On the Hit and Run Process January 1998

by Persi Diaconis, Mathematics Department, Cornell University, Ithaca, NY 14853
and David Freedman, Statistics Department, U.C. Berkeley, CA 94720

Abstract

Given a density f on Euclidean k-space R^k and a starting point x, choose a line L at random through x and move to a point on L chosen at random from f restricted to L. This procedure defines a Markov chain—the "hit and run process." The given density f is stationary; for almost all starting points x, the distribution of the chain at time n converges to the stationary distribution as n gets large. This expository paper proves some of the convergence theorems, and gives examples.

1. Introduction

This paper is largely expository. We explore convergence properties of the "hit and run" process, as defined below. The mathematical context is (1); heuristics will be discussed later.

(1) Let $f \geq 0$ be a non-negative, real-valued function on R^k, with $\displaystyle\int_{R^k} f(x)\,dx = 1$.

If $x \neq x'$ are two points in R^k, let $L_{xx'}$ be the line through them, and let $M_{xx'} = M_{xx'}(f)$ be the integral of f along $L_{xx'}$. Generally, $0 < M_{xx'} < \infty$, but there are exceptions (see below). If $0 < M_{xx'} < \infty$, let $\phi_{xx'}$ be the probability measure on $L_{xx'}$ whose (linear) density is f restricted to $L_{xx'}$ and renormalized. The "hit-and-run" kernel $Q_x(dy)$ is an algorithm for choosing y given x:

(i) Choose a point x' at random on the sphere of radius 1 centered at x.
(ii) If $0 < M_{xx'} < \infty$, choose a point at random from $\phi_{xx'}$ and move to that point.
(iii) Otherwise, stay at x.

A little more notation is needed. If μ is a probability and $K = K_x(A)$ is a kernel, the probability μK is defined by the relation $\mu K(A) = \int K_x(A)\,\mu(dx)$. If K and L are kernels, then KL is the kernel $(KL)_x = K_x L$. Multiplication of kernels is associative. If μ, ν are probabilities, then $\|\mu - \nu\| = \sup_A |\mu(A) - \nu(A)|$. If $x \in R^k$, then $\|x\|$ is the Euclidean norm of x. Our objective in this paper is to provide relatively self-contained proofs of the following three facts:

Theorem 1. Assume (1). Let $\phi(dx) = f(x)\,dx$. Then ϕ is stationary under Q, that is, $\phi Q = \phi$.

Theorem 2. If f satisfies condition (1), then $\|Q_x^n - \phi\| \to 0$ as $n \to \infty$, for Lebesgue-almost all x in R^k.

Theorem 3. Suppose f vanishes off some compact set C, and $\int_C f(x)\,dx = 1$. Suppose there is a finite constant c with $0 \leq f \leq c$ on C. Then there is a $\delta = \delta_c \in (0,1)$ such that $\sup_{x\in C}\|Q_x^n-\phi\|<(1-\delta)^n$ for all $n=1,2,\ldots$.

These theorems will be proved in sections 2–4. Under condition (1), convergence can be arbitrarily slow; examples will be given in section 5. Variations on the hit-and-run algorithm will be discussed in section 6, which also has a literature review, and explains the motivation for this line of research.

2. Proof of Theorem 1

Throughout this section, we assume (1). The easy proof of Lemma 1 is omitted.

Lemma 1. $M_{xy}=M_{yx}$ for all $x\neq y$ in R^k. Moreover, $M_{xy}=M_{xz}$ provided $z-x=a(y-x)$ for some real number $a\neq 0$.

Part of the reasoning for Lemma 3 may be easier to see in the abstract. Let $(\Omega_i,\mathcal{F}_i)$ be measurable spaces; $\mathcal{P}_1$ is a probability on $\mathcal{F}_1$ and μ_2 is a σ-finite measure on $\mathcal{F}_2$; V is a measurable mapping from Ω_1 to Ω_2, whose distribution $\mathcal{P}_2=\mathcal{P}_1V^{-1}$ happens to be absolutely continuous with respect to μ_2, having density g.

Lemma 2. $\mathcal{P}_1\{g(V)>0\}=1$.

Proof. Plainly, $g>0$ almost surely with respect to $g\,d\mu_2$; that is, $\mathcal{P}_2\{g>0\}=1$. But $\mathcal{P}_1\{g(V)>0\}=\mathcal{P}_1V^{-1}\{g>0\}=\mathcal{P}_2\{g>0\}$. QED

The "unit sphere" in R^k is $S_k=\{v:v\in R^k \text{ and } \|v\|=1\}$; this is a $(k-1)$-dimensional object in k-dimensional space.

Lemma 3. For each x, for ϕ-almost all y, $M_{xy}>0$.

Proof. Without real loss of generality, take $x=0$. Let v be a point in the unit sphere S_k. Let $\gamma_v=\int_0^\infty r^{k-1}f(rv)\,dr$. Then $\gamma_v\geq 0$ and $\int\gamma_v\,dv=1$, as one sees by changing to polar coordinates: indeed, $v\to\gamma_v$ is the density of $V(y)=y/\|y\|$, where y is picked at random from f. Thus, $\gamma_{V(y)}>0$ for ϕ-almost all y, by Lemma 2. Consequently, $0<\int_0^\infty f\big(rV(y)\big)\,dr=M_{0y}$ for ϕ-almost all y. QED.

Lemma 4. $M_{xy}(f)$ is a jointly measurable function of x and y.

Proof. If f is continuous and vanishes off a compact set, then $x,y\to M_{xy}(f)$ is continuous. The argument can be completed in the usual way, via repeated monotone passages to the limit. QED

Let λ_k be the uniform distribution on the unit sphere S_k; λ_k is normalized to have total mass 1. Let $\mathcal{X}_1$ consist of all the x in R^k such that $M_{x,x+v}<\infty$ for λ_k-almost all $v\in S_k$. If A,B are sets, we write $A-B$ for the set of points in A but not in B.

Lemma 5. $\mathcal{X}_1$ is Borel, $R^k-\mathcal{X}_1$ has Lebesgue measure 0, and $\phi(\mathcal{X}_1)=1$.

Proof. That $\mathcal{X}_1$ is Borel measurable follows from Lemma 4 and Fubini's theorem. Let ξ run through a fixed $(k-1)$-dimensional linear subspace $\mathcal{L}$ of R^k, for instance, all vectors whose kth coordinate vanishes. Let t run through the real line R. For each $v\in S_k-\mathcal{L}$, there is a constant c_v such that for any measurable $g \geq 0$ on R^k,

$$\int_{R^k} g(x)\, dx = c_v \int_{\mathfrak{L}} \int_R g(\xi + tv)\, dt d\xi. \tag{2}$$

This is an "oblique" version of Fubini's theorem; c_v is the sine of the angle between v and $\mathfrak{L}$: the more familiar version has $v \perp \mathfrak{L}$ and $c_v = 1$. Let $\mathfrak{L}_0$ be the set of $\xi \in \mathfrak{L}$ with $M_{\xi,\xi+v}(f) = \infty$: this is a Borel set with Lebesgue measure 0, as one sees by using (2) with f in place of g. Since M_{xy} depends only on the line through x and y, the set $\mathfrak{L}_v$ of $x \in R^k$ with $M_{x,x+v} = \infty$ is the set of lines parallel to v running through points in $\mathfrak{L}_0$. So $\mathfrak{L}_v$ has Lebesgue measure 0 by another application of (2). In other words, for each v, $M_{x,x+v} < \infty$ for almost all x. By Fubini's theorem—the ordinary measure-theoretic one—for almost all x, $M_{x,x+v} < \infty$ for almost all $v \in S_k$. Lemma 3 shows that for all x, $\lambda_k\{M_{x,x+v} > 0\} > 0$. Hence, $R^k - \mathfrak{X}_1$ is Lebesgue-null, and $\phi(\mathfrak{X}_1) = 1$ because ϕ is absolutely continuous. QED

Lemma 6. Let $Q_x(dy)$ be the kernel for the hit and run process. Let $s_k = 2\pi^{k/2}/\Gamma(k/2)$ be the "surface area" of the unit sphere S_k. And let

$$Q_x^0(dy) = \frac{2f(y)\, dy}{s_k \|y - x\|^{k-1} M_{xy}},$$

with the understanding that the right hand side is 0 for y with $M_{xy} = 0$ or ∞. (It is possible that $Q_x^0(R^k) < 1$, even for a set of x's of positive Lebesgue measure.)

(a) If $0 < M_{xy} < \infty$ for Lebesgue-almost all y, then $Q_x = Q_x^0$.

(b) Otherwise, $Q_x(dy) = Q_x^0(dy) + p_x\delta_x$, where p_x is the chance of choosing $x' \in S_k$ with $M_{xx'} = 0$ or ∞, and δ_x is point mass at x.

Proof. We follow the convention that $0/0 = 0/\infty = 0$, and keep $x' \in S_k$:

$$\begin{aligned} Q_x(A) &= \frac{1}{s_k} \int_{S_k} \phi_{xx'}(A)\, dx' + p_x\delta_x \\ &= \frac{2}{s_k} \int_{S_k} \int_0^\infty \frac{1_A(x + rx')f(x + rx')}{M_{x,x+x'}}\, dr dx' + p_x\delta_x \\ &= \frac{2}{s_k} \int_{R^k} \frac{1_A(y)f(y)}{\|y - x\|^{k-1} M_{xy}}\, dy + p_x\delta_x. \end{aligned}$$

The 1st equality holds by the definition of the hit-and-run process; the 2nd, by the definition of $\phi_{xx'}$. The factor of 2 comes in because r is restricted to the positive half-line, while $\phi_{x,x'}$ covers the whole line. The last equality is obtained by changing from polar to rectangular coordinates: $dy = r^{k-1} dr dx'$. Polar coordinates are centered at x, so $r = \|y - x\|$. QED

Corollary 1. $\displaystyle\int_{R^k} \frac{2f(y)}{s_k \|y - x\|^{k-1} M_{xy}}\, dy = 1 - p_x.$

This and Lemma 1 prove Theorem 1. Indeed,

$$
\begin{aligned}
\int_{R^k} f(x)Q_x(A)\,dx &= \int_{R^k} f(x)p_x 1_A(x)\,dx + \int_{R^k} f(x)Q_x^0(A)\,dx \\
&= \int_{R^k} f(x)p_x 1_A(x)\,dx + \int_{R^k}\int_{R^k} \frac{2f(x)f(y)1_A(y)}{s_k\|y-x\|^{k-1}M_{xy}}\,dydx \\
&= \int_{R^k} f(x)p_x 1_A(x)\,dx + \int_{R^k}\int_{R^k} \frac{2f(x)f(y)1_A(y)}{s_k\|x-y\|^{k-1}M_{yx}}\,dxdy \\
&= \int_{R^k} f(x)p_x 1_A(x)\,dx + \int_{R^k} f(y)1_A(y)(1-p_y)\,dy \\
&= \int_{R^k} f(y)1_A(y)p_y\,dy + \int_{R^k} f(y)1_A(y)(1-p_y)\,dy \\
&= \int_{R^k} f(y)1_A(y)\,dy.
\end{aligned}
$$

Corollary 2. Let $\mathcal{X}$ consist of all the x in R^k such that $0 < M_{x,x+v} < \infty$ for a set of $v \in S_k$ of positive λ_k-measure, with S_k the unit sphere in R^k, and λ_k the uniform distribution on S_k. Then $\mathcal{X}$ is Borel, $R^k - \mathcal{X}$ has Lebesgue measure 0, and $\phi(\mathcal{X}) = 1$. For every $x \in \mathcal{X}$, $Q_x = p_x\delta_x + (1-p_x)Q'_x$ where $0 \le p_x < 1$, δ_x is point mass at x, and $Q'_x \equiv \phi$; in particular, $Q_x(\mathcal{X}) = 1$. Moreover, Q' is a kernel.

Remarks.

(i) Lemma 5 cannot be improved to "all x." For instance, let $f(x) \sim e^{-\|x\|}/\|x\|^{k-1}$. Then $M_{0y} = \infty$ for all $y \neq 0$.

(ii) In case (b) of Lemma 6, there are some $x' \in S_k$ with $\int_{L_{xx'}} f(y) = \infty$, and some x' with $f = 0$ almost everywhere on $L_{xx'}$. The first sort of x' generate a cone with apex at x; for y in this cone, $M_{xy} = \infty$. This cone has Lebesgue measure 0, for Lebesgue-almost all x; that is the content of Lemma 5. (A "cone" has a measurable base, but may be quite irregular in shape.) Likewise, the second sort of x' generate a cone with apex at x; for y in this cone, $M_{xy} = 0$. This second cone may well have positive Lebesgue measure, but its ϕ-measure is 0 for all x; that is the content of Lemma 3. For a concrete example, take $k = 2$; let $x \in R^2$ have coordinates x_1 and x_2. Let $f(x) \sim e^{-\|x\|}$ provided $x_1 > 0$ and $x_2 > 0$; otherwise, $f(x) = 0$. If $x_1 < 0$ and $x_2 < 0$, then $M_{xy} = 0$ if $(y_1 - x_1)(y_2 - x_2) < 0$.

(iii) If x is a Lebesgue point of $\{f > 0\}$, then $M_{xy} > 0$ for Lebesgue-almost all y. In particular, $\phi\{p_x = 0\} = 1$.

(iv) For $x \in \mathcal{X}$, Q_x is reversible relative to ϕ.

(v) In proving Theorem 2, we will show that $\|Q_x^n - \phi\| \to 0$ for all $x \in \mathcal{X}$. If on the other hand $x \notin \mathcal{X}$, the hit-and-run process stagnates at x.

3. Proof of Theorem 3

We begin with a result of Doeblin's, stated a bit more abstractly than is needed for Theorem 3. Let $(\mathcal{X}, \mathcal{B})$ be a measurable space. Let $Q_x(A)$ be a kernel. That is, $A \to Q_x(A)$

is a probability on $\mathcal{B}$ while $x \to Q_x(A)$ is $\mathcal{B}$-measurable. Let φ be an auxiliary probability on $(\mathcal{X}, \mathcal{B})$, and ϵ a positive real number. We assume that for all $A \in \mathcal{B}$ and all $x \in \mathcal{X}$,

$$Q_x(A) \geq \epsilon\varphi(A). \tag{3}$$

Theorem 4. Assume (3).
(a) There is a unique stationary probability μ.
(b) $\mu \geq \epsilon\varphi$.
(c) $\|Q_x^n - \mu\| \leq (1-\epsilon)^n$ for all $x \in \mathcal{X}$.

Proof. If α, β are probabilities on $(\mathcal{X}, \mathcal{B})$, and $0 \leq f \leq 1$ is a measurable function, then

$$|\textstyle\int f\, d\alpha - \int f\, d\beta| \leq \|\alpha - \beta\| \leq 1;$$

these inequalities will be used below. Let $R_x = (Q_x - \epsilon\varphi)/(1-\epsilon)$, which is also a kernel.

Step 1. If α and β are two probabilities on $\mathcal{B}$, then

$$\|\alpha Q^n - \beta Q^n\| \leq (1-\epsilon)^n \|\alpha - \beta\|.$$

Indeed,

$$\begin{aligned}
\alpha Q^n &= \int_{\mathcal{X}} \int_{\mathcal{X}} Q_y^{n-1} Q_x(dy)\alpha(dx) \\
&= \epsilon \int_{\mathcal{X}} \int_{\mathcal{X}} Q_y^{n-1} \varphi(dy)\alpha(dx) + (1-\epsilon) \int_{\mathcal{X}} \int_{\mathcal{X}} Q_y^{n-1} R_x(dy)\alpha(dx) \\
&= \epsilon\varphi Q^{n-1} + (1-\epsilon)\alpha R Q^{n-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\alpha Q^n - \beta Q^n\| &= (1-\epsilon)\|\alpha R Q^{n-1} - \beta R Q^{n-1}\| \\
&= (1-\epsilon)^2 \|\alpha R^2 Q^{n-2} - \beta R^2 Q^{n-2}\| \\
&\ \ \vdots \\
&= (1-\epsilon)^n \|\alpha R^n - \beta R^n\| \\
&\leq (1-\epsilon)^n \|\alpha - \beta\|.
\end{aligned}$$

Step 2. If α is any probability on $\mathcal{B}$, then $\{Q^n : n = 1, 2, \ldots\}$ is a Cauchy sequence of probabilities. Indeed, by Step 1,

$$\|\alpha Q^n - \alpha Q^{n+m}\| = \|\alpha Q^n - (\alpha Q^m) Q^n\| \leq (1-\epsilon)^n.$$

Step 3. If α is any probability on $\mathcal{B}$, then αQ^n converges to a limiting probability μ as $n \to \infty$; convergence is at a geometric rate, and the limit does not depend on α. Indeed,

$$\|\alpha Q^n - \alpha Q^{n+m}\| \leq (1-\epsilon)^n.$$

Let $m \to \infty$, so $Q^{n+m} \to \mu$:

$$\|\alpha Q^n - \mu\| \leq (1-\epsilon)^n.$$

In principle, μ could depend on α, but that possibility is eliminated by Step 1.

Step 4. μ is stationary. Indeed,

$$\mu - \mu Q = \lim (\mu Q^n - \mu Q Q^n) = \lim \mu Q^n - \lim \mu Q^{n+1} = 0.$$

Step 5. μ is unique, by Step 1.

Step 6. It remains only to prove claim (b). Now

$$\mu = \mu Q = \mu(\epsilon\varphi + R) \geq \epsilon\varphi.$$

This completes the proof of Theorem 4, and Theorem 3 follows. Indeed, M_{xy} is positive for ϕ-almost all y; it and $\|x-y\|$ are bounded above. Theorem 4 applies, with φ a small multiple of ϕ: see Lemma 6.

4. More on Doeblin's theory

References on Doeblin's theory of Markov chains include Doob (1953) and Orey (1971). This theory does not seem to be so accessible, even today; we will prove here the results we need, taking advantage of special features which permit a simplification of the argument. We begin by extending Theorem 4, replacing (3) by a "local" condition. Again, $(\mathcal{X}, \mathcal{B})$ is a measurable space; $Q_x(A)$ is a kernel; φ is an auxiliary probability on $\mathcal{B}$; and $\epsilon > 0$. Now, C is a fixed set in $\mathcal{B}$, with $\varphi(C) > 0$. The local condition is that

$$Q_x(A) \geq \epsilon\varphi(A) \text{ for } x \in C. \tag{4}$$

Suppose further that, from any starting positions, a pair of independent chains will hit C almost surely at the same time. More specifically, if X and Y are independent Markov chains moving according to Q, with $X_0 = x$ and $Y_0 = y$, we assume the coupling condition

$$\text{Prob}\{X_n \in C \text{ and } Y_n \in C \text{ for some } n = 1, 2, \ldots\} = 1, \quad \text{for any } x \text{ and } y. \tag{5}$$

Condition (4) is weaker than (3), because nothing is assumed for $x \notin C$. As written, the condition is to hold for all A, whether or not $A \subset C$. That is of course immaterial, because φ can be restricted to C and then renormalized. The set C is like "C-sets" in the literature; compare p. 7 in Orey (1971). Condition (5) is a strong type of recurrence condition.

Theorem 5. Let α and β be two starting probabilities on $(\mathcal{X}, \mathcal{B})$. The local Doeblin and coupling conditions (4–5) imply $\|\alpha Q^n - \beta Q^n\| \to 0$ as $n \to \infty$.

Proof. We construct two independent chains X_n and Y_n starting from α and β respectively. Let τ_1 be the first time that both chains hit C. We can redefine X and Y so that with probability ϵ their next move is from the φ in (4), hence the redefined processes agree from $\tau_1 + 1$ forwards; with the remaining probability $1 - \epsilon$ the processes continue to move independently. Now by strong Markov there will be another time at which both chains are simultaneously in C, and so forth. In short, given any $\delta > 0$, we can construct chains X_n^δ and Y_n^δ which start from α and β respectively, such that

$$\text{Prob}\{X_n^\delta = Y_n^\delta \text{ for all sufficiently large } n\} > 1 - \delta.$$

Hence

$$\liminf_{n\to\infty} \text{Prob}\{X_n^\delta = Y_n^\delta\} > 1 - \delta,$$

which completes the proof. QED

Remark. Theorem 5 applies to nonstationary processes, for instance, the random walk on the integers.

To use Theorem 5, it is necessary to verify the conditions. The hit-and-run kernel has two features which make verification relatively easy:

(6a) a stationary probability ϕ is given;

(6b) for all x, $Q_x = p_x\delta_x + (1 - p_x)Q'_x$, where $0 \le p_x < 1$, δ_x is point mass at x, and $Q'_x \equiv \phi$.

(For the hit-and-run kernel, the second condition holds after deleting a null set of x's.)

To exploit the stationarity, we use the following lemma, which is well known in ergodic theory. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability triple, and let T be a measure-preserving transformation: that is, T is a measurable mapping of Ω into itself, and $\mathcal{P}T^{-1} = \mathcal{P}$. If $F \in \mathcal{F}$, let τ_F be the time of the first return to F: that is, $\tau_F(\omega)$ is the least $n = 1, 2, \ldots$ if any with $T^n(\omega) \in F$, and $\tau_F(\omega) = \infty$ if none.

Lemma 7. If $\mathcal{P}(F) > 0$ then $\mathcal{P}\{\, \tau_F < \infty \mid F \,\} = 1$.

Proof. Let $F_0 = F - (T^{-1}F \cup T^{-2}F \cup \cdots)$. Starting from $F_0 \subset F$, the process never returns to F. Suppose by way of contradiction that $\mathcal{P}\{F_0\} > 0$. Now $F_0, T^{-1}F_0, \ldots$ are pairwise disjoint sets with the same positive probability, which is impossible. QED

Remark. Suppose $\mathcal{P}(F) > 0$. Let $\mathcal{P}_F = \mathcal{P}\{\, \bullet \mid F \,\}$. Define T_F almost surely on F as $T_F = T^{\tau_F}$. Then T_F is measure-preserving relative to $\mathcal{P}_F$.

The kernel Q is "φ-irreducible" if, for all x in $\mathcal{X}$ and all $A \in \mathcal{B}$ with $\varphi(A) > 0$, there is a positive probability that a chain starting from x at time 0 hits A in positive time: in other words, $\varphi(A) > 0$ implies $V_x(A) > 0$, where

$$V_x(A) = \sum_{n=1}^{\infty} Q_x^n(x, A). \tag{7}$$

Generally, irreducibility is much weaker than recurrence. However, condition (6)—which implies irreducibility and stationarity—is enough to get recurrence, in the form of the coupling condition (5).

Lemma 8. Suppose the kernel Q on $(\mathcal{X}, \mathcal{B})$ satisfies (6). Suppose $F \in \mathcal{B}$ and $\phi(F) > 0$. For any $x \in \mathcal{X}$, a chain starting from x and moving according to Q visits F i.o. a.s.

Proof. Let $\tilde{F}$ be the set of x such that a chain starting from x and moving according to Q is almost sure to hit F infinitely often. We begin by showing that $\phi(\tilde{F}) = 1$. Now $Q_x(F) > 0$ for all $x \in \mathcal{X}$ by (6b), because $\phi(F) > 0$. Thus, $\mathcal{X} = \bigcup_m F_m$, where $F_m = \{\, x : Q_x(F) > 1/m \,\}$. Therefore, it suffices to prove that $Q_x\{\,\text{hit } F \text{ i.o.}\} = 1$ for ϕ-almost all $x \in F_m$.

In view of Lemma 7, for ϕ-almost all $x \in F_m$, a chain that starts from x and moves according to Q will return to F_m i.o. a.s. On each return, the chain has chance at least $1/m$ to hit F, and the conditional form of the Borel-Cantelli lemma completes the proof that $\phi(\tilde{F}) = 1$. (Of course, some of the F_m may be empty or ϕ-null; for such F_m, the argument is vacuous.)

Now $\mathcal{N} = \mathcal{X} - \tilde{F}$ is ϕ-null; there seems to remain the possibility that $\mathcal{N}$ is non-empty. However, condition (6b) shows that $\mathcal{N}$ is actually empty. Indeed, starting from x, the process is almost sure to move eventually, and when it moves, it almost surely moves to some $y \in \tilde{F}$: from there it will almost surely visit F infinitely often. QED

For reasons that will be clear in a moment, we need to consider the chain at times $0, 2, 4, \ldots$.

Lemma 9. Suppose the kernel Q on $(\mathcal{X}, \mathcal{B})$ satisfies (6). Then Q^j satisfies the coupling condition (5), for any positive integer j, and any set C of positive ϕ-measure.

Proof. To begin with, Q^j satisfies (6), so we might as well take $j = 1$. Now define $R = Q \times Q$ as a kernel on $(\mathcal{X} \times \mathcal{X}, \mathcal{B} \times \mathcal{B})$:

$$R_{xy}(D) = (Q_x \times Q_y)(D)$$

for any product-measurable set D. It is easy to see that R satisfies (6), the stationary probability being $\phi \times \phi$. Lemma 7—with $F = C \times C$—completes the proof. QED

The next project is to construct a C-set as in (4). That may not be possible for Q itself, but is possible for Q^2. (This wrinkle necessitates the k in the previous lemma.) We adapt the argument from Orey (1971). For the measure-theoretic preliminaries, let $(\mathcal{X}_i, \mathcal{B}_i)$ be measurable spaces, and let ψ_i be a probability measure on $(\mathcal{X}_i, \mathcal{B}_i)$. The setting for Lemma 10 is the product space $(\mathcal{X}_1 \times \mathcal{X}_2,\ \mathcal{B}_1 \times \mathcal{B}_2,\ \psi_1 \times \psi_2)$; the leading special case, of course, is Lebesgue measure on the unit square. If $A \subset \mathcal{X}_1 \times \mathcal{X}_2$, then $A_{x\bullet}$ is the vertical section of A through x, namely, $\{y : y \in \mathcal{X}_2 \text{ and } (x, y) \in A\}$. Likewise, $A_{\bullet y}$ is the horizontal section through y. We use ϵ for the generic small positive number, and N for the generic large positive number.

Lemma 10. Let g be a non-negative, measurable function on $\mathcal{X}_1 \times \mathcal{X}_2$, with

$$\int g \, d\psi_1 d\psi_2 \leq \epsilon.$$

Then

$$\psi_1\{\, x : x \in \mathcal{X}_1 \text{ and } \textstyle\int g(x,y)\psi_2(dy) \geq \sqrt{\epsilon}\,\} \leq \sqrt{\epsilon}.$$

Proof. Let $U(x) = \int g(x,y)\,\psi_2(dy)$; consider U as a random variable on the probability triple $(\mathcal{X}_1, \mathcal{B}_1, \psi_1)$. By Fubini's theorem, $E(U) = \int g\, d\psi_1 d\psi_2$. This is at most ϵ by the conditions of the lemma, so the chance that $U \geq N\epsilon$ is at most $1/N$ by Markov's inequality. Put $N = 1/\sqrt{\epsilon}$ to complete the proof. QED

Corollary 3. If $A \in \mathcal{B}_1 \times \mathcal{B}_2$ and $(\psi_1 \times \psi_2)(A) \leq \epsilon$, then

$$\psi_1\{\, x : x \in \mathcal{X}_1 \text{ and } \psi_2(A_{x\bullet}) \geq \sqrt{\epsilon}\,\} \leq \sqrt{\epsilon}.$$

Proof. Use Lemma 10, with $g = 1_A$. QED

Corollary 4. If $B \in \mathcal{B}_1 \times \mathcal{B}_2$ and $(\psi_1 \times \psi_2)(B) \geq 1-\epsilon$, then

$$\psi_1\{\, x : x \in \mathcal{X}_1 \text{ and } \psi_2(B_{x\bullet}) > 1-\sqrt{\epsilon}\,\} \geq 1-\sqrt{\epsilon}.$$

Proof. Set $A = \mathcal{X}_1 \times \mathcal{X}_2 - B$. Then $(\psi_1 \times \psi_2)(A) \leq \epsilon$. By Corollary 3,

$$\psi_1\{\, x : x \in \mathcal{X}_1 \text{ and } \psi_2(A_{x\bullet}) \geq \sqrt{\epsilon}\,\} \leq \sqrt{\epsilon}.$$

so that

$$\psi_1\{\, x : x \in \mathcal{X}_1 \text{ and } \psi_2(A_{x\bullet}) < \sqrt{\epsilon}\,\} \geq 1-\sqrt{\epsilon}.$$

But $\psi_2(A_{x\bullet}) < \sqrt{\epsilon}$ iff $\psi_2(B_{x\bullet}) > 1-\sqrt{\epsilon}$, which completes the proof. QED

Our next topic is lower bounds on transition densities. Recall that $(\mathcal{X}, \mathcal{B})$ is a measurable space; let ψ is a probability on $\mathcal{B}$. Let $p(x,y)$ be a non-negative measurable function on $\mathcal{X} \times \mathcal{X}$, with $\int p(x,y)\,\psi(dy) = 1$. Then p is a transition density with corresponding kernel $P(x,dy) = p(x,y)\,\psi(dy)$. If p and q are transition densities, so is

$$(p \star q)(x,y) = \int p(x,u)q(u,y)\,\psi(du).$$

Now $p^{\star n}$ can be defined in the obvious way. Even if p is positive everywhere, $p \geq \delta$ may include no positive rectangle—see Example 1 below. However, the two-step density does admit such rectangles; that is the content of the next result.

Proposition 1. Suppose p is a transition density and $p(x,y) > 0$ for all x,y. For any δ with $0 < \delta < 1$, there are measurable sets G_δ, H_δ and a positive real number δ' such that $\psi(G_\delta) > 1-\delta$, $\psi(H_\delta) > 1-\delta$, and $p^{\star 2}(x,y) \geq \delta' > 0$ for all $x \in G_\delta$ and $y \in H_\delta$.

Proof. Without loss, we take $0 < \delta < 1/3$. There is a finite positive N so large that $\int p \wedge N\, d\psi^2 \geq 1-\delta^2$. (As usual, for non-negative a and b, $a \wedge b$ is the smaller of the two.) Let G be the set of $x \in \mathcal{X}$ with $\int p(x,u) \wedge N\,\psi(du) > 1-\delta$. Lemma 10 may be applied to $p - p\wedge N$, to see that $\psi(G) \geq 1-\delta$. Let ϵ be a small positive number, to be chosen later, and define $\epsilon' = 1 - \psi^2\{p \geq \epsilon\}$; thus $\epsilon' \downarrow 0$ as $\epsilon \downarrow 0$. Let H be the set of y in $\mathcal{X}$ for which $\psi\{p(\bullet,y) \geq \epsilon\} > 1-\sqrt{\epsilon'}$. Then $\psi(H) \geq 1-\sqrt{\epsilon'}$, by Corollary 4 applied to the set B of pairs (x,y) in $\mathcal{X}^2$ with $p(x,y) \geq \epsilon$; horizontal and vertical directions have been interchanged.

Fix $x \in G$ and $y \in H$. Now

$$\begin{aligned}
p^{\star 2}(x,y) &= \int_{\mathcal{X}} p(x,u)p(u,y)\,\psi(du) \\
&\geq \epsilon \int_{\{p(u,y)\geq\epsilon\}} p(x,u)\,\psi(du) \\
&\geq \epsilon \int_{\{p(u,y)\geq\epsilon\}} p(x,u)\wedge N\,\psi(du) \\
&\geq \epsilon\Big[\int p(x,u)\wedge N\,\psi(du) - N\psi\{\,u : p(u,y) < \epsilon\,\}\Big] \\
&> \epsilon\Big[1 - \delta - N\psi\{\,u : p(u,y) < \epsilon\,\}\Big] \\
&> \epsilon\Big[1 - \delta - N\sqrt{\epsilon'}\Big];
\end{aligned}$$

the second-to-last line holds because $x \in G$, and the last line holds because $y \in H$. We now choose ϵ so small that $\epsilon' < \delta^2$ and $N\sqrt{\epsilon'} < 1/3$. That completes the proof of Proposition 1. QED

Corollary 5. If the kernel Q satisfies (6), then Q^2 satisfies the local Doeblin condition (4).

Proof. Use Proposition 1 on the transition density of the kernel Q'_x, with ϕ in place of ψ. Let $C = G_\delta \cap F_\epsilon$, where $F_\epsilon = \{x : p_x < 1 - \epsilon\}$ and ϵ is chosen so that $\phi(F_\epsilon) > 1 - \delta$. let φ be ϕ retracted to H_δ and renormalized to have mass 1. QED

Remark. The C-set may be chosen to have ϕ-measure arbitrarily close to 1; the auxiliary measure φ may be chosen to be arbitrarily close to ϕ.

Corollary 6. If the kernel Q satisfies (6), then $\|Q_x^n - \phi\| \to 0$ as $n \to \infty$, for all $x \in \mathcal{X}$.

Proof. This is immediate from Corollary 5 and Theorem 5; even and odd n can be considered separately. QED.

Proof of Theorem 2

We can restrict $x \in R^k$ to the $\mathcal{X}$ of Corollary 2. The probability ϕ on $\mathcal{X}$ is stationary by Theorem 1, and condition (6) holds by Corollary 2. Corollary 6 completes the proof.

Example 1. Let $\mathcal{X} = [0,1]$ and let $\mathcal{B}$ be the σ-field of Borel sets in $\mathcal{X}$. Let φ be Lebesgue measure on $\mathcal{B}$. There is a positive measurable function f on $\mathcal{X}^2$ with the following property. If (i) $\delta > 0$, (ii) A and B are Borel sets, and (iii) $A \times B \subset \{f \geq \delta\}$ up to a φ^2-null set, then $\varphi(A) \times \varphi(B) = 0$.

Construction. Let $\mathcal{R}$ consist of the rationals in the open interval (0, 2). Let x be the point (x_1, x_2) in $\mathcal{X}^2$. The basic idea is to set $f(x) = 1$ unless $x_1 + x_2 \in \mathcal{R}$, otherwise $f(x) = 0$. Then $f > 0$ includes no measurable rectangle of positive measure. However,

$f(x) = 0$ for some x, and $f \geq 1/2$ includes the whole unit square—up to a null set. To remove these blemishes, we redefine f, as follows. Order $\mathcal{R}$ as $r_1, r_2, \ldots$. Fix a sequence of small positive numbers ϵ_i such that $\sum_i \epsilon_i < 1/10$. Let V_i consist of all $x = (x_1, x_2) \in \mathcal{X}^2$ such that $r_i - \epsilon_i < x_1 + x_2 < r_i + \epsilon_i$.

Define f_n inductively: $f_0(x) = 1$ for all x; $f_{n+1}(x) = f_n(x)$ except on V_{n+1}, where $f_{n+1}(x) = f_n(x)/2$. Of course, $f_n(x)$ is non-negative, and non-decreasing with n. Let $f_\infty(x) = \lim_n f_n(x)$. Plainly, $f_\infty(x) > 0$ unless $x \in V_i$ for infinitely many i; this set of x has Lebesgue measure 0. Let $f = f_\infty$ where $f_\infty > 0$, and $f = 1$ where $f_\infty = 0$. To verify the claimed properties of f, write $\star$ for convolution. Let A and B be Borel sets of positive measure. Then $g = 1_A \star 1_B$ is a continuous, non-negative, and non-vanishing function on (0, 2). In particular, $g > 0$ on $(r_i - \epsilon_i, r_i + \epsilon_i)$ for infinitely many i. Consequently, there must be infinitely many i for which $\varphi^2\{(A \times B) \cap V_i\} > 0$. Further details are omitted.

For the hit-and-run kernel in R^2, the density of $Q_x(dy)$ with respect to $\phi(dy)$ is

$$\text{const.}/\|y - x\| M_{xy}.$$

There is an issue as to whether $\big\{(x, y) : x \in R^2, y \in R^2, \|y - x\| M_{xy} < N < \infty\big\}$ includes a "rectangle" $A \times B$, where A and B are planar Borel sets of positive Lebesgue measure. The construction in Example 1 can be modified to show that this is not necessarily so. Let λ be Lebesgue measure on the Borel subsets of $[0, 1]$.

Example 2. There is a positive measurable function f on $[0, 1]^2$ which has $\int f \, d\lambda^2 = 1$, and which has the following property. If A and B are any two Borel subsets of $[0, 1]^2$ with $\lambda(A) > 0$ and $\lambda(B) > 0$, and N is any finite positive number, then

$$\lambda^4\big\{(x, y) : x \in A \text{ and } y \in B \text{ and } M_{xy}(f) > N\big\} > 0.$$

Construction. Let $\mathcal{R}$ consist of all pairs of points with rational coordinates on the perimeter of $[0, 1]^2$; each point in a pair is to be on a different edge of the unit square. Order $\mathcal{R}$ as $r_1, r_2, \ldots$; each r_i consists of a pair of points. Fix a sequence of small positive numbers ϵ_i such that $\sum_i i\epsilon_i < \infty$. Let L_i be the line through r_i, and let V_i be the set of points in the unit square within a distance ϵ_i of L_i, measured perpendicular to L_i. Define the continuous function f_i on $[0, 1]^2$ as follows: $f_i(x) = i$ if x is on L_i; $f_i(x) = 0$ if x is outside V_i; and f_i is linearly interpolated in between, along the line perpendicular to L_i. Let $f = \sum_i f_i$, normalized so that $\int f(x) \, dx = 1$. We will refer to the V_i as "tubes."

Let A and B be Borel sets in $[0, 1]^2$. Suppose for now that their symmetric difference has positive Lebesgue measure. Let a be a Lebesgue point of $A' = A - B$; let b be a Lebesgue point of $B' = B - A$. Make sure that a, b are interior points of the unit square. Fix a positive number $\delta < \|a - b\|/8$. Find an open disk D_a around a, of radius less than δ, such that $\lambda^2(D_a \cap A') > .99\lambda^2(A')$. Likewise for D_b with $b \in B'$. Make sure that D_a and D_b are wholly within the unit square.

Let ρ be the smaller of the two radii (of D_a and D_b). There will be infinitely many n such that $\epsilon_n < \rho/16$, while L_n passes within $\rho/16$ of a and of b. Take any such n. The tube V_n runs through D_a; however, at least 1/4 of the area of D_a lies above the tube. There must therefore be a set of x, of positive Lebesgue measure, with $x \in D_a \cap A'$, and x above the tube. Likewise, there must be a set of y, of positive Lebesgue measure, with

$y \in D_b \cap B'$, and y below the tube. In particular, the line L_{xy} cuts through the tube V_n, on a line segment of length at least $\|a - b\| - 2\delta > 6\delta$. Consequently, $f > n/2$ on a segment of L_{xy} of length at least 3δ, and $M_{xy} > n\delta$. Now $\delta > 0$ is fixed, while n is free; for sufficiently large n, $n\delta > N$, which completes the argument under the assumption that the symmetric difference between the two sets has positive Lebesgue measure. Essentially the same argument goes through if A coincides with B, provided the set has positive Lebesgue measure: just take two different Lebesgue points a, b in A.

Remarks.

(i) The function f is lower semi-continuous.

(ii) We can make $\sum_i \epsilon_i$ small. Then the unit square less the tubes—$[0,1]^2 - \bigcup_i V_i$—is large. And $\{M_{xy} > N\}$ wholly includes no tube. The "tube-crossing" argument in the proof is needed to overcome this difficulty.

5. The Saturn Construction

For simplicity, take $k = 2$. We are going to construct a density f for which Theorem 2 holds, but convergence will be arbitrarily slow. Let $A_n : n = 1, 2, \ldots$ be annuli centered at the origin; A_n has inner radius r_n which is large, outer radius $r_n + w_n$ where w_n is small, and mass μ_n spread uniformly over the area. The μ_n decrease to 0 slowly, but $\sum_n \mu_n = 1$. There is radial symmetry around 0. We are going to take $r_n = 10^n$, say, and show that $Q_x(A_{n+j}) \le 2/10^j$ for all $n, j = 1, 2, \ldots$, and all $x \in A_n$.

Consider an auxiliary random walk with IID increments X_i, where $P\{X_i = j\} = 2/10^j$ for $j = 1, 2, \ldots$ and $P\{X_i\} = 0$ with the remaining probability. We can couple the Saturn process and the auxiliary walk so that if Saturn moves forward, so does the walk: the coupling has to be done in time; results in space follow. The Saturn process will always be closer to 0 than the walk. (We view the Saturn process as being on the index set of the rings, which is permissible by radial symmetry; the walk and the process start from the same place, i.e., if the Saturn process starts in ring A_{n_0}, the walk start at n_0.)

Of course, $E(X_j) = \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} 2/10^j = 20/81$. The auxiliary walk therefore moves forward at rate 20/81: in other words, for any finite constant $C > 20/81$, there is a positive, finite D and ρ with $0 < \rho < 1$ such that

$$P\{X_1 + X_2 + \cdots + X_n > Cn\} < D\rho^n.$$

The variation distance between Q_x^n and the stationary distribution is bounded above by

$$D\rho^n + \sum_{i=Cn}^{\infty} \mu_i.$$

The displayed sum tends to 0 with arbitrary slowness.

Some geometric preliminaries

Fix r large and w small. Consider the annulus bounded by an inner circle of diameter r and an outer circle of diameter $r+w$; these circles have a common center at the origin of coordinates. Consider a line which cuts the annulus; we are interested in the length of the cut. A moment's reflection shows that we may assume the line to be above the horizontal axis and parallel to it; let s be the height of the line above the horizontal axis. Let $\phi(s)$ denote the length cut off on the annulus by the line. By Pythagoras' theorem,

$$\phi(s) = 2\left(\sqrt{(r+w)^2 - s^2} - \sqrt{r^2 - s^2}\right) \quad \text{for} \quad 0 \le s \le r.$$

Clearly, ϕ' is positive and increasing, so ϕ is convex increasing; alas, $\phi'(r) = \infty$. If $r \le s \le r+w$, then $\phi(s) = 2\sqrt{(r+w)^2 - s^2}$; further details are omitted. In particular, we have the following result.

Lemma 11. ϕ is convex increasing as s increases from 0 to r; then ϕ is convex decreasing.

Let $r_0 \le r/10$, and let w_0 be another small, positive number. Consider the small annulus bounded by circles of radius r_0 and $r_0 + w_0$. Suppose a line cuts the small annulus and the big annulus. How long is the cut on big annulus?

Lemma 12. The length of the cut is at least $2w$ and at most $2.02w$, provided w_0 and w are smaller than certain small positive constants.

Proof. This follows from Lemma 11. The lower bound is immediate on putting $s = 0$. For the upper bound, put $s = r_0 + w_0$. Let

$$\psi(x) = \sqrt{(r+x)^2 - (r_0+w_0)^2} - \sqrt{r^2 - (r_0+w_0)^2}.$$

Plainly, $\psi(0) = 0$. For $0 \le x \le w$ and w small,

$$\frac{\partial}{\partial x}\psi(x) = \frac{r+x}{\sqrt{(r+x)^2 - (r_0+w_0)^2}} < 1.01,$$

because $\limsup \psi'(x) \le 1/\sqrt{.99}$ as $x, w_0 \to 0$. In particular, if w_0, w are small and $x < w$, then $\psi(x) < 1.01x$.

Some estimates

We need the following estimates, uniform in $n = 1, 2, \ldots$; ϵ is a small positive number, which does not depend on n; we will require $\epsilon < 1/10$. It is understood that we may choose w_n small.

(a) The area of A_n is bounded between $2\pi r_n w_n$ and $(1+\epsilon)2\pi r_n w_n$ provided $w_n \le 2\epsilon r_n$.

(b) Then the density f_n on A_n is bounded between $\mu_n/[(1+\epsilon)2\pi r_n w_n]$ and $\mu_n/(2\pi r_n w_n)$.

(c) Draw a line through any point in A_n. The length cut off on A_{n+j} is bounded between $2w_{n+j}$ and $2.02w_{n+j}$ for $j = 1, 2, \ldots$. This follows from Lemma 12 above.

(d) Fix $x \in A_n$. Without loss of generality, put x on the horizontal axis. Draw a backward-sloping line through x, making acute angle α with the horizontal axis; thus, $0 \le \alpha \le \pi/2$. The length cut off on A_n is at least $2w_n$ unless $\cos\alpha < w_n/(r_n + w_n)$. The latter α are "bad"; the "good" α have $\cos\alpha > w_n/(r_n + w_n)$. The worst x in the present regard is on the outer boundary of A_n: see Lemma 11. (Then you consider the isosceles triangle with one long edge being the radius from 0 to x; the base of the triangle is the chord whose length is to be computed.)

(e) The chance of picking a bad α is at most w_n/r_n. Indeed, the construction starts at a point x in A_n, on the horizontal axis, and chooses a backward sloping angle α at random between 0 and $\pi/2$. Of course, $\cos\alpha = \sin(\pi/2 - \alpha)$; so the chance of picking a bad α is the chance that $\sin\alpha < w_n/(r_n + w_n)$. But $\sin\alpha > 2\alpha/\pi$, so the chance of getting a bad α is bounded above by $w_n/(r_n + w_n) < w_n/r_n$. (This only does quadrant #2; however, #3 follows by symmetry—and the other two quadrants follow because lines are bidirectional.)

(f) Recall that M_{xy} is the integral of f along the line through x and y. Fix $x \in A_n$ and $y \in A_{n+j}$: later, we will choose an angle α defining a line through x, and move to a random y on that line. For the good α,

$$M_{xy} \ge 2w_n f_n \ge \frac{\mu_n}{(1+\epsilon)\pi r_n}.$$

For the bad α,

$$M_{xy} \ge 2w_{n+1} f_{n+1} \ge \frac{\mu_{n+1}}{(1+\epsilon)\pi r_{n+1}}.$$

Fix $x \in A_n$ and $j = 1, 2, \ldots$. We compute $Q_x(A_n)$ by picking an angle $\alpha \in (0, \pi/2)$ and then a point along the corresponding line; let $L_{\alpha,n+j}$ be the length cut off by the line on A_{n+j}, and write $M_{x\alpha}$ for M_{xy}, the latter being constant for y on the α-line through x:

$$\begin{aligned}
Q_x(A_{n+j}) &= \frac{2}{\pi}\int_0^{\pi/2} \frac{f_{n+j}L_{\alpha,n+j}}{M_{x\alpha}}\,d\alpha \\
&= \frac{2}{\pi}\int_{\{\text{good }\alpha\}} \frac{f_{n+j}L_{\alpha,n+j}}{M_{x\alpha}}\,d\alpha + \frac{2}{\pi}\int_{\{\text{bad }\alpha\}} \frac{f_{n+j}L_{\alpha,n+j}}{M_{x\alpha}}\,d\alpha \\
&< 1.01(1+\epsilon)\frac{\mu_{n+j}}{\mu_n}\frac{r_n}{r_{n+j}} + 1.01(1+\epsilon)\frac{w_n r_{n+1}}{r_n^2}\frac{\mu_{n+j}}{\mu_{n+1}}\frac{r_n}{r_{n+j}} \\
&< 2\frac{r_n}{r_{n+j}}.
\end{aligned}$$

The last inequality holds provided

$$1.01(1+\epsilon)\Big(1 + \frac{w_n r_{n+1}}{r_n^2}\Big) < 2;$$

by assumption, μ_n is monotone decreasing with n, so e.g. $\mu_{n+j}/\mu_n \leq 1$.

Details for bounding the integral over the bad α are omitted; details for the good α follow below.

$$\frac{2}{\pi}\int_{\{\text{good } \alpha\}} d\alpha < \frac{w_n}{r_n} \quad \text{by (e).}$$

$$\frac{1}{M_{x\alpha}} < \frac{(1+\epsilon)\pi r_n}{\mu_n} \quad \text{by (f).}$$

$$f_{n+j} < \frac{\mu_{n+j}}{2\pi r_{n+j} w_{n+j}} \quad \text{by (b).}$$

$$L_{\alpha,n+j} < 2.02 w_{n+j} \quad \text{by (c).}$$

6. Literature Review

Extending results in Smith (1984), Bélisle, Romeijn and Smith (1993) show convergence for a density on a compact set; the density is positive everywhere, bounded above, and continuous a.e. No rates are established. A fairly general probability on the unit sphere S_k is used to choose the direction in which to move; our arguments may handle this case. Athreya, Doss and Sethuraman (1996) show convergence and slow convergence for other algorithms; Athreya and Ney (1978) formulate the "local Doeblin condition" in their definition (2.2), and derive convergence from the renewal theorem. Bélisle (1998) shows slow convergence for the Gibbs sampler. Eaton (1992) proves convergence theorems using reversibility; also see Smith (1984). Lamperti (1960) gives martingale recurrence conditions. The most accessible reference on Doeblin's general theory is perhaps Orey (1973); other references are Asmussen (1987), Doob (1953), Lindvall (1992), Meyn and Tweedie (1993), and Revuz (1984); Cohn (1993) gives an overview of the history. Doeblin (1940) and Harris (1956) should be mentioned.

The hit-and-run construction, as we have defined it, is feasible for choosing a point uniformly in a high-dimensional bounded, convex set. Otherwise, a Metropolis step could be taken along L_{xy}; the arguments would be about the same, because condition (6) would still hold. The Gibbs sampler may also be covered by our argument, although the proofs of Lemmas 8–9 would need to be reworked: $Q_x \perp \phi$, but in many circumstances Q_x^k has a large component that is equivalent to ϕ. In Diaconis and Freedman (1997), we consider more general versions of Doeblin's theory. Elsewhere, we hope to give a more abstract definition of the hit and run process, with additional examples.

If f has compact support but is unbounded, it is natural to ask whether there a rate of convergence. Bélisle (1998a) answers this in the negative. The idea is the following. In the unit square, put a sequence of points along the 45-degree line, converging to $(1,1)$. Put a disk around each point, with radius converging very rapidly to 0. Given any sequence $p_n > 0$ with $\sum_n p_n = 1$, spread mass p_n uniformly on the nth disk. From disk n, the angle subtended by disk $n+1$ is minute, so, it takes a long time to get from disk n to disk $n+1$.

References

Asmussen, S. (1987). *Applied Probability and Queues.* Wiley, New York.

Athreya, K.B., and Ney, P. (1978). A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* vol. 245 pp.493–501.

Athreya, K.B., Doss, H., and Sethuraman, H. (1996). On the convergence of the Markov chain simulation method. *Annals of Statististics* 24 69–100.

Bélisle, C. (1998). Slow convergence of the Gibbs sampler. Technical report, Department of Mathematics and Statistics, University of Laval at Quebec, Canada. To appear in the *Canadian Journal of Statistics.*

Bélisle, C. (1998a). Slow convergence of the hit-and-run sampler. Technical report, Department of Mathematics and Statistics, University of Laval at Quebec, Canada.

Bélisle, C., Romeijn, H.E., and Smith, R.L. (1993). Hit-and-run algorithms for generating multivariate distributions. *Mathematics of Operations Research* 18 255–266.

Cohn, H., ed. (1993). *Doeblin and Modern Probability: Proceedings of the Doeblin Conference.* American Mathematical Society, vol. 149.

Diaconis, P. and Freedman, D. (1997). On Markov chains with continuous state space. Technical Report No. 501, Department of Statistics, U.C. Berkeley.

Doeblin, W. (1940). Eléments d'une théorie générale des chaînes simple constantes de Markoff. *Ann. Sci. Ecole Norm. Sup.* 57 61–111.

Doob, J.L. (1953). *Stochastic Processes.* Wiley, New York.

Eaton, M.L. (1992). A statistical diptych: admissible inferences—recurrence of symmetric Markov chains. *Annals of Statistics* 20 1147–1179.

Harris, T.E. (1956). The existence of stationary measures for certain Markov processes. *Third Berkeley Symposium on Mathematical Statistics and Probability,* vol. II pp. 113–24.

Lamperti, J. (1960). Criteria for the recurrence or transience of stochastic processes, I. *J. Math. Anal. Appl.* 1 314–330.

Lindvall, T. (1992). *Lectures on the Coupling Method.* New York: Wiley.

Meyn, S.P. and Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability.* Springer, London.

Orey, S. (1971). *Lecture notes on limit theorems for Markov chain transition probabilities.* Mathematical Studies, 34. Van Nostrand Reinhold Co., New York.

Revuz, D. (1984). *Markov Chains.* North-Holland, Amsterdam.

Smith, R.L. (1984). Efficient Monte Carlo procedures for generating points uniformly distributed over random regions. *Operations Research* 32 1296–1308.

Technical Report no. 497
Department of Statistics
University of California
Berkeley, CA 94270