Approximation and concrete error bounds for acim density approximation for $C^{1+\alpha}$ expanding mappings.

Rua Murray*

August 13, 1996

Abstract

We consider the calculation of absolutely continuous invariant measures for $\mathcal{C}^{1+\alpha}$ piecewise onto, expanding maps of a Riemannian manifold. An approximation to this density can be computed by a well known Markov approximation, which we call the *area overlaps Markov chain*. We obtain concrete estimates on the accuracy of this density approximation by comparing an asymptotic formula for the measure of the subsets of a given partition with the invariant vector of the area overlaps Markov chain. The constants in the bounds can be calculated explicitly from columns of the associated transition matrix.

1 Introduction

Absolutely continuous invariant measures are frequently of interest in dynamical systems. However, work on how to approximate them accurately seems relatively scarce. A well known method [Li76, Fro95, DZ96] is to approximate the dynamical system in question by a Markov chain, and use the invariant probability vector to approximate the invariant density (see the above references for details). The idea here is that the transition matrix of the Markov chain (which we call the *area overlaps Markov chain*) can be thought of as a finite dimensional approximation to the Perron–Frobenius operator for the dynamical system in question.

In the case of expanding transformations, Ding and Zhou [DZ96] show that as the Markov approximation is refined, the approximate densities converge strongly to the unique density of the absolutely continuous invariant measure (acim). Their arguments involve careful estimates on how the Perron–Frobenius operator alters the *variation* of densities, and do not lead to bounds on the rate of convergence.

Our approach is different. Rather than characterising the invariant density as the unique fixed point of the Perron-Frobenius operator, we write the acim as the limit of the push–forward of Lebesgue measure. This furnishes us with an asymptotic formula for the invariant measure of any given set (Proposition 2).

*Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, UNITED KINGDOM. *Email:* R.Murray@statslab.cam.ac.uk

We describe how to use this in Section 3. In Section 4, we use a standard bounded distortion argument to show that the stochastic matrices appearing in Proposition 2 are uniformly close to the transition matrix for the area overlaps Markov chain. Finally, in Section 5, we put this together with an observation about the exponential decay of correlations to show how the invariant vector of the area overlaps Markov chain provides a uniformly good approximation to the invariant density. This is Theorem 7 and is the main result. In Appendix A, we describe how to compute the constants appearing in Theorem 7. The paper concludes in Section 6 with a numerical illustration of the result for a family of one dimensional maps.

2 The problem setup

Let $f : X \circlearrowleft$ be a continuous, piecewise–onto, finite to one, expanding self mapping of a compact Riemannian manifold (with metric ρ). The restrictions we put on f can be stated as follows. There exist constants $c > 0$, $\lambda > 1$ such that for each $n \in \mathcal{Z}^+$

$$\|D(f^n)(x)\, v\| \geq c\lambda^n \|v\|$$

for every tangent vector $v \in T_x X$, $x \in X$. We also require that the derivative map $Df : T_x X \to T_{f(x)} X$ is Hölder continuous. In particular, we require Hölder continuity of the Jacobian determinant: that is, there exists $0 < \alpha \leq 1$ and a constant K such that

$$|\det(Df(x)) - \det(Df(y))| \leq K\rho(x,y)^\alpha$$

for all $x, y \in X$.

In fact, it will be more useful for us to rewrite this another way. The expanding property of the Jacobian implies that each branch of the inverse map f^{-n} is exponentially contracting. Renaming c and λ (if necessary), we write

$$c\lambda^n \rho(f_a^{-n}(x), f_a^{-n}(y)) \leq \rho(x,y) \tag{1}$$

for every pair of points $x, y \in X$ and branch f_a^{-n} of f^{-n}. Moreover, the facts that $\det(Df)$ is Hölder continuous and bounded below imply that $\log|\det(Df)|$ is also Hölder continuous. That is, there exists a constant K_α such that

$$\Big| \log|\det\, Df(x)| - \log|\det\, Df(y)| \Big| \leq K_\alpha \rho(x,y)^\alpha \tag{2}$$

for all $x, y \in X$.

Now, X is equipped with a normalised Lebesgue measure (corresponding to the Riemannian metric ρ) which we denote by $|\cdot|$. We seek to approximate μ, an absolutely continuous invariant measure (*acim*)—with respect to Lebesgue—for the dynamical system (X, f).

Theorem 1 (Folklore) *For f as above, there exists an acim μ. Moreover, for each Borel set A,*

$$\mu(A) = \frac{1}{n}\sum_{k=0}^{n-1} |f^{-k}(A)|.$$

Remark 2.1. The suggestion that this characterisation of the acim might be of practical use arose from a discussion with Sebastian van Strien. □

3 Measures and probability vectors

First of all, choose a partition $\xi = \{X_1, \ldots, X_m\}$ of the manifold X. We define the *diameter* of the partition to be the largest diameter of any element of ξ. That is

$$\gamma = \mathrm{diam}(\xi) = \max_{i=1\ldots m} \sup_{x,y \in X_i} \rho(x,y). \tag{3}$$

Without loss of generality, assume that each element of the partition is a connected set. Later on, the diameter of a (general) partition element will refer to the maximal diameter of a connected component. For technical reasons, we will add a non–degeneracy assumption on the partition. The partition ξ will be called *non–degenerate* if for every $A \in \xi$

$$|A \setminus \partial A| > 0. \tag{4}$$

Every probability measure ν will be represented by a probability vector $(\nu_1, \ldots, \nu_m)$ where each $\nu_i = \nu(X_i)$. Consequently, our representation of the measure ν is based on the choice of partition ξ, and the only sets from the full σ–algebra on X whose measure we can recover exactly are those that can be written exactly as unions of elements of ξ. We assume that for any practical purpose, explicit knowledge of the invariant measure of all the elements of some such partition is sufficient.

We are going to show that the vector (μ_i) corresponding to the acim μ arises as the limit of a certain asymptotic product of stochastic matrices. We will then show how a uniformly good approximation to this limit can be obtained by the invariant vector of the area overlaps Markov chain studied extensively by Li [Li76], Froyland [Fro95], Ding and Zhou [DZ96] and others. This current work is different in character to that reported in the above papers in that we investigate the closeness of the actual invariant measure to the approximation obtained for a given partition, rather than proving that the approximations converge strongly as the diameter γ of the partition is refined to zero.

Let $\mathcal{R}^m_{prob} = \{x \in R^m : \sum_{i=1}^m x_i = 1 \text{ and } x_i \geq 0 \, \forall i = 1 \ldots m\}$ denote the set of all probability vectors in R^m. Let the partition ξ be fixed, and for each $n > 0$, $i, j = 1 \ldots m$ put

$$P^{(n)}_{ij} = \frac{|f^{-n}X_j \cap f^{-(n-1)}X_i|}{|f^{-(n-1)}X_i|},$$

unless $|f^{-n}X_j \cap f^{-(n-1)}X_i| = 0$ in which case we put $P^{(n)}_{ij} = 0$. Let $P^{(n)}$ be the matrix with entries $P^{(n)}_{ij}$. By summing over j for each i, it follows immediately that each $P^{(n)}$ is a stochastic matrix. Consequently each $P^{(n)} : R^m_{prob} \circlearrowleft$ by left multiplication $x \mapsto xP^{(n)}$.

For each $n \geq 0$, let $x_n \in R^m_{prob}$ be defined component-wise by

$$(x_n)_i = |f^{-n}X_i|.$$

It is an easy exercise to check that $x_n = x_{n-1}P^{(n)}$. Moreover, it is an immediate corollary of the folklore theorem above that

$$\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} x_n$$

where $\mu \in R^m_{prob}$ is the probability vector with components $\mu(X_i)$. We state this in the form of a proposition.

Proposition 2 *For f, X, ξ, μ as above*

$$\mu(X_i) = \Big(\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} x_0 P^{(1)} \dots P^{(k)}\Big)_i.$$

We are interested in approximating this asymptotic form. We will do this by proving uniform bounds on the growth of the coefficients of the matrices $P^{(n)}$.

4 Growth of the matrices $P^{(n)}$

We seek to control the growth of the elements of $P^{(n)}$ as n increases. This section follows the standard bounded distortion argument used in the proof of the folklore theorem. We use the regularity of f and exploit the following key inequality:

Proposition 3 *Let X be a Riemannian manifold with metric ρ and corresponding normalised Lebesgue measure $|\cdot|$. Let $B \subset A \subset X$. Let $f : X \circlearrowleft$, and let α, K_α be as in equation (2). Then*

$$e^{-K_\alpha (diam(f^{-1}A))^\alpha} \frac{|B|}{|A|} \le \frac{|f^{-1}B|}{|f^{-1}A|} \le e^{K_\alpha (diam(f^{-1}A))^\alpha} \frac{|B|}{|A|}.$$

Note that here $diam(A)$ refers to the largest diameter of a connected component of A.

Proof: We prove only the second inequality, since the first is similar. Assume that A is connected; the proof extends easily to the general case. Next, let $\{C_i\}_{i=1}^N$ be the partition of X into components such that $f|_{C_i} : C_i \to X$ is *onto* for each $i = 1 \dots N$. Consider the component $f^{-1}(A) \cap C_i$. From the change of variables formula for integration it follows that

$$\inf_{x\in f^{-1}A\cap C_i} |\det(Df(x))|\, |f^{-1}B \cap C_i| \le |B| \le \sup_{x\in f^{-1}A\cap C_i} |\det(Df(x))|\, |f^{-1}B\cap C_i|$$

for every subset $B \subset A$ (including A). A simple lemma gives us the necessary bound on the distortion of f.

Lemma 3.1. *If $g : X \to \mathcal{R}^+$ is a function satisfying*

$$|\log g(x) - \log g(y)| \le K_\alpha \rho(x, y)^\alpha$$

for some constants $K_\alpha > 0$ and $0 < \alpha \le 1$, then

$$e^{-K_\alpha \rho(x,y)^\alpha} \le \frac{g(x)}{g(y)} \le e^{K_\alpha \rho(x,y)^\alpha}.$$

Hence,

$$\frac{|f^{-1}B \cap C_i|}{|f^{-1}A \cap C_i|} \le \frac{\sup_{x\in f^{-1}A\cap C_i} |\det(Df(x))|}{\inf_{x\in f^{-1}A\cap C_i} |\det(Df(x))|} \frac{|B|}{|A|} \le e^{K_\alpha(\mathrm{diam}(f^{-1}A\cap C_i))^\alpha} \frac{|B|}{|A|},$$

by Lemma 3.1, and the definition in equation (2). But $\mathrm{diam}(f^{-1}A \cap C_i) \le \mathrm{diam}(f^{-1}A)$, so that on each component C_i

$$\frac{|f^{-1}B \cap C_i|}{|f^{-1}A \cap C_i|} \le e^{K_\alpha(\mathrm{diam}(f^{-1}A))^\alpha} \frac{|B|}{|A|}.$$

The full result follows since

$$\frac{|f^{-1}B|}{|f^{-1}A|} = \frac{\sum_{i=1}^N |f^{-1}B \cap C_i|}{\sum_{i=1}^N |f^{-1}A \cap C_i|} \le \max_{i=1\ldots N} \frac{|f^{-1}B \cap C_i|}{|f^{-1}A \cap C_i|}$$

(and the same is true for all connected components). □

Corollary 4 *For each $n > 0$, $i, j = 1 \ldots m$,*

$$e^{-c^{-\alpha}\lambda^{-n\alpha}K_\alpha\gamma^\alpha} \le \frac{P_{ij}^{(n+1)}}{P_{ij}^{(n)}} \le e^{c^{-\alpha}\lambda^{-n\alpha}K_\alpha\gamma^\alpha}.$$

The constants $c, \lambda, \alpha, K_\alpha$ are as in equations (1) and (2) and γ is the diameter of the partition ξ, as defined in (3).

Proof: We apply Proposition 3 to each component of $f^{-(n-1)}X_i$. Observe that each element of the partition ξ has diameter at most γ. Consequently, each connected component of the partition $f^{-n}\xi$ has diameter at most $c^{-1}\lambda^{-n}\gamma$ by equation (1). Now, each

$$(f^{-n}X_j \cap f^{-(n-1)}X_i) \subset f^{-(n-1)}X_i,$$

so we can put $A = f^{-(n-1)}X_i$ and $B = f^{-n}X_j \cap f^{-(n-1)}X_i$ to obtain $B \subset A$. Moreover, $P_{ij}^{(n+1)} = \frac{|f^{-1}B|}{|f^{-1}A|}$ and $P_{ij}^{(n)} = \frac{|B|}{|A|}$, so that Proposition 3 implies the result. □

Finally, we want to compare the matrices $P^{(n)}$ to the matrix $P = P^{(1)}$.

Theorem 5 *In the notation established so far, we have for each $n > 0$, $i, j = 1 \ldots m$,*

$$e^{-\frac{K_\alpha\gamma^\alpha}{c^\alpha(\lambda^\alpha - 1)}} \le \frac{P_{ij}^{(n)}}{P_{ij}} \le e^{\frac{K_\alpha\gamma^\alpha}{c^\alpha(\lambda^\alpha - 1)}}. \tag{5}$$

Proof: First of all, apply Corollary 4 to the product

$$\frac{P_{ij}^{(n)}}{P_{ij}} = \frac{P_{ij}^{(n)}}{P_{ij}^{(n-1)}} \cdots \frac{P_{ij}^{(2)}}{P_{ij}^{(1)}}.$$

The result follows since $\sum_{k=1}^{n-1} \lambda^{-\alpha k} \leq \frac{1}{\lambda^{\alpha}-1}$. □

Now, notice that by choosing the initial partition to be sufficiently fine (γ sufficiently small), the exponential in inequality (5) can be made arbitrarily close to 1. Let us recap what this means. The invariant measure (the vector with components $\mu(X_i)$) arises as the limit of a sequence of iteratively multiplied stochastic matrices. We have a uniform bound on the elements of these matrices, so the problem of approximating the acim now consists in estimating what that asymptotic limit might be. In the next section, we exploit inequality (5) to compare the limit of the asymptotic product with the invariant vector of P.

Remark 4.1. In the case where the map f instead of being $C^{1+\alpha}$ is a piecewise linear Markov map, and the partition ξ is a Markov partition such that f is linear on each $A \in \xi$, then the above argument shows that each $P^{(n)} = P$. To see this, note that Hölder continuity of $\det(Df)$ is used only to compare the value of $\det(Df)$ at points which lie within a connected component of elements of $f^{-n}\xi$. Since f is linear on all such components, one can assume that the Hölder constant $K_\alpha = 0$. Corollary 4 now implies that $P^{(n)} = P^{(n-1)} = \ldots = P^{(1)} = P$. Thus, the components of the acim vector are precisely the invariant vector for P. □

Remark 4.2. The problem of the asymptotic matrix product we describe is similar to a situation considered by Diamond et. al. [DKP95]. They define an *interval stochastic matrix* to be the collection of all stochastic matrices Q which satisfy $A_{ij} \leq Q_{ij} \leq B_{ij}$ for some fixed positive matrices A and B (we could take these to be $e^{\pm \frac{K_\alpha \gamma^\alpha}{c^\alpha(\lambda^\alpha - 1)}} P$). They then characterise the set of invariant vectors for stochastic matrices in the collection by a series of inequalities. Unfortunately, it is not obvious how to turn their combinatorial relations into concrete bounds on the diameter of the set of invariant vectors. Neither is it clear whether their results help characterise probability vectors which arise as limits of products of a countable family of matrices in the collection (our situation), or whether it is only useful for vectors which are limits of asymptotic products by just one matrix from the collection (their situation). Their methods are not based on asymptotic analysis, but may offer a promising direction. □

5 Area overlaps Markov chain

The Markov chain with transition probabilities given by the matrix $P = P^{(1)}$ will be called the *area overlaps Markov chain*. We give it this name because P_{ij} is the proportion of mass from the set B_i which overlaps the set B_j under one iteration of f. Our main result is that provided the matrix P has a unique invariant probability vector π, then π can be shown to be within a certain distance of the probability vector corresponding to the acim.

Proposition 6 *Let f be expanding in the sense described, and let ξ be a non–degenerate partition of X. The matrix $P = P^{(1)}$ with components*

$$P_{ij} = \frac{|f^{-1}X_j \cap X_i|}{|X_i|}$$

has a unique invariant probability vector.

Proof: First of all, we show that if $A, B \in \xi$ are such that $|A \cap f^{-n}B| > 0$ for some n, then $(P^n)_{ij} > 0$ (where $A = X_i$ and $B = X_j$). We then complete the proof by observing that there always exists such an n.

The partition $\vee_{i=0}^{n} f^{-i}\xi$ partitions $A \cap f^{-n}B$ into finitely many components. Since $|A \cap f^{-n}B| > 0$, there exists at least one such component C such that $|C| > 0$. Since the map f is expanding, $|f^k C| > 0$ for each integer k, and since $C \in \vee_{i=0}^{n} f^{-i}\xi$, $f^k C \subset X_{i_k} \in \xi$ for some i_k for each $0 \le k \le n$. Note that $i_0 = i, i_n = j$ and $f^k C \subset f^{-1}X_{i_{k+1}} \cap X_{i_k}$ for $0 \le k < n$. Consequently,

$$|f^{-1}X_{i_{k+1}} \cap X_{i_k}| \ge |f^k C| > 0,$$

so that $P_{i_k i_{k+1}} > 0$ for $0 \le k < n$. Hence

$$(P^n)_{ij} \ge P_{ii_1} \ldots P_{i_{n-1}j} > 0,$$

as required.

Next, since ξ is a non–degenerate partition, each X_i contains an open ball (this follows from equation (4)). Let $\epsilon > 0$ be small enough that each X_i contains a ball of radius ϵ, and let n be large enough that $c\lambda^n\epsilon > diam(X)$. Expansivity implies that $f^n A = X$ for each $A \in \xi$, and hence that $|A \cap f^{-n}B| > 0$ for each $A, B \in \xi$.

These two parts imply that $(P^m)_{ij} > 0$ for every i, j and $m > n$. Thus, the matrix P is ergodic, so has a unique invariant probability vector. $\square$

Theorem 7 *Let $f, c, \lambda, \alpha, K_\alpha, \mu, \xi, \gamma, P$ be as above. Suppose that π is the unique invariant probability vector for P. Then there exists a constant $C_P > 0$ such that*

$$\|\pi - \mu\|_1 \le C_P(\Lambda - 1)$$

where $\Lambda = e^{\frac{K_\alpha \gamma^\alpha}{c^\alpha(\lambda^\alpha - 1)}}$, and μ is the vector with components $\mu(X_i)$. The constant C_P can be calculated from the columns of P.

Proof: Throughout the proof, P^k denotes the kth power of $P = P^{(1)}$ and should be distinguished from the matrix $P^{(k)}$. First of all, recall from Proposition 2 that the vector μ arises as

$$\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} x_0 P^{(1)} \ldots P^{(k)}.$$

Now, x_0 is the vector with components $|X_i|$, and we define inductively

$$x_n = x_{n-1} P^{(n)}.$$

We will estimate

$$\|e_n\|_1 = \left\| \frac{1}{n} \sum_{k=0}^{n-1} x_k - \pi \right\|_1.$$

Write

$$\begin{aligned} x_k - \pi &= \sum_{j=0}^{k-1} (x_{k-j} P^j - x_{k-j-1} P^{j+1}) + x_0 P^k - \pi \\ &= \sum_{j=0}^{k-1} ((x_{k-j} - x_{k-j-1} P) P^j) + x_0 P^k - \pi. \end{aligned}$$

Then (since π is an invariant vector for P)

$$\begin{aligned} \|e_n\|_1 &= \left\| \frac{1}{n} \sum_{k=1}^{n-1} \sum_{j=0}^{n-1-k} (x_k - x_{k-1} P) P^j + \frac{1}{n} \sum_{k=0}^{n-1} (x_0 - \pi) P^k \right\|_1 \\ &\leq \frac{1}{n} \sum_{k=1}^{n-1} \sum_{j=0}^{\infty} \|(x_k - x_{k-1} P) P^j\|_1 + \frac{1}{n} \sum_{k=0}^{\infty} \|(x_0 - \pi) P^k\|_1. \qquad (6) \end{aligned}$$

We now estimate bounds for the asymptotic sums.

Lemma 7.1. *There exist constants $C > 0$ and $0 < \beta < 1$ such that for any two probability vectors y, z, $n > 0$,*

$$\|(y - z) P^n\|_1 \leq C \, \beta^n \, \|y - z\|_1.$$

Proof of lemma 1: Since the stochastic matrix P has a unique invariant probability vector (by Proposition 6), it has exponential convergence to equilibrium. The constant β is just the modulus of the second largest eigenvalue of P, and C is a constant depending only the columns of P. □

Lemma 7.2. *For each n, we have*

$$\|x_n - x_{n-1} P\|_1 \leq \Lambda - 1.$$

Proof of lemma 2: Note that $x_n = x_{n-1} P^{(n)}$. By Theorem 5,

$$\frac{1}{\Lambda} P_{ij} \leq P_{ij}^{(n)} \leq \Lambda P_{ij}.$$

Consequently,

$$-(\Lambda - 1) P_{ij} \leq (\frac{1}{\Lambda} - 1) P_{ij} \leq P_{ij}^{(n)} - P_{ij} \leq (\Lambda - 1) P_{ij},$$

so that

$$|(x_{n-1})_i P_{ij}^{(n)} - (x_{n-1})_i P_{ij}| \leq (x_{n-1})_i P_{ij} (\Lambda - 1).$$

The Lemma follows by summing over i and j. □

We can complete the proof of Theorem 7 by observing that the rhs of (6) can be bounded by

$$\frac{1}{n}\sum_{k=1}^{n-1}\sum_{j=0}^{\infty} C\beta^{n}(\Lambda-1)+\frac{1}{n}\sum_{j=0}^{\infty} C\beta^{j}\|x-z\|_1 \leq \frac{C(\Lambda-1)}{1-\beta}+O(\frac{1}{n}).$$

The Theorem follows by letting $n \to \infty$. □

The strength of Theorem 7 depends on the constant C_P. While it is not *a priori* clear that C_P does not grow too badly as the partition size m is increased, computational evidence is encouraging. In Appendix A, we briefly describe how to compute an upper bound for the constant C_P. Numerically, (for X one dimensional) C_P seems to scale like $\log m$, where m is the size of the partition. If this turns out to be a generally occuring feature of the matrices P, then Theorem 7 is useful. This is because once m is sufficiently large, the $\Lambda - 1$ part of the error bound will scale like $1/m$ (for suitably chosen partitions of a one dimensional phase space—like $1/m^{1/d}$ for a d–dimensional phase space). Together with the numerical results for the constant C_P, this suggests a general scaling law for error of $O(\log(m)/m)$ for a partition of size m. Consequently, the area overlaps Markov chain is a good way to approximate the acim for this class of maps.

This discussion is a first approach to analysing the asymptotics of the matrix product $P^{(1)}P^{(2)}P^{(3)}\ldots$. It is hoped that a more "dynamical" method of analysis can be developed which does not rely on simply comparing the exact asymptotic product with the sequence $P^{(1)}P^{(1)}P^{(1)}\ldots$.

6 A numerical example

For each $l \in [2,\infty)$, let $f_l : [0,1] \circlearrowleft$ be given by

$$f_l(x) = 1 - |2x-1|^l.$$

These maps are one of the simplest one–parameter families studied by van Strien and others [NvS91, Now93, dMvS93]. In these references, it is shown (among other things) that each such map supports a unique ergodic acim. For the value $l = 2$, the map f_2 is the well known fully developed logistic map. The other maps in this family are also piecewise onto unimodal maps, and also have a critical point at $\frac{1}{2}$ of lth order. Consequently, the maps have neither expansion, nor bounded distortion. However, experimentation shows that a certain conjugate family does have these properties. For each l, put

$$h_l(x) = \frac{1 + x^{1/l} - (1-x)^{1/l}}{2} \quad \text{and} \quad g_l(x) = h_l \circ f_l \circ h_l^{-1}(x).$$

Each conjugacy is smooth (except its inverse loses differentiability at the endpoints) and has the effect of "straightening out" the singularity at $x = \frac{1}{2}$. For four selected values of l, the straightening resulting from the conjugacy is illustrated in Figure 1.

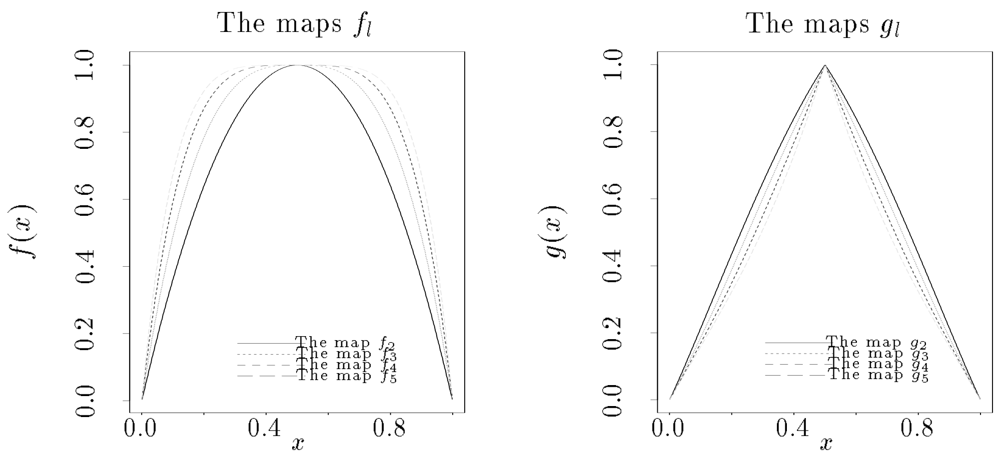

Figure 1: The maps f_l and their conjugates $g_l = h_l^{-1} \circ f_l \circ h_l$.

Each g_l has a unique acim (this can be seen by pulling back the acim for each f_l by the conjugacy h_l^{-1}), but its density is not known *a priori* (since no explicit formula is known for the acim of f_l unless $l = 2$). This makes $\{f_l\}_{l \geq 2}$ a good candidate family for numerical experimentation. We pick the values $l = 2, 3, 4, 5$. For $l = 2$, we can compare a numerical approximation of the invariant density explicitly with the pullback by h_2 of the invariant density for the logistic map; for the other values, we rely on the estimates provided by Theorem 7.

For each l, we estimate the constants λ and K_α by a fine grid search. (Recall that λ is the minimum modulus derivative of g_l, and—in this case—K_α is the Lipschitz constant for $\log |g_l'|$; that is, we have both c and λ equal to 1 in equations (1) and (2)). Next, we partition $[0, 1]$ into 1600 equal sized subintervals. This means that $\gamma = 1/1600$, where γ is the maximum diameter of any cell in the partition. Finally, we compute the invariant probability vector corresponding to the area overlaps Markov chain for this partition, and calculate the constant C_P as described in Appendix A. The density approximations are displayed in Figure 2, and the numerically determined constants in Table 1.

Theorem 7 implies that the last column of Table 1 provides an upper bound on the variational distance from our approximations to the exact invariant measures. For the record, we note that the distance from our computed approximation to the invariant density for g_2 is approximately 3.4×10^{-5}, substantially within the error bound in the first row of Table 1; we have no way to compare the distance for the other values of l.

In the discussion at the end of Section 5 the claim is made that the constants in the final column of Table 1 scale as $O(\frac{\log(m)}{m})$, where m is the number of cells in the partition. It is easy to see that since $\gamma = 1/m$, the $e^{\frac{K\gamma}{\lambda - 1}} - 1$ part scales like $1/m$, and numerical experiment suggests that the constant C_P—an upper bound on the decay of correlations for the matrix P—scales like $\log(m)$ (P is

l	λ	K	$e^{\frac{K\gamma}{\lambda-1}}-1$	C_P	$C_P(\Lambda-1)$
2.0	1.417	2.814	4.219×10^{-3}	22.07	0.0931
3.0	1.817	0.987	7.552×10^{-4}	23.44	0.0177
4.0	1.681	1.203	1.104×10^{-3}	26.19	0.0289
5.0	1.584	2.246	2.403×10^{-3}	28.91	0.0695

Table 1: Expansivity constants and acim approximation error bounds for certain maps g_l. The expansivity and distortion constants are computed by a fine numerical search (so are really just accurate estimates), and the constant C_P calculated as described in Appendix A. The final column gives a bound on $\|\pi-\mu\|_1$, where π is the computed approximation to the invariant μ.

an $m\times m$ matrix). For each $l=2,3,4,5$, Figure 3 shows how C_P scales as the partition is increased. The partitions used were $25,50,100,200,400,800,1600$ equal subintervals of $[0,1]$. I do not have a convincing explanation for this phenomenon, but the numerical facts seem to have intrinsic interest.

A Appendix: Bounding the constant C_P

In the proof of Theorem 7, the exponential convergence to equilibrium is used to prove bounds on

$$\left\|\frac{1}{n}\sum_{k=1}^{n-1}\sum_{j=0}^{n-1-k}(x_k-x_{k-1}P)P^j\right\|_1.$$

We now show how to numerically obtain an upper bound for these. Let $g_k = x_k - x_{k-1}P$, where P is an $m\times m$ stochastic matrix. Let f_i denote the usual basis for $\mathcal{R}^m$, and for each $1\le i<m$ put $\Gamma_i = f_i - f_m$. We make two observations. First of all, since each g_k is the difference of two probability vectors, it follows that

$$g_k = \sum_{i=1}^{m-1}(g_k)_i\Gamma_i.$$

Secondly, since each Γ_i is also the difference of two probability vectors, $\|\Gamma_i P^n\|\to 0$ exponentially fast as $n\to\infty$. Consequently, the number C defined as

$$C = \max_{i=1}^{m-1}\sum_{j=0}^{\infty}\|\Gamma_i P^j\|_1$$

is finite, and because of the exponential convergence of the series, it is very easy to compute accurately.

Putting all this together, we obtain

$$\left\|\sum_{j=0}^{n-1-k}(x_k-x_{k-1}P)P^j\right\|_1 = \left\|\sum_{j=0}^{n-1-k}\sum_{i=1}^{m-1}(g_k)_i\Gamma_i P^j\right\|_1$$

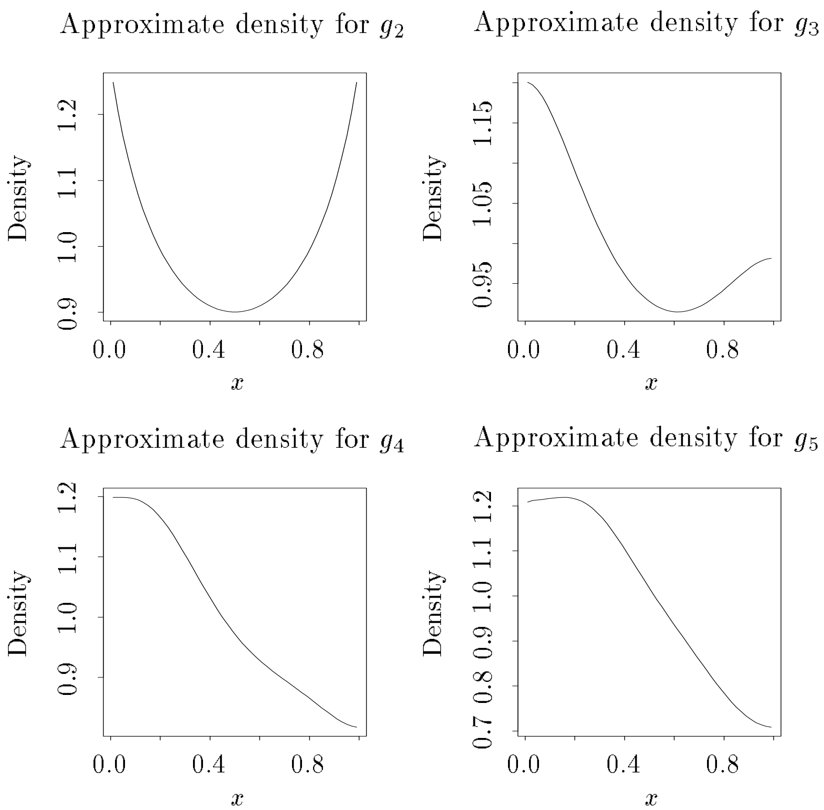

Figure 2: Density approximations for the maps $g_{2,3,4,5}$. The area overlaps method is used on a partition of 1600 equal cells, and the density approximations are drawn as linearly interpolated histograms over 50 bins.

$$\begin{aligned} &\leq \sum_{i=1}^{m-1} |(g_k)_i| \sum_{j=0}^{n-1-k} \left\| \Gamma_i P^j \right\|_1 \\ &\leq \|g_k\|_1 C. \end{aligned}$$

We can replace C_P by this C in the conclusion of Theorem 7, to obtain an *explicit numerical bound* on the accuracy of our invariant measure approximation.

References

[DKP95] P Diamond, P Kloeden, and A Pokrovskii. Interval stochastic matrices: a combinatorial lemma and the computation of invariant measures of dynamical systems. *J. Dynamics and Differential Equations*, 7(2), 1995.

[dMvS93] W de Melo and S van Strien. *One–dimensional dynamics*. Springer, 1993.

Correlation decay constants as partition is refined.

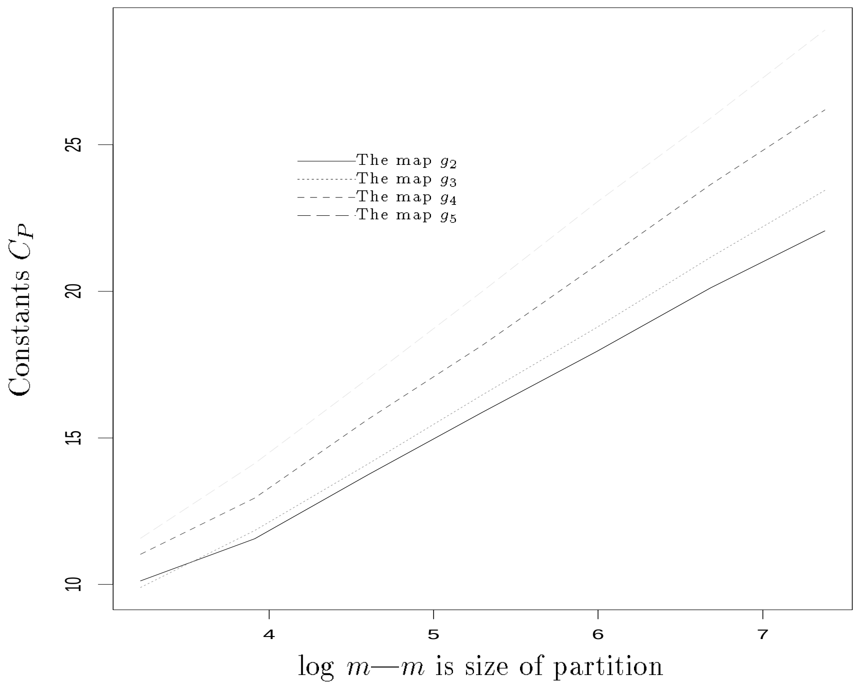

Figure 3: Scaling of decay of correlation constants for the area overlaps matrices corresponding to maps g_l as the partition is refined (size of matrix increased).

[DZ96] Jiu Ding and Aihui Zhou. Finite approximations of Frobenius–Perron operators. a solution of Ulam's conjecture to multi–dimensional transformations. *Physica D*, 92:61–68, 1996.

[Fro95] G Froyland. Finite approximation of Sinai–Bowen–Ruelle measures for Anosov systems in two dimensions. *Random & Computational Dynamics*, 3(4):251–264, 1995.

[Li76] Tien-Yien Li. Finite approximation for the Perron–Frobenius operator. a solution to Ulam's conjecture. *Journal of approximation theory*, 17:177–186, 1976.

[Now93] T Nowicki. Some dynamical properties of S–unimodal maps. *Fundamenta Mathematicae*, 142:45–57, 1993.

[NvS91] T Nowicki and S van Strien. Invariant measures exist under a summability condition for unimodal maps. *Invent. Math.*, 105:123–136, 1991.