alpha > 0$, and is differentiable at 0 precisely when $\alpha > 1$, that is, when it is of bounded variation.)

The total variation as defined in (3) is a semi-norm on BV. We want to think of BV as a subspace of L^1, but we must be careful to remember that elements of L^1 are *equivalence classes* of functions (mod zero w.r.t. Lebesgue measure), and note that the quantity in (3) depends on which representative of the equivalence class we choose. Thus to define $|\cdot|_{BV}$ on L^1 we put (abusing notation slightly)

$$|\varphi|_{BV} = \inf\{|\hat{\varphi}|_{BV} \mid \varphi = \hat{\varphi} \text{ Lebesgue-a.e.}\}. \tag{4}$$

An alternate approach that allows us to avoid this step is to define the BV-semi-norm through integration: it can be shown that (3) is equivalent to

$$|\varphi|_{BV} = \sup\left\{\left|\int_{[0,1]} \varphi \cdot g' \, dx\right| \,\Big|\, g \in \mathcal{G}\right\}, \tag{5}$$

where $\mathcal{G} = \{g \in C^1([0,1], \mathbb{C}) \mid \|g\|_\infty \leq 1, g(0) = g(1) = 0\}$. The idea behind this equivalence is the following.

- When φ is differentiable, (3) is equivalent to $|\varphi|_{BV} = \int_{[0,1]} |\varphi'| \, dx$.

- Choosing $g \in \mathcal{G}$ such that $\varphi' \cdot g \approx |\varphi'|$, one gets $\int |\varphi'| \, dx \approx \varphi' \cdot g \, dx$.

- Integrating by parts yields the expression in (5).

Although the expression (5) does not make the heuristic interpretation of "total variation" as obvious as (3) does, it nevertheless has two important advantages over that definition:

1. it does not depend on the choice of representative function in an equivalence class of L^1, and so allows us to define $|\cdot|_{BV}$ on L^1 without an extra step along the lines of (4);

2. it generalises more readily to functions on higher-dimensional domains.

As with the Lipschitz semi-norm that we used last time for the doubling map, we can define a BV-norm by adding the L^1-norm to the BV-semi-norm:

$$\|\varphi\|_{BV} = \|\varphi\|_1 + |\varphi|_{BV}.$$

The space of BV functions is appropriate for us to study because its unit ball is relatively compact in L^1 – this is Helly's selection theorem, which states that if $\varphi_n \in BV$ is such that $\|\varphi_n\|_{BV}$ is uniformly bounded, then there is $\varphi \in BV$ such that $\varphi_{n_j} \xrightarrow{L^1} \varphi$ for some subsequence n_j.

In particular, if we can show that the sequence φ_n defined in (2) is uniformly bounded in the BV norm, then Helly's theorem will yield a BV limit point φ, and the measure μ defined by $d\mu = \varphi\, dx$ will be an acip for T.

3   A Lasota–Yorke inequality

In order to proceed further, we must investigate the properties of the transfer operator $\mathcal{P}_T$ with respect to the BV norm. Along the way we will see that BV is invariant under $\mathcal{P}_T$. We give an argument using the definition (5) to derive a bound that was first given by A. Lasota and J. Yorke in a 1974 paper – the argument there is equivalent to the one here, but uses the definition (3).

Given a function $g \in \mathcal{G}$, we need to estimate $\int (\mathcal{P}_T\varphi) \cdot g'\, dx$. To this end we recall that by the definition of the transfer operator, we have

$$\int (\mathcal{P}_T\varphi) \cdot g'\, dx = \int \varphi \cdot (g' \circ T)\, dx = \int \varphi \cdot (g \circ T)' \cdot (T')^{-1}\, dx,$$

where the second equality is valid because T is differentiable at all but finitely many points. Recalling the definition (5), this gives

$$|\mathcal{P}_T\varphi|_{BV} \leq \sup \left\{ \left| \int \varphi \cdot (g \circ T)' \cdot (T')^{-1}\, dx \right| \, \Big| \, g \in \mathcal{G} \right\}. \tag{6}$$

It is tempting to try and use the bound $|T'(x)| \geq \lambda$ to conclude that this quantity is $\leq \lambda^{-1} \sup \left\{ \left| \int \varphi \cdot (g \circ T)'\, dx \right| \, \big| \, g \in \mathcal{G} \right\}$, but we must take care – the argument of the integrand may vary, and so we cannot proceed quite so directly. Rather, we use the identity

$$\frac{d}{dx}\left(\frac{g \circ T}{T'} \right) = (g \circ T)'(T')^{-1} - (g \circ T)\frac{T''}{(T')^2}$$

to obtain

$$\left| \int \varphi \cdot (g \circ T)' \cdot (T')^{-1}\, dx \right| \leq \left| \int \varphi \left(\frac{g \circ T}{T'} \right)'\, dx \right| + \int |\varphi| \cdot |g \circ T| \cdot \frac{|T''|}{|T'|^2}\, dx$$

$$\leq \lambda^{-1} \left| \int \varphi \tilde{g}'\, dx \right| + K\|\varphi\|_1,$$

where $\tilde{g} = \lambda \frac{g \circ T}{T'}$ has $\|g\|_\infty \leq 1$ and $K = \max(|T''|/|T'|^2)$. (Note that it is at this point that we use the hypothesis that T is C^2 – elsewhere only C^1 is used.)

If the map T were differentiable on the entire interval $[0, 1]$ and fixed the endpoints, then we would have $\tilde{g} \in \mathcal{G}$ and so (6) would immediately imply $|\mathcal{P}_T \varphi|_{BV} \leq \lambda^{-1}|\varphi|_{BV} + K\|\varphi\|_1$. Unfortunately, as shown in Figure 2, $\tilde{g}$ is discontinuous at each of the discontinuity points of T, and moreover does not vanish at the endpoints of $[0, 1]$ if those endpoints are not fixed by T. Thus we must be more careful.
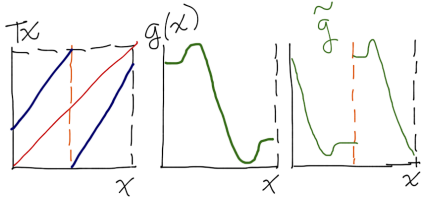

Figure 2: $\tilde{g}$ may not be in $\mathcal{G}$.

The idea is to approximate $\tilde{g}$ with functions from $\mathcal{G}$, as shown in Figure 3. Let $0 = b_0 < b_1 < \cdots < b_n = 1$ be the endpoints of the intervals on which the map T is C^2. Given $\varepsilon > 0$, let $h\colon [0, 1] \to \mathbb{C}$ be a continuous function such that $h(0) = h(1) = 0$, $h(x) = \tilde{g}(x)$ when $|x - b_k| \geq \varepsilon$ for each k, and h is linear on $B(b_k, \varepsilon)$.
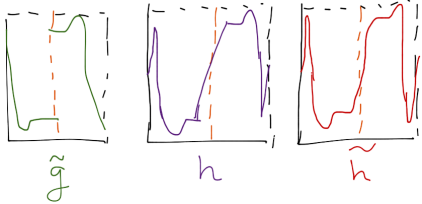

Figure 3: Approximating $\tilde{g}$ with elements of $\mathcal{G}$.

Finally, let $\tilde{h} \in \mathcal{G}$ be close to h in the uniform metric and agree with h except on an ε^2-neighbourhood of each point where h is non-differentiable.

6

We get

$$\int \varphi \cdot \tilde{g}' \, dx \leq \int \varphi \cdot \tilde{h}' \, dx + \int \varphi \cdot |\tilde{h}' - \tilde{g}'| \, dx$$

$$\leq |\varphi|_{BV} + \sum_{k=0}^{n} \left(\int_{B(b_k,\varepsilon)} \varphi \cdot |\tilde{h}'| \, dx \right. \tag{7}$$

$$\left. + \int_{B(b_k,\varepsilon)} \varphi \cdot |\tilde{g}'| \, dx \right).$$

The second integral in the sum goes to 0 as $\varepsilon \to 0$. (This uses the assumption that $T' \in L^1$.) For the first integral, we use the fact that $h' = \frac{1}{2\varepsilon}(\tilde{g}(b_k + \varepsilon) - \tilde{g}(b_k - \varepsilon))$ to conclude that as $\varepsilon \to 0$, the integral goes to $\varphi(b_k)|\tilde{g}(b_k^+) - \tilde{g}(b_k^-)|$, where $\varphi(b_k)$ is understood as $\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{B(b_k,\varepsilon)} \varphi \, dx$, so that in particular we choose the representative of the L^1-equivalence class that minimises the total variation, as in (4).

Since $\|g\|_\infty \leq 1$ and $g(0) = g(1) = 0$, we conclude that

$$\sum_{k=0}^{n} \int_{B(b_k,\varepsilon)} \varphi \cdot |\tilde{h}'| \, dx \leq \sum_{k=1}^{n} |\varphi(b_{k-1})| + |\varphi(b_k)|. \tag{8}$$

We can bound this sum in terms of $|\varphi|_{BV}$ and $\|\varphi\|_1$. Let $m_k = \inf_{x \in [b_{k-1}, b_k]} |\varphi(x)|$, then

$$|\varphi(b_{k-1})| + |\varphi(b_k)| \leq 2m_k + \left|\varphi|_{[b_{k-1},b_k]}\right|_{BV},$$
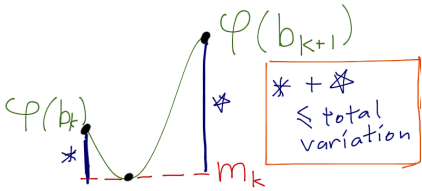
as suggested by Figure 4.

Figure 4: Bounding $|\varphi(b_{k-1})| + |\varphi(b_k)|$.

Moreover, $\int_{[b_{k-1},b_k]} |\varphi| \, dx \geq m_k(b_k - b_{k-1}) \geq m_k \Delta$, where $\Delta = \min_k(b_k - b_{k-1})$, and so we can sum over k to get

$$\sum_{k=1}^{n} |\varphi(b_{k-1})| + |\varphi(b_k)| \leq 2\Delta^{-1} \|\varphi\|_1 + |\varphi|_{BV}.$$

7

Together with (7) and (8), this gives

$$\int \varphi \cdot \tilde{g}' \, dx \leq 2|\varphi|_{BV} + 2\Delta^{-1}\|\varphi\|_1,$$

so that (6) and the discussion following it gives us

$$|\mathcal{P}_T\varphi|_{BV} \leq 2\lambda^{-1}|\varphi|_{BV} + (2\Delta^{-1} + K)\|\varphi\|_1.$$

In terms of the BV norm we have

$$\|\mathcal{P}_T\varphi\|_{BV} \leq 2\lambda^{-1}\|\varphi\|_{BV} + (2\Delta^{-1} + K + 1)\|\varphi\|_1;$$

using the assumption that $\lambda > 2$, we can write this in the form

$$\|\mathcal{P}_T\varphi\|_{BV} \leq r\|\varphi\|_{BV} + R\|\varphi\|_1 \tag{9}$$

for $r \in (0,1)$ and $R > 0$. This is a *Lasota–Yorke inequality*, and turns out to have important implications for the statistical properties of the map T.

4 Existence of an acip

Now we can return to the sequence φ_n defined in (2) as $\frac{1}{n}\sum_{k=0}^{n-1}\mathcal{P}_T^k\mathbf{1}$, and show that it is uniformly bounded in BV. Indeed, iterating the Lasota–Yorke inequality (9) gives

$$\begin{aligned}\|\mathcal{P}_T^2\varphi\|_{BV} &\leq r\|\mathcal{P}_T\varphi\|_{BV} + R\|\mathcal{P}_T\varphi\|_1 \\ &\leq r^2\|\varphi\|_{BV} + (1+r)R\|\varphi\|_1,\end{aligned}$$

where we use the inequality $\|\mathcal{P}_T\varphi\|_1 \leq \|\varphi\|_1$ from (1). Writing $\bar{R} = R(1 + r + r^2 + \cdots) = R(1-r)^{-1}$, we have by induction

$$\|\mathcal{P}_T^k\varphi\|_{BV} \leq r^k\|\varphi\|_{BV} + \bar{R}\|\varphi\|_1. \tag{10}$$

In particular, we conclude that the sequence φ_n is uniformly bounded in BV, since

$$\|\varphi_n\|_{BV} \leq r^n + \bar{R} \leq 1 + \bar{R}.$$

As discussed above, Helly's theorem shows that there is $\varphi \in BV$ such that $\varphi_{n_j} \xrightarrow{L^1} \varphi$ for some subsequence n_j, and the measure μ defined by $d\mu = \varphi \, dx$ is an acip for T.

Note that this proves the *existence* of an acip for T, but it does not prove *uniqueness*. For the doubling map there is only one acip, Lebesgue measure, but for other piecewise expanding interval maps there may be more than one. For example, the map shown in Figure 5 has two ergodic acips, one supported on $[0, 1/2]$ and the other supported on $[1/2, 1]$.
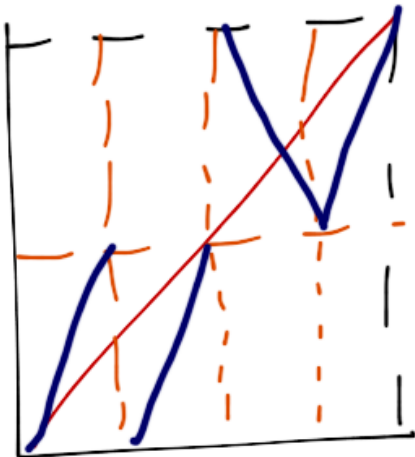

Figure 5: Non-uniqueness of an acip.

5   The spectrum of the transfer operator

The Lasota–Yorke inequality (9) also lets us deduce spectral information about $\mathcal{P}_T$. First we observe that by the spectral radius formula and the iterated inequality (10), the spectral radius of $\mathcal{P}_T \colon BV \to BV$ is bounded above by the inequality

$$\rho(\mathcal{P}_T) = \lim_{n \to \infty} \|\mathcal{P}_T^n\|^{1/n} \leq \lim_{n \to \infty} (r^n + \bar{R})^{1/n} = 1,$$

where we use the fact that $\|\varphi\|_1 \leq \|\varphi\|_{BV}$. The previous section shows that $1 \in \sigma(\mathcal{P}_T)$, and we conclude that $\rho(\mathcal{P}_T) = 1$. Now we would like to show that the Lasota–Yorke inequality also gives the existence of a spectral gap and exponential decay of correlations, and this will be the subject of next week's talk.

9