RANDOM & COMPUTATIONAL DYNAMICS, 1(3), 261-276 (1992-93)

REPRESENTATION OF MARKOV CHAINS ON TORI

Anthony N. Quas
Maths Institute
University of Warwick
Coventry, CV4 7AL, ENGLAND

ABSTRACT

We consider the problem of representing Markov chains on smooth Riemannian manifolds by smooth maps, and show that in the case where the manifold is an n-torus, any sufficiently smooth Markov chain may be represented by a collection of homotopic N-to-1 local diffeomorphisms for some N. We then go on to consider which possible values of N can occur. For this, we specialise to the circle, where we provide a necessary and sufficient condition on N for the existence of a 'nice' representation of the Markov chain by degree N maps. We use this to construct maps which cannot be represented by degree N local homeomorphisms, and finally to construct an example of a Markov chain which cannot be represented by homeomorphisms.

§1. INTRODUCTION AND STATEMENT OF RESULTS

Throughout this paper, by smooth, we will mean C^∞. The problem which we consider is the problem of representing a Markov chain by smooth

AMS classification: 60J05

maps. This problem was originally raised in [1]. A Markov chain on a measurable space $(M, \mathcal{B})$ may be described by a map $P : M \times \mathcal{B} \to [0,1]$ where $P(x, A)$ is the probability of going from a point x to a point in the measurable set A. The map P is then called the *transition map* of the Markov chain.

Note that given a collection $\mathcal{F}$ of maps from the set M to itself and a measure m on the collection $\mathcal{F}$, we may create a Markov chain by setting $P(x, A) = m\{f \in \mathcal{F} : f(x) \in A\}$ (ignoring for the time being problems of measurablity). Representation is the reverse of this process. Given a Markov chain $\mathcal{M}$, with transition map P, a *representation* of $\mathcal{M}$ is a collection $\mathcal{F}$ of maps and a probability measure m on them such that $P(x, A) = m(\{f : f(x) \in A\})$. Note that for fixed x, the map $A \mapsto P(x, A)$ is a probability measure. We write P_x for this map.

Definition. *A smooth Markov chain on a smooth compact Riemannian manifold is one whose transition map P is given by the equation $P(x, A) = \int_A h(x, y)dV(y)$ where V is the Riemannian volume measure and $h : M \times M \to (0, \infty)$ is a smooth map.*

Note that in this definition, the transition densities are all taken to be strictly positive. It would be possible to allow the densities to take the value 0, but in that case, some of the results would fail.

In [3], the following Theorem was proved.

Theorem 1. *Let M be a smooth, orientable, compact, connected Riemannian manifold with Riemannian volume measure V. Then if $\mathcal{M}$ is a smooth Markov chain on M, then $\mathcal{M}$ may be represented by a collection of smooth maps on M.*

Further we have that the collection $\mathcal{F}$ of maps may be smoothly labelled by the points of the manifold, so as to satisfy the following:

(i) $\mathcal{F} = \{f_y : y \in M\}$

(ii) The map $(x, y) \mapsto f_y(x)$ is smooth in x and y

(iii) For fixed x the map $y \mapsto f_y(x)$ is a diffeomorphism $M \to M$

(iv) The measure m on $\mathcal{F}$ is equal to a volume form μ on the labels (That is $m(\{f_y : y \in A\}) = \mu(A)$).

This provides an affirmative answer to the question raised in [1] about representing Markov chains on manifolds by smooth maps. In what follows, we look at the specialised situation in which the manifold is an n-torus. In this case, we can extend the above result to show that any smooth Markov chain on an n-torus may be represented by a collection of homotopic N-to-1 surjective local diffeomorphisms for some N. This is the result of §2. In §3, we restrict attention to the circle and introduce two quantities δ_+ and δ_- with $\delta_- \leq 0 \leq \delta_+$, such that a smooth Markov chain on the circle may be 'nicely' represented by degree N local diffeomorphisms if and only if $N > \delta_+$ or $N < \delta_-$. We also show that if a Markov chain may be represented by degree N local homeomorphisms, then $N \geq \delta_+$ or $N \leq \delta_-$, so since we can construct Markov chains with arbitrarily large values of δ_+, we can construct Markov chains which may not be represented by degree N maps for any positive N smaller than some N_0. This goes part of the way to negatively answering another question from [1]: Can every smooth Markov chain on a manifold be represented by homeomorphisms?

In §4, we explicitly construct a Markov chain which cannot be represented by homeomorphisms. That is an example of a Markov chain which cannot be represented by a collection $\mathcal{F}$ of maps which consists partly of degree 1 homeomorphisms and partly of degree -1 homeomorphisms.

§2. REPRESENTATION OF MARKOV CHAINS ON TORI

Theorem 2. *Let $\mathcal{M}$ be a smooth Markov Chain on T^n with transition map P. Then $\mathcal{M}$ may be represented by a collection of homotopic N-to-1 local diffeomorphisms for some N.*

Proof. By Theorem 1, there exists a smoothly parameterised collection $\{f_y\}_{y\in\mathsf{T}^n}$ of maps with the property that for each pair $x, z \in \mathsf{T}^n$, there is exactly one y such that $f_y(x) = z$. In this case, write $y = \phi(x,z)$. For fixed x, the map $z \mapsto \phi(x,z)$ is a diffeomorphism of T^n. We also have that $P(x, A) = \mu\{y : f_y(x) \in A\}$ where there exists an $x_0 \in \mathsf{T}^n$ such that for all $B \in \mathcal{B}$, $\mu(B) = P(x_0, B)$. As such, μ is a smooth volume form on T^n, so by

Moser's Theorem ([2]), there exists a smooth diffeomorphism $\alpha : \mathsf{T}^n \to \mathsf{T}^n$ such that $\mu(B) = \lambda(\alpha^{-1}B)$ for Borel sets B, where λ is Haar measure.

Now define $e_y(x) = f_{\alpha(y)}(x)$. Then

$$P(x, A) = \mu\{y : f_y(x) \in A\} = \lambda\{\alpha^{-1}(y) : f_y(x) \in A\}$$
$$= \lambda\{y : f_{\alpha(y)}(x) \in A\} = \lambda\{y : e_y(x) \in A\}.$$

Note also that $e_{\alpha^{-1}(\phi(x,z))}(x) = f_{\phi(x,z)}(x) = z$, so the collection e_y has exactly the properties of the collection f_y except that the measure on the parameters is just Haar measure. We may therefore assume without loss of generality that the original measure μ was in fact Haar measure.

Next, write $\phi_z(x) = \phi(x, z)$. Then ϕ_z has lift $\Phi_z : \mathsf{R}^n \to \mathsf{R}^n$, say. As usual, we write $\Phi_z(x) = A_z x + P_z(x)$ where A_z is an integer matrix and P_z is periodic (that is $P_z(x + \mathbf{m}) = P_z(x)$ for $\mathbf{m} \in \mathsf{Z}^n$.) Since the collection ϕ_z is smoothly parameterised, it follows that the linear part A_z is continuously dependent on z, so since A_z is an integer matrix, A_z must be constant, say $A_z = A$.

Fix a norm, $\|\cdot\|$ on R^n. This induces the operator norm $\|\cdot\|$ on $M_n(\mathsf{R})$ satisfying $\|Ax\| \le \|A\|\|x\|$. Consider T^n as $[0,1)^n$ mod 1, pick $M \in \mathsf{N}$ such that $M > \|A\| + \sup_{x,z\in\mathsf{T}^n} \|D_x P_z\|$ and set $\theta(x, z) = \phi(x, z) + Mx$. For fixed x, the map $z \mapsto \theta(x, z)$ remains a diffeomorphism of T^n. Set $g_y(x) = {\theta_x}^{-1}(y)$. Clearly $g_y(x)$ is continuously dependent on y for fixed x, and so the maps g_y are certainly homotopic. Note also that

$$P(x, A) = \lambda\{y : f_y(x) \in A\} = \lambda\{\phi(x, z) : z \in A\}$$
$$= \lambda\{\theta(x, z) : z \in A\} = \lambda\{y : g_y(x) \in A\},$$

so the collection $\{g_y\}_{y\in\mathsf{T}^n}$ represents $\mathcal{M}$. It therefore remains to show that all the maps g_y are N-to-1 surjections for some uniform N.

To prove this, consider ${g_y}^{-1}\{z\} = \{x : \theta_x(z) = y\}$. Setting $\gamma_z(x) = \theta(x, z)$, we see ${g_y}^{-1}\{z\} = {\gamma_z}^{-1}\{y\}$, so it is sufficient to show that γ_z is an N-to-1 surjection for the some N which is independent of z. But $\gamma_z(x) = \phi(x, z) + Mx$, which has lift $\Gamma_z(x) = Ax + Mx + P_z(x)$. Write L for the matrix $A + MI$ where I is the identity matrix and suppose $x \ne y$. Then

$$\begin{aligned}\|\Gamma_z(x)-\Gamma_z(y)\| &= \|M(x-y)+A(x-y)+P_z(x)-P_z(y)\| \\ &\geq M\|x-y\|-(\|A\|+\sup_{z,x\in T^n}\|D_xP_z\|)\|x-y\|>0.\end{aligned}$$

So Γ_z is injective.

We now show Γ_z is surjective. Since $\|A\|<M$, we see that the matrix L is invertible, so given $y\in\mathbb{R}^n$, define the map

$$F:\mathbb{R}^n\to\mathbb{R}^n\ ;x\mapsto L^{-1}(y-P_z(x)).$$

The image of F is a bounded subset of $\mathbb{R}^n$ and so is contained in some closed ball $B(0,R)$. Now consider F as a map from $B(0,R)$ into itself. By the Brouwer fixed point theorem, there exists a point $x_0\in B(0,R)$ such that $F(x_0)=x_0$. Then $x_0=L^{-1}(y-P_z(x_0))$, so we see that $\Gamma_z(x_0)=y$. It follows then that Γ_z is surjective. We then show that this implies that γ_z is a $|detL|$-to-1 surjection. Note that $\mathbb{Z}^n$ is the disjoint union of cosets $L\mathbb{Z}^n+x_i,\ 1\leq i\leq m$ where $m=|detL|$ by standard theory of maps on tori. Denote by π the standard projection from $\mathbb{R}^n$ to $\mathbb{T}^n$ and pick $\zeta\in\mathbb{T}^n$. Then $\pi^{-1}(\zeta)=\mathbb{Z}^n+x$ for some $x\in\mathbb{R}^n$. Let $\rho_i=\pi({\Gamma_z}^{-1}(x+x_i))$. These are distinct, for if $\rho_i=\rho_j$, then

$${\Gamma_z}^{-1}(x+x_i)={\Gamma_z}^{-1}(x+x_j)+\mathbf{m},\text{ where }\mathbf{m}\in\mathbb{Z}^n.$$

Applying Γ_z, we get $x_i=x_j+L\mathbf{m}$, which implies $i=j$. This shows that $\rho_1,\ldots,\rho_m$ are distinct as claimed and we have $\gamma_z^{-1}\{\zeta\}\supset\{\rho_1,\ldots,\rho_m\}$.

Conversely, suppose $\gamma_z(\rho)=\zeta$, then pick $w\in\pi^{-1}(\rho)$. So, $\pi(\Gamma_z(w))=\zeta$ and $\Gamma_z(w)\in\mathbb{Z}^n+x$, so in particular $\Gamma_z(w)=L\mathbf{m}+x+x_i$ for some $\mathbf{m}\in\mathbb{Z}^n$ and some i. So

$$\Gamma_z(w)=\Gamma_z(\mathbf{m}+\Gamma_z^{-1}(x+x_i)).$$

From this, we deduce $w=\mathbf{m}+\Gamma_z^{-1}(x+x_i)$ and $\rho=\rho_i$. Then $\gamma_z^{-1}\{\zeta\}=\{\rho_1,\ldots,\rho_m\}$ and γ_z is $|detL|$-to-1 as required. This number is clearly independent of z. □

Further, we may characterise the homotopy class of the maps as follows. By standard results on the theory of maps of the torus, the map θ has a lift $\Theta : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$. The lift Θ may then, as usual, be split up into linear and periodic parts: $\Theta(x,z) = Ax + Bz + C(x,z)$, where A and B are integer matrices, and C is periodic in x and z (that is $C(x+\mathbf{n}, y+\mathbf{m}) = C(x,z)$ for all $\mathbf{m}$ and $\mathbf{n}$). Note that $|detB| = 1$. We have, however, that $g_y(x) = \theta_x^{-1}(y)$. Let $G_y(x)$ be the lift of g_y. By definition, we see $\theta(x, g_y(x)) = y$. Lifting this, we get $\Theta(x, G_y(x)) = Y$, where Y is a preimage of y under the natural projection. Substitution gives $Ax + BG_y(x) + C(x, G_y(x)) = Y$. As x varies, the right hand side must remain a preimage of y, so by continuity, we have that the right hand side is constant. As x moves through an integer displacement $\mathbf{m} \in \mathbf{Z}^n$, Ax moves through $A\mathbf{m}$ and $C(x, G_y(x))$ remains constant as C is periodic. It therefore follows that $BG_y(x+\mathbf{m}) = BG_y(x) - A\mathbf{m}$, so in particular

$$G_y(x+\mathbf{m}) = G_y(x) - B^{-1}A\mathbf{m}.$$

The linear part of G_y therefore has matrix $-B^{-1}A$, where A is the matrix of the linear part of θ for fixed z (considered as a map of x) and B is the corresponding matrix for fixed x.

§3. MARKOV CHAINS ON THE CIRCLE

Definition. *A smooth Markov chain $\mathcal{M}$ on a smooth Riemannian manifold M is said to be nicely represented by a collection $\{f_y\}_{y \in M}$ of maps and a volume form μ on M if the following properties hold:*

(i) For all points x and z in M, there exists a unique y in M such that $f_y(x) = z$. In this case, write $y = Y(x,z)$.

(ii) $Y(x,z)$ as defined above is smooth in both variables and for fixed x, the map $z \mapsto Y(x,z)$ is a diffeomorphism of M.

(iii) For each Borel set A and each point $x \in M$, $P(x,A) = \mu\{y : f_y(x) \in A\}$.

We now restrict ourselves to the case where $M = S^1$.

Definition. *The positive and negative degrees of $\mathcal{M}$ are given by*

$$\delta_+ = \int dx \sup_z \frac{\partial}{\partial x} P(x, [z_0, z])$$

$$\delta_- = \int dx \inf_z \frac{\partial}{\partial x} P(x, [z_0, z]).$$

Note that these degrees are independent of the point z_0 as

$$\frac{\partial}{\partial x}P(x,[z_0,z]) = \frac{\partial}{\partial x}P(x,[z_0,z_0']) + \frac{\partial}{\partial x}P(x,[z_0',z]).$$

Taking suprema over z, the first term of the right hand side is unaffected, and then integrating, this term drops out altogether. It follows that the degree δ_+ is independent of the point z_0. The same obviously holds for δ_-. We therefore fix a point $z_0 \in S^1$ for the rest of this section.

Theorem 3. *Let $N > 0$. Let $\mathcal{M}$ be a smooth Markov Chain on the circle. $\mathcal{M}$ may be nicely represented by degree N local diffeomorphisms if and only if $N > \delta_+$.*

Theorem 4. *Let $N > 0$. Suppose $\mathcal{M}$ is a smooth Markov chain on the circle, which may be represented by degree N local homeomorphisms. Then $N \geq \delta_+$.*

Corollary. *Let $N > 0$. There exist smooth Markov chains on the circle which cannot be represented by degree N local homeomorphisms.*

Note that the above statements have corresponding versions for $N < 0$. These are as follows.

Theorem 3'. *Let $N < 0$. Let $\mathcal{M}$ be a smooth Markov Chain on the circle. $\mathcal{M}$ may be nicely represented by degree N local diffeomorphisms if and only if $N < \delta_-$.*

Theorem 4'. *Let $N < 0$. Suppose $\mathcal{M}$ is a smooth Markov chain on the circle, which may be represented by degree N local homeomorphisms. Then $N \leq \delta_-$.*

Corollary'. *Let $N < 0$. There exist smooth Markov chains on the circle which cannot be represented by degree N local homeomorphisms.*

Proof of Theorem 3. Suppose $\mathcal{M}$ is nicely represented by degree N local diffeomorphisms. We may then assume the measure on the parameter space to be Haar measure as in §2. Write $Y(x,z)$ or $Y_x(z)$ for the parameter value of the unique map taking x into z and write $h(x,z)$ for the probability density of going from x to z. Y_x sends the image point of the map to its parameter value for fixed point x. It therefore follows that the density has to be related to the map Y by $|\frac{\partial}{\partial z}Y_x| = h(x,z)$. We may however, without loss of generality remove the moduli signs as these can only effect a change of parametrization. We therefore have the relation

$$\frac{\partial}{\partial z}Y(x,z) = h(x,z).$$

In what follows, we treat the circle as the interval $[0,1)$ mod 1. We then see that $Y(x,z) = P(x,[z_0,z]) - X(x)$ where X is some map $S^1 \to S^1$. For $\mathcal{M}$ to be represented by local diffeomorphisms, by the implicit function theorem, we require

$$\frac{d}{dx}X(x) > \sup_{z\in S^1} \frac{\partial}{\partial x}P(x,[z_0,z]) \text{ or } \frac{d}{dx}X(x) < \inf_{z\in S^1} \frac{\partial}{\partial x}P(x,[z_0,z]).$$

To ensure that the maps of the representation are orientation-preserving, we require the former condition to hold. We therefore see that X has degree greater than δ_+. But we have

$$N = |\{x : Y(x,z) = y\}| = |\{x : P(x,[z_0,z]) - X(x) = y\}|.$$

By the requirements placed on X, we have that the expression $P(x,[z_0,z]) - X(x)$ is monotonic in x. It follows that the cardinality in question is just the degree of the expression, but this is just the degree of X, so we see that $N > \delta_+$.

Conversely, suppose $N > \delta_+$. Set

$$\alpha(x) = \sup_z \frac{\partial}{\partial x}P(x,[z_0,z]).$$

Then we have $\int \alpha(x)dx = \delta_+$, so we can find an $\epsilon > 0$ and a smooth function $\beta(x)$ such that

$$(i)\ \beta(x) \geq \alpha(x) + \epsilon$$

$$(ii) \int \beta(x)dx = N.$$

Then set $X(x) = \int_{x_0}^{x} \beta(x)dx$ and finally, let $Y(x,z) = \pi(P(x,[z_0,z]) - X(x))$, where π is the standard projection of the real line onto the circle. For fixed $x \in S^1$, the map $z \mapsto Y(x,z)$ is a diffeomorphism $S^1 \to S^1$. Y is also smooth in x. Write $Y_x(z) = Y(x,z)$ and define $f_y(x) = Y_x^{-1}(y)$. By the implicit function theorem, we have $D_x f_y = -D_x Y / D_z Y$, so since $D_x Y < 0$ and $D_z Y > 0$, we have $D_x f_y > 0$. The map f_y is a smooth local diffeomorphism and

$$f_y^{-1}\{z\} = \{x : f_y(x) = z\} = \{x : y = Y_x(z)\}.$$

From this, we see that f_y is N-to-1 and so f_y has degree N. Finally, we check that with μ taken to be Haar measure, this does indeed provide a nice representation of the Markov chain $\mathcal{M}$.

$$\begin{aligned}\mu\{y : f_y(x) \in A\} &= \mu\{y : Y_x^{-1}(y) \in A\} = \mu\{Y_x(z) : z \in A\} \\ &= \mu\Big\{P\big(x,[z_0,z]\big) : z \in A\Big\} = P(x,A).\end{aligned}$$

Note that we are using the translation invariance of Haar measure for the third equality. The Markov chain $\mathcal{M}$ is therefore nicely represented by degree N local diffeomorphisms. □

Before embarking on the proof of Theorem 4, we need some lemmas and definitions.

Definition. *A map $p : S^1 \times \mathcal{B} \to [0,1]$ is an S-map if*

(i) For each $x \in S^1$, the map $A \mapsto p(x,A)$ is a measure,

(ii) $p(x,S^1)$ is independent of x and

(iii) For fixed $A \in \mathcal{B}$, the map $x \mapsto p(x,A)$ is measurable.

Note that an S-map is just a constant multiple of a transition map.

Further, if p_1 and p_2 are S-maps, then we say p_1 is subordinate to p_2 if $p_1(x,A) \le p_2(x,A)$ for each $x \in S^1$ and $A \in \mathcal{B}$.

The weight of an S-map p is denoted by $w(p)$ and is defined to be $p(x,S^1)$ (which is independent of x).

Definition. *Suppose the Markov chain $\mathcal{M}$ is represented by the collection of maps $\mathcal{F}$ and a measure ν on them. A subrepresentation of this is defined* *by a measurable subset $\mathcal{F}'$ of $\mathcal{F}$ and the restriction of the measure ν to a measure ν' defined on $\mathcal{F}'$.*

Note that in this case the induced S-map (defined by $p(x,A) = \nu'\{f \in \mathcal{F}' : f(x) \in A\}$) is subordinate to P.

Let Σ be the collection of all S-maps. Then define the map

$$V : \Sigma \to \mathbf{R}^+;\ p \mapsto \lim_{m\to\infty} \sup_{z_1\ldots z_m} \sum_{i=1}^{m} \Big[p\big(\pi(\tfrac{i}{m}), [z_0, z_i]\big) - p\big(\pi(\tfrac{i-1}{m}), [z_0, z_i]\big)\Big].$$

Lemma 1. *Let p be a smooth S-map. Then we have*

$$\int \sup_z \frac{\partial}{\partial x} p\big(x, [z_0, z]\big)\, dx = V(p). \tag{1}$$

Proof. Set $A(x,z) = p(x,[z_0,z])$. This is smooth in x and z. Write A_x for $\frac{\partial}{\partial x}A$. Then

$$p\big(\pi(\tfrac{i}{m}), [z_0, z_i]\big) - p\big(\pi(\tfrac{i-1}{m}), [z_0, z_i]\big) = \int_{\pi(i-1/m)}^{\pi(i/m)} dx \frac{\partial}{\partial x} p\big(x, [z_0, z_i]\big)$$
$$\leq \int_{\pi(i-1/m)}^{\pi(i/m)} dx \sup_z \frac{\partial}{\partial x} p\big(x, [z_0, z]\big).$$

From this, we see

$$\sup_{z_1\ldots z_m} \sum_{i=1}^{m} \Big[p\big(\pi(\tfrac{i}{m}), [z_0, z_i]\big) - p\big(\pi(\tfrac{i-1}{m}), [z_0, z_i]\big)\Big] \leq \int dx \sup_z \frac{\partial}{\partial x} p\big(x, [z_0, z]\big).$$

This shows the right hand side of (1) is bounded above by the left hand side.

Now we show they are equal. Pick $\epsilon > 0$. By uniform continuity, there exists a $\delta > 0$ such that $|x_1 - x_2| < \delta \Rightarrow \forall z,\ |A_x(x_1,z) - A_x(x_2,z)| < \epsilon$. Now pick $m \in \mathbf{N}$ such that $m > \delta^{-1}$ and choose i with $1 \leq i \leq m$. Then let $y \in S^1$ be such that $A_x(\frac{i-1}{m}, y) = \sup_z A_x(\frac{i-1}{m}, z)$. Then

$$x \in [\tfrac{i-1}{m}, \tfrac{i}{m}) \Rightarrow A_x(x,y) \geq A_x(\tfrac{i-1}{m}, y) - \epsilon.$$

Also,

$$x \in [\tfrac{i-1}{m}, \tfrac{i}{m}) \Rightarrow A_x(x,z) \leq A_x(\tfrac{i-1}{m}, z) + \epsilon \leq A_x(\tfrac{i-1}{m}, y) + \epsilon, \ \forall z.$$

It follows that $A_x(x,y) \geq \sup_z A_x(x,z) - 2\epsilon$. Integrating between $\pi(\frac{i-1}{m})$ and $\pi(\frac{i}{m})$, we get

$$A(\tfrac{i}{m}, y) - A(\tfrac{i-1}{m}, y) \geq \int_{\frac{i-1}{m}}^{\frac{i}{m}} dx \Big[\sup_z \frac{\partial}{\partial x} p(x, [z_0, z]) - 2\epsilon\Big].$$

Finally, adding gives

$$\sup_{z_1 \dots z_m} \Big[\sum_{i=1}^{m} p\big(\pi(\tfrac{i}{m}), [z_0, z_i]\big) - p\big(\pi(\tfrac{i-1}{m}), [z_0, z_i]\big)\Big] \geq$$

$$\int dx \sup_z \frac{\partial}{\partial x} p(x, [z_0, z]) - 2\epsilon.$$

This completes the proof of the lemma. □

Lemma 2. *Suppose p is a smooth S-map arising from some measure ν on some collection $\mathcal{F}$ of degree N orientation-preserving local homeomorphisms (possibly with $\nu(\mathcal{F}) \neq 1$), then $V(p) \leq Nw(p)$.*

Proof.

$$V(p) = \lim_{m \to \infty} \sup_{z_1 \dots z_m} \sum_{i=1}^{m} \Big[p\big(\pi(\tfrac{i}{m}), [z_0, z_i]\big) - p\big(\pi(\tfrac{i-1}{m}), [z_0, z_i]\big)\Big]$$

$$= \lim_{m \to \infty} \sup_{z_1 \dots z_m} \int d\nu(f) \sum_{i=1}^{m} \Big[\chi_{[z_0, z_i]}\big(f(\tfrac{i}{m})\big) - \chi_{[z_0, z_i]}\big(f(\tfrac{i-1}{m})\big)\Big]$$

$$\leq \lim_{m \to \infty} \int d\nu(f) \sum_{i=1}^{m} \sup_{z_i} \Big[\chi_{[z_0, z_i]}\big(f(\tfrac{i}{m})\big) - \chi_{[z_0, z_i]}\big(f(\tfrac{i-1}{m})\big)\Big]$$

where χ_A is the characteristic function of the set A. But we have

$$\sum_{i=1}^{m} \sup_{z_i} \Big[\chi_{[z_0, z_i]}\big(f(\tfrac{i}{m})\big) - \chi_{[z_0, z_i]}\big(f(\tfrac{i-1}{m})\big)\Big]$$

$$= \Big|\big\{i : 1 \leq i \leq m, \ z_0 \leq f(\tfrac{i}{m}) < f(\tfrac{i-1}{m})\big\}\Big|.$$

where by inequalities on the circle, we mean that there is a continuous choice of argument on a connected subset of the circle including the specified points on which the order of the values of the argument is that specified. The cardinality above is however bounded above by N as there can be at most one such i between any adjacent pair of preimages of z_0 under f. It therefore follows that $V(p) \leq N\nu(\mathcal{F}) = Nw(p)$. □

Theorem 4 then follows as a straightforward application of the Lemmas.

Proof of Theorem 4. Applying Lemmas 1 and 2 to the transition map P of the Markov chain, we see $\delta_+ = V(P) \leq Nw(P)$. But, we also have that $w(P) = 1$, so the Theorem is proved. □

Proof of the Corollary. To prove the corollary, it is sufficient to construct a Markov chain with $\delta_+ > N$. As an example of such a Markov chain, consider the following:

Considering the circle as the interval $[0, 1)$ mod 1, and given positive constants α and β, such that $\alpha < 1$ and $\beta < \frac{1}{2}$, pick a smooth function f such that

$$(i)\ f(x) > 0,\ \forall x \in S^1$$

$$(ii)\ \int f(x)dx = 1$$

$$(iii)\ \int_{1/2-\beta}^{1/2+\beta} f(x)dx = \alpha$$

$$(iv)\ f(x) = \frac{1-\alpha}{1-2\beta} \text{ for } x \notin (\tfrac{1}{2} - \beta, \tfrac{1}{2} + \beta).$$

Next, set $h(x, z) = f(z + rx)$, where $r \in \mathbb{N}$, and using this, define P to be the transition map with probability density h: For a Borel set A and a point $x \in S^1$, $P(x, A)$ is defined to be $\int_A h(x, z)dz$. This clearly defines a smooth Markov chain.

We then estimate the value of $\sup_z \frac{\partial}{\partial x} P(x, [z_0, z])$ as follows:

$$\begin{aligned}\frac{\partial}{\partial x} P(x, [z_0, z]) &= \frac{\partial}{\partial x} \int_{z_0}^{z} h(x, y)dy = \frac{\partial}{\partial x} \int_{z_0+rx}^{z+rx} f(y)dy \\ &= r\big(f(z + rx) - f(z_0 + rx)\big)\end{aligned}$$

Note that defining $\zeta(x) = \frac{1}{2} - rx$, $|z_0 - \zeta(x)| \geq \beta \Rightarrow h(x, z_0) = \frac{1-\alpha}{1-2\beta}$. Clearly, however, we have $\sup_z h(x, z) > \frac{\alpha}{2\beta}$, so we have

$$|z_0 - \zeta(x)| \geq \beta \Rightarrow \sup_z \frac{\partial}{\partial x} P\big(x, [z_0, z]\big) > r\Big(\frac{\alpha}{2\beta} - \frac{1-\alpha}{1-2\beta}\Big).$$

We also have, however, that $\sup_z \frac{\partial}{\partial x} P(x, [z_0, z]) \geq 0$. But $|z_0 - \zeta(x)| \geq \beta$ on a set of measure $1 - 2\beta$, so we deduce

$$\delta_+ > r\Big(\frac{\alpha}{2\beta} - \frac{1-\alpha}{1-2\beta}\Big)(1-2\beta) = r\Big(\frac{\alpha}{2\beta} - 1\Big).$$

Then taking $\alpha = \frac{1}{2}$, $\beta = \frac{1}{8}$ and $r = N$, we get $\delta_+ > N$. This proves the corollary. □

§4. A MARKOV CHAIN WHICH CANNOT BE REPRESENTED BY HOMEOMORPHISMS

The above shows that there exist Markov chains which cannot be represented by orientation-preserving homeomorphisms. It remains an interesting question to ask if there are Markov chains which cannot be represented by a combination of orientation-preserving and orientation-reversing homeomorphisms. Such Markov chains do in fact exist, and we will modify the above example to show this.

Theorem 5. *There exists a smooth Markov chain on the circle which cannot be represented by homeomorphisms.*

Proof. The strategy of the proof will be to construct a Markov chain with transition map P and to show that there can be no S-map induced by a collection of degree 1 homeomorphisms of weight $\frac{1}{2}$ which is subordinate to P and the same thing for degree -1 homeomorphisms. This will then complete the proof of the Theorem as, if the result did not hold, there would be a representation of the Markov chain which would be composed of degree 1 and -1 homeomorphisms. In particular, the measure of one of these subsets would have to be at least $\frac{1}{2}$, and taking the S-map induced by a subrepresentation of this would contradict the above.

In the course of the proof, we will take λ to be Haar measure on the circle. The Markov chain which we will use is that which we constructed in the Corollary above. The parameters α, β and r are to be determined.

Write P for the transition map of this Markov chain. Let $\zeta(x)$ be given by $\frac{1}{2} - rx$.

Suppose then that $\mathcal{F}_+$ is a collection of orientation-preserving homeomorphisms and ν_+ is a measure on them such that $\nu_+(\mathcal{F}_+) = \frac{1}{2}$ and such that the induced S-map is subordinate to P. We perform two estimates: First, fix $f \in \mathcal{F}_+$ and consider the set $\{x : |f(x) - \zeta(x)| < \beta\}$. The function $f(x) - \zeta(x)$ is monotonic of degree $r+1$, so the set above has $r+1$ components. Take a lift G of the function $f(x) - \zeta(x)$ and suppose $G(y) = n - \beta$ where $n \in \mathbb{N}$. Then $G(y + \frac{2\beta}{r}) > n + \beta$, so the measure of each component of the set is less than $\frac{2\beta}{r}$, so we get

$$\lambda\big(\{x : |f(x) - \zeta(x)| > \beta\}\big) > 1 - \frac{2\beta}{r}(r+1).$$

By construction of P however, we have $P\big(x, S^1 \backslash [\zeta(x) - \beta, \zeta(x) + \beta]\big) = 1 - \alpha$, so in order for the induced S-map to be subordinate to P, we must also have

$$\nu_+\big(\{f \in \mathcal{F}_+ : |f(x) - \zeta(x)| > \beta\}\big) \leq 1 - \alpha.$$

Integrating these inequalities with respect to f and x respectively and applying Fubini's theorem, we see that for consistency, we are forced to have

$$1 - \alpha > \frac{1}{2}\Big(1 - \frac{2\beta}{r}(r+1)\Big).$$

Instead of this, suppose we have a collection $\mathcal{F}_-$ of orientation-reversing homeomorphisms and that ν_- is a measure on them such that $\nu_-(\mathcal{F}_-) = \frac{1}{2}$ and such that the induced S-map is subordinate to P. Then, as before, we have

$$\nu_-\big(\{f \in \mathcal{F}_- : |f(x) - \zeta(x)| > \beta\}\big) \leq 1 - \alpha.$$

We will also require an estimate of the measure of the set $\{x : |f(x) - \zeta(x)| > \beta\}$ for fixed $f \in \mathcal{F}_-$. This time, $f(x) - \zeta(x)$ has degree $r - 1$, but we can no longer say that the function is monotonic. We consider a lift G of the function $f(x) - \zeta(x)$. As we noted above, this has degree $r-1$. Then pick a point y such that $G(y) = \frac{1}{2} + \beta$ and for $1 \leq i \leq r-1$ set

$$a_i = \sup\{z \in [y, y+1) : G(z) = \tfrac{1}{2} + (i-1) + \beta\}$$
$$b_i = \inf\{z \in [a_i, y+1) : G(z) = \tfrac{1}{2} + i - \beta\}.$$

Note that $G(y+1) = \frac{1}{2} + \beta + (r-1)$, so that each of the above exists. We have also however that $\pi(a_i, b_i) \subset \{x : |f(x) - \zeta(x)| > \beta\}$, but $b_i - a_i > \frac{1-2\beta}{r}$ as $G(a_i) = \frac{1}{2} + \beta + (i-1)$ and $G(a_i + \sigma) < G(a_i) + r\sigma$. So since the sets $\pi(a_i, b_i)$ are disjoint, we get

$$\lambda(\{x : |f(x) - \zeta(x)| > \beta\} > (r-1)\frac{1-2\beta}{r}.$$

Integrating and using Fubini's theorem as before, we find that we require for consistency that

$$1 - \alpha > \tfrac{1}{2}(r-1)\frac{1-2\beta}{r}.$$

We may then choose α, β and r, so taking $r = 2, \beta = \frac{1}{4}$ and $\alpha > \frac{7}{8}$, we find that neither of the above inequalities is satisfied, and so we have a smooth Markov chain which cannot be represented by homeomorphisms. □

ACKNOWLEDGEMENTS

For the period during which this research was done, I was supported by an SERC studentship.

I would like to thank Peter Walters for the encouragement he has offered in the preparation of this paper, and also for his help in improving the proofs in §2.

REFERENCES

1. Kifer, Y., *Ergodic Theory of Random Transformations*, Birkhäuser, Boston, 1986.

2. Moser, J., On the volume elements on a manifold, Trans. Amer. Math. Soc. **120**(1965), 286–294.

3. Quas, A.N., On representations of Markov chains by random smooth maps, Bull. London Math. Soc. **23**(1991), 487–492.