Crossing Estimates and Convergence of Dirichlet Functions Along Random Walk and Diffusion Paths

by Alano Ancona, Russell Lyons and Yuval Peres

Abstract. Let $\{X_n\}$ be a transient reversible Markov chain and let f be a function on the state space with finite Dirichlet energy. We prove crossing inequalities for the process $\{f(X_n)\}_{n\geq 1}$ and show that it converges almost surely and in L^2. Analogous results are also established for reversible diffusions on Riemannian manifolds.

§1. Introduction.

Benjamini and Schramm (1996) proved that transient planar graphs with bounded degree support nonconstant harmonic Dirichlet functions. In the course of their argument, they showed that on such graphs, certain non-harmonic Dirichlet functions converge almost surely along random walk paths. A question of O. Schramm (personal communication) led us to investigate whether a.s. convergence holds for all Dirichlet functions and to study L^p convergence as well.

For discrete transient Markov chains, Theorem 1.1 below gives a.s. and L^2 convergence. The same simple argument works also for quite general reversible diffusions (Theorem 9.1). One way to make a.s. convergence quantitative is to prove a *crossing inequality.* We present such an inequality for the discrete time processes $\{f(X_n)\}$ of Theorem 1.1 in Theorem 1.3; the crossing inequalities in continuous time are Theorems 1.4 and 8.1. Our first proof of the a.s. convergence in Theorem 1.1 was based on the crossing inequality; this inequality may have other applications as well.

For diffusions on manifolds, the almost sure convergence result (but not L^2 convergence, nor the crossing inequality) is covered by classical results of J.L. Doob (1957, 1962),

1991 Mathematics Subject Classification. Primary 60J45. Secondary 31C25, 60F15.
Key words and phrases. Dirichlet energy, random walk, almost sure convergence, Markov chain, diffusions, manifolds, crossing.
Research partially supported by the Institute for Advanced Studies, Jerusalem (Lyons) and NSF grant DMS-94-04391 (Peres).

obtained in the framework of the Green's spaces of Brelot and Choquet. Indeed, Doob (1962) showed that a Dirichlet function f, on a transient Riemann surface say, admits a fine limit almost everywhere with respect to harmonic measure on the Martin boundary of the surface. By Doob (1957), this means that f converges to a finite limit along almost all Brownian paths. Doob's arguments can be adapted to the discrete setting as well. Some other related earlier results are mentioned below.

We now describe our results for the discrete setting. Let $\{p(x,y)\,;\ x,y \in \mathsf{V}\}$ be transition probabilities on the countable set V, and let $\{X_n\,;\ n \ge 0\}$ be a Markov chain with these transition probabilities and an arbitrary initial state. We assume that the Markov chain is **reversible** with stationary measure π, i.e., $\pi(x)p(x,y) = \pi(y)p(y,x)$ for all $x,y \in \mathsf{V}$. We also assume that $\pi(x) > 0$ for all $x \in \mathsf{V}$, and that the transition matrix $\{p(x,y)\}$ is stochastic and **irreducible**, i.e., for any partition $\mathsf{V} = A_1 \cup A_2$ of the state space V into two nonempty subsets, there exist $x \in A_1$ and $y \in A_2$ such that $p(x,y) > 0$.

The **Dirichlet energy** of a function $f : \mathsf{V} \to \mathbb{R}$ is

$$D(f) := \frac{1}{2} \sum_{x,y \in \mathsf{V}} \pi(x)p(x,y)\big(f(x) - f(y)\big)^2 .$$

If $D(f) < \infty$, then we call f a **Dirichlet function** and write $f \in \mathbf{D}$.

THEOREM 1.1. *Suppose that the reversible Markov chain $\{X_n\}$ is transient. Then for any function $f \in \mathbf{D}$, the sequence $\{f(X_n)\}$ converges a.s. and in L^2.*

This theorem readily extends to reducible Markov chains by considering each irreducible component separately. An example showing that L^q-convergence may fail for all $q > 2$ is given in Section 2. If f in Theorem 1.1 is also assumed to be harmonic, then a simple calculation (see (2.2)) shows that $\{f(X_n)\}$ is an L^2-martingale, whence convergence a.s. and in L^2 is immediate. The point of Theorem 1.1 is that harmonicity is not assumed.

Fix $o \in \mathsf{V}$ and define the **Dirichlet norm** of $f \in \mathbf{D}$ by $\|f\|^2 := D(f) + \pi(o)f(o)^2$. Recall that any $f \in \mathbf{D}$ can be written in a unique way as a sum of a harmonic Dirichlet function $f_{\mathbf{HD}}$ and a function f_0 in $\mathbf{D}_0$, the class of functions that are limits, in the Dirichlet norm, of functions with finite support. This is called the **Royden decomposition**; see Soardi (1994), Theorem 3.69.

COROLLARY 1.2. *Under the assumptions of Theorem 1.1, if $f = f_{\mathbf{HD}} + f_0$ is the Royden decomposition of $f \in \mathbf{D}$, then* $\lim f(X_n) = \lim f_{\mathbf{HD}}(X_n)$ *a.s. and the limit is a random variable in L^2.*

Let $G(x,y)$ denote the Green function for the given chain, i.e., the expected number of visits to y for the Markov chain started at x. Let $C(a,b)$ denote the number of crossings of the interval $[a,b]$ by the sequence $\{f(X_n)\}$.

THEOREM 1.3. *Under the assumptions of Theorem 1.1, for any $a < b$, we have*

$$\pi(x)\mathbf{E}_x[C(a,b)] \leq 2G(x,x)\frac{D(f)}{(b-a)^2},$$

where the subscript in $\mathbf{E}_x$ indicates the initial state.

This is proved as a consequence of another result, namely, that the Dirichlet energy of f can only be decreased by inducing on a subset of the state space. These ancillary results are given in Section 3.

Theorem 1.1 is reminiscent of a well-known theorem of Yamasaki (1986) that Dirichlet functions converge along almost all paths in the sense of extremal length; see Soardi (1994), Theorem 3.85. For nearest-neighbor random walks on trees, the two theorems are equivalent. In general, however, neither of these theorems implies the other: even on $\mathbb{Z}^3$, simple random walk is supported on a set of paths with infinite extremal length, namely the set of paths that satisfy the law of the iterated logarithm.

Various special cases of the almost sure convergence result in Theorem 1.1 are already known. P.M. Soardi (personal communication) noted that for any group-invariant random walk $\{X_n\}$ on a free group, a.s. convergence of $\{f(X_n)\}$ for $f \in \mathbf{D}$ follows from the main result of Cartwright and Soardi (1989). More generally, O. Schramm (personal communication) knew a simple proof for reversible Markov chains whose Green function is uniformly bounded. In the continuous setting and under stronger hypotheses on the transition kernel, convergence of related stochastic integrals was recently established by T. Lyons and L. Stoica (1999).

In Section 4, we prove two extensions of Theorem 1.1.

In Sections 5 to 9, we establish a crossing inequality and a convergence theorem for a large class of symmetric diffusions in Riemannian manifolds. To illustrate these results, we state here a special case of the crossing inequality for Brownian motion; we refer to Section 5 for definitions of the terminology used in the continuous setting. Consider a Riemannian manifold M and the Brownian motion $\{X_t\}_{t\geq 0}$ in M. Assume that $\{X_t\}_{t\geq 0}$ is transient. Denote by N the number of crossings realized by X_t between two smooth bounded disjoint compact sets A and B in M.

THEOREM 1.4. *Let F be the function of minimal Dirichlet energy on M that satisfies $F = 1$ a.e. on A and $F = 0$ a.e. on B. Let G denote the Green function of M and let σ denote* *the Riemannian volume in M. Then the following inequality holds for all $x \in M$:*

$$\mathbf{E}_x(N) \le 2 \int_M G(x,y)\, |\nabla F(y)|^2 \, d\sigma(y) \, .$$

Note that $2G$ is the Green function with respect to $\frac{1}{2}\Delta_M$ if Δ_M is the Laplace-Beltrami operator of M. A more general statement is given in Section 8 (Theorem 8.1). The continuous version of Theorem 1.1 is in Section 9.

There are some parallels between our results and those of T. Lyons and W.A. Zheng (1988) who established a crossing inequality for a bounded time interval. The setting is quite different, however: they were concerned with stationary Markov processes, where there is no hope for convergence, and their proof relies on time invariance.

§2. Majorization and Convergence in the Discrete Case.

The relevant terminology was defined before the statement of Theorem 1.1. The following notation will also be handy:

$$\mathcal{E}_f(x) := \sum_y p(x,y)[f(y) - f(x)]^2 = \mathbf{E}_x[|f(X_1) - f(X_0)|^2] \, .$$

Thus $D(f) = \frac{1}{2} \sum_{x \in \mathsf{V}} \pi(x) \mathcal{E}_f(x)$.

LEMMA 2.1. *Suppose that the underlying Markov chain is transient, so that the Green function $G(x,y)$ is finite everywhere. Let f be a function in $\mathbf{D}_0$. For any $x \in \mathsf{V}$, we have*

$$\pi(x) f(x)^2 \le D(f) G(x,x) \, . \tag{2.1}$$

In particular, $\|f\|^2 \le [G(o,o)+1]D(f)$. Furthermore, there exists a superharmonic function $h \in \mathbf{D}_0$ such that $h \ge |f|$ pointwise and $D(h) \le D(f)$.

Proof. For finitely supported f, the inequality (2.1) is a consequence of Dirichlet's principle (see Soardi (1994), Theorem 3.41). The general case ($f \in \mathbf{D}_0$) follows by approximation. Thus $D(\cdot)^{1/2}$ is a Hilbert-space norm on $\mathbf{D}_0$, equivalent to $\|\cdot\|$ there. Note that for $f \in \mathbf{D}_0$, also $|f| \in \mathbf{D}_0$. The set of functions $\{g \in \mathbf{D}_0 \; ; \; g \ge |f|\}$ is convex and closed in $\mathbf{D}_0$; let h be the function with minimal energy $D(h)$ in this set. Clearly $D(h) \le D(|f|) \le D(f)$, and it is easy to verify that h is superharmonic. ∎

Proof of Theorem 1.1 and Corollary 1.2. Since o is arbitrary, we take $X_0 = o$. We first observe that for any $f \in D$, it is easy to bound the sum of squared increments along the random walk. For any Markov chain, we have $G(y,o) = \mathbf{P}_y[\tau_o < \infty]G(o,o) \le G(o,o)$. By reversibility, we have $\pi(o)G(o,y) = \pi(y)G(y,o) \le \pi(y)G(o,o)$, whence

$$\begin{aligned}\sum_{k=1}^{\infty} \mathbf{E}\big[|f(X_k) - f(X_{k-1})|^2\big] &= \sum_{y\in\mathsf{V}} G(o,y)\mathcal{E}_f(y) \le \frac{G(o,o)}{\pi(o)} \sum_{y\in\mathsf{V}} \pi(y)\mathcal{E}_f(y) \\ &= 2\frac{G(o,o)}{\pi(o)} D(f)\,. \end{aligned} \tag{2.2}$$

It follows from (2.2) that when f is harmonic, $\{f(X_n)\}$ is a martingale bounded in L^2, and thus converges a.s. and in L^2. Therefore, to prove Theorem 1.1 and Corollary 1.2, it suffices to show that for any $f \in \mathbf{D}_0$, the sequence $\{f(X_n)\}$ converges a.s. and in L^2 to 0.

Given $f \in \mathbf{D}_0$, Lemma 2.1 yields a superharmonic $h \in D_0$ that satisfies $h \ge |f|$ and $D(h) \le D(f)$. If $\{h(X_n)\}$ converges a.s. and in L^2 to 0, then so does $\{f(X_n)\}$.

The sequence $\{h(X_n)\}$ is a nonnegative supermartingale, so it converges a.s. and $\mathbf{E}[h(X_k) - h(X_{k-1}) \mid X_1, \ldots, X_{k-1}] \le 0$. Therefore

$$\begin{aligned}\mathbf{E}\big[h(X_k)^2 - h(X_{k-1})^2\big] &= \mathbf{E}\big[|h(X_k) - h(X_{k-1})|^2\big] + 2\mathbf{E}\big[\big(h(X_k) - h(X_{k-1})\big)h(X_{k-1})\big] \\ &\le \mathbf{E}\big[|h(X_k) - h(X_{k-1})|^2\big]\,. \end{aligned}$$

Summing these inequalities for $k = 1, \ldots, n$ gives

$$\mathbf{E}\big[h(X_n)^2 - h(X_0)^2\big] \le \sum_{k=1}^{n} \mathbf{E}\big[|h(X_k) - h(X_{k-1})|^2\big] \le 2\frac{G(o,o)}{\pi(o)} D(f)$$

by (2.2). Since $h(X_0)^2 = h(o)^2 \le \frac{G(o,o)}{\pi(o)} D(f)$ by Lemma 2.1, we infer that

$$\mathbf{E}\big[h(X_n)^2\big] \le 3\frac{G(o,o)}{\pi(o)} D(f)\,. \tag{2.3}$$

Next, we show that $h(X_n) \to 0$ in L^2. Given $\epsilon > 0$, write $h = f_1 + f_2$, where f_1 is finitely supported and $D(f_2) < \epsilon$. Lemma 2.1 applied to $f_2 \in \mathbf{D}_0$ yields a superharmonic function $h_2 \in \mathbf{D}_0$ that satisfies $h_2 \ge |f_2|$ and $D(h_2) \le D(f_2)$. By (2.3), for any $n \ge 1$,

$$\mathbf{E}\big[f_2(X_n)^2\big] \le \mathbf{E}\big[h_2(X_n)^2\big] \le 3\frac{G(o,o)}{\pi(o)}\epsilon\,.$$

Since $\mathbf{E}\big[f_1(X_n)^2\big] = \sum_y f_1(y)^2 \mathbf{P}[X_n = y] \to 0$ as $n \to \infty$, it follows that $h(X_n) \to 0$ in L^2. Thus the a.s. limit of $\{h(X_n)\}$ must also be 0. This completes the proof. ∎

EXAMPLE 2.2. We give an example of a Dirichlet function f with a transient random walk $\{X_n\}$ such that $\{f(X_n)\}$ is unbounded in L^q for every $q > 2$. Let $\{X_n\}$ be simple random walk on a graph that is the union of a transient graph T and the positive integers $\mathbb{N}$, with one node of T identified with $1 \in \mathbb{N}$. Suppose that $f(2^n) = 2^{n/2}/(n+1)$ for all $n \geq 0$, that f is extended by linear interpolation to the other positive integers, and that f is constant on T. Then $D(f) \leq \sum_n 2^{1-n} f(2^n)^2 < \infty$. With probability bounded below by $c2^{-n}$ for some $c > 0$, the random walk started at 1 stays in $\mathbb{N}$ for the first 4^n steps, and is at an integer greater than 2^n at time 4^n. Thus $\mathbf{E}[f(X_{4^n})^q] \geq c2^{-n} f(2^n)^q$, so for any $q > 2$, the sequence $\{f(X_n)\}$ is unbounded in L^q.

§3. The Induced Markov Chain and the Crossing Inequality.

For $A \subseteq \mathsf{V}$, write $\tau_A := \inf\{n \geq 0 ;\ X_n \in A\}$ and $\tau_A^+ := \inf\{n \geq 1 ;\ X_n \in A\}$.

LEMMA 3.1. *Let $\{X_n\}$ be a reversible irreducible Markov chain with stationary measure π. Let A be a nonempty subset of the state space V. For any $f \in \mathbf{D}$, we have*

$$D(f) \geq \frac{1}{2} \sum_{x \in A} \pi(x) \mathbf{E}_x \Big[|f(X_{\tau_A^+}) - f(x)|^2 ;\ \tau_A^+ < \infty \Big] , \tag{3.1}$$

with equality if f is harmonic off A and $\mathbf{P}_x[\tau_A < \infty] = 1$ for all x.

REMARK. When $\mathbf{P}_x[\tau_A < \infty] = 1$ for all x, the right-hand side is the Dirichlet energy of the restriction $f{\restriction}A$ for the Markov chain that is obtained by inducing on A.

Proof. Choose $o \in A$. The set of Dirichlet functions that agree with f on A is convex and closed, whence it has a unique element $\widetilde{f}$ of minimal norm. Since $o \in A$, this element $\widetilde{f}$ is also the unique one of minimal energy. Since $D(f) \geq D(\widetilde{f})$, it suffices to prove (3.1) in the case that $f = \widetilde{f}$. Note that in this case, f is harmonic off A.

Since $\mathbf{E}_x \Big[f(X_{k \wedge \tau_A^+}) - f(X_{(k-1) \wedge \tau_A^+}) \mid X_0, \ldots, X_{k-1} \Big] = 0$ for $k \geq 2$, we have that

$$\begin{aligned} \mathbf{E}_x \left[|f(X_{n \wedge \tau_A^+}) - f(x)|^2 \right] &= \sum_{k=1}^{n} \mathbf{E}_x \left[|f(X_{k \wedge \tau_A^+}) - f(X_{(k-1) \wedge \tau_A^+})|^2 \right] \\ &= \mathbf{E}_x \left[\sum_{k=1}^{n \wedge \tau_A^+} \mathcal{E}_f(X_{k-1}) \right] \leq \sum_{y \in \mathsf{V}} G_A(x, y) \mathcal{E}_f(y) , \end{aligned}$$

where

$$G_A(x, y) := \sum_{k=1}^{\infty} \mathbf{P}_x[X_{k-1} = y ;\ \tau_A^+ \geq k]$$

is the expected number of visits to y for a random walk started at x and stopped at time $\tau_A^+ - 1$. (In particular, $G_A(x,x) = 1$ for $x \in A$.) In case $\tau_A^+ = \infty$, widen the definition of $f(X_{\tau_A^+})$ to mean $\lim_{n\to\infty} f(X_{n\wedge\tau_A^+})$. Then the sequence $\{f(X_{n\wedge\tau_A^+})\,;\ n \ge 1\}$ is an L^2-bounded martingale, whence

$$\mathbf{E}_x\left[|f(X_{\tau_A^+}) - f(x)|^2\right] = \sum_{y\in\mathsf{V}} G_A(x,y)\mathcal{E}_f(y)\,.$$

By reversibility, $\pi(x)G_A(x,y) = \pi(y)G'_A(y,x)$, where

$$G'_A(y,x) := \sum_{k=0}^{\infty} \mathbf{P}_y[X_k = x\,;\ \tau_A \ge k]\,.$$

Therefore,

$$\begin{aligned}
\sum_{x\in A} \pi(x)\mathbf{E}_x\left[|f(X_{\tau_A^+}) - f(x)|^2\right] &= \sum_{x\in A} \pi(x)\left[\sum_{y\in\mathsf{V}} G_A(x,y)\mathcal{E}_f(y)\right] \\
&= \sum_{y\in\mathsf{V}} \pi(y)\mathcal{E}_f(y) \sum_{x\in A} G'_A(y,x) \\
&= \sum_{y\in\mathsf{V}} \pi(y)\mathcal{E}_f(y)\mathbf{P}_y[\tau_A < \infty] \\
&\le 2D(f)\,.
\end{aligned}$$

This gives (3.1) and, clearly, equality holds if $\mathbf{P}_x[\tau_A < \infty] = 1$ for all x. ∎

Proof of Theorem 1.3. Let $A := \{x\,;\ f(x) \le a \text{ or } f(x) \ge b\}$. Let A_k be the event that A is visited at least k times and let τ_k be the kth time that a random walk visits A. We have

$$\begin{aligned}
\mathbf{E}_x[C(a,b)](b-a)^2 &\le \sum_{k\ge 1} \mathbf{E}_x\Big[|f(X_{\tau_{k+1}}) - f(X_{\tau_k})|^2\,;\ A_{k+1}\Big] \\
&= \sum_{y\in A} G(x,y)\mathbf{E}_y\Big[|f(X_{\tau_A^+}) - f(y)|^2\,;\ \tau_A^+ < \infty\Big]\,.
\end{aligned}$$

Reversibility gives $\pi(x)G(x,y) = \pi(y)G(y,x) \le \pi(y)G(x,x)$. Therefore,

$$\begin{aligned}
\pi(x)\mathbf{E}_x[C(a,b)](b-a)^2 &\le \sum_{y\in A} \pi(y)G(x,x)\mathbf{E}_y\Big[|f(X_{\tau_A^+}) - f(y)|^2\,;\ \tau_A^+ < \infty\Big] \\
&\le 2G(x,x)D(f)
\end{aligned}$$

by Lemma 3.1. ∎

§4. Extensions and Applications.

Theorem 1.1 can be extended in two somewhat different ways.

COROLLARY 4.1. *Let $\{X_n\}$ be a transient reversible Markov chain on V with stationary measure π. Let H denote the space of anti-symmetric functions $\theta : \mathsf{V} \times \mathsf{V} \to \mathbb{R}$ (i.e., $\theta(x, y) = -\theta(y, x)$) such that $\sum_{x,y\in\mathsf{V}} \pi(x)p(x, y)\theta(x, y)^2 < \infty$. Then the partial sums $S_n := \sum_{k=1}^{n} \theta(X_{k-1}, X_k)$ converge a.s. and in L^2.*

Proof. The inner product $\langle\theta_1, \theta_2\rangle := \sum_{x,y\in\mathsf{V}} \pi(x)p(x, y)\theta_1(x, y)\theta_2(x, y)$ makes H a Hilbert space. For any $f \in \mathbf{D}$, the antisymmetric function $\nabla f(x, y) := f(x) - f(y)$ is in H; Corollary 4.1 for $\theta \in \nabla\mathbf{D}$ reduces to Theorem 1.1, so it suffices to consider $\theta \perp \nabla\mathbf{D}$. Any such θ satisfies $\sum_{y\in\mathsf{V}} p(x, y)\theta(x, y) = 0$ for all $x \in \mathsf{V}$, since the indicator $\mathbf{1}_{\{x\}}$ is in $\mathbf{D}$. Consequently, the partial sums S_n form a martingale when $\theta \perp \nabla\mathbf{D}$; they are bounded in L^2 since

$$\sum_{n=1}^{\infty} \mathbf{E}_x[\theta(X_{n-1}, X_n)^2] = \sum_{y,z\in\mathsf{V}} G(x, y)p(y, z)\theta(y, z)^2 = \sum_{y,z\in\mathsf{V}} \frac{G(y, x)}{\pi(x)}\pi(y)p(y, z)\theta(y, z)^2$$

and $G(y, x) \leq G(x, x)$. ∎

The following proposition answers another question due to O. Schramm (personal communication, 1994).

PROPOSITION 4.2. *Let $\{X_n\}$ be a transient reversible Markov chain on V with stationary measure π. Let $\delta(x, y) \equiv \delta(y, x) \geq 0$ satisfy $\sum_{x,y\in\mathsf{V}} \pi(x)p(x, y)\delta(x, y)^2 < \infty$. Define the induced metric on V by*

$$d(x, y) := \inf\left\{\sum_{k=1}^{n} \delta(x_k, x_{k-1})\,;\ x = x_0,\ y = x_n,\ \ \forall k \in [1, n]\ p(x_k, x_{k-1}) > 0\right\}.$$

Then $\{X_n\}$ is a Cauchy sequence in this metric with probability 1.

This is an immediate consequence of Theorem 1.1 and the following general lemma:

LEMMA 4.3. *If (Z, d) is a complete separable metric space and $\{X_n\}$ is a process in Z such that for each bounded Lipschitz function $f : Z \to \mathbb{R}$, the sequence $\{f(X_n)\}$ is a.s. convergent, then the process $\{X_n\}$ is already a.s. convergent in Z.*

Proof. Fix a dense sequence $\{a_k\}$ in Z. Let $B(x, r)$ denote the ball of radius r about x. Given $\epsilon > 0$ and $J < J' \leq \infty$, set

$$W(J, J') := \bigcup\left\{B(a_k, \epsilon)\,;\ J \leq k < J' \text{ and } \forall j < J\ \ d(a_j, a_k) \geq 3\epsilon\right\}.$$

Define the event

$$A_\epsilon := \{\forall J\ \exists n\ X_n \in W(J,\infty)\}\ .$$

We claim that $\mathbf{P}[A_\epsilon] = 0$. Otherwise, since $W(J,\infty) = \bigcup_{J'<\infty} W(J,J')$, we could find an increasing sequence of integers $\{J_p\ ;\ p \geq 1\}$ such that for all p,

$$\mathbf{P}\big[A_\epsilon \cap \big\{\exists n\ X_n \in W(J_p, J_{p+1})\big\}\big] \geq (1-3^{-p})\mathbf{P}[A_\epsilon]\,.$$

Therefore,

$$\mathbf{P}[\forall p\ \exists n\ X_n \in W(J_p, J_{p+1})] > 0\,. \tag{4.1}$$

But it is easy to construct a (bounded) Lipschitz function f on Z such that $(-1)^p f \geq 1$ on $W(J_p, J_{p+1})$. Thus $\{f(X_n(\omega))\}$ does not converge when ω lies in the event in (4.1), which contradicts the assumptions.

Since $P[A_\epsilon] = 0$ for each $\epsilon > 0$, it follows that $\{X_n\ ;\ n \geq 1\}$ is totally bounded a.s. and thus admits a cluster point a.s. Using the fact that almost surely $d(X_n, a_k) \wedge 1$ converges when $n \to \infty$ for all $k \geq 1$, it is then easily checked that $\{X_n\}$ converges a.s. ∎

§5. The Continuous Setting: Terminology.

In this section, we describe our continuous setting. Some relevant (well-known) basic definitions and facts are also recalled here and in Section 6.

In all that follows, M denotes a connected noncompact C^2 Riemannian manifold (*that is not necessarily complete*), σ_M is the Riemannian volume in M, and $\mathcal{A}$ is a Borel measurable section of the bundle $\mathrm{End}(T(M))$ such that for all $x \in M$ and all $\xi \in T_x(M)$,

$$c_0(x)^{-1}\,\|\xi\|^2 \leq \langle \mathcal{A}_x(\xi), \xi\rangle \leq c_0(x)\,\|\xi\|^2\,,$$

where c_0 is a positive locally bounded function in M and $c_0 \geq 1$. It is also assumed that each $\mathcal{A}_x$ is symmetric, that is, $\langle \mathcal{A}_x(\xi), \eta\rangle = \langle \mathcal{A}_x(\eta), \xi\rangle$ when $x \in M$, $(\xi,\eta) \in T_x(M)^2$. We let $\mathcal{L}$ denote the divergence-type elliptic operator $\mathcal{L}(\bullet) := \mathrm{div}\,(\mathcal{A}(\nabla\bullet))$ (where div and ∇ are the usual divergence and gradient operators for M) and we set $\widetilde{\nabla} := \mathcal{A}^{\frac{1}{2}}\nabla$.

Related to $\mathcal{L}$, there is a well-known "local" potential theory on M whose harmonic functions are the continuous versions of the local weak $\mathcal{L}$-solutions (see Brelot (1969), Stampacchia (1965), and Hervé and Hervé (1969)). Hence, one may speak of $\mathcal{L}$-superharmonic functions, $\mathcal{L}$-potentials and so forth. We shall assume (this is our final assumption) that $\mathcal{L}$ is *transient*, i.e., there are nontrivial positive $\mathcal{L}$-superharmonic functions in M. Equivalently, the $\mathcal{L}$-Green function $G = G_{\mathcal{L}}$ exists in M, or the inequality

$$\int_M |u(m)|^2\,\pi(m)\,d\sigma_M(m) \leq \int_M |\widetilde{\nabla} u(m)|^2\,d\sigma_M(m)$$

holds for all $u \in C_0^1(M)$ and some continuous positive function π in M (see Fukushima, Oshima, and Takeda (1994), §1.5, or Ancona (1990)).

An element $u \in L^1_{\mathrm{loc}}(\sigma_M)$ is an **$\mathcal{L}$-Dirichlet finite function** if

(i) it admits a weak gradient ∇u (i.e., there is a locally integrable vector field ∇u such that $\int_M \nabla u \cdot W \, d\sigma_M = -\int_M u \operatorname{div}(W) \, d\sigma_M$ for all C_0^1 vector fields W in M), and

(ii) $|\widetilde{\nabla} u|$ is square integrable on M.

The Dirichlet form $a_{\mathcal{L}}$ associated to $\mathcal{L}$ is given by the formula

$$a_{\mathcal{L}}(u,v) := \int_M \mathcal{A}(\nabla u(x)) \cdot \nabla v(x) \, d\sigma_M(x) = \int_M \widetilde{\nabla} u(x) \cdot \widetilde{\nabla} v(x) \, d\sigma_M(x)$$

when u and v are $\mathcal{L}$-Dirichlet finite functions in M. Denote by $\mathbf{D}$ the vector space of all $\mathcal{L}$-Dirichlet finite functions over M equipped with the semi-norm $u \mapsto \|u\|_e := \sqrt{a_{\mathcal{L}}(u,u)}$.

Since $\mathcal{L}$ is transient, the completion of $C_0^1(M)$ in the norm $\|\bullet\|_e$ can be identified with a subspace $\mathbf{D}_0$ of $\mathbf{D}$; the subspace $\mathbf{D}_0$ contains all compactly supported elements of $\mathbf{D}$; and, equipped with the inner product $a_{\mathcal{L}}$, $\mathbf{D}_0$ is a regular Dirichlet space with reference measure σ_M (see Deny (1970) or Fukushima (1980)).

For $A \subset M$, define the capacity $\operatorname{cap}(A)$ in the usual manner,

$$\operatorname{cap}(A) := \inf \left\{ \|v\|_e^2 ;\ v \in \mathbf{D}_0,\ v \geq 1\ \sigma_M\text{-a.e. in a neighborhood of } A \right\}.$$

A set $A \subset M$ is **polar** if $\operatorname{cap}(A) = 0$. A property is said to hold **quasi-everywhere (q.e.)** if it holds outside a polar set. A function $u : M \to \mathbb{R}$ is **quasi-continuous** (with respect to $\mathbf{D}_0$) if there is a sequence of open subsets ω_n of M such that the restrictions $f \restriction F_n$ are continuous in $F_n = M \setminus \omega_n$ and $\lim_{n\to\infty} \operatorname{cap}(\omega_n) = 0$. Each $u \in \mathbf{D}$ admits a quasi-continuous version and two such versions are equal quasi-everywhere. In the sequel, when we speak of an element of $\mathbf{D}$, we always mean the class of its quasi-continuous versions.

The semi-normed space $\mathbf{D}$ is **complete** in the sense that if $\{f_n\}$ is a Cauchy sequence in $\mathbf{D}$, then there is a sequence of reals $\{c_n\}$ such that $\{f_n + c_n\}$ converges in $L^2_{\mathrm{loc}}(\sigma_M)$ to some $f \in \mathbf{D}$ and moreover $\|f_n - f\|_e \to 0$. This is easily seen by using local L^2-Poincaré inequalities. We also note that $\{f_n + c_n\}$ has a subsequence that converges to f quasi-everywhere.

By minimizing $\|f - f_0\|_e$, $f_0 \in \mathbf{D}_0$, one obtains the familiar Royden decomposition:

LEMMA 5.1. *For each $f \in \mathbf{D}$, there is a (unique) decomposition $f = f_0 + h$, where h is a (Dirichlet finite) $\mathcal{L}$-harmonic function and $f_0 \in \mathbf{D}_0$.*

Denote by $\widehat{M} := M \cup \{\partial\}$ the one-point compactification of M. Associated to $\mathcal{L}$ (with Dirichlet condition at infinity), there is a standard Markov process (Blumenthal and

Getoor (1968), §2)

$$\left(\Omega, \{\mathcal{F}_t\}_{0\le t\le\infty}, \{\mathbf{P}_x\}_{x\in\widehat{M}}, \{X_t\}_{0\le t\le\infty}\right)$$

on M with continuous paths on $[0,\infty]$ and cemetery point ∂ (see, e.g., Fukushima (1980), §7). The **lifetime** of the process X in M is denoted by τ, i.e.,

$$\tau := \inf\{t \ge 0\,;\ X_t = \partial\}\,.$$

A quasi-Borel set $A \subset M$ is **polar** with respect to the process X if for all $x \in M$, we have $\mathbf{P}_x[\exists t > 0\ X_t \in A] = 0$; recall that this holds iff A is polar with respect to $\mathbf{D}_0$ and iff there exists an $\mathcal{L}$-superharmonic function f in M such that $A \subset \{f = +\infty\} \ne M$. Also, a quasi-everywhere finite function in M is quasi-continuous iff it is finely continuous q.e. with respect to X (Fukushima, Oshima, and Takeda (1994), p. 168). Note that (nonnegative) $\mathcal{L}$-superharmonic functions in M are excessive and hence quasi-continuous.

§6. The Continuous Setting: Background.

We collect here some known facts that enter directly in the proof of Proposition 7.1 below.

1. *Weak supersolutions.* If $f \in \mathbf{D}$ and $a_{\mathcal{L}}(f,v) \ge 0$ for every nonnegative $v \in C_0^1(M)$ (that is, $-\mathcal{L}(f)$ is a positive measure μ in M), then the map $v \mapsto a_{\mathcal{L}}(f,v)$, $v \in C_0^1(M)$, extends (obviously) to a continuous linear form on $\mathbf{D}_0$. It follows that μ does not charge polar sets and that $a_{\mathcal{L}}(f,v) = \int_M v\,d\mu$ for $v \in \mathbf{D}_0$. If, moreover, $f \in \mathbf{D}_0$, then $f \ge 0$ q.e. (Deny (1970) or Fukushima (1980)). In fact, if $f \in \mathbf{D}_0$, we have $f = G(\mu)$ q.e., that is, $f(\bullet) = \int_M G(\bullet, y)\,d\mu(y)$ q.e. (The reduced functions f_{K^c} (Fukushima, Oshima, and Takeda (1994), p. 92) verify $\lim_{j\to\infty} \|f_{K_j^c}\|_e = 0$ if $\{K_j\}$ is an exhaustion of M.)

2. *Composed Processes.* **(a)** If $f \in \mathbf{D}$ and $x \in M$, then $t \mapsto f(X_t)$ is $\mathbf{P}_x$-a.s. continuous on $(0,\tau)$. This follows from the quasi-continuity of f.

(b) Let $f : U \to \mathbb{R}_+$ be a bounded nonnegative $\mathcal{L}$-superharmonic function in an open subset U of M. Extend f to be zero in $\widehat{M} \setminus U$ and let $Z_t := f(X_{t\wedge T})$ for $0 \le t \le \infty$, where T is the first entry time in $\widehat{M} \setminus U$. Then $\{Z_t\}_{0\le t\le\infty}$ is, under $\mathbf{P}_x$, $x \in U$, an $\{\mathcal{F}_t\}_{t\ge 0}$ right-continuous supermartingale (hence càdlàg $\mathbf{P}_x$-a.s.), the only possible left discontinuity arising at time T. Thus, if $Z'_t := Z_{t-}$ and $Y_t := Z'_{t\wedge T}$, the process Y is a continuous $\mathbf{P}_x$-supermartingale on $[0,\infty]$ such that $Y_s = Z_s$ for $s < T$ (Blumenthal and Getoor (1968) or Fukushima, Oshima, and Takeda (1994)). If f is, moreover, a **potential** in U (i.e., every nonnegative $\mathcal{L}$-harmonic minorant of f in U vanishes), then $Z'_T = 0$ $\mathbf{P}_x$-a.s. since $x \mapsto \mathbf{E}_x(Z'_T)$ is $\mathcal{L}$-harmonic. Recall that f is a potential in U iff $f \not\equiv \infty$ and $f = G_U(\mu)$ for some positive measure μ in U. Here and below, G_U denotes the $\mathcal{L}$-Green function in U (Ancona (1990)).

§7. The Probability of a Crossing.

We need the following "condenser" construction (compare Deny (1970)). Let A and B be two subsets in M and let

$$\Gamma := \Gamma(A, B) := \{f \in \mathbf{D}\,;\ f = 1 \text{ q.e. in } A, f = 0 \text{ q.e. in } B\}\,.$$

Assume that Γ is not empty and that $A \cup B$ is not polar. Then, by a standard Hilbert-space argument, there is a unique minimizer $F := F_{A,B} \in \Gamma$ of $\|u\|_e$. If $A \cup B$ is polar, set $F_{A,B} := \frac{1}{2}$ q.e. in M.

Clearly $0 \le F \le 1$ since $g = (F \wedge 1) \vee 0 \in \Gamma$ and $\|g\|_e \le \|F\|_e$. Note that $\Gamma(A, B) = \Gamma(A', B')$ if A' and B' are the fine closures of A and B, respectively. Since $\Gamma \neq \varnothing$, the intersection $A' \cap B'$ is polar.

The definition of F implies that $a_{\mathcal{L}}(F, v) \ge 0$ for every $v \in \mathbf{D}$ such that $v = 0$ q.e. in B and $v \ge 0$ q.e. in A. (Note that $\|F\|_e \le \|1 \wedge [F + \lambda v]\|_e \le \|F + \lambda v\|_e$ for $\lambda \ge 0$.) By Section 6.1, it follows that if B is closed, then F is equal quasi-everywhere in $U := M \setminus B$ to an $\mathcal{L}$-superharmonic function in U; moreover, the positive measure $\mu := -\mathcal{L}(F)$ in U does not charge polar sets and $\mu(U \setminus A') = 0$. The latter follows from the fact that there exists $\varphi \in \mathbf{D}_0$ which vanishes q.e. in $A \cup B$ and which is > 0 q.e. in $M \setminus (A' \cup B)$ (see, e.g., Fukushima, Oshima, and Takeda (1994), p. 168 and Theorem 4.4.1 there).

Let A and B be two finely closed sets in M such that $\Gamma(A, B) \neq \varnothing$. Let

$$S := \inf\{t \ge 0\,;\ X_t \in B\}$$

and let $F := F_{A,B}$. We then have the following inequality:

PROPOSITION 7.1. *For every $x \in A \setminus B$ that is not finely isolated in A,*

$$\mathbf{P}_x(S < \tau) \le 2\mathbf{E}_x \left(\int_0^{S \wedge \tau} |\widetilde{\nabla} F(X_s)|^2 \, ds \right) .$$

The proposition will be proved in two stages: We first consider the case where B is closed in M, and later establish the general case by a limit argument. Obviously, we may assume that both A and B are nonpolar.

Proof for B closed. Let $U := M \setminus B$ and choose $F = F_{A,B}$ so as to be $\mathcal{L}$-superharmonic in U. Thus, $F(x) = 1$. By the Riesz decomposition in U, we have

$$F = p + h$$

in U, where h is a bounded (nonnegative) $\mathcal{L}$-harmonic function in U and p is a bounded $\mathcal{L}$-potential in U. We know that $\lim_{t\uparrow S\wedge\tau} p(X_t) = 0$ $\mathbf{P}_x$-a.s. and that there is a bounded continuous $\mathbf{P}_x$-martingale $\{Y_s\}_{0\le s\le\infty}$ such that $Y_s = h(X_s)$ for $s < S\wedge\tau$. Thus,

$$\mathbf{E}_x(Z) = F(x) - p(x) = 1 - p(x)\,,$$

where $Z := \lim_{s\uparrow S\wedge\tau} F(X_s)$.

Similarly, $F^2 = q + k$ in U, where q is a difference of two bounded $\mathcal{L}$-potentials in U and k is bounded and $\mathcal{L}$-harmonic in U. (Proof: Observe that $F^2 = 2F - [1-(1-F)^2]$ is a difference of two bounded positive $\mathcal{L}$-superharmonic functions in U since $(1-F)^2$ is subharmonic.) Hence, as before,

$$\mathbf{E}_x(Z^2) = F(x)^2 - q(x) = 1 - q(x)\,.$$

The continuity of $t \mapsto F(X_t)$ on $[0,\tau)$ yields $Z = F(X_S) = 0$ $\mathbf{P}_x$-a.s. on $\{S<\infty\} = \{S<\tau\}$. It follows that $2p(x) - q(x) = 1 - \mathbf{E}_x[(2Z - Z^2)\mathbf{1}_{\{S=\infty\}}]$ and

$$\mathbf{P}_x(S<\infty) = \mathbf{E}_x(1 - \mathbf{1}_{\{S=\infty\}}) \le 2p(x) - q(x)\,.$$

To conclude, consider the positive measure $\mu := -\mathcal{L}(F)$ in U. It is easily checked that $\mathcal{L}(F^2) = 2|\widetilde{\nabla} F|^2\cdot\sigma_M - 2F\cdot\mu$ in the weak sense in U. Thus, if G_U denotes, as before, the $\mathcal{L}$-Green function in U, we have $p = G_U(\mu)$ in U and $q = 2G_U\big(-|\widetilde{\nabla} F|^2\big) + 2G_U(F\cdot\mu)$ in U. Note that $G_U(F\cdot\mu) = G_U(\mu)$ because $F=1$ q.e. in A and μ is supported by A, so that the preceding expression of q is meaningful everywhere in U; moreover $2p(x) - q(x) = 2G_U\big(|\widetilde{\nabla} F|^2\big)(x)$. The result follows. ∎

For the second part of the proof, we need the following observations.

Remark 7.2. If ν is a positive and compactly supported measure in M, then obviously $\{G(\nu)=\infty\}\ne M$; hence, $G(\nu)<\infty$ q.e. Since, for any $a\in M$, the map $x\mapsto G(a,x)$ is bounded outside any neighborhood of a, it follows that, more generally, for each finite positive measure μ in M, the potential $G(\mu)$ is finite q.e. in M.

Remark 7.3. Let U be a finely open subset in M and let f be a nonnegative function in U that we extend by setting $f=0$ outside U. Assume that f is excessive for the process X stopped outside U. Since f is finely continuous with respect to the stopped process (Blumenthal and Getoor (1968), p.74), it is clear that f is also finely continuous in U. This applies, in particular, to $f(x) := \mathbf{E}_x\left(\int_0^S \varphi(X_t)\,dt\right)$, $x\in U$, where $\varphi\ge 0$ is Borel measurable in U and where S is the first entry time in $\widehat{M} \setminus U$. Another example is the function $f : x \mapsto \mathbf{P}_x(S < \tau)$, $x \in U$. (See also Fuglede (1972), Thm. 9.10.)

Proof of Proposition 7.1 for general B. We may write $B = N \cup [\bigcup_{n \geq 1} B_n]$ with N polar and $\{B_n\}$ an increasing sequence of nonpolar closed sets. This follows, for example, from the existence of a quasi-continuous function f in M such that $f = 0$ in B and $f > 0$ in $U := M \setminus B$.

Let $F_n := F_{A,B_n}$. Since $\Gamma(A,B) = \bigcap_{n \geq 1} \Gamma(A, B_n)$, it is easily seen that $F_n \to F = F_{A,B}$ in $\mathbf{D}$, which means that $\widetilde{\nabla} F_n \to \widetilde{\nabla} F$ in $L^2(M, T(M))$. Moreover, since $|\widetilde{\nabla} F_n|^2 \to |\widetilde{\nabla} F|^2$ in $L^1(\sigma_M)$, we may, after extracting a subsequence, assume that $\big|\, |\widetilde{\nabla} F_n|^2 - |\widetilde{\nabla} F|^2 \,\big| \leq g_n$, where $\{g_n\}$ is a decreasing sequence of nonnegative integrable functions such that $\int_M g_n \, d\sigma_M \to 0$.

Clearly, if $S_n := \inf\{t \geq 0\,;\; X_t \in B_n\}$, then $\mathbf{P}_z(S < \tau) = \lim_{n \to \infty} \mathbf{P}_z(S_n < \tau)$, $z \in U$.

On the other hand, we claim that

$$\mathbf{E}_z \left(\int_0^{S_n \wedge \tau} |\widetilde{\nabla} F_n(X_s)|^2 \, ds \right) \to \mathbf{E}_z \left(\int_0^{S \wedge \tau} |\widetilde{\nabla} F(X_s)|^2 \, ds \right)$$

for quasi-every $z \in U$. In fact, if $z \in U$ is such that $G(g_1)(z) + G(|\widetilde{\nabla} F|^2)(z) < \infty$, then the Lebesgue dominated convergence theorem shows that $G\left(\big|\, |\widetilde{\nabla} F_n|^2 - |\widetilde{\nabla} F|^2 \,\big| \right)(z) \to 0$. Also,

$$\begin{aligned}
&\left| \mathbf{E}_z \left(\int_0^{\tau \wedge S} |\widetilde{\nabla} F(X_s)|^2 \, ds \right) - \mathbf{E}_z \left(\int_0^{\tau \wedge S_n} |\widetilde{\nabla} F_n(X_s)|^2 \, ds \right) \right| \\
&\qquad \leq \mathbf{E}_z \left(\int_0^{\tau} \big|\, |\widetilde{\nabla} F(X_s)|^2 - |\widetilde{\nabla} F_n(X_s)|^2 \big| \, ds \right) + \mathbf{E}_z \left(\int_{\tau \wedge S}^{\tau \wedge S_n} |\widetilde{\nabla} F(X_s)|^2 \, ds \right) \\
&\qquad = G\left(\big|\, |\widetilde{\nabla} F|^2 - |\widetilde{\nabla} F_n|^2 \,\big| \right)(z) + \mathbf{E}_z \left(\int_{\tau \wedge S}^{\tau \wedge S_n} |\widetilde{\nabla} F(X_s)|^2 \, ds \right) .
\end{aligned}$$

The last expectation is easily seen to converge to zero as $n \to \infty$ and our claim follows since $G(g_1)(z) + G(|\widetilde{\nabla} F|^2)(z) < \infty$ q.e. in M by Remark 7.2.

We conclude from the above that $\mathbf{P}_x(S < \tau) \leq 2\mathbf{E}_x \left(\int_0^{S \wedge \tau} |\widetilde{\nabla} F(X_s)|^2 \, ds \right)$ q.e. in A. Since each side of this inequality is a finely continuous function of $x \in U$ by Remark 7.3, it must hold for every $x \in A \setminus B$ that is not finely isolated in A. The proof of Proposition 7.1 is complete. ∎

§8. The Crossing Estimate on a Riemannian Manifold.

Let A and B be two disjoint finely closed subsets of M *without finely isolated points* such that $\Gamma(A,B) \neq \varnothing$. We define by induction an increasing sequence of stopping times as follows: Let

$$S_0 := \inf\{t \geq 0\,;\ X_t \in A \cup B\}\,.$$

Define

$$S_1 := \begin{cases} \inf\{t \geq S_0\,;\ X_t \in B\} & \text{if } S_0 < \tau \text{ and } X_{S_0} \in A, \\ \inf\{t \geq S_0\,;\ X_t \in A\} & \text{if } S_0 < \tau \text{ and } X_{S_0} \in B, \\ \infty & \text{if } S_0 = \infty. \end{cases}$$

In general, let

$$S_{n+1} := \begin{cases} \inf\{t \geq S_n\,;\ X_t \in B\} & \text{if } S_n < \tau \text{ and } X_{S_n} \in A, \\ \inf\{t \geq S_n\,;\ X_t \in A\} & \text{if } S_n < \tau \text{ and } X_{S_n} \in B, \\ \infty & \text{if } S_n = \infty. \end{cases}$$

Let $N := \sup\{n \geq 1\,;\ S_n < \tau\}$ be the number of crossings between A and B realized by the process. Proposition 7.1 and the Markov property yield the following estimate of the expectation of N. As mentioned in the introduction, T. Lyons and Zheng (1988) give an energy estimate of the expectation of the number of crossings on the time interval $[0,1]$ when the initial distribution is a stationary reference measure. The corresponding bound for the time interval $[0,T]$ explodes as $T \to \infty$. Compare Corollary 8.5 below.

THEOREM 8.1. *Let $F := F_{A,B}$ be the element of $\mathbf{D}$ with minimum energy $\|F\|_e$ such that $F = 1$ q.e. in A and $F = 0$ q.e. in B. Then the following inequality holds for all $x \in M$:*

$$\mathbf{E}_x(N) \leq 2\,G(\widetilde{|\nabla F|^2})(x)\,.$$

Note that in our notation, $2G$ is the Green function with respect to $\frac{1}{2}\mathcal{L}$. Similarly, the factor 2 in Proposition 7.1 vanishes if one replaces $\{X_t\}$ by the diffusion driven by $\frac{1}{2}\mathcal{L}$.

REMARK 8.2. If A and B are only assumed to be finely closed with $\Gamma(A,B) \neq \varnothing$ (but not necessarily disjoint or finely perfect), then the inequality in Theorem 8.1 holds for q.e. x in M. This follows from the polarity of $A \cap B$ and the fact that every set $N \subset M$ that is finely discrete must be polar (a semipolar set is polar).

REMARK 8.3. If A or B is empty, the result is trivial.

Proof. Assume first that $x \in A$. Given n, let $S' := \inf\{t \geq 0\,;\ X_t \in A\}$ if n is odd and $S' := \inf\{t \geq 0\,;\ X_t \in B\}$ otherwise.

Then, by the Markov property, Proposition 7.1, and the fact that $F_{B,A} = 1 - F_{A,B}$, we have

$$\begin{aligned}
\mathbf{E}_x\left(\mathbf{1}_{\{S_{n+1}<\tau\}}\right) &= \mathbf{E}_x\left(\mathbf{1}_{\{S_n<\tau\}}\mathbf{E}_x\left(\mathbf{1}_{\{S_{n+1}<\tau\}}\,|\,\mathcal{F}_{S_n}\right)\right) \\
&= \mathbf{E}_x\left(\mathbf{1}_{\{S_n<\tau\}}\mathbf{E}_{X_{S_n}}\left(\mathbf{1}_{\{S'<\tau\}}\right)\right) \\
&\le 2\mathbf{E}_x\left(\mathbf{1}_{\{S_n<\tau\}}\mathbf{E}_{X_{S_n}}\left(\int_0^{\tau\wedge S'} |\widetilde{\nabla}F(X_s)|^2\,ds\right)\right) \\
&= 2\mathbf{E}_x\left(\int_{\tau\wedge S_n}^{\tau\wedge S_{n+1}} |\widetilde{\nabla}F(X_s)|^2\,ds\right).
\end{aligned}$$

It follows that $\mathbf{E}_x\left(\sum_{n\ge 1}\mathbf{1}_{\{S_n<\tau\}}\right) \le 2\mathbf{E}_x\left(\int_0^\tau |\widetilde{\nabla}F(X_s)|^2\,ds\right) = 2G(|\widetilde{\nabla}F|^2)(x)$.

The case where $x \in B$ is, of course, treated similarly. Finally, if $x \notin A \cup B$, then we may write, using the cases already proved and the Markov property,

$$\begin{aligned}
\mathbf{E}_x(N) &= \mathbf{E}_x(N\mathbf{1}_{\{S_0<\infty;X_{S_0}\in A\}}) + \mathbf{E}_x(N\mathbf{1}_{\{S_0<\infty;X_{S_0}\in B\}}) \\
&= \mathbf{E}_x(\mathbf{1}_{\{S_0<\infty;X_{S_0}\in A\}}\mathbf{E}_{X_{S_0}}(N)) + \mathbf{E}_x(\mathbf{1}_{\{S_0<\infty;X_{S_0}\in B\}}\mathbf{E}_{X_{S_0}}(N)) \\
&\le 2\mathbf{E}_x\left(\mathbf{1}_{\{S_0<\infty;X_{S_0}\in A\}}\int_{S_0}^\tau |\widetilde{\nabla}F(X_t)|^2\,dt + \mathbf{1}_{\{S_0<\infty;X_{S_0}\in B\}}\int_{S_0}^\tau |\widetilde{\nabla}F(X_t)|^2\,dt\right) \\
&\le 2\mathbf{E}_x\left(\int_0^\tau |\widetilde{\nabla}F(X_t)|^2\,dt\right) = 2G(|\widetilde{\nabla}F|^2)(x)\,. \qquad \blacksquare
\end{aligned}$$

Theorem 8.1 yields easily the following corollaries.

COROLLARY 8.4. *Let $u \in \mathbf{D}$ and let a, b be reals such that $a < b$. Set $A := \{x \in M\,;\ u(x) \le a\}$ and $B := \{x \in M\,;\ u(x) \ge b\}$ and let $F := F_{A,B}$ (clearly, $\Gamma(A,B) \ne \varnothing$). Then the expectation of the number N of crossings over $[a,b]$ of $t \mapsto u(X_t)$, $0 \le t < \tau$, satisfies*

$$\mathbf{E}_x(N) \le 2G(|\widetilde{\nabla}F|^2)(x)$$

for quasi-every $x \in M$. Moreover, the number N' of crossings over $[a,b]$ of $u(X_t)$ on the time interval $(0,\tau)$ satisfies $\mathbf{E}_x(N') \le 2G(|\widetilde{\nabla}F|^2)(x)$ for all *$x \in M$.*

(Note that we may have $N = N' + 1$.) Recall that $G(|\widetilde{\nabla}F|^2) < \infty$ q.e. in M since $|\widetilde{\nabla}F|^2 \in L^1(\sigma_M)$ (see Remark 7.2).

Proof. Since u is quasi-finely continuous, A and B differ from their fine closures by at most polar subsets in M. The first claim then follows from Theorem 8.1 and Remark 8.2.

The second claim results from the first, the Markov property, and the fact that $\{X_t\}_{t>0}$ avoids almost surely the polar set C where the first claim fails. That is, if $\varepsilon > 0$ and if

N^ε is the number of crossings over $[a,b]$ of $t \mapsto u(X_t)$, $\varepsilon \le t < \tau$, we have, since $X_\varepsilon \notin C$ $\mathbf{P}_x$-a.s.,

$$\begin{aligned}\mathbf{E}_x(N^\varepsilon) &= \mathbf{E}_x\left(\mathbf{1}_{\{\tau>\varepsilon\}}\mathbf{E}_x\left(N^\varepsilon \mid \mathcal{F}_\varepsilon\right)\right) = \mathbf{E}_x\left(\mathbf{1}_{\{\tau>\varepsilon\}}\mathbf{E}_{X_\varepsilon}(N)\right)\\ &\le 2\mathbf{E}_x\left(\mathbf{1}_{\{\tau>\varepsilon\}}G(|\widetilde{\nabla}F|^2)(X_\varepsilon)\right) \le 2G(|\widetilde{\nabla}F|^2)(x)\,,\end{aligned}$$

where we have used the fact that $G(|\widetilde{\nabla}F|^2)$ is an excessive function. Letting ε go to zero, we obtain the desired estimate. ∎

COROLLARY 8.5. *Let μ be any probability measure in M. Then, with the notations and assumptions of Corollary 8.4, we have*

$$\mathbf{E}_\mu(N') \le 2\,(b-a)^{-2}\|u\|_e^2\|G\mu\|_\infty\,.$$

If μ has finite energy (or does not charge polar sets), we may replace N' by N in this inequality.

Proof. Integrate with respect to μ the first inequality in Corollary 8.4. Use Fubini's theorem and the fact that by the very definition of $F = F_{A,B}$ (and the contraction property of the Dirichlet norm $a_{\mathcal{L}}$), we have $\|\widetilde{\nabla}F\|^2_{L^2(M)} \le (b-a)^{-2}\,\|u\|_e^2$. ∎

Let V be a relatively compact open subset in M, let $x \in V$ and let T be the first entry time in ∂V. Let N''_V be the number of crossings over $[a,b]$ of $t \mapsto u(X_t)$, $T \le t < \tau$. The distribution μ of X_T under $\mathbf{P}_x$ (the harmonic measure with respect to x in V) satisfies $G(\mu)(\bullet) = G(x,\bullet)$ q.e. outside V (in potential-theory terminology, $G(\mu)$ is the regularized "réduite" of G_x on $M \setminus V$; see Ancona (1990)). Thus, we obtain the following corollary.

COROLLARY 8.6. *With the above notations, we have*

$$\mathbf{E}_x(N''_V) \le \frac{2\,\|u\|_e^2}{(b-a)^2}\,\sup\{G(x,z)\,;\,z\notin V\}\,.$$

Note that in general, $\sup\{G(x,z)\,;\; z \notin V\} \ge [\operatorname{cap}(V)]^{-1}$; if $G(x,\bullet)$ is constant on ∂V, then $\sup\{G(x,z)\,;\; z\notin V\} = [\operatorname{cap}(V)]^{-1}$.

REMARK 8.7. It is not difficult to construct an example (with $M := \mathbb{R}^d$, $d \ge 3$ and $\mathcal{L} := \Delta$) where, in the notations of Corollary 8.4, $N' = +\infty$ $\mathbf{P}_x$-a.s. and hence $G(|\widetilde{\nabla}F|^2)(x) = +\infty$.

§9. L^2 and Almost-Sure Convergence as $t \uparrow \tau$.

Fix $a \in M$ and let $\{T_j\}_{j\geq 1}$ be an increasing sequence of stopping times $< \tau$ such that $T_j \uparrow \tau$ when $j \to \infty$ $\mathbf{P}_a$-a.s. We also assume that $\mathbf{E}_a(s(X_{T_1})) < \infty$ for all potentials s; this is the case if $T_1 \geq t_0$ $\mathbf{P}_a$-a.s. for some $t_0 \in (0, \infty)$ or if $T_1 \geq S_V := \inf\{t \geq 0\,;\ X_t \notin V\}$ for some neighborhood V of a. The aim of this section is to prove the following statement.

THEOREM 9.1. *Let $f \in \mathbf{D}$ and let $a \in M$. Then $\{f(X_t)\}$ converges $\mathbf{P}_a$-almost surely as $t \uparrow \tau$. Moreover, $\{f(X_{T_j})\}_{j\geq 1}$ converges in $L^2(\Omega, \mathcal{F}, \mathbf{P}_a)$.*

Recall that the asserted almost sure convergence was already obtained by Doob (1962) within the framework of Green's spaces and using Martin boundary theory. It also follows from Corollary 8.6 above and we shall indicate below a simple direct argument.

We shall rely on the decomposition $f = f_0 + h$ from Lemma 5.1. Notice that if f is $\mathcal{L}$-subharmonic, then $h \geq f$ since every supersolution in $\mathbf{D}_0$ is nonnegative (Section 6.1). Applying this to $|h|$ for h harmonic and $\mathcal{L}$-Dirichlet finite in M, we obtain the next well-known lemma.

LEMMA 9.2. *Every $\mathcal{L}$-harmonic function in $\mathbf{D}$ is the difference of two nonnegative Dirichlet finite harmonic functions.*

It is then easy to settle the case of Theorem 9.1 where f is $\mathcal{L}$-harmonic.

LEMMA 9.3. *Assume that f is $\mathcal{L}$-harmonic. Then $\mathbf{P}_a$-a.s., $\{f(X_t)\}$ converges to a square integrable random variable Z as $t \uparrow \tau$. Moreover, for each stopping time $T < \tau$, $f(X_T) = \mathbf{E}_a(Z \mid \mathcal{F}_T)$ $\mathbf{P}_a$-a.s.*

Proof. To prove a.s. convergence, assume, as we may by Lemma 9.2, that $f \geq 0$ and set $f(\partial) := 0$. Clearly, $\{f(X_t)\}_{t\in\mathbb{Q}_+}$ is a supermartingale and must have a left-hand limit at each $t > 0$, $\mathbf{P}_a$-a.s. Since f is continuous in M, the a.s. convergence follows.

Set $Z := \limsup_{t\uparrow\tau} f(X_t)$. We have $\mathcal{L}(f^2) = 2|\widetilde{\nabla} f|^2 = g \in L^1_+(\sigma_M)$ and $G(g)$ is a potential (recall Remark 7.2). Note that $v = f^2 + G(g)$ is ≥ 0 and $\mathcal{L}$-harmonic. In particular, $G(g)$ is finite and continuous. Now $\mathbf{E}_a(v(X_T)) \leq v(a)$ (apply, e.g., Section 6.2 to $v \wedge n$, $n \geq 1$) and $\mathbf{E}_a(f^2(X_T)) \leq f^2(a) + G(g)(a)$, which shows that $\{f(X_R)\,;\ R$ a stopping time, $R < \tau\}$ is bounded in $L^2(\mathbf{P}_a)$ and $\mathbf{P}_a$-uniformly integrable.

If $K \subset M$ is compact and if $S_K := \inf\{t \geq 0\,;\ X_t \notin K\}$, then the process $\{f(X_{S_K\wedge t})\}_{t\geq 0}$ is a bounded $\mathbf{P}_a$-martingale. For a given exhaustion $\{K_j\}$, $\{f(X_{S_{K_j}})\}$ is an L^2-bounded $\mathbf{P}_a$-martingale (so $Z \in L^2(\mathbf{P}_a)$) and simple limit arguments yield

$$\mathbf{E}_a(Z \mid \mathcal{F}_T) = \lim_{j\to\infty} \mathbf{E}_a(Z \mid \mathcal{F}_{T\wedge S_{K_j}}) = \lim_{j\to\infty} f(X_{T\wedge S_{K_j}}) = f(X_T) \quad \mathbf{P}_a\text{-a.s.} \qquad \blacksquare$$

Lemma 9.4. *Let $s = G(\lambda)$ be an $\mathcal{L}$-potential in M generated by a positive measure λ in M that does not charge polar sets. Then* $\lim_{j\to\infty} \mathbf{E}_a(s(X_{T_j})) = 0$.

Proof. The point here is that by the assumption on λ, the potential s is a sum $\sum_{p\geq 1} s_p$ of bounded nonnegative potentials (see, e.g., Fukushima, Oshima, and Takeda (1994), p. 82). For each p, $\lim_{j\to\infty} \mathbf{E}_a(s_p(X_{T_j})) = 0$ by Section 6.2. Thus, $\limsup_{j\to\infty} \mathbf{E}_a(s(X_{T_j})) \leq \sum_{p\geq N} \mathbf{E}_a(s_p(X_{T_1}))$ for each integer $N > 1$. Since $\mathbf{E}_a(s(X_{T_1})) = \sum_{p\geq 1} \mathbf{E}_a(s_p(X_{T_1})) < \infty$, the last expression tends to zero when $N \to \infty$. ∎

The next lemma is the well-known interpretation of "réduites" in the framework of Dirichlet spaces. The proof is straightforward.

Lemma 9.5. *Let $u \in \mathbf{D}_0$ and let s be a minimizer of $\|s\|_e$ for $s \in \mathbf{D}_0$, $s \geq u$. Then s admits an $\mathcal{L}$-superharmonic version.*

In particular, for each $u \in \mathbf{D}_0$, there is an $\mathcal{L}$-potential s such that $s \in \mathbf{D}_0$ and $|u| \leq s$ q.e.

Lemma 9.6. *If f is $\mathcal{L}$-superharmonic in M and $f \in \mathbf{D}_0$, then $\lim_{t\uparrow\tau} f(X_t) = 0$ $\mathbf{P}_a$-a.s. and $\{f(X_{T_j})\}_{j\geq 1}$ converges in $L^2(\mathbf{P}_a)$. Moreover, there is a constant C such that $\mathbf{E}_a(f(X_T)^2) \leq C$ for all stopping times T such that $T \geq T_1$.*

Proof. We know that $f(X_t)$ is a nonnegative càdlàg $\mathbf{P}_a$-supermartingale on $(0, \infty]$. Moreover, we have $\lim_{t\uparrow\tau} f(X_t) = 0$ $\mathbf{P}_a$-a.s. because f is a potential (or by Lemma 9.4). Note that, thanks to the previous lemmas, the first claim of Theorem 9.1 is now proved.

We have $\mathcal{L}(f^2) = 2|\widetilde{\nabla} f|^2 + 2f\mathcal{L}(f)$ and $g = 2|\widetilde{\nabla} f|^2 \in L^1_+(\sigma_M)$. Thus, $w = G(g) + f^2$ is nonnegative and $\mathcal{L}$-superharmonic. It follows that for T as in the statement,

$$\mathbf{E}_a(|f(X_T)|^2) \leq \mathbf{E}_a(w(X_T)) \leq \mathbf{E}_a(w(X_{T_1})) = C < \infty\,.$$

To get the L^2-convergence, we show that w is a quasi-bounded $\mathcal{L}$-potential. To see this, set $\varphi_k(t) := t^2$ if $0 \leq t \leq k$, $\varphi_k(t) := k^2 + 2k(t-k)$ if $t \geq k$, where $k \in \mathbb{N}$. Thus, $\varphi_k'' = 2\cdot\mathbf{1}_{[0,k]}$ (if we define $\varphi_k''(k) := 2$). By a well-known contraction principle, $u_k = \varphi_k(f) \in \mathbf{D}_0$ and standard arguments show that if $\mu = -\mathcal{L}(f)$, then $\mathcal{L}(u_k) = -\varphi_k'(f)\cdot\mu + \varphi_k''(f)|\widetilde{\nabla} f|^2\sigma_M$ in the weak sense (see Remark 9.7 below). Thus,

$$u_k = G[\varphi_k'(f)\mu] - G[\varphi_k''(f)|\widetilde{\nabla} f|^2]\,.$$

When $k \to \infty$, the last term, $G[\varphi_k''(f)|\widetilde{\nabla} f|^2]$, increases to $2G(|\widetilde{\nabla} f|^2)$, which is an $\mathcal{L}$-potential since $|\widetilde{\nabla} f|^2 \in L^1(\sigma_M)$. It follows that $G(f\,\mu)$ is not ∞ everywhere and that q.e. in M,

$$f^2 = 2G(f\mu) - 2G(|\widetilde{\nabla} f|^2)\,.$$

Now, f^2 is majorized by the $\mathcal{L}$-potential $p = 2G(f\mu)$ and $\mathbf{E}_a(p(X_{T_j}))$ must decrease to zero when $j \to \infty$ by Lemma 9.4. The lemma is proved. ∎

Theorem 9.1 follows now from Lemma 9.5 and Lemma 5.1.

REMARK 9.7. To justify the computation in Lemma 9.6, notice that $\operatorname{div}(u \cdot \mathcal{A}\nabla f) = -u \cdot \mu + (\nabla u \cdot \mathcal{A}\nabla f) \cdot \sigma_M$ in the weak sense for all $u \in \mathbf{D}_0$. This is clear when $u \in C_0^1(M)$. In general, $u = \lim u_n$ q.e. and in $\mathbf{D}$ (hence in $L^1(\mu)$) with $u_n \in C_0^1(M)$ and the formula follows.

Acknowledgement. We are grateful to Itai Benjamini and Oded Schramm for useful discussions. We are also grateful to the organizers of the meeting on "Random Walks and Discrete Potential Theory" in Cortona, Italy, June 1997, where we were able to combine our methods and results.

REFERENCES

ANCONA, A. (1990). Théorie du potentiel sur les graphes et les variétés, in *École d'Été de Probabilités de Saint-Flour XVIII—1988*, Lect. Notes Math. **1427**, Springer, Berlin, pp. 1–112.

BENJAMINI, I. and SCHRAMM, O. (1996). Random walks and harmonic functions on infinite planar graphs using square tilings, *Ann. Probab.* **24**, 1219–1238.

BLUMENTHAL, R.M. and GETOOR, R.K. (1968). *Markov Processes and Potential Theory*, Academic Press, New York.

BRELOT, M. (1969). *Axiomatique des Fonctions Harmoniques.* 2nd ed. Les Presses de l'Université de Montréal, Montreal, Quebec.

CARTWRIGHT, D. and SOARDI, P.M. (1989). Convergence to ends for random walk on the automorphism group of a tree, *Proc. Amer. Math. Soc.* **107**, 817–823.

DENY, J. (1970). Méthodes hilbertiennes en théorie du potentiel, in *Potential Theory (C.I.M.E., I Ciclo, Stresa, 1969)*, Edizioni Cremonese, Rome, pp. 121–201.

DOOB, J.L. (1957). Conditional Brownian motion and the boundary limits of harmonic functions, *Bull. Soc. Math. France* **85**, 431–458.

DOOB, J.L. (1962). Boundary properties of functions with finite Dirichlet integrals, *Ann. Inst. Fourier* **12**, 573–621.

FUGLEDE, B. (1972). *Finely Harmonic Functions.* Lecture Notes in Math. **289**, Springer, Berlin.

FUKUSHIMA, M. (1980). *Dirichlet Forms and Markov Processes.* North-Holland, Amsterdam.

FUKUSHIMA, M., OSHIMA, Y., and TAKEDA, M. (1994). *Dirichlet Forms and Symmetric Markov Processes*, de Gruyter, Berlin.

HERVÉ, M. and HERVÉ R.-M. (1969). Les fonctions surharmoniques associées à un opérateur elliptique du second ordre à coefficients discontinus, *Ann. Inst. Fourier* **19**, 305–359.

LYONS, T.J. and STOICA, L. (1999). The limits of stochastic integrals of differential equations, *Ann. Probab.* **27**, 1–49.

LYONS, T.J. and ZHANG, T.S. (1994). Decomposition of Dirichlet processes and its application, *Ann. Probab.* **22**, 494–524.

LYONS, T.J. and ZHENG, W.A. (1988). A crossing estimate for the canonical process on a Dirichlet space and tightness result, *Astérisque* **157–158**, 249–271.

SOARDI, P.M. (1994). *Potential Theory on Infinite Networks.* Lect. Notes Math. **1590**, Springer, Berlin.

STAMPACCHIA, G. (1965). Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus, *Ann. Inst. Fourier* **15** (1), 189–258.

YAMASAKI, N. (1986). Ideal limit of discrete Dirichlet functions, *Hiroshima Math. J.* **16**, 353–360.

BÂTIMENT 425, MATHÉMATIQUE, UNIVERSITÉ PARIS-SUD, ORSAY 91405, FRANCE
ancona@matups.math.u-psud.fr

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY, BLOOMINGTON, IN 47405-5701, USA
rdlyons@indiana.edu
http://php.indiana.edu/~rdlyons/

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY, GIVAT RAM, JERUSALEM 91904, ISRAEL AND DEPARTMENT OF STATISTICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720-3860, USA
peres@stat.berkeley.edu
http://www.stat.berkeley.edu/~peres/