Hit-and-Run Algorithms for Generating Multivariate Distributions*

Claude J.P. Belisle
Department of Statistics
The University of Michigan
Ann Arbor, MI-48109-1027

H. Edwin Romeijn
Department of Operations Research and Tinbergen Institute
Erasmus University Rotterdam
3000 DR Rotterdam, The Netherlands

Robert L. Smith
Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor, MI 48109-2117

Technical Report 90-18

Hit–and–Run Algorithms for Generating Multivariate Distributions*

CLAUDE J. P. BÉLISLE
Department of Statistics
The University of Michigan
Ann Arbor, MI 48109-1027

H. EDWIN ROMEIJN
Department of Operations Research and Tinbergen Institute
Erasmus University Rotterdam
3000 DR Rotterdam, The Netherlands

ROBERT L. SMITH
Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109-2117

ABSTRACT

We introduce a general class of Hit–and–Run algorithms for generating absolutely continuous distributions on $\mathbf{R}^d$. They include the Hypersphere Directions algorithm and the Coordinate Directions algorithm that have been proposed for identifying nonredundant linear constraints and for generating uniform distributions over subsets of $\mathbf{R}^d$.

Given a bounded open set S in $\mathbf{R}^d$, an absolutely continuous probability distribution π on S (the target distribution) and an arbitrary probability distribution ν on the boundary ∂D of the d-dimensional unit sphere D centered at the origin (the direction distribution), the (ν, π)–Hit–and–Run algorithm produces a sequence of iteration points as follows. Given the n^{th} iteration point x, choose a direction θ according to the distribution ν and choose the $(n+1)^{st}$ iteration point according to the conditionalization of the distribution π along the line $\{x+\lambda\theta\,;\,\lambda \in \mathbf{R}\}$. (The Hypersphere Directions algorithm corresponds to the case where π is the uniform distribution on S and where ν is the uniform distribution on ∂D. The Coordinate Directions algorithm corresponds to the case where π is the uniform distribution on S and where ν is the distribution that assigns equal probability $1/2d$ to each of the $2d$ coordinate directions).

The (ν, π)–Hit–and–Run algorithm defines a Markov chain on S. Our first main result is that under some mild conditions on the density of π, this (ν, π)–Hit–and–Run Markov chain is time reversible with respect to π. It then follows that π is a stationary distribution for the chain. Our second main result is that under an appropriate condition on ν and S, the (ν, π)–Hit–and–Run Markov

*This work was supported in part by NATO Collaborative Research Grant No. 0119/89.
Key words and phrases. Hit–and–Run algorithms, Monte-Carlo methods, Markov chains, time reversibility, Harris-recurrence.

chain is Harris-Recurrent with respect to Lebesgue measure. It then follows from a theorem of Orey that for every initial distribution the (ν, π)–Hit–and–Run Markov chain converges in total variation, and hence in distribution, to its stationary distribution π.

1. Introduction

Let S be a bounded open subset of $\mathbf{R}^d$ and let π be an absolutely continuous probability measure on S. Let $f(x)$ be a probability density function for π and assume that it is bounded, almost everywhere continuous (with respect to Lebesgue measure on S), and strictly positive. Let D denote the d-dimensional unit sphere centered at the origin and let ∂D denote its topological boundary. Thus

$$D = \{x \in \mathbf{R}^d : \|x\| < 1\} \qquad \text{and} \qquad \partial D = \{x \in \mathbf{R}^d : \|x\| = 1\}.$$

Finally, let ν be an arbitrary probability measure on ∂D. The *Hit–and–Run algorithm* with *direction distribution* ν and with *target distribution* π (for short, the (ν, π)–Hit–and–Run algorithm on S) can be described as follows:

Step 0. Choose a starting point $x_0 \in S$ and set $k = 0$.

Step 1. Choose a direction θ_k on ∂D, with distribution ν.

Step 2. Choose $\lambda_k \in \Lambda_k = \{\lambda \in \mathbf{R} : x_k + \lambda\theta_k \in S\}$, from the distribution with density

$$f_k(\lambda) = \frac{f(x_k + \lambda\theta_k)}{\int_{\Lambda_k} f(x_k + r\theta_k)dr} \qquad \lambda \in \Lambda_k.$$

Step 3. Set $x_{k+1} = x_k + \lambda_k\theta_k$ and set $k = k + 1$.
Go to 1.

Figure 1 illustrates $f_k(\lambda)$, the (reparametrized) density function of the conditionalization of the distribution π along the line Λ_k. Geometrically, f_k is a cross section of the multivariate density function f, normalized to be a density function.

If ν is the uniform distribution on ∂D and if π is the uniform distribution on S the algorithm is known as the Hypersphere Directions algorithm. This special case was first suggested by Boneh and Golan [1979] in the context of non-redundant constraint identification and later independently by Smith [1980] for generating points uniformly distributed over S. Later Smith [1984] showed that, for S open and bounded, the sequence of iteration points of the Hypersphere Directions algorithm converges to the uniform distribution on S. If ν is the discrete distribution that assigns mass $1/2d$ to each of the $2d$ coordinate directions and if π is the uniform distribution on S, the algorithm is known as the Coordinate Directions algorithm. This special case was suggested by Telgen [1980]. Later Berbee et

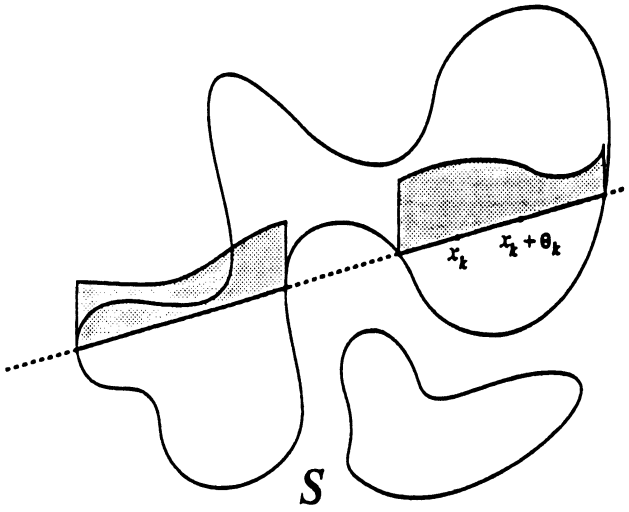

Figure 1. The conditionalization of π along the line Λ_k

al. [1987] showed that if S is a convex polyhedron then the sequence of iteration points of the Coordinate Directions algorithm converges to the uniform distribution on S. Recently Boender et al. [1989] have introduced a related class of algorithms, known as Shake–and–Bake algorithms, for generating points that are (asymptotically) uniformly distributed on the boundary of a full-dimensional convex polyhedron. The Shake–and–Bake algorithms will not be investigated here.

Clearly the (ν, π)–Hit–and–Run algorithm defines a discrete time Markov chain on S with stationary transition probabilities. The main purpose of this paper is to investigate the convergence properties of this Markov chain and to generalize the results of Smith [1984] and Berbee et al. [1987] in the following three directions: a) the region S will be an arbitrary bounded open subset of $\mathbf{R}^d$, b) the direction distribution ν will be completely arbitrary, c) the target distribution π will be an arbitrary absolutely continuous distribution on S with a bounded, almost everywhere continuous, and strictly positive density.

The paper is organized as follows. In section 2 we give a precise analytical description of the (ν, π)–Hit–and–Run Markov chain. In section 3 we show that the (ν, π)–Hit–and–Run Markov chain is time reversible with respect to π. This implies that π is a stationary distribution for the (ν, π)–Hit–and–Run Markov chain. In section 4 we show that under an appropriate communication structure, described in terms of ν and S, the stationary distribution π is unique and the (ν, π)–Hit–and–Run Markov chain is Harris-recurrent (with respect to Lebesgue measure on S). Using a theorem of Orey, we then conclude that for every starting point x, the (ν, π)–Hit–and–Run Markov chain $(X_n; n \geq 0)$ converges in total variation to the stationary distribution π, i.e.

$$\lim_{n \to \infty} \mathbf{P}\left[X_n \in B | X_0 = x\right] = \pi(B) \qquad \forall x \in S,\ \forall B \in \mathcal{B}_S,$$

where $\mathcal{B}_S$ denotes the Borel σ-field on S. Section 5 is a brief discussion of some generalizations and practical implications of our results.

2. Hit–and–Run Markov kernels

We begin by recalling some standard definitions. Let $(S, \mathcal{B})$ be a measurable space, i.e. S is an arbitrary set and $\mathcal{B}$ is a σ-field on S. A *kernel* on $(S, \mathcal{B})$ is a nonnegative function, say K, defined on $S \times \mathcal{B}$, such that

(i) $\forall x \in S$, $K(x, \cdot)$ is a σ-finite measure on $\mathcal{B}$.

(ii) $\forall A \in \mathcal{B}$, $K(\cdot, A)$ is a measurable function on S.

A *substochastic kernel* is a kernel K satisfying $K(x, S) \leq 1$ for every x. A *stochastic* (or *Markov*) *kernel* is a kernel K satisfying $K(x, S) = 1$ for every x. See e.g. Nummelin [1984] section 1, Orey [1971] chapter 1 section 0, Revuz [1975] chapter 1 section 1.

Now let S, π, f and ν be as described at the beginning of section 1 and let $\mathcal{B}_S$ be the Borel σ-field on S. Let Θ and U be random variables, defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, such that Θ has distribution ν, such that U is uniformly distributed over the interval $(0,1)$, and such that Θ and U are independent. For $x, y \in S$, with $x \neq y$, let

$$\Lambda(x,y) = \left\{ \lambda \in \mathbf{R} : x + \lambda \frac{y-x}{\|y-x\|} \in S \right\}$$

and let $f_{(x,y)}$ be the p.d.f. on $\mathbf{R}$ defined by

$$f_{(x,y)}(\lambda) = \frac{f(x + \lambda(y-x)/\|y-x\|)}{\int_{\Lambda(x,y)} f(x + r(y-x)/\|y-x\|)dr} \qquad \text{if } \lambda \in \Lambda(x,y)$$

and $f_{(x,y)}(\lambda) = 0$ if $\lambda \notin \Lambda(x,y)$. Let $F_{(x,y)}$ denote the c.d.f. of $f_{(x,y)}$, i.e.,

$$F_{(x,y)}(\lambda) = \int_{-\infty}^{\lambda} f_{(x,y)}(u)du.$$

The function

$$P(x,A) = \mathbf{P}\left[x + \left(F^{-1}_{(x,x+\Theta)}(U)\right)\Theta \in A\right]$$

defines a Markov kernel on $(S, \mathcal{B}_S)$. Here $F^{-1}_{(x,y)}$ is the usual right continuous inverse of $F_{(x,y)}$. Since Θ has distribution ν and since U and Θ are independent, the right hand side of the last equation can be written as

$$\int_{\partial D} \mathbf{P}\left[x + \left(F^{-1}_{(x,x+\theta)}(U)\right)\theta \in A\right] \nu(d\theta).$$

Furthermore, since $F^{-1}_{(x,x+\theta)}(U)$ is a $\Lambda(x, x+\theta)$-valued random variable with p.d.f. $f_{(x,x+\theta)}$, the above Markov kernel is the one-step transition probability of the Hit–and–Run algorithm described in section 1. This motivates the following formal definition:

DEFINITION 1. *The Markov kernel*

$$P(x, A) = \int_{\partial D} \mathbf{P}\left[x + \left(F^{-1}_{(x,x+\theta)}(U)\right)\theta \in A\right] \nu(d\theta) \qquad x \in S,\ A \in \mathcal{B}_S$$

will be called the (ν, π)-Hit-and-Run Markov kernel on S. *The probability measure* ν *will be called the* direction distribution. *The probability measure* π *will be called the* target distribution.

3. SYMMETRY, TIME REVERSIBILITY AND STATIONARY DISTRIBUTIONS

Let P be a Markov kernel on an arbitrary measurable space $(S, \mathcal{B})$ and let μ be a probability measure on $(S, \mathcal{B})$.

DEFINITION 2. *The Markov kernel* P *is said to be* time reversible *with respect to the probability measure* μ *if*

$$\int_A P(x, B)\mu(dx) = \int_B P(x, A)\mu(dx) \qquad \forall A, B \in \mathcal{B}.$$

DEFINITION 3. *The probability measure* μ *is said to be a* stationary distribution *for the Markov kernel* P *if*

$$\mu(A) = \int_S P(x, A)\mu(dx) \qquad \forall A \in \mathcal{B}.$$

Definition 3 is standard. Definition 2 is a straightforward generalization of the concept of time reversibility for discrete state space Markov chains (See e.g. Ross [1983], page 126) and it is consistent with the concept of time reversibility for general stochastic processes as discussed e.g. in Kelly [1979], page 5. A function $p(x, y)$ defined on S^2 will be called *symmetric* if $p(x, y) = p(y, x)$ for every x and y in S. The following propositions are elementary consequences of the definitions.

PROPOSITION 1. *Suppose that* P *is of the form*

$$P(x, A) = \int_A p(x, y)\mu(dy) \qquad \forall x \in S, \forall A \in \mathcal{B}$$

for some jointly measurable symmetric *function* $p(x, y)$ *on* S^2. *Then* P *is time reversible with respect to* μ.

PROOF: Under the given assumptions, an application of Fubini's theorem yields

$$\begin{aligned}\int_A P(x,B)\mu(dx) &= \int_A \int_B p(x,y)\mu(dy)\mu(dx) \\ &= \int_B \int_A p(x,y)\mu(dx)\mu(dy) \\ &= \int_B \int_A p(y,x)\mu(dx)\mu(dy) \\ &= \int_B P(y,A)\mu(dy) \qquad \forall A, B \in \mathcal{B}.\end{aligned}$$

Thus P is time reversible with respect to μ. ∎

PROPOSITION 2. *If P is time reversible with respect to μ then μ is a stationary distribution for P.*

PROOF: Assuming time reversibility with respect to μ yields

$$\begin{aligned}\int_S P(x,A)\mu(dx) &= \int_A P(x,S)\mu(dx) \\ &= \int_A \mu(dx) \\ &= \mu(A) \qquad \forall A \in \mathcal{B}.\end{aligned}$$

Thus μ is stationary for P. ∎

We now return to the case where S is a bounded Borel subset of $\mathbf{R}^d$, where ν is an arbitrary probability measure on ∂D and where π is an absolutely continuous probability measure on S with a bounded, almost everywhere continuous, and strictly positive density $f(x)$. For every Borel set $G \subset \partial D$, and for every $0 \leq r_1 \leq r_2 < \infty$, let

$$G \otimes (r_1, r_2) = \{y \in \mathbf{R}^d : y = \lambda\theta \text{ for some } \theta \in G \text{ and } \lambda \in (r_1, r_2)\}.$$

Figure 2 illustrates the set $G \otimes (r_1, r_2)$.

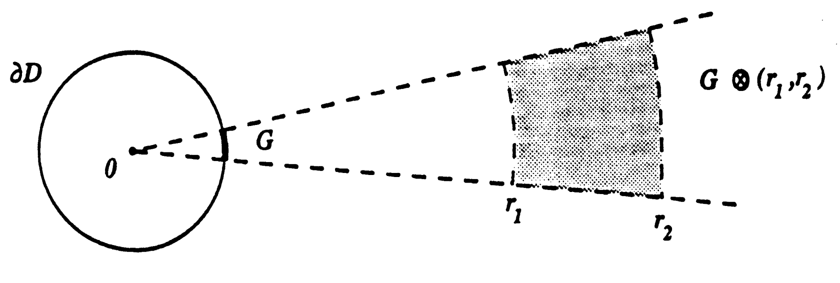

Figure 2. An illustration of the set $G \otimes (r_1, r_2)$

Let $m(x,A)$ be the unique kernel, on $(\mathbf{R}^d, \mathcal{B}_{\mathbf{R}^d})$, satisfying

$$m(x, x + G \otimes (r_1, r_2)) = (\nu(-G) + \nu(G))(r_2 - r_1). \tag{1}$$

Thus $m(0,\cdot)$ is the measure with polar infinitesimal volume element $(\nu(-d\theta) + \nu(d\theta))dr$ and $m(x,\cdot)$ is just a translation of the measure $m(0,\cdot)$. The next proposition gives us an explicit analytical expression for the Hit–and–Run Markov kernel.

PROPOSITION 3.

(a) *The (ν,π)–Hit–and–Run Markov kernel can be written as*

$$P(x,A) = \int_A \frac{f(y)}{\int_{\Lambda(x,y)} f(x + r(y-x)/\|y-x\|)dr} m(x,dy).$$

(b) *If ν is absolutely continuous with respect to the uniform probability distribution on ∂D then the (ν,π)–Hit–and–Run Markov kernel can be written as*

$$P(x,A) = \int_A \frac{h((y-x)/\|y-x\|) + h((x-y)/\|x-y\|)}{\int_{\Lambda(x,y)} f(x + r(y-x)/\|y-x\|)dr \; \|y-x\|^{d-1} C_d} \pi(dy). \tag{2}$$

where $h(\theta)$ denotes the density of ν with respect to the uniform probability distribution on ∂D and where $C_d = 2\pi^{d/2}/\Gamma(d/2)$ (i.e. C_d is the surface area of ∂D).

PROOF: Part (a) is an immediate consequence of the definitions:

$$\begin{aligned}
P(x,A) &= \int_{\partial D} \mathbf{P}\left[x + (F^{-1}_{(x,x+\theta)}(U))\theta \in A\right] \nu(d\theta) \\
&= \int_{\partial D} \int_{\Lambda(x,x+\theta)} 1_A(x+\lambda\theta) \frac{f(x+\lambda\theta)}{\int_{\Lambda(x,x+\theta)} f(x+r\theta)dr} d\lambda\, \nu(d\theta) \\
&= \int_A \frac{f(y)}{\int_{\Lambda(x,y)} f(x + r(y-x)/\|y-x\|)dr} m(x,dy). \qquad (3)
\end{aligned}$$

Here, and throughout the rest of this paper, 1_A denotes the indicator function of the set A. The first equality is Definition 1. The second equality follows from the fact that $F^{-1}_{(x,x+\theta)}(U)$ is a $\Lambda(x, x+\theta)$-valued random variable with p.d.f.

$$f_{(x,x+\theta)}(\lambda) = \frac{f(x+\lambda\theta)}{\int_{\Lambda(x,x+\theta)} f(x+r\theta)dr}.$$

The last equality can be justified as follows. If B is a Borel subset of S of the form $B = x + G \otimes (r_1, r_2)$ then

$$\int_{\partial D} \int_{\Lambda(x,x+\theta)} 1_B(x+\lambda\theta) d\lambda \nu(d\theta) = \int_S 1_B(y) m(x,dy)$$

since both sides are equal to $m(x, x+G\otimes(r_1,r_2))$. A standard measure theoretic argument (see e.g. Breiman [1968] Proposition 2.23) then yields

$$(4) \qquad \int_{\partial D}\int_{\Lambda(x,x+\theta)} g(x+\lambda\theta)d\lambda\nu(d\theta) = \int_S g(y)m(x,dy)$$

for every non-negative measurable function g defined on S. This holds in particular for the function

$$g(y) = \frac{1_A(y)f(y)}{\int_{\Lambda(x,y)} f(x+r(y-x)/\|y-x\|)dr}$$

in which case (4) reduces to (3).

Now consider part (b). If ν is of the form

$$\nu(G) = \int_G h(\theta)\nu_o(d\theta)$$

where ν_o is the uniform distribution on ∂D, then the kernel $m(x, A)$ can be written as

$$m(x,A) = \int_A \frac{h((y-x)/\|y-x\|) + h((x-y)/\|y-x\|)}{\|y-x\|^{d-1}C_d}dy$$

and therefore the (ν,π)–Hit–and–Run Markov kernel can be written as

$$\begin{aligned} P(x,A) &= \int_A \frac{h((y-x)/\|y-x\|) + h((x-y)/\|y-x\|)}{\int_{\Lambda(x,y)} f(x+r(y-x)/\|y-x\|)dr\,\|y-x\|^{d-1}C_d}f(y)dy \\ &= \int_A \frac{h((y-x)/\|y-x\|) + h((x-y)/\|y-x\|)}{\int_{\Lambda(x,y)} f(x+r(y-x)/\|y-x\|)dr\,\|y-x\|^{d-1}C_d}\pi(dy). \quad \blacksquare \end{aligned}$$

Proposition 3 is the key step towards our first main result:

THEOREM 1. *The (ν,π)–Hit–and–Run Markov kernel is time reversible with respect to π.*

PROOF: In the absolutely continuous case the result follows from Proposition 1 and part (b) of Proposition 3 (since the integrand in (2) is symmetric).

Now consider the general case. Let ν be an arbitrary probability measure on ∂D. Let $\nu_1, \nu_2, \nu_3, \ldots$ be a sequence of probability measures on ∂D such that ν_n converges weakly to ν as $n \to \infty$ (for a definition of weak convergence, see e.g. Billingsley [1968]) and such that ν_n is absolutely continuous with respect to the uniform distribution on ∂D. (The fact that such a sequence exists can be seen as follows. Let Θ be a ∂D-valued random variable with distribution ν, let U_n be an $\mathbf{R}^d$-valued random variable with uniform distribution over the open ball of radius $1/n$ centered at the origin and assume that Θ and U_n are independent. Let ν_n be the distribution of $(\Theta + U_n)/\|\Theta + U_n\|$. Then ν_n has the desired properties). Let $m_n(x, A)$ be the kernel induced by ν_n, as in (1), and let $P_n(x, A)$ denote the (ν_n, π)–Hit–and–Run Markov kernel. Since ν_n is absolutely continuous, $P_n(x, A)$ is time reversible:

$$\text{(5)} \qquad \int_A P_n(x, B)\pi(dx) = \int_B P_n(x, A)\pi(dx) \qquad \forall A, B \in \mathcal{B}_S.$$

Furthermore, since ν_n converges weakly to ν, the measure $m_n(x, \cdot)$ converges weakly to the measure $m(x, \cdot)$, for every x in S. Now from part (a) of Proposition 3 we have

$$\text{(6)} \qquad P_n(x, G) = \int_S \frac{1_G(y) f(y)}{\int_{\Lambda(x,y)} f(x + r(y-x)/\|y - x\|)dr} m_n(x, dy) \qquad \forall G \in \mathcal{B}_S.$$

If G is a finite intersection of open balls then one can show that for almost every $x \in S$ the set of discontinuity points of the function

$$y \to \frac{1_G(y) f(y)}{\int_{\Lambda(x,y)} f(x + r(y-x)/\|y - x\|)dr}$$

has $m(x, \cdot)$ measure zero. This implies that the right hand side of (6) converges, as $n \to \infty$, to

$$\int_S \frac{1_G(y) f(y)}{\int_{\Lambda(x,y)} f(x + r(y-x)/\|y - x\|)dr} m(x, dy)$$

(see e.g. Billingsley [1968], section 5). Thus, if G is a finite intersection of open balls then

$$\lim_{n \to \infty} P_n(x, G) = P(x, G) \qquad \text{for almost all } x \in S.$$

Thus, letting $n \to \infty$ in (5) and using the Lebesgue dominated convergence theorem (see e.g. Billingsley [1986] Theorem 16.4), we obtain

$$\text{(7)} \qquad \int_A P(x, B)\pi(dx) = \int_B P(x, A)\pi(dx)$$

whenever A and B are finite intersections of open balls in S. To complete the proof, we need to show that (7) holds for every A and $B \in \mathcal{B}_S$. This can be done via a standard extension argument. Fix A, a finite intersection of open balls in S, and let $\mathcal{L}_A = \{B \in \mathcal{B}_S : \int_A P(x, B)\pi(dx) = \int_B P(x, A)\pi(dx)\}$. Using the monotone convergence theorem (see e.g. Billingsley [1986] Theorem 16.2) one can easily verify that $\mathcal{L}_A$ is a λ-system (see Billingsley [1986] page 36). Since $\mathcal{L}_A$ contains all finite intersections of open balls, Dynkin's π-λ-theorem (see Billingsley [1986] Theorem 3.2) implies that $\mathcal{L}_A = \mathcal{B}_S$. Thus (7) holds whenever A is a finite intersection of open balls in S and B is Borel set in S. Now fix $B \in \mathcal{B}_S$ and let $\mathcal{L}_B = \{A \in \mathcal{B}_S : \int_A P(x, B)\pi(dx) = \int_B P(x, A)\pi(dx)\}$. The same argument yields $\mathcal{L}_B = \mathcal{B}_S$. Thus (7) holds for every $A, B \in \mathcal{B}_S$. ∎

In view of Proposition 2, the following result follows from Theorem 1.

THEOREM 2. *The probability measure π is stationary for the (ν, π)-Hit-and-Run Markov kernel.* ∎

It is easy to see that the stationary distribution of a (ν, π)-Hit-and-Run Markov kernel is not necessarily unique. The next section will present a necessary and sufficient condition for π to be the unique stationary distribution for the (ν, π)-Hit-and-Run Markov kernel.

4. THE MAIN LIMIT THEOREM

As before, let S be a bounded open set in $\mathbf{R}^d$, let π be an absolutely continuous probability measure on $(S, \mathcal{B}_S)$, assume that π possesses a density $f(x)$ which is bounded, almost everywhere continuous, and strictly positive on S, and let ν be a probability measure on ∂D. By Theorem 2, π is stationary for the (ν, π)-Hit-and-Run Markov kernel $P = (P(x, B); x \in S, B \in \mathcal{B}_S)$. The purpose of this section is to show that under an appropriate condition on ν and S, π is the unique stationary distribution for the Markov kernel P and for every initial distribution the Hit-and-Run Markov chain converges to π. This will be achieved by showing that the Hit-and-Run Markov chain is Harris-recurrent (see the definition below) and by using a well known theorem of Orey. The Hit-and-Run Markov chain will be denoted $(X_n; n \geq 0)$ and we will write $\mathbf{P}_x$ to denote conditional probability given $X_0 = x$. We begin with some preliminary results on the communicating structure of the Hit-and-Run Markov kernel. For every non-negative integer n, let $P^n = (P^n(x, B); x \in S, B \in \mathcal{B}_S)$ denote the n-step Hit-and-Run Markov kernel. Thus

$$P^0(x, B) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \notin B \end{cases}$$

$$P^1(x, B) = P(x, B)$$

and for $n \geq 2$, $$P^n(x, B) = \int_S P^{n-1}(y, B) P(x, dy) = \int_S P(y, B) P^{n-1}(x, dy).$$

Recall that $P^n(x, B)$ has a unique decomposition

$$\begin{aligned} P^n(x, B) &= P^n_{\text{sing}}(x, B) + P^n_{\text{abs}}(x, B) \\ &= P^n_{\text{sing}}(x, B) + \int_B p_n(x, y) dy \end{aligned}$$

where $P^n_{\text{sing}}(x, B)$ is a singular substochastic kernel, where $P^n_{\text{abs}}(x, B)$ is an absolutely continuous substochastic kernel and where $p_n(x, y)$ is jointly measurable. (See Orey [1971] section 1.1). If $P^n_{\text{abs}}(x, S) > 0$ then we say that the probability measure $P^n(x, \cdot)$ has an absolutely continuous part. If the density $p_n(x, y)$ is positive for almost all y's in some Borel set B, then we say that the probability measure $P^n(x, \cdot)$ *spreads* over B.

PROPOSITION 4. *Let G be an open subset of S. Then*

$$P(y,G) > 0 \qquad \forall y \in G.$$

PROOF: From Proposition 3,

$$P(y,G) = \int_G \frac{f(z)}{\int_{\Lambda(y,z)} f(y + r(z-y)/\|z-y\|)dr} m(y,dz)$$

The above integrand is strictly positive. Furthermore, if G is open and $y \in G$, then $m(y,G) > 0$. Thus the above integral is strictly positive. ∎

Recall that the *support* of a probability measure μ on $\mathbf{R}^d$, to be denoted $\mathrm{Supp}(\mu)$, is defined as

$$\mathrm{Supp}(\mu) = \{x \in \mathbf{R}^d : \mu(B(x,\epsilon)) > 0, \forall \epsilon > 0\}$$

where $B(x,\epsilon)$ is the open ball of radius ϵ centered at x.

PROPOSITION 5.

$$Supp(P^n_{\mathbf{abs}}(x,\cdot)) \subset Supp(P^{n+1}_{\mathbf{abs}}(x,\cdot)) \qquad \forall x \in S,\ \forall n \geq 1.$$

PROOF: Fix $n \geq 1$ and $x \in S$, and suppose that $v \in \mathrm{Supp}(P^n_{\mathbf{abs}}(x,\cdot))$. Fix $\epsilon > 0$ small enough so that $B = B(v,\epsilon) \subset S$. Then

$$\begin{aligned} P^{n+1}(x,B) &= \int_S P(y,B)P^n(x,dy) \\ &= \int_S P(y,B)P^n_{\mathbf{sing}}(x,dy) + \int_S P(y,B)P^n_{\mathbf{abs}}(x,dy). \end{aligned}$$

By Proposition 4 we have $P(y,B) > 0$ for every y in B and since $v \in \mathrm{Supp}(P^n_{\mathbf{abs}}(x,\cdot))$, we obtain $P^n_{\mathbf{abs}}(x,B) > 0$. Thus the second integral on the right hand side of the last equality is strictly positive. This implies that $P^{n+1}_{\mathbf{abs}}(x,B) > 0$. Since this holds for every $\epsilon > 0$, we conclude that $v \in \mathrm{Supp}(P^{n+1}_{\mathbf{abs}}(x,\cdot))$. Thus $\mathrm{Supp}(P^n_{\mathbf{abs}}(x,\cdot)) \subset \mathrm{Supp}(P^{n+1}_{\mathbf{abs}}(x,\cdot))$. ∎

DEFINITION 4. *The probability measure ν is said to be* full dimensional *if $Supp(\nu)$ contains a set of vectors that span $\mathbf{R}^d$ (in other words if the set $\{0\} \cup Supp(\nu)$ is not contained in any linear subspace of dimension $d-1$).*

PROPOSITION 6. *Suppose that ν is full dimensional. Then* *(a) $\forall n \geq d$ and $\forall x \in S$, the probability measure $P^n(x,\cdot)$ has an absolutely continuous part.*

(b) $\forall x \in S$, $\lim_{n\to\infty} P^n_{\text{abs}}(x,S) = 1$.

(c) There exists an $r > 0$, depending only on ν, such that for all $x \in S$ and for all $n \geq d$, the probability measure $P^n_{\text{abs}}(x,\cdot)$ has a density which is strictly positive on $B(x, r\gamma_x)$, where $\gamma_x = \inf_{y\in S^c} \|x-y\|$ (and where S^c denotes the complement of S).

PROOF: Let L_n denote the subset of $(\partial D)^n$ consisting of those points $(\theta_1, \theta_2, ..., \theta_n)$ for which there exists integers $1 \leq i_1 < i_2 < ... < i_d \leq n$ such that $\theta_{i_1}, \theta_{i_2}, ..., \theta_{i_d}$ are linearly independent in $\mathbf{R}^d$. Let $\Theta_1, \Theta_2, \Theta_3, ...$ be the successive directions taken by the (ν, π)–Hit–and–Run Markov chain. Since $\Theta_1, \Theta_2, \Theta_3, ...$ are independent random variables with distribution ν and since ν is full dimensional,

$$\begin{aligned} &\mathbf{P}\left[(\Theta_1, \Theta_2, ..., \Theta_n) \in L_n\right] > 0 \qquad \forall n \geq d \\ \text{and} \qquad &\mathbf{P}\left[(\Theta_1, \Theta_2, ..., \Theta_n) \in L_n\right] \to 1 \qquad \text{as } n \to \infty. \end{aligned}$$

One of the most important features of the Hit–and–Run Markov chain is that if $(\theta_1, \theta_2, ..., \theta_n) \in L_n$ then for every x in S the conditional distribution

$$\mathbf{P}_x\left[X_n \in G | (\Theta_1, \Theta_2, ..., \Theta_n) = (\theta_1, \theta_2, ..., \theta_n)\right]$$

is absolutely continuous with respect to Lebesgue measure on S. Thus parts (a) and (b) follow from the fact that

$$P^n(x,G) = \int_{(\partial D)^n} \mathbf{P}_x\left[X_n \in G | (\Theta_1, \Theta_2, ..., \Theta_n) = (\theta_1, \theta_2, ..., \theta_n)\right] \nu(d\theta_1)\nu(d\theta_2)...\nu(d\theta_n).$$

Now consider linearly independent $\theta_1, \theta_2, ..., \theta_d$ in Supp(ν). By continuity, we can choose $r > 0$ and $\epsilon > 0$ small enough so that the sets

$$A_k = B(\theta_k, \epsilon) \cap \partial D \qquad k = 1, 2, ..., d$$

are linearly independent (in the sense that $u_1, u_2, ..., u_d$ are linearly independent whenever $u_k \in B(\theta_k, \epsilon) \cap \partial D$, $k = 1, 2, ..., d$) and

$$B(0,r) \subset \left\{ \sum_{k=1}^{d} \alpha_k u_k : -1 < \alpha_k < 1, k = 1, 2, ..., d \right\}$$

for every $(u_1, u_2, ..., u_d) \in A_1 \times A_2 \times ... \times A_d$. Then for every $(u_1, u_2, ..., u_d) \in A_1 \times A_2 \times ... \times A_d$, the conditional distribution

$$\mathbf{P}_x\left[X_d \in G | (\Theta_1, \Theta_2, ..., \Theta_d) = (u_1, u_2, , ..., u_d)\right]$$

is absolutely continuous and it has a density which is strictly positive on $B(x, r\gamma_x)$. Since $\Theta_1, \Theta_2, ..., \Theta_d$ are independent random vectors with common distribution ν and since $\theta_1, \theta_2, ..., \theta_d$ are in the support of ν,

$$\mathbf{P}\left[(\Theta_1, \Theta_2, ..., \Theta_d) \in A_1 \times A_2 \times ... \times A_d\right] = \Pi_{i=1}^{d} \mathbf{P}\left[\Theta_i \in A_i\right] = \Pi_{i=1}^{d} \nu(A_i) > 0.$$

Thus $P^d_{\text{abs}}(x, \cdot)$ has a density which is strictly positive on $B(x, r\gamma_x)$. Combined with Proposition 5, this proves part (c). ▮

Let φ be a σ-finite measure on $(S, \mathcal{B}_S)$. Recall that a Markov kernel P is called *φ-irreducible* if

$$\sum_{n=1}^{\infty} P^n(x, B) > 0 \qquad \forall x \in S$$

whenever $\varphi(B) > 0$. In terms of the associated Markov chain $(X_n; n \geq 0)$, φ-irreducibility is equivalent to the statement that

$$\mathbf{P}_x\left[X_n \in B \text{ for some } n \geq 1\right] > 0 \qquad \forall x \in S$$

whenever $\varphi(B) > 0$. Throughout the rest of this section φ will denote the Lebesgue measure on $(S, \mathcal{B}_S)$.

Recall that a topological space T is said to be *connected* if the only subsets of T which are both open and closed are the empty set and T itself. It is said to be *pathwise connected* if for every $a, b \in T$ there exists a continuous function $f : [0,1] \to T$ such that $f(0) = a$ and $f(1) = b$. If T is an open subset of $\mathbf{R}^d$, then connectedness and path connectedness are equivalent. Furthermore every open subset T of $\mathbf{R}^d$ is a union of at most countably many disjoint open connected sets called the connected components of T.

PROPOSITION 7. *Suppose that ν is full dimensional. Let A be a connected component of S. Let G be a measurable subset of A with positive Lebesgue measure. Then*

(a) *For all $x \in A$, there exists an integer n such that $P^n(x, G) > 0$.*

(b) *If $m(\cdot)$ is a measure on $(S, \mathcal{B}_S)$ such that $m(A) > 0$, then there exists an integer n such that*

$$\int_A P^n(x, G) m(dx) > 0.$$

PROOF: Fix $x \in A$. Let G be a measurable subset of A with positive Lebesgue measure. Choose $y \in A$ such that for every $\epsilon > 0$ the set $B(y, \epsilon) \cap G$ has positive Lebesgue measure. Let $g : [0,1] \longrightarrow S$ be such that $g(0) = x$, $g(1) = y$, and g is continuous on $[0,1]$. Let

$$\gamma = \inf\{\|u - v\|;\ u \in g([0,1]), v \in A^c\}.$$

Observe that γ is strictly positive since it is the distance between the compact set $g([0,1])$ and the closed set A^c. Let r be as in proposition 6 and let $\epsilon_* = r\gamma$. Now consider open balls

$$B_i = B(x_i, \epsilon_*) \qquad i = 1, 2, \ldots, k$$

where $x_1 = x$, $x_k = y$, and $x_i \in B_{i-1} \cap g([0,1])$ for $i = 2, 3, \ldots, k$, as shown in Figure 3.

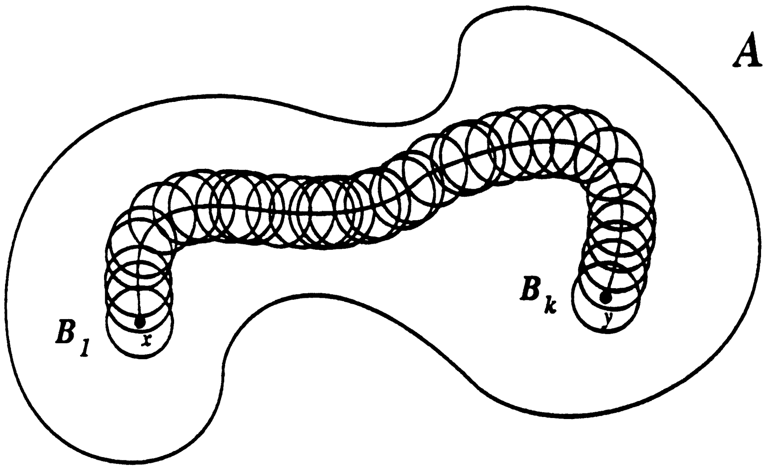

Figure 3. A path from x to y

Such a construction is possible since the set $g([0,1])$ is compact. It follows from Proposition 6 that $P^{id}(x,\cdot)$ spreads over B_i. In particular, $P^{kd}(x,\cdot)$ spreads over B_k. This implies that $P^{kd}(x,G) > 0$. This proves part (a). The argument actually shows that $P^{kd}(x',G) > 0, \forall x' \in B_1$. Thus if in the proof of part (a) we take $x \in \mathrm{Supp}(m)$, then we obtain

$$\int_A P^{kd}(x',G)m(dx') \geq \int_{B_1} P^{kd}(x',G)m(dx') > 0.$$

This proves part (b). ∎

DEFINITION 5. *The connected components of S are said to be ν-*communicating *if for every pair of connected components A and B there exists an $x \in A$ and an $n \geq 1$ such that*

$$P^n(x,B) > 0 \tag{8}$$

In particular if S is connected or if $\mathrm{Supp}(\nu) = \partial D$ then the connected components of S ν-communicate. If $d = 2$, $S = B((0,0),1) \cup B((10,10),1)$ and $\mathrm{Supp}(\nu) = \{(1,0),(0,1)\}$ then the connected components of S do not ν-communicate. If $d = 2$, $S = B((0,0),1) \cup$

$B((10,1),1) \cup B((10,10),1)$ and $\mathrm{Supp}(\nu) = \{(1,0),(0,1)\}$ then the connected components of S do ν-communicate.

Using part (a) of Proposition 3 and Theorems 5.2 and 5.5 of Billingsley [1968], one concludes that the Hit–and–Run Markov kernels and their iterates are continuous in the sense that if $x_k \to x$ then, for $n \geq 1$, the probability measure $P^n(x_k, \cdot)$ converges weakly, as $k \to \infty$, to the probability measure $P^n(x, \cdot)$. This implies that if B is an open set and if $k \to \infty$ then

$$\liminf_{k \to \infty} P^n(x_k, B) \geq P^n(x, B).$$

(See e.g. Billingsley [1968], Theorem 2.1). A consequence of this observation is that condition (8) is equivalent to the condition that for some $n \geq 1$

$$\varphi(\{x \in A : P^n(x, B) > 0\}) > 0. \tag{9}$$

PROPOSITION 8. *Suppose that ν is full dimensional and that the connected components of S are ν-communicating. Then the (ν, π)-Hit–and–Run Markov kernel P is φ-irreducible.*

PROOF: Fix $x \in S$. Fix $G \in \mathcal{B}_S$ with $\varphi(G) > 0$. We need to show that

$$P^n(x, G) > 0 \qquad \text{for some } n \geq 1. \tag{10}$$

Let A be the open connected component of S to which x belongs. Let B be an open connected component of S such that $\varphi(G \cap B) > 0$. If $A = B$, then (10) follows from part (a) of Proposition 7. If $A \neq B$, then let $n_2 \geq 1$ and $y \in A$ be such that $P^{n_2}(y, B) > 0$. By the remark following Definition 5, the set

$$C = \{z \in A : P^{n_2}(z, B) > 0\}$$

has positive Lebesgue measure. By part (a) of Proposition 7, there exists an integer $n_1 \geq 1$ such that $P^{n_1}(x, C) > 0$. We then obtain

$$\begin{aligned} P^{n_1+n_2}(x, B) &= \int_S P^{n_2}(z, B) P^{n_1}(x, dz) \\ &\geq \int_C P^{n_2}(z, B) P^{n_1}(x, dz) > 0 \end{aligned}$$

and therefore by part (b) of Proposition 7 there exists an integer $n_3 \geq 1$ such that

$$\int_B P^{n_3}(z, G \cap B) P^{n_1+n_2}(x, dz) > 0$$

Now if we take $n = n_1 + n_2 + n_3$, then

$$P^n(x,G) \geq P^n(x, G \cap B) \\ \geq \int_B P^{n_3}(z, G \cap B) P^{n_1+n_2}(x, dz) > 0,$$

i.e. (10) holds. ∎

Recall that a Markov kernel P is said to be *indecomposible* if there are no disjoint nonempty sets A and B such that

$$P(x,A) = 1 \ \forall x \in A \qquad \text{and} \qquad P(x,B) = 1 \ \forall x \in B$$

(See Breiman [1968]). It is easy to verify that φ-irreducibility implies indecomposability. Thus the following proposition is an immediate consequence of Proposition 8.

PROPOSITION 9. *Suppose that ν is full dimensional and that the connected components of S are ν-communicating. Then the (ν, π)–Hit–and–Run Markov kernel is indecomposible.* ∎

Theorem 7.16 of Breiman [1968] says that an indecomposible Markov chain posseses at most one stationary probability distribution. The result is stated for the case where the state space is a Borel subset of $\mathbf{R}$ but the proof that Breiman presents is also valid for the case where the state space is a Borel subset of $\mathbf{R}^d$. Thus Theorem 2 combined with Proposition 9 implies that if ν is full dimensional and if the connected components of S are ν-communicating then π is the unique stationary distribution for the (ν, π)–Hit–and–Run Markov kernel. It is easy to see that these two conditions are actually necessary for uniqueness of the stationary distribution. If the connected components of S do not ν-communicate then S is a union of two connected open sets, say A and B, that do not ν-communicate, and the probability measures defined by $\pi_A(G) = \pi(G \cap A)/\pi(A)$ and $\pi_B(G) = \pi(G \cap B)/\pi(B)$ are both stationary. If ν is not full dimensional then the linear subspace of $\mathbf{R}^d$ spanned by Supp(ν), say H, is less than d dimensional and any conditionalization of the probability measure π to a set of the form $x + H$, with $x \in S$, will be stationary. Thus we have proved the following refinement of Theorem 2:

THEOREM 3. *The probability distribution π is stationary for the (ν, π)–Hit–and–Run Markov kernel. Furthermore, it is the unique stationary distribution if and only if ν is full dimensional and the connected components of S are ν-communicating.* ∎

Orey [1971] introduced a notion of periodicity for general φ-irreducible Markov kernels. The next proposition states that Hit–and–Run Markov kernels are aperiodic in Orey's sense.

PROPOSITION 10. *Suppose that ν is full dimensional and that the connected components of S are ν-communicating. Then the (ν, π)-Hit–and–Run Markov kernel is aperiodic.*

PROOF: By Proposition 8 we have φ-irreducibility. Thus Theorem 3.1 of Orey (1971) applies. By part (c) of Proposition 6, we cannot have a cycle of length $k > 1$ (See Definition 3.2 of Orey (1971)). Thus the (ν, π)-Hit–and–Run Markov kernel is aperiodic. ∎

Finally, we recall that the Markov kernel P is said to be *φ-recurrent* (or *Harris-recurrent with respect to φ*) if the corresponding Markov chain $(X_n; n \geq 0)$ satisfies

$$\mathbf{P}_x \left[X_n \in B \text{ for some } n \geq 1\right] = 1 \qquad \forall x \in S$$

whenever $\varphi(B) > 0$. This is equivalent to the statement that

$$\mathbf{P}_x \left[\sum_{n=1}^{\infty} 1_B(X_n) = \infty\right] = 1 \qquad \forall x \in S$$

whenever $\varphi(B) > 0$.

THEOREM 4. *Suppose that ν is full dimensional and that the connected components of S are ν-communicating. Then the (ν, π)-Hit–and–Run Markov kernel P is Harris recurrent with respect to Lebesgue measure on S.*

PROOF: By Theorem 2, π is invariant for P and by Proposition 9, P is indecomposible. Thus, under P_π, the Hit–and–Run Markov chain $(X_n; n \geq 0)$ is a stationary ergodic sequence (see Breiman [1968], Theorem 7.16, page 136). Now fix $B \in \mathcal{B}_S$. Then under P_π the sequence $(1_B(X_n); n \geq 0)$ is also a stationary ergodic sequence (see Breiman [1968], Proposition 6.31, page 119; here again Breiman states the result for the case where the state space is a Borel subset of $\mathbf{R}$ but his proof is also valid for the case where the state space is a Borel subset of $\mathbf{R}^d$). Thus by the ergodic theorem (see e.g. Breiman [1968], Theorem 6.28, page 118) we obtain

$$\mathbf{P}_\pi \left[\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B)\right] = 1$$

$$\text{i.e.} \quad \int_S \mathbf{P}_x \left[\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B)\right] f(x)dx = 1.$$

Since $f(x) > 0$ $\forall x \in S$, this implies that

$$\mathbf{P}_x \left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B) \right] = 1 \tag{11}$$

for almost all x in S. We will now show that (11) holds for all x in S. Fix x in S and fix $\epsilon > 0$. By part (b) of Proposition 6 we can choose k so that $P^k_{\mathbf{abs}}(x, S) > 1 - \epsilon$. Now let

$$S_o = \left\{ y \in S : \mathbf{P}_y \left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B) \right] = 1 \right\}.$$

Then

$$\begin{aligned}
&\mathbf{P}_x \left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B) \right] \\
&= \int_S \mathbf{P}_x \left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B) \mid X_k = y \right] P^k(x, dy) \\
&= \int_S \mathbf{P}_y \left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B) \right] P^k(x, dy) \\
&\geq \int_{S_o} \mathbf{P}_y \left[\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_B(X_i) = \pi(B) \right] P^k_{\mathbf{abs}}(x, dy) \\
&= P^k_{\mathbf{abs}}(x, S_o) = P^k_{\mathbf{abs}}(x, S) > 1 - \epsilon.
\end{aligned}$$

The last equality follows from the fact that $\varphi(S_o) = \varphi(S)$. This holds for every $\epsilon > 0$. Thus (11) holds for all $x \in S$. Now if $B \in \mathcal{B}_S$ and $\varphi(B) > 0$, then $\pi(B) > 0$ and therefore (11) yields

$$\mathbf{P}_x \left[\sum_{i=1}^{\infty} 1_B(X_i) = \infty \right] = 1.$$

Thus the Markov chain is Harris-recurrent with respect to Lebesgue measure on S. ∎

Recall that if $\pi, \pi_1, \pi_2, \pi_3, \ldots$ are probability measures on S then, by definition, π_n is said to converge in total variation to π if

$$\lim_{n\to\infty} \pi_n(B) = \pi(B) \qquad \forall B \in \mathcal{B}_S.$$

(In comparison, π_n is said to converge in distribution to π (or converge weakly to π) if

$$\lim_{n\to\infty} \pi_n(B) = \pi(B) \qquad \forall B \in \mathcal{B}_S \text{ with } \pi(\partial B) = 0;$$

see Billingsley [1968] Theorem 2.1). Orey ([1971], page 30, part (i) of Theorem 7.1) has shown that if an aperiodic φ-recurrent Markov chain possesses a stationary probability distribution π, then for every initial distribution μ, the distribution of X_n converges in total variation, and hence in distribution, to the stationary distribution π. In view of Orey's theorem, the following result is a consequence of Theorem 2, Theorem 4, and Proposition 10.

THEOREM 5. *Let S be a bounded open subset of $\mathbf{R}^d$. Let π be an absolutely continuous probability measure on S. Assume that π has a density which is bounded, almost everywhere continuous, and strictly positive on S. Let ν be a probability measure on the boundary ∂D of the d-dimensional unit sphere centered at the origin. Assume that ν is full dimensional and that the connected components of S are ν-communicating. Then π is the unique stationary distribution for the (ν, π)-Hit–and–Run Markov kernel. Furthermore, for every initial distribution the (ν, π)-Hit–and–Run Markov chain converges in total variation, and hence in distribution, to the stationary distribution π.* ∎

REMARK. We proved the convergence part of Theorem 5 via Harris-recurrence and Orey's theorem. If the direction distribution ν is absolutely continuous (with respect to the uniform distribution on ∂D) then by part (b) of Proposition 3 the probability measures $P(x, \cdot)$ are absolutely continuous with respect to the stationary distribution π. In this case the convergence part of Theorem 5 can be obtained via a theorem of Doob [1948] which says that if an aperiodic and indecomposible Markov chain possesses a stationary distribution π and if its transition probability distributions $P(x, \cdot)$ are absolutely continuous with respect to π then, for every initial distribution, this Markov chain converges in total variation to the stationary distribution π. (See Theorem 7.18 of Breiman [1968].) With this approach one doesn't need to prove Harris-recurrence. But if ν is not absolutely continuous (with respect to the uniform distribution on ∂D), as in the Coordinate Direction algorithm for instance, then the probability measures $P(x, \cdot)$ are not absolutely continuous with respect to π and this approach via Doob's theorem fails.

5. CONCLUSION

Theorem 5 asserts that

$$\lim_{n \to \infty} \mathbf{P}_x [X_n \in B] = \pi(B) \qquad \forall B \in \mathcal{B}_S,\ \forall x \in S,$$

where $(X_n; n \geq 0)$ is the (ν, π)-Hit–and–Run Markov chain on S. This result has a significant practical implication for Monte-Carlo methods. It says that one can use the Hit–and–Run algorithm to generate points that are asymptotically π-distributed.

Historically, the Coordinate Directions and the Hypersphere Directions Hit–and–Run algorithms have been used for generating uniform distributions over bounded open subsets of $\mathbf{R}^d$. In that context they are known to be efficient when compared to standard methods such as the acceptance-rejection method, specially when d is large. In this paper we have generalized these traditional Hit–and–Run algorithms to allow for an arbitrary direction distribution ν and an essentially arbitrary absolutely continuous target distribution π. Although not addressed here, the boundedness conditions on the open set S and on the density function f can both be removed. Moreover, the condition that f be strictly positive on S and the condition that S be open can both be substantially relaxed.

As a final comment, the general (ν, π)-Hit–and–Run algorithms introduced in this paper show considerable promise for solving certain global optimization problems. For example, by setting the density f to be a suitable increasing function of the objective function g, the target distribution π will concentrate near the global maximum of g and therefore the Hit–and–Run Markov chain will, with high probability, find itself near that global maximum. This idea is currently being investigated.

REFERENCES

Boneh A., and Golan A. [1979]. *Constraints' redundancy and feasible region boundedness by random feasible point generator (RFPG).* Third European Congress on Operations Research, EURO III, Amsterdam.

Berbee H.C.P., Boender C.G.E., Rinnooy Kan A.H.G., Scheffer C.L., Smith R.L., Telgen J. [1987]. *Hit-and-Run algorithms for the identification of nonredundant linear inequalities.* Mathematical Programming, **37**, 184–207.

Billingsley P. [1968]. *Convergence of probability measures*, Wiley, New York.

Billingsley P. [1986]. *Probability and measure*, Second Edition, Wiley, New York.

Boender C.G.E., Caron R.J., Rinnooy Kan A.H.G., McDonald J.F., Romeijn H.E., Smith R.L., Telgen J. and Vorst A.C.F. [1989]. *Shake–and–Bake algorithms for generating uniform points on the boundary of bounded polyhedra.* Technical Report 89-24, Department of Industrial and Operations Engineering, The University of Michigan. Submitted to Operations Research.

Breiman L. [1968]. *Probability*, Addison-Wesley, Reading, Massachusetts.

Doob J.L. [1948]. *Asymptotic properties of Markov transition probabilities*, Trans. Amer. Math. Soc. **63**, 393–421.

Kelly F.P. [1979]. *Reversibility and stochastic networks*, Wiley, New York.

Nummelin E. [1984]. *General irreducible Markov chains and non-negative operators*, Cambridge University Press, Cambridge.

Orey S. [1971]. *Limit theorems for Markov chain transition probabilities*, Van Nostrand Reinhold, New York.

Revuz D. [1975]. *Markov chains*, North-Holland Publishing Company, Amsterdam.

Ross, S. [1983]. *Stochastic processes*, Wiley, New York.

Smith R.L. [1980]. *Monte Carlo procedures for generating random feasible solutions to mathematical programs*, ORSA/TIMS Conference, Washington D.C.

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

Telgen J. [1980]. Private communication with A. Boneh.