Z. Wahrscheinlichkeitstheorie verw. Gebiete
62, 73–112 (1983)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Constructions of Local Time for a Markov Process

Bert Fristedt*[1] and S.J. Taylor*[2]

[1] School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455, USA
[2] Department of Pure Mathematics, Liverpool University, Box 147, Liverpool L69 3BX, England

Summary. A sequence of functions defined on the space of excursions of a Markov process from a fixed point is considered. For each of the functions the sum over the excursions that begin by time t is normalized in an appropriate manner. Conditions are obtained for the convergence of the sequence of normalized sums to the local time evaluated at time t. We obtain a unified structure for convergence theorems which includes some new constructions of local time as well as many constructions previously obtained by quite varied techniques.

1. Introduction

When $X(t,\omega)$ is a Markov process and x is a suitable point in the state space, there is a continuous additive functional $A(x,t)$ which grows only when $X(t)=x$. Such a functional is called a version of the local time at x: we use the framework and terminology of Blumenthal and Getoor [1968] which will be explained more fully in the next section. The general theory ensures the existence of $A(x,t)$, and its uniqueness apart from a multiplicative constant. However, much effort has been put into the construction of $A(x,t)$ by limiting processes based on the sample path properties of $X(t)$. In fact, the results of Lévy [1948] for Brownian motion on the line were obtained before there was any general theory. Let us recall these: we write $A(t)$ for $A(0,t)$.

The first concerns the density of the occupation time of $(-\varepsilon,\varepsilon)$;

$$\frac{1}{2\varepsilon}\int_0^t \mathbf{1}_{(-\varepsilon,\varepsilon)}(X(s))\,ds \to A(t) \text{ as } \varepsilon\downarrow 0. \tag{1.1}$$

Now let $D(\varepsilon,t)$ denote the number of times $X(t)$ crosses from ε to 0 before t;

$$\varepsilon^{1/2} D(\varepsilon,t)\to A(t) \text{ as } \varepsilon\downarrow 0. \tag{1.2}$$

* Partially supported by National Science Foundation Grant MCS 78-01168

0044-3719/83/0062/0073/$08.00

The zero set $Z(t)=\{s\in[0,t): X(s)=0\}$ is closed and its complement $[0,t)\setminus Z(t)$ is a countable union of disjoint open intervals of lengths ρ_i where $\sum_{i=1}^{\infty}\rho_i=t$. Set $N(\varepsilon,t)$ equal to the number of $\rho_i\geqq\varepsilon$ and

$$S(\varepsilon,t)=\sum_{\rho_i<\varepsilon}\rho_i.$$

Then

$$\varepsilon^{1/2}N(\varepsilon,t)\to A(t) \text{ as } \varepsilon\downarrow 0 \tag{1.3}$$

and

$$\varepsilon^{-1/2}S(\varepsilon,t)\to A(t) \text{ as } \varepsilon\downarrow 0. \tag{1.4}$$

At first sight all the above constructions look very different and it is puzzling that they should all work. In fact, not only have the above methods all been shown to work in quite general circumstances, but there are in the literature many more such constructions some of which we will consider later. Our main object is to set up a general "umbrella" construction for which as many as possible of the methods of construction are special cases.

The key tool is the theory of Poisson point processes which was first applied to the excursions of $X(t)$ from a fixed point x by Itô [1970]. We apply a family of functions to the excursions, sum over the excursions which start before t, normalize suitably and then proceed to the limit to obtain $A(t)$. We consider three distinct modes of convergence to $A(t)$. General theorems are developed in Sect. 3 which give necessary and sufficient conditions for convergence in probability and L^2-convergence. In Sect. 4 we obtain general sufficient conditions for a.s. convergence which turn out to be surprisingly powerful. Necessary and sufficient conditions for a.s. convergence are only available in one special case. Maisonneuve [1980, 1981] has also observed that general theorems can be proved by using the Poisson point process, the points of which are excursions of a strong Markov process.

As pointed out by Blumenthal and Getoor [1968], the local time $A(x,t)$ at x will, as $t\to\infty$, be unbounded a.s. when $\{x\}$ is recurrent for $X(t)$, but it remains constant for $t\geqq\tau_0$ if $\{x\}$ is a transient set and τ_0 is the last exit time from $\{x\}$. All our theorems are valid whether $\{x\}$ is recurrent or transient, but we give the detailed proofs in Sects. 2 to 4 for the recurrent case and then, in Sect. 5, outline the changes needed to deal with the transient case.

In Sect. 6 we give constructions, a general one of which is due to Maisonneuve [1974], that involve counting excursions; (1.3) was the first such result. Here very little is needed beyond the requirement that the number of excursions counted tend to infinity. In Sect. 7 we discuss intrinsic constructions in which we work directly with the complementary intervals of the zero set and are not concerned with the behavior of $X(t)$ away from 0. Both (1.3) and (1.4) are of this type. The generalized version of (1.3) works for every X and all modes of convergence. We discuss the differing conditions under which (1.4) works for each mode of convergence. The curious thing (Corollary 7.2) is that a minor modification of (1.4) will always work. This result is due to Kingman [1973]. It is worth noting that the strong Markov sets of Hoffmann-Jørgensen [1969] are the zero sets of a suitable Markov process, so that we could construct a local time for such a random set via the sample paths of such an X. For this context, however, it is natural to use only intrinsic constructions, as described in Sect. 7.

In Sect. 8 we generalize a result of Knight [1971] concerning the zero set of the process $\sup\{Y(s): s \leq t\} - Y(t)$ where $Y(t)$ is a symmetric stable process of index α to the case where $Y(t)$ is any strictly stable process. Interest centers on working out the precise effect of the lack of symmetry for differing values of α. For the most part we are content to show that many existing constructions are corollaries of our procedure, though in many cases the constructions can be generalized or their validity extended. However, in Sect. 9 we consider briefly a construction which is new even for Brownian motion – based on the area $\int_0^t (\varepsilon - \varepsilon \wedge |X(s)|)\, ds$. It turns out that ε^{-2} is the correct normalizing factor for every strictly stable process of index $\alpha > 1$.

Although we have developed a structure which seems to include most of the known constructions we should point out two kinds of difficulties which do not allow all constructions of local time to fit under our umbrella. For any Markov process satisfying the conditions which guarantee the existence of local time, Taylor [1973] showed that there is an appropriate Hausdorff measure function $\varphi(s)$ such that

$$\varphi - m(Z \cap (0, t)) = A(t).$$

This is an intrinsic construction based on the zero set, but it does not appear to be covered by our umbrella because Hausdorff measure depends on the order as well as the lengths of the complementary intervals. The second problem arises when the construction is not intrinsic, and cannot be carried out using one excursion at a time. For example, the construction of Getoor [1976] is of this kind; we can only deal with a special case of this, which we discuss in Sect. 6.

Even for particular cases covered by our umbrella, computational problems may arise – firstly when calculating the normalizing constants and secondly when checking the conditions of the theorems. Lemma 3.4 is a surprisingly helpful computational tool. The main result in [Getoor and Millar, 1972] is an example of a result we have not been able to obtain as a corollary of our results because we have been unable to do the computations necessary to check the conditions of the theorem. Getoor and Millar hypothesize the existence of a reference measure; it is not clear how to use that hypothesis as a computational aid in our approach.

2. Preliminaries

We use the framework of Blumenthal and Getoor [1968]. Let X be a standard Markov process with state space S. We are interested in a particular state x such that

$$P^x(\inf\{s > 0: X(s) = x\} = 0) = 1 \tag{2.1}$$

and

$$P^x(\inf\{s>0: X(s)\neq x\}=0)=1. \tag{2.2}$$

Let $Z=\{s: X(s)=x\}$. Condition (2.2) implies that $\bar{Z}$ is nowhere dense a.s. Condition (2.1) implies that $\bar{Z}$ is perfect a.s., that $\bar{Z}-Z$ is at most countable a.s., and that there exists a continuous additive functional $A(t)$, not identically zero, of the process X whose support is $\{x\}$, that is, which is constant on intervals outside of $\bar{Z}$. Such a continuous additive functional, called a *local time at* $\{x\}$, is determined up to a positive multiplicative constant and it is unique if normalized so that

$$E^x\left\{\int_0^\infty e^{-t}\,dA(t)\right\}=1. \tag{2.3}$$

We will not be considering local times at more than the single state x, so the normalization (2.3) is not important to us; it merely ensures that the local time, $A(x,t)$, at x can be integrated in x over a set $\subset S$ to give the occupation time of $\subset$. When we state our results we suppress x and always mean that A is some version of the local time at x and we only require that $E^x\left\{\int_0^\infty e^{-t}\,dA(t)\right\}$ be finite and positive, rather than equal to 1. We assume $X(0)=x$. The reader will see that, for our results, this assumption entails no loss generality.

The inverse function

$$T(\tau)=\inf\{t: A(t)>\tau\} \tag{2.4}$$

is right continuous and, considered as a process in the variable τ, is a subordinator. Thus, for $\theta>0$,

$$E(e^{-\theta T(\tau)})=\exp\left[-\tau\kappa\theta-\tau\int_{(0,\infty]}(1-e^{-\theta s})\,\mu_0(ds)\right] \tag{2.5}$$

where $\kappa\geqq 0$ is the *drift* and μ_0 is the *Lévy measure.* The condition (2.2) ensures that $\mu_0(0,\infty]=+\infty$. We recall that μ_0 is σ-finite and, in fact, $\int_{(0,\infty]}(s\wedge 1)\,\mu_0(ds)<\infty$. We note in passing that if μ_0 is any Borel measure on $(0,\infty]$ satisfying these conditions, there is a standard Markov process X with state space $\mathbb{R}$ for which the local time at zero has an inverse satisfying (2.5) (see, for example, [Horowitz, 1972]). This allows us to construct the local time for a strong Markov set by using the excursions of such a Markov process. We have included the possibility of a finite atom at $+\infty$ which corresponds to an infinite jump in the subordinator T; this can arise when the set in (2.4) is empty which will happen when $\tau\geqq\zeta$, the lifetime of T which is an exponentially distributed random variable. Thus, μ_0 has an atom at $+\infty$ if and only if $\{x\}$ is a transient set for X. In the remainder of this section and in Sect. 3 and 4 we follow Itô [1970] in giving a detailed discussion only for the case where $\{x\}$ is recurrent so that the integral in (2.5) can be taken over $(0,\infty)$. In Sect. 5 we indicate the changes needed in case $\{x\}$ is transient.

Itô's contribution was to think of the excursions of X from $\{x\}$ as a Poisson point process. Suppose $(\alpha(\omega),\beta(\omega))$ is a component of the complement of the closure of the time set $Z(\omega)=\{t: X(t,\omega)=x\}$. Then the piece of the sample path of X for $\alpha \leqq t < \beta$ is called an *excursion*, which we denote by W. Thus,

$$W(t,\omega)=\begin{cases} X(t,\omega) & \text{for } \alpha(\omega)\leqq t<\beta(\omega) \\ \Delta & \text{for } t\geqq \beta(\omega) \end{cases}$$

denotes a process starting at time $\alpha = W^-$ and leaving S for a terminal state Δ at $\beta = W^+$. For each such W we translate the time axis to obtain $V=V(W)$ given by

$$V(t)=W(t+W^-) \qquad \text{for } t\geqq 0. \tag{2.6}$$

If $\mathscr{V}$ is the space of all right continuous functions with left limits from $[0,\infty)\to S\cup\{\Delta\}$ which are "trapped" when they reach Δ, we can endow $\mathscr{V}$ with the Skorohod topology, and consider measures on $\mathscr{V}$ defined at least on the Borel sets.

Now note that $A(t)$ remains constant during the excursion W so we can define

$$\tau_W = A(W^-) = A(W^+) \tag{2.7}$$

and think of τ_W as the time at which the excursion occurs (on the τ time scale). Itô [1970] showed that there is a σ-finite measure ν on $\mathscr{V}$, defined at least on the smallest σ-field $\mathscr{F}$ generated by the finite-dimensional cylinder sets in function space such that:

(i) for any measurable subset D of $[0,\infty)\times\mathscr{V}$, $\#\{(\tau_W, V(W))\in D\}$, where τ_W is defined by (2.7) and $V(W)$ by (2.6), is a Poisson random variable with mean $(\lambda\times\nu)(D)$ where λ denotes Lebesgue measure. (We use the convention that a Poisson variable with infinite mean is a.s. $+\infty$);

(ii) if $\{D_\gamma\}$ is any disjoint family of measurable subsets of $[0,\infty)\times\mathscr{V}$, then the corresponding family of Poisson random variables is an independent family.

We now consider a measurable function $f:\mathscr{V}\to\mathbb{R}^+$ and extend it to excursions by

$$f(W)=f(V(W)).$$

For any Borel set $B\subset\mathbb{R}^+$ we put

$$\mu(B)=\nu\{V: f(V)\in B\}. \tag{2.8}$$

This implies that

$$\int_{\mathscr{V}} f(V)\nu(dV)=\int_{(0,\infty)} r\mu(dr).$$

The key result which we use in all our constructions is the following.

Lemma 2.1. *Suppose X is a standard Markov process satisfying* (2.1) *and* (2.2) *and $f:\mathscr{V}\to\mathbb{R}^+$ is a measurable function on the excursion of X from x satisfying*

$$\int_{\mathscr{V}} [f(V)\wedge 1]\,\nu(dV)<\infty. \tag{2.9}$$

Then, for fixed $b>0$,

$$G(\tau)=\sum_{W^-\leqq T(\tau)}\frac{f(W)}{b}$$

is a subordinator in τ *with* 0 *drift and Lévy measure* η *where* $\eta(s,\infty)=\mu(bs,\infty)$ *and* μ *is defined by* (2.8). *If* (2.9) *does not hold, then* $G(\tau)=\infty$ *a.s. for each* $\tau>0$.

Proof. It is clear that $G(\tau)$ is the sum of its jumps occuring before time τ and that the expected number of jumps occuring during $(\tau_1,\tau_2]$ and having magnitudes belonging to B equals

$$(\tau_2-\tau_1)\,\nu\{V: f(V)/b\in B\}=(\tau_2-\tau_1)\,\eta(B).$$

Suppose (2.9) holds. Then $\int(s\wedge b^{-1})\,\eta(ds)<\infty$; so, corresponding to η there is a corresponding subordinator with drift 0. In fact, $G(\tau)$ is such a subordinator for its jumps have the correct distributions for a subordinator with Lévy measure η; in particular it has independent increments because the excursions corresponding to τ's belonging to one interval are independent of those for τ's belonging to another interval disjoint from the first interval.
ing to another interval disjoint from the first interval.

Suppose (2.9) does not hold. Write

$$G(\tau)\geqq\sum_{n=-\infty}^{\infty}\ \sum_{\substack{W^-\leqq T(\tau)\\ 2^{n-1}<f(W)\leqq 2^n}}\frac{2^{n-1}\wedge 1}{b},$$

an infinite sum of Poisson random variables which, for $\tau>0$, is, according to the Three Series Theorem, equal to $+\infty$ a.s. (since the sum of expected values is $+\infty$ and $2^{n-1}\wedge 1\leqq 1$). □

We adopt the convention that, if f has a subscript, then the corresponding μ and η inherit the same subscript. For consistency we reserve f_0 for the particular function $f_0(V)=V^+$, so $f_0(W)=W^+-W^-$ denotes the duration of the excursion W. This is consistent with (2.5) because the jump of the subordinator $T(\tau)$ at τ_W is exactly the duration of W.

For $m=1,2,\ldots$, let $f_m:\mathscr{V}\to\mathbb{R}^+$ be measurable functions on the excursions of X from $\{x\}$ and let b_m be positive constants. Set

$$F_m(t)=\sum_{W^-\leqq t}\frac{f_m(W)}{b_m}. \tag{2.10}$$

Our main objective is to obtain results of the form $F_m(t)\to A(t)$ as $m\to\infty$ in one or more of three modes of convergence – in probability, L^2, and a.s. To do this we study a sequence of subordinators:

$$G_m(\tau)=\sum_{W^-\leqq T(\tau)}\frac{f_m(W)}{b_m}. \tag{2.11}$$

Clearly,

$$F_m(T(\tau))=G_m(\tau) \tag{2.12}$$

and, since both A and F_m are constant on excursions,

$$G_m(A(t)) = F_m(t), \tag{2.13}$$

even though $T(A(t))$ is not usually equal to t. The following lemma states that $F_m(t) \to A(t)$ is equivalent to $G_m(\tau) \to \tau$.

Lemma 2.2. *Let F_m and G_m be defined by* (2.10) *and* (2.11). *For each of the three modes of convergence the following are equivalent:*

(A) *for all $t > 0$, $F_m(t) \to A(t)$;*
(B) *for each $t_0 > 0$, $F_m \to A$ uniformly on $[0, t_0]$;*
(C) *there exists $\tau_0 > 0$ such that $G_m(\tau_0) \to \tau_0$;*
(D) *for all $\tau > 0$, $G_m(\tau) \to \tau$;*
(E) *for each $\tau_0 > 0$, $G_m(\tau) \to \tau$ uniformly on $[0, \tau_0]$.*

Proof. Part 1. From Lemma 2.1 we see that none of (A)–(E) is true if (2.9) does not hold for all but finitely many f_m; so in the remaining parts of the proof we assume that (2.9) is true for each f_m.

Part 2. The implications (B) ⇒ (A) and (E) ⇒ (D) ⇒ (C) are clear.

Part 3. The implications (A) ⇒ (B) and (D) ⇒ (E), in all modes of convergence, follow from the monotonicity of all the functions involved and the continuity of the limit functions.

Part 4. We prove (C) ⇒ (D) in each mode of convergence. Since each G_m is a subordinator, $G_m((k+1)\tau_0) - G_m(k\tau_0) \to \tau_0$ for each positive integer k. Hence,

$$G_m(k\tau_0) \to k\tau_0 \quad \text{as } m \to \infty. \tag{2.14}$$

Since $G_m(\tau_0/2)$ and $[G_m(\tau_0) - G_m(\tau_0/2)]$ are independent, identically distributed random variables, it follows from (C) that $G_m(\tau_0/2) \to \tau_0/2$. By induction and an application of (2.14) we obtain, for all positive integers k and n, $G_m(k2^{-n}\tau_0) \to k2^{-n}\tau_0$. Now (D) follows from the monotonicity of each G_m.

Part 5. We complete the proof in the L^2-case by showing (A) ⇒ (C) and (D) ⇒ (A). We first want to show that $E(G_m(1)) \to 1$ is a consequence of (A) and also of (D) so that, without loss of generality, we may assume that b_m is chosen so that $E(G_m(1)) = 1$. It is obviously a consequence of (D) [L^2-case]. Let $(\mathscr{F}_t : t \geqq 0)$ denote the family of σ-fields to which the Markov process X is adapted. Then, by (2.11), G_m is a subordinator adapted to the family $(\mathscr{F}_{T(\tau)} : \tau \geqq 0)$; and, for a fixed $t_0, A(t_0)$ is a stopping time with respect to this family. Hence, by (2.13) and Wald's Identity for Lévy processes [Hall, 1970],

$$E(F_m(t_0)) = E(A(t_0)\, E(G_m(1)).$$

Therefore, (A) implies that $E(G_m(1)) \to 1$.

Assume $E(G_m(1)) = 1$. Then $G_m(\tau) - \tau$ is a martingale adapted to the family $(\mathscr{F}_{T(\tau)} : \tau \geqq 0)$, so by (2.13) and [Hall, 1970]

$$E([F_m(t) - A(t)]^2) = E([G_m(A(t)) - A(t)]^2) = E(A(t))\, E([G_m(1) - 1]^2). \tag{2.15}$$

Since

$$E(A(t)) = E\{\inf\{\tau\colon T(\tau) \geqq t\}\} < \infty,$$

we conclude from (2.15) that (A) $\Rightarrow$ (C) and (D) $\Rightarrow$ (A).

Part 6. We prove (D) $\Rightarrow$ (A) in the cases of convergence in probability and a.s. convergence. Fix $t>0$ and $\varepsilon>0$ and use the monotonicity of F_m, G_m, and A to obtain

$$\begin{aligned}\{|F_m(t) - A(t)| > \varepsilon\} &= \bigcup_{j=1}^{\infty} \left\{|F_m(t) - A(t)| > \varepsilon, \frac{j-1}{2^k} \leqq A(t) < \frac{j}{2^k}\right\} \\ &\subset \bigcup_{j=1}^{N} \left\{\left|G_m\left(\frac{j-1}{2^k}\right) - \frac{j}{2^k}\right| > \varepsilon\right\} \cup \bigcup_{j=1}^{N} \left\{\left|G_m\left(\frac{j}{2^k}\right) - \frac{j-1}{2^k}\right| > \varepsilon\right\} \cup \left\{A(t) \geqq \frac{N}{2^k}\right\}.\end{aligned}$$

First choose k such that $2^{-k} < \varepsilon/2$. For convergence in probability choose N so that the last event has probability less than ε and m large enough so that each of the events $\{|G_m(j2^{-k}) - j2^{-k}| > \varepsilon/2\}$, $1 \leqq j \leqq N$, has probability at most ε/N, giving probability less than 3ε for $\{\omega\colon |F_m(t,\omega) - A(t,\omega)| > \varepsilon\}$. For a.s. convergence we see that each of the events $\{|G_m(j2^{-k}) - j2^{-k}| > \varepsilon/2\}$ occurs only finitely often. This gives a.s. convergence of $F_m(t)$ outside the event $\{A(t) \geqq N2^{-k}\}$. Now let $N \to \infty$.

Part 7. To prove (A) $\Rightarrow$ (D) in the cases of convergence in probability and a.s. convergence one uses

$$\{|G_m(\tau) - \tau| > \varepsilon\} = \bigcup_{j=1}^{\infty} \left\{|G_m(\tau) - \tau| > \varepsilon, \frac{j-1}{2^k} \leqq \tau < \frac{j}{2^k}\right\}.$$

We omit the details which are similar to the details in Part 6 of this proof. □

We will not systematically investigate the possibility of replacing "$W^- \leqq t$" by "$W^+ \leqq t$" in (2.10). Often there are no problems in doing so as in Corollary 7.2. However, in other cases the corresponding results are false – an explicit example can be constructed by allowing excursions of long duration to make a contribution that does not vanish in the limit.

We suspect that the statement "There exists $t_0 > 0$ such that $F_m(t_0) \to A(t_0)$" is, for each of the three modes of convergence, equivalent to the statements (A)–(E) in Lemma 2. A proof for the L^2 case is essentially contained in Part 5 of the preceding proof. We are not able to find an argument for either of the other two modes of convergence. A proof for the a.s. case will, via a subsequence argument, yield a proof for the in-probability case.

Our object in the rest of the paper is to obtain analytic conditions which imply the truth of (A) in Lemma 2.2, and to consider whether these conditions are also necessary. However, it is easier to examine these questions in terms of (C), and we can then appeal to Lemma 2.2. In the sequel we will state our results in terms of condition (A) of Lemma 2.2 and will no longer repeat the other equivalent statements.

It is helpful to state a general result about convergence of the ratio of two monotone functions. We will only have need of it for a.s. convergence.

Lemma 2.3. *Suppose Y is a random variable and f and g are random monotonic functions defined on $(0,\gamma)$, $\gamma>0$, such that g is continuous a.s. and either*

$$\text{(i) } f(\varepsilon) \quad \textit{and} \quad g(\varepsilon) \quad \textit{both} \ \to +\infty \quad \textit{as} \ \varepsilon \downarrow 0 \ \textit{a.s.}$$

or

$$\text{(ii) } f(\varepsilon) \quad \textit{and} \quad g(\varepsilon) \quad \textit{both} \quad \downarrow 0 \quad \textit{as} \ \varepsilon \downarrow 0 \ \textit{a.s.}$$

For each of the three modes of convergence the following two statements are equivalent:

$$\frac{f(\varepsilon)}{g(\varepsilon)} \to Y \quad \textit{as} \ \varepsilon \downarrow 0;$$

$$\frac{f(\varepsilon_n)}{g(\varepsilon_n)} \to Y \quad \textit{as} \ n \to \infty$$

for each sequence ε_n defined by

$$\varepsilon_n = \inf\{\varepsilon : g(\varepsilon) = \rho^n\}, \quad \begin{cases} \rho > 1 & \textit{in case (i)} \\ \rho < 1 & \textit{in case (ii)} \end{cases}. \tag{2.16}$$

We omit the easy proof based on the fact that ρ may be taken arbitrarily close to 1.

Throughout the paper c will stand for a finite, positive constant whose value is unimportant and may change from line to line.

3. Convergence in Probability and in L^2

We can now obtain necessary and sufficient conditions for convergence in probability.

Theorem 3.1. *For a standard Markov process X satisfying (2.1) and (2.2), let $(f_m : m=1,2,\ldots)$ be a sequence of nonnegative measurable functions on the space $\mathscr{V}$ of excursions from $\{x\}$, $(b_m : m=1,2,\ldots)$ a sequence of finite, positive constants, and η_m the measure corresponding, as in Lemma 2.1, to f_m and b_m. Then, as $m \to \infty$,*

$$\sum_{W^- \le t} \frac{f_m(W)}{b_m} \to A(t) \quad \textit{in probability} \tag{3.1}$$

for each $t>0$ if and only if, as $m \to \infty$,

$$\eta_m(s,\infty) \to 0 \quad \textit{for} \ s>0 \tag{3.2}$$

and

$$\int_{(0,\beta]} s\,\eta_m(ds) \to 1 \quad \textit{for} \ 0<\beta<\infty. \tag{3.3}$$

Proof. According to Lemma 2.2 the convergence in (3.1) is equivalent to

$$G_m(1) \to 1 \quad \text{in probability.} \tag{3.4}$$

Since convergence in distribution to a constant implies convergence in probability to that constant, we conclude from the continuity theorem for Laplace transforms that (3.4) is equivalent to

$$\int_{(0,\infty)} (1-e^{-\lambda s})\,\eta_m(ds)\to\lambda \qquad \text{for } \lambda>0. \tag{3.5}$$

Integration by parts shows that (3.5) is equivalent to

$$\int_0^\infty \eta_m(s,\infty)\,e^{-\lambda s}\,ds\to 1 \qquad \text{for } \lambda>0$$

which happens if and only if

$$\int_0^\beta \eta_m(s,\infty)\,ds\to 1 \qquad \text{for } 0<\beta<\infty \tag{3.6}$$

(that is, if and only if the measure $\eta(s,\infty)\,ds$ converges vaguely to the atom of size 1 at 0). In case (3.2) holds, integration by parts shows (3.3) to be equivalent to (3.6). On the other hand, suppose (3.6) holds. Then (3.2) follows from

$$\int_\alpha^\beta \eta_m(s,\infty)\,ds\to 0, \qquad 0<\alpha<\beta<\infty,$$

and the monotonicity of $\eta_m(s,\infty)$. □

Remark. The conditions (3.2) and (3.3) are not independent. The above arguments show that (3.2) implies that $\int_{(0,\beta]} s\,\eta_m(ds)\to 1$ for all or no β; (3.3) implies that $\eta_m(s,\infty)\to 0$ for all or no s.

Corollary 3.2. *Suppose that b_m in Theorem 3.1 satisfies*

$$b_m=\int_{(0,\infty)} r\,\mu_m(dr), \tag{3.7}$$

where μ_m, corresponding to f_m, is defined via (2.8). Then (3.1) is true if and only if, as $m\to\infty$,

$$\frac{1}{b_m}\int_{(\beta b_m,\infty)} r\,\mu_m(dr)\to 0 \qquad \textit{for } \beta>0. \tag{3.8}$$

Proof. By (3.7)

$$\int_{(0,\infty)} s\,\eta_m(ds)=1$$

and, hence,

$$\frac{1}{b_m}\int_{(\beta b_m,\infty)} r\,\mu_m(dr)=\int_{(\beta,\infty)} s\,\eta_m(ds)=1-\int_{(0,\beta]} s\,\eta_m(ds)$$

from which the equivalence of (3.3) and (3.8) follows. Since

$$\eta_m(s,\infty)=\mu_m(s\,b_m,\infty)\leqq\frac{1}{s\,b_m}\int_{(sb_m,\infty)} r\,\mu_m(dr),$$

(3.8) ⇒ (3.2). □

Remark. (3.7) $\Rightarrow E(G_m(\tau))=\tau$. This follows from (3.22) appearing in the proof of the forthcoming Theorem 3.5.

Example 7.8 shows that (3.1) may hold with an appropriate normalizing sequence even though it does not hold with the normalizing sequence $(\int r\,\mu_m(dr): m=1,2,\dots)$. In order to avoid the need to check (3.2) and (3.3) for every sequence $(b_m: m=1,2,\dots)$ we now find a canonical sequence which will give convergence provided there is any sequence that does so.

Proposition 3.3. *The assertion* (3.1) *of Theorem* 3.1 *holds for some sequence* $(b_m: m=1,2,\dots)$ *if and only if it is true when* $b_m=a_m$, *where*

$$a_m=\sup\left\{b>0: \frac{1}{b}\int_{(0,b]} r\,\mu_m(dr)\geqq 1\right\}. \tag{3.9}$$

In case $b_m=a_m$, (3.1) *is true if and only if, as* $m\to\infty$,

$$\mu_m(t\,a_m,\infty)\to 0 \quad \textit{for } t>0. \tag{3.10}$$

Remark. The proposition should be interpreted as asserting that if (3.2) and (3.3) hold for some sequence $(b_m: m=1,2,\dots)$, then, for all sufficiently large m, a_m is defined, via (3.10), and is finite and positive.

Proof. Part 1. Let $(b_m: m=1,2,\dots)$ be a sequence satisfying (3.1) and, thus, (3.2) and (3.3); whence,

$$\mu_m(t\,b_m,\infty)\to 0 \quad \text{for } t>0 \tag{3.11}$$

and

$$\frac{1}{b_m}\int_{(0,\beta b_m]} r\,\mu_m(dr)\to 1 \quad \text{for } 0<\beta<\infty. \tag{3.12}$$

Let $0<\varepsilon<\frac{1}{2}$. From (3.12) we deduce that, for sufficiently large m,

$$\frac{1}{(1-2\varepsilon)\,b_m}\int_{(0,(1-2\varepsilon)b_m]} r\,\mu_m(dr)>\frac{1}{1-\varepsilon} \tag{3.13}$$

and

$$\frac{1}{(1+2\varepsilon)\,b_m}\int_{(0,(1+2\varepsilon)b_m]} r\,\mu_m(dr)<\frac{1}{1+\varepsilon}. \tag{3.14}$$

From (3.13) we see that the set in (3.9) is nonempty so that $a_m(\leqq\infty)$ is defined and $a_m\geqq(1-2\varepsilon)\,b_m$.

Let $b\geqq(1+2\varepsilon)\,b_m$. From (3.14) we see that

$$\frac{1}{b}\int_{(0,b]} r\,\mu_m(dr)<\frac{1}{1+\varepsilon}+\int_{((1+2\varepsilon)b_m,\,b]}\mu_m(dr)$$

which, by (3.11), is less than 1 for sufficiently large m (independent of b). For such m, $a_m\leqq(1+2\varepsilon)\,b_m$.

Since ε was arbitrary we have shown $a_m/b_m\to 1$ as $m\to\infty$. So (3.1) holds with $b_m=a_m$.

Part 2. Let $b_m = a_m$ at least for all sufficiently large m. Since (3.2) and (3.3) are equivalent to (3.11) and (3.12), we want to show that (3.11) and (3.12) both hold if and only if (3.10) holds. But (3.10) is identical to (3.11) (for $b_m = a_m$), so it remains to prove that (3.9) and (3.10) imply (3.12).

From (3.9) we obtain

$$\frac{1}{a_m} \int_{(0, a_m]} r\,\mu_m(dr) = 1 \tag{3.15}$$

which implies the case $\beta = 1$ of (3.12). The case $\beta > 1$ of (3.12) follows from (3.15) and

$$\frac{1}{a_m} \int_{(a_m, \beta a_m]} r\,\mu_m(dr) \leqq \beta\,\mu_m(a_m, \infty) \to 0,$$

a consequence of (3.10). A similar argument works in case $\beta < 1$.

It is clearly useful to be able to compute $\int_{(0,\infty)} r\,\mu(dr)$ from the sample path properties of X. The following is a surprisingly potent weapon.

Lemma 3.4. *Let* $\varphi(\theta) = -\log E(\exp(-\theta T(1)))$. *Then, for each* $\theta > 0$,

$$\int_{(0,\infty)} r\,\mu(dr) = \varphi(\theta)\, E\left\{ \int_0^\infty e^{-\theta t}\, d\Big(\sum_{W^- \leqq t} f(W) \Big) \right\}.$$

Proof. Note that (2.5) can be written as

$$E(e^{-\theta T(\tau)}) = e^{-\tau \varphi(\theta)}. \tag{3.16}$$

For p a positive integer

$$\begin{aligned}
&\varphi(\theta)\, E\left\{ \int_0^\infty e^{-\theta t}\, d\Big(\sum_{W^- \leqq t} f(W) \Big) \right\} \\
&= \varphi(\theta) \sum_{k=0}^{\infty} E\left\{ \int_{T(k/p)}^{T((k+1)/p)} e^{-\theta t}\, d\Big(\sum_{W^- \leqq t} f(W) \Big) \right\} && (3.17) \\
&= \varphi(\theta) \sum_{k=0}^{\infty} E\left\{ e^{-\theta T(k/p)} \int_0^{T((k+1)/p) - T(k/p)} e^{-\theta t}\, d\Big(\sum_{\substack{T(k/p) < W^- \\ \leqq t + T(k/p)}} f(W) \Big) \right\}. && (3.18)
\end{aligned}$$

Write the expectation in (3.18) as the expectation of the conditional expectation with respect to $\mathscr{F}_{T(k/p)}$ (Recall $\mathscr{F}_t$ is the σ-field with respect to which $X(t)$ is measurable) and use the fact that $T(k/p)$ is $\mathscr{F}_{T(k/p)}$-measurable to get (3.18) equal to

$$\begin{aligned}
&\varphi(\theta) \sum_{k=0}^{\infty} E\{ e^{-\theta T(k/p)} \}\, E\left\{ \int_0^{T(1/p)} e^{-\theta t}\, d\Big(\sum_{W^- \leqq t} f(W) \Big) \right\} \\
&= \varphi(\theta) \sum_{k=0}^{\infty} e^{-\varphi(\theta) k/p}\, E\left\{ \int_0^{T(1/p)} e^{-\theta t}\, d\Big(\sum_{W^- \leqq t} f(W) \Big) \right\} && \text{[by (3.16)]} \\
&\sim p E\left\{ \int_0^{T(1/p)} e^{-\theta t}\, d\Big(\sum_{W^- \leqq t} f(W) \Big) \right\} \quad (\text{as } p \to \infty) && (3.19)
\end{aligned}$$

$$= \sum_{j=0}^{p-1} E\left\{ \int_{T(j/p)}^{T((j+1)/p)} e^{-\theta(t-T(j/p))} d\left(\sum_{W^- \leqq t} f(W) \right) \right\} \tag{3.20}$$

$$= E\left\{ \int_{0}^{T(1)} e^{-\theta(t-T(j(t)/p))} d\left(\sum_{W^- \leqq t} f(W) \right) \right\} \tag{3.21}$$

where the integer $j(t)$ is defined by

$$T(j(t)/p) \leqq t < T((j(t)+1)/p).$$

If $t = W^-$ for some excursion W, then, as $p \to \infty$, $T(j(t)/p) \to t$. Hence, a.s. it is true that $t - T(j(t)/p) \to 0$ almost everywhere with respect to the measure

$$d\left(\sum_{W^- \leqq t} f(W) \right).$$

By the Lebesgue Dominated Convergence Theorem, (3.21) approaches, as $p \to \infty$,

$$E\left(\sum_{W^- \leqq T(1)} f(W) \right) = \int_{(0,\infty)} r\,\mu(dr). \quad \square \tag{3.22}$$

Remark. It is not true that $t - T(j(t)/p) \to 0$ almost everywhere with respect to the measure

$$d\left(\sum_{W^+ \leqq t} f(W) \right).$$

The above proof and, in fact, the lemma itself break down if W^- is replaced by W^+.

We are also able to obtain necessary and sufficient conditions for L^2 convergence.

Theorem 3.5. *For a standard Markov process X satisfying* (2.1) *and* (2.2), *let $(f_m\colon m=1,2,\ldots)$ be a sequence of nonnegative measurable functions on the space $\mathscr{V}$ of excursions from $\{x\}$, $(b_m\colon m=1,2,\ldots)$ a sequence of finite, positive constants, and μ_m the measure corresponding to f_m via* (2.8). *Then, as $m \to \infty$,*

$$\sum_{W^- \leqq t} \frac{f_m(W)}{b_m} \to A(t) \quad \text{in } L^2 \tag{3.23}$$

for each $t > 0$ if and only if, as $m \to \infty$,

$$b_m \sim \int_{(0,\infty)} r\,\mu_m(dr) \tag{3.24}$$

and

$$\frac{1}{b_m^2} \int_{(0,\infty)} r^2\,\mu_m(dr) \to 0. \tag{3.25}$$

Proof. By Lemma 2.2 the convergence in (3.23) is equivalent to $G_m(1) \to 1$ in L^2. For this it is necessary and sufficient that $E(G_m(1)) \to 1$ and $\operatorname{var}(G_m(1)) \to 0$.

From Lemma 2.1 we obtain

$$E(\exp(-\theta G_m(1))) = \exp\left(-\int_{(0,\infty)} (1 - e^{-\theta s})\,\eta_m(ds)\right).$$

Two differentiations with respect to θ yield

$$E(G_m(1)) = \int_{(0,\infty)} s\,\eta_m(ds) = \frac{1}{b_m} \int_{(0,\infty)} r\,\mu_m(dr) \tag{3.26}$$

and

$$\operatorname{var}(G_m(1)) = \int_{(0,\infty)} s^2\,\eta_m(ds) = \frac{1}{b_m^2} \int_{(0,\infty)} r^2\,\mu_m(dr),$$

from which we see that $E(G_m(1)) \to 1$ and $\operatorname{var}(G_m(1)) \to 0$ are equivalent to (3.24) and (3.25), respectively. □

Proposition 3.6. *The assertion* (3.23) *of Theorem* 3.5 *holds for some sequence* $(b_m : m = 1, 2, \ldots)$ *if and only if it is true when* $b_m = a_m$, *where*

$$a_m = \int_{(0,\infty)} r\mu_m(dr).$$

In case $b_m = a_m$, (3.23) *is true if and only if,* $a_m < \infty$ *for all but finitely many* m *and, as* $m \to \infty$,

$$\frac{1}{a_m^2} \int_{(\beta a_m,\infty)} r^2\,\mu_m(dr) \to 0 \quad \textit{for } \beta > 0. \tag{3.27}$$

Proof. The first assertion is obvious. For the second, let $b_m = a_m$ and observe that (3.25) implies (3.27). It remains to prove that (3.27) implies (3.25). This is the case since

$$\begin{aligned} &a_m^{-2} \int_{(0,\infty)} r^2\,\mu_m(dr) \\ &= a_m^{-2} \int_{(0,ta_m]} r^2\,\mu_m(dr) + a_m^{-2} \int_{(ta_m,\infty)} r^2\,\mu_m(dr) \\ &\leqq t\,a_m^{-1} \int_{(0,ta_m]} r\,\mu_m(dr) + a_m^{-2} \int_{(ta_m,\infty)} r^2\,\mu_m(dr) \\ &\leqq t + a_m^{-2} \int_{(ta_m,\infty)} r^2\,\mu_m(dr) \end{aligned}$$

and t may be taken arbitrarily small. □

Example 7.7 shows that L^2-convergence can fail for a monotonic sequence $(f_m : m = 1, 2, \ldots)$ even when we have a.s. convergence. However, if $(f_m : m = 1, 2, \ldots)$ is bounded but $b_m \to \infty$, we always get L^2-convergence, even without monotonicity.

Corollary 3.7. *Suppose that for some constant* K, $f_m(V) \leqq K$ *for all* m *and* V, *that* $b_m = \int_{(0,\infty)} r\,\mu_m(dr) < \infty$, *and that* $b_m \to \infty$ *as* $m \to \infty$. *Then* (3.23) *holds.*

Proof. By Proposition 3.6 we need only verify (3.27), since $b_m = a_m$. But (3.27) is obvious since its left side equals 0 for $b_m \geqq K/\beta$. □

Remark. All results of this section are valid when applied to an uncountable family $\{f_\varepsilon : \varepsilon > 0\}$ with, say, $\varepsilon \downarrow 0$.

4. Almost Sure Convergence

In general we cannot hope to obtain necessary and sufficient conditions for a.s. convergence. Since a.s. convergence can occur without finite moments any conditions involving moments cannot be necessary because they can be violated by adding a second term which converges to zero a.s. but has large finite moments. Unless we can use monotonicity or other dependence properties of the family (f_m) the sufficient conditions for a.s. convergence needs to be strong enough for a Borel-Cantelli argument to work. We start with a crude result, which is relevant when we already know there is L^2-convergence.

Suppose we are in the situation of Theorem 3.5, so that, with

$$G_m(1)=\sum_{W^-\leqq T(1)}\frac{f_m(W)}{b_m},$$
$$\operatorname{var} G_m(1)=\int_{(0,\infty)} r^2\,\mu_m(dr)<\infty.$$

We can apply Chebychev to deduce, for each $\delta>0$,

$$\sum_{m=1}^{\infty} \mathrm{P}\left\{\left|\frac{\sum_{W^-\leqq T(1)} f_m(W)}{\int_{(0,\infty)} r\,\mu_m(dr)}-1\right|>\delta\right\}<\infty \tag{4.1}$$

from

$$\sum_{m=1}^{\infty}\frac{\int_{(0,\infty)} r^2\,\mu_m(dr)}{\left(\int_{(0,\infty)} r\,\mu_m(dr)\right)^2}<\infty. \tag{4.2}$$

Using Borel Cantelli, (4.1) implies that for each δ only finitely many of the events in braces occur a.s., and taking a sequence of values of $\delta\downarrow 0$ gives

$$G_m(1)\to 1 \quad \text{a.s.}$$

If we now apply Lemma 2.2 we complete the proof of:

Theorem 4.1. *For a standard Markov process X satisfying* (2.1) *and* (2.2), *let $(f_m\colon m=1,2,\ldots)$ be a sequence of nonnegative measurable functions on the space $\mathscr{V}$ of excursions from $\{x\}$ and μ_m the measure corresponding to f_m via* (2.8). *If condition* (4.2) *is satisfied, then, for each t,*

$$\frac{\sum_{W^-\leqq t} f_m(W)}{\int_{(0,\infty)} r\,\mu_m(dr)}\to A(t) \quad a.s.$$

as $m\to\infty$.

It is surprising that this apparently weak theorem yields a strong result when applied to a bounded family $(f_\varepsilon\colon \varepsilon>0)$ that is increasing as $\varepsilon\downarrow 0$.

Theorem 4.2. *Let X be a standard Markov process satisfying* (2.1) *and* (2.2). *Let K be a finite constant and let $(f_\varepsilon\colon \varepsilon>0)$ be a family of measurable functions from*

$\mathscr{V}$ to $\mathbb{R}^+$ such that $f_\varepsilon(V) \leqq K$ for $\varepsilon > 0$ and $V \in \mathscr{V}$ and $f_\varepsilon(V)\uparrow$ as $\varepsilon \downarrow 0$ for $V \in \mathscr{V}$. Suppose further that, as $\varepsilon \downarrow 0$,

$$+\infty > b_\varepsilon = \int_{(0,\infty)} r\,\mu_\varepsilon(dr) \to +\infty.$$

Then, as $\varepsilon \downarrow 0$,

$$b_\varepsilon^{-1} \sum_{W^- \leqq t} f_\varepsilon(W) \to A(t) \quad a.s. \tag{4.3}$$

Proof. We first suppose that $\varepsilon \mapsto b_\varepsilon$ is continuous. Preparing to use Lemma 2.3 we let $\rho > 1$ and define ε_n by (2.16) so that

$$b_{\varepsilon_n} = \rho^n.$$

Since

$$\frac{\int_{(0,\infty)} r^2 \mu_{\varepsilon_n}(dr)}{b_{\varepsilon_n}^2} \leqq \frac{K}{b_{\varepsilon_n}} = \frac{K}{\rho^n},$$

(4.2) holds; so an application of first Theorem 4.1 and then Lemma 2.3 yields the result in case $\varepsilon \mapsto b_\varepsilon$ is continuous.

We now drop the assumption that $\varepsilon \mapsto b_\varepsilon$ is continuous and proceed to complete the proof by "patching" $\varepsilon \mapsto f_\varepsilon(W)$ at each discontinuity of $\varepsilon \mapsto b_\varepsilon$. Suppose the discontinuities of $\varepsilon \mapsto b_\varepsilon$ in $(0,1)$ occur on the countable set $\{\varepsilon_i : i = 1, 2, \ldots\}$. Consider the mapping σ: $(0,1) \mapsto (0,3)$:

$$\sigma(\varepsilon) = \varepsilon + 2 \sum_{\varepsilon_i < \varepsilon} 2^{-i} + \sum_{\varepsilon_i = \varepsilon} 2^{-i}.$$

Then σ is monotone increasing in ε with a double jump at each ε_i. Let $\sigma(\varepsilon_i) = \delta_i$ and note that

$$\delta_i - \lim_{\varepsilon \uparrow \varepsilon_i} \sigma(\varepsilon) = 2^{-i} = \lim_{\varepsilon \downarrow \varepsilon_i} \sigma(\varepsilon) - \delta_i.$$

For each excursion W we define $g_\delta(W)$ for $\delta > 0$ as follows, remembering that $\varepsilon \mapsto f_\varepsilon(W)$ is monotonic:

$g_{\sigma(\varepsilon)}(W) = f_\varepsilon(W)$ if σ is continuous at ε;

$g_{\delta_i}(W) = f_{\varepsilon_i}(W)$;

$g_{\delta_i - 2^{-i}}(W) = \lim_{\varepsilon \uparrow \varepsilon_i} f_\varepsilon(W)$;

$g_{\delta_i + 2^{-i}}(W) = \lim_{\varepsilon \downarrow \varepsilon_i} f_\varepsilon(W)$;

by linear interpolation in the intervals $(\delta_i - 2^{-i}, \delta_i)$ and $(\delta_i, \delta_i + 2^{-i})$.

Let η_δ correspond to g_δ as μ corresponds to f in (2.8) and let

$$c_\delta = \int_{(0,\infty)} s\,\mu_\delta(ds).$$

Since $\delta \mapsto g_\delta(W)$ is continuous on each $[\delta_i - 2^{-i}, \delta_i + 2^i]$ and $\delta \mapsto \mu_\delta$ is continuous elsewhere, $\delta \mapsto c_\delta$ is continuous everywhere. Also, $\delta \mapsto g_\delta(W)$ is monotonic, and, hence, the continuous case of the theorem, already proved, implies

$$c_\delta^{-1} \sum_{W^- \leqq t} g_\delta(W) \to A(t) \quad \text{a.s. as} \quad \delta \downarrow 0. \tag{4.4}$$

Now let $\delta \downarrow 0$ omitting the intervals

$$(\delta_i - 2^{-i}, \delta_i) \quad \text{and} \quad (\delta_i, \delta_i + 2^{-i}).$$

The convergence holds on the restricted set, the closure of the image of σ. Since

$$f_\varepsilon(W) = g_{\sigma(\varepsilon)}(W), \qquad b_\varepsilon = c_{\sigma(\varepsilon)},$$

we conclude from (4.4) that, as $\varepsilon \downarrow 0$, (4.3) holds. □

Remark. Unfortunately, the result in which $f_\varepsilon \uparrow$ is replaced by $f_\varepsilon \downarrow$ as $\varepsilon \downarrow 0$ in Theorem 4.2 is in general false; we will return to this case later.

To obtain Theorem 4.1 we used Chebychev to estimate the tails in (4.1). If we use the special nature of the distribution we can obtain better estimates for these tails.

Lemma 4.3. *Suppose X is a standard Markov process satisfying* (2.1) *and* (2.2), *and $f: \mathscr{V} \to \mathbb{R}^+$ is measurable with $f(V) \leqq M$ for all V. Let*

$$b = \int s\, \mu(ds), \qquad c = \int s^2\, \mu(ds), \qquad a = b\, c^{-1/2}.$$

Then, for each $\delta > c$,

$$P\{|b^{-1} \sum_{W^- \leqq T(1)} f(W) - 1| > \delta\} \leqq \begin{cases} 2\exp[-\frac{1}{2}\delta^2 a^2 (1 - \frac{1}{2}\delta a M)] & \text{if } \delta a M \leqq 1 \\ 2\exp\left[-\dfrac{1}{4M}\delta a\right] & \text{if } \delta a M \geqq 1. \end{cases}$$

Remark. These inequalities look very like the negative exponential bounds in [Loève, 1960, p. 254], which is hardly surprising since we use these in the proof.

Proof. With τ_W as defined at (2.7),

$$\sum_{W^- \leqq T(1)} f(W) = \sum_{(\tau_W, f(W)) \in (0,1] \times (0, M]} f(W). \tag{4.5}$$

Write $(0,1] \times (0, M]$ as the union of a sequence $(R_n: n = 1, 2, \dots)$ of disjoint rectangles, an *infinite* sequence being required so that $(\lambda \times \mu)(R_n) < \infty$ for each n. The quantity at (4.5) equals

$$\sum_{n=1}^{\infty} \sum_{(\tau_W, f(W)) \in R_n} f(W). \tag{4.6}$$

The probability that at least one of the inner sums has move than one term is bounded by

$$\sum_{n=1}^{\infty} [1-e^{-(\lambda\times\mu)(R_n)}-(\lambda\times\mu)(R_n)\,e^{-(\lambda\times\mu)(R_n)}]$$
$$\leqq \sum_{n=1}^{\infty} [(\lambda\times\mu)(R_n)]^2. \tag{4.7}$$

where λ denotes Lebesgue measure. The quantity at (4.7) can be made as small as desired by an appropriate choice of $(R_n\colon n=1,2,\ldots)$ and, therefore, with probability as close to one as desired (4.6) equals

$$\sum_{n=1}^{\infty} \sup\{f(W)\colon(\tau_w, f(W))\in R_n\} \tag{4.8}$$

where $\sup\varnothing=0$.

The summands in (4.8) are bounded by M, so the bounds in [Loève, 1960, p. 254] are applicable (a slight extension being needed for the infinite sum). The mean and variance of (4.8) are not exactly b and c; b and c are the mean and variance of (4.6). It can be shown, because (4.7) can be made arbitrarily small, that the ratios between the two means and between the two variances can be made as close to one as desired. In this way we obtain the inequalities in the statement of Lemma 4.3, since the errors introduced on both sides can be made as small as we please. □

By using Lemma 4.3 to obtain a sufficient condition for (4.1) and then using Borel-Cantelli we immediately obtain:

Theorem 4.4. *For a standard Markov process X satisfying* (2.1) *and* (2.2), *let* $(f_m\colon m=1,2,\ldots)$ *be a sequence of nonnegative measurable functions on the space* $\mathscr{V}$ *of excursions from* $\{x\}$ *and* μ_m *the measure corresponding to* f_m *via* (2.8). *For each m suppose that* $b_m=\int_{(0,\infty)} s\,\mu_m(ds)<\infty$ *and, for each W,* $f_m(W)\leqq M_m$. *Put, for each m.*

$$c_m=\int_{(0,\infty)} s^2\,\mu_m(ds), \qquad a_m=b_m c_m^{-1/2}.$$

If, for all $\gamma>0$,

$$\sum_{m=1}^{\infty} e^{-\gamma a_m/M_m}+\sum_{m=1}^{\infty} e^{-\gamma a_m^2}<\infty, \tag{4.9}$$

then, for each $t>0$,

$$b_m^{-1}\sum_{W^-\leqq t} f_m(W)\to A(t) \quad a.s.$$

as $m\to\infty$.

Corollary 4.5. *In the special case where*

$$f_m=\mathbf{1}_{E_m}$$

for some measurable $E_m \subset \mathscr{V}$,

$$\frac{\#\{W: W^- \leqq t \quad and \quad V(W) \in E_m\}}{\nu(E_m)} \to A(t) \quad a.s.$$

if, for every $\gamma > 0$,

$$\sum_{m=1}^{\infty} e^{-\gamma \nu(E_m)^{1/2}} < \infty. \tag{4.10}$$

Proof. In this case $c_m = b_m = \nu(E_m)$, $M_m = 1$, and $a_m = \nu(E_m)^{1/2}$. Clearly, the second term in (4.9) is irrelevant and (4.9) is equivalent to (4.10). □

In Theorem 4.4 there is no monotonicity condition so we cannot replace m by a continuous parameter – and there is no hope of getting a.s. convergence with a continuous parameter without some smoothness condition. We now consider one important special case where a sharp theorem is possible.

Theorem 4.6. *Suppose X is a standard Markov process satisfying* (2.1) *and* (2.2) *and f is a nonnegative measurable function on the space $\mathscr{V}$ of excursions from $\{x\}$. Let $(\mathscr{E}_\varepsilon: \varepsilon > 0)$ be a family of measurable subsets of $\mathscr{V}$ with $\mathscr{E}_\varepsilon \downarrow \varnothing$ as $\varepsilon \downarrow 0$. Put $f_\varepsilon = f \mathbf{1}_{\mathscr{E}_\varepsilon}$ and let μ_ε correspond to f_ε via* (2.8). *Suppose $\varepsilon \mapsto b_\varepsilon$ is a nondecreasing continuous function such that $b_\varepsilon \to 0$ as $\varepsilon \downarrow 0$. For $\rho < 1$ define*

$$\varepsilon_n = \inf\{\varepsilon: b_\varepsilon = \rho^n\}, \qquad n = 1, 2, \ldots.$$

Then

$$b_\varepsilon^{-1} \sum_{W^- \leqq t} f_\varepsilon(W) \to A(t) \quad a.s. \tag{4.11}$$

as $\varepsilon \downarrow 0$ for each $t > 0$ if and only if

$$\sum_{n=1}^{\infty} P\left\{\left|\frac{\sum_{W^- \leqq T(1)} [f_{\varepsilon_n}(W) - f_{\varepsilon_{n+1}}(W)]}{b_{\varepsilon_n} - b_{\varepsilon_{n+1}}} - 1\right| > \delta\right\} < \infty \tag{4.12}$$

for every $\rho < 1$ and $\delta > 0$.

Proof. By Lemma 2.2, (4.11) holds for each t if and only if

$$b_\varepsilon^{-1} \sum_{W^- \leqq T(1)} f_\varepsilon(W) \to 1 \quad \text{a.s.}$$

which, by Lemma 2.3, is equivalent to

$$b_{\varepsilon_n}^{-1} \sum_{W^- \leqq T(1)} f_{\varepsilon_n}(W) \to 1 \quad \text{a.s.} \tag{4.13}$$

for every $\rho < 1$. (Recall: ε_n depends on ρ.)

To begin the proof of the equivalence of (4.12) for every δ and (4.13), assume that (4.12) is true for every δ. By Borel-Cantelli there exists for almost every ω and each rational δ an n_0 such that, for $n \geqq n_0$,

$$\begin{aligned}(1-\delta)(b_{\varepsilon_n} - b_{\varepsilon_{n+1}}) \leqq \sum_{W^- \leqq T(1)} [f_{\varepsilon_n}(W) - f_{\varepsilon_{n+1}}(W)] \\ \leqq (1+\delta)(b_{\varepsilon_n} - b_{\varepsilon_{n+1}}).\end{aligned} \tag{4.14}$$

Sum over $n \geqq N \geqq n_0$ to obtain

$$(1-\delta) b_{\varepsilon_N} \leqq \sum_{W^- \leqq T(1)} f_{\varepsilon_N}(W) \leqq (1+\delta) b_{\varepsilon_N}$$

and let $\delta \downarrow 0$ to obtain (4.13).

Now suppose (4.13) is true and let $\delta > 0$. For almost every ω there exists an n_0 such that, for $n \geqq n_0$,

$$\Big| \sum_{W^- \leqq T(1)} f_{\varepsilon_n}(W) - b_{\varepsilon_n} \Big| < \delta \left(\frac{1-\rho}{1+\rho} \right) b_{\varepsilon_n}.$$

Adding these inequalities for n and $n+1$ and using that

$$\frac{b_{\varepsilon_n} - b_{\varepsilon_{n+1}}}{b_{\varepsilon_n} + b_{\varepsilon_{n+1}}} = \frac{1-\rho}{1+\rho},$$

we obtain (4.14) for $n \geqq n_0$. Since the events in braces in (4.12) are independent, a consequence of the sets $\mathscr{E}_{\varepsilon_n} - \mathscr{E}_{\varepsilon_{n+1}}$, $n = 1, 2, 3, \ldots$, being disjoint, (4.12) follows by Borel-Cantelli. □

We cannot get a nice version of Theorem 4.6 without assuming that $\varepsilon \mapsto b_\varepsilon$ is continuous. However, the same arguments yield the following three results.

Corollary 4.7. *Assume the definitions and conditions of Theorem* 4.6 *except for the continuity of* $\varepsilon \mapsto b_\varepsilon$. *If there are* $\delta > 0$, *and* $\rho < 1$ *and a sequence* $\varepsilon_n \downarrow 0$ *such that* $b_{\varepsilon_{n+1}} \leqq \rho\, b_{\varepsilon_n}$ *for all n and* (4.12) *fails, then* (4.11) *is false.*

Corollary 4.8. *Using the definitions and conditions of Theorem* 4.6 *except for the continuity of* $\varepsilon \mapsto b_\varepsilon$ *and the definition of* ε_n; *suppose that for every* $\delta > 0$ *and* $\rho < 1$ *there is a sequence* $\varepsilon_n \downarrow 0$ *for which* $b_{\varepsilon_{n+1}} \geqq \rho\, b_{\varepsilon_n}$ *for all n and* (4.12) *holds. Then* (4.11) *is also true.*

A defect in Corollary 4.8 is that there is no chance of it being applicable if $\varepsilon \mapsto b_\varepsilon$ has infinitely many jumps that are too large, a situation that arises naturally when μ_0 has large atoms. We will not, for instance, be able to use Corollary 4.8 in Examples 7.7 and 7.9, although the next result, which can also be proved by the same arguments, can easily be applied in these examples.

Corollary 4.9. *Assume, except for the continuity of* $\varepsilon \mapsto b_\varepsilon$ *and the definition of* ε_n, *the conditions and definitions of Theorem* 4.6. *If for every* $\delta > 0$ *and* $\rho < 1$ *there is a sequence* $\varepsilon_n \downarrow 0$ *such that:*

$$b_{(\varepsilon_{n+1})+} \geqq \rho\, b_{\varepsilon_n -} \quad \textit{for all } n, \tag{4.15}$$

$$b_{\varepsilon -} \leqq \rho(b_{\varepsilon +} - b_{\varepsilon -}) \Rightarrow \varepsilon = \varepsilon_n \quad \textit{for some } n,$$

$$\sum_{n=2}^{\infty} P\left\{ \left| \frac{\sum_{W^- \leqq T(1)} [f_{(\varepsilon_{n-1})-}(W) - f_{\varepsilon_n -}(W)]}{b_{(\varepsilon_{n-1})-} - b_{\varepsilon_n -}} - 1 \right| > \delta \right\} < \infty, \tag{4.16}$$

$$\sum_{n=1}^{\infty} P\left\{\left|\frac{\sum_{W^- \leqq T(1)} [f_{\varepsilon_n}(W) - f_{\varepsilon_n -}(W)]}{b_{\varepsilon_n} - b_{\varepsilon_n -}} - 1\right| > \delta\right\} < \infty, \tag{4.17}$$

$$\sum_{n=1}^{\infty} P\left\{\left|\frac{\sum_{W^- \leqq T(1)} [f_{\varepsilon_n +}(W) - f_{\varepsilon_n}(W)]}{b_{\varepsilon_n +} - b_{\varepsilon_n}} - 1\right| > \delta\right\} < \infty; \tag{4.18}$$

then (4.11) *is true; where, in* (4.17) *and* (4.18), $c/0 = \infty$ *if* $c > 0$ *and* $0/0 = 0$.

Examples 7.9 and 7.10 will further illustrate the difficulty of obtaining necessary and sufficient conditions for a.s. convergence.

5. The Transient Case

Most of the conditions which we have imposed on the process X are needed to make the problem non-trivial. For example, if x is not regular for $\{x\}$, the set $Z = \{s: X(s) = x\}$ is discrete. If $\mu_0(0, \infty) < \infty$ and $\delta > 0$, which is the case if and only if (2.2) fails, then Z is a union of intervals whose lengths are independent, identically distributed exponential random variables. The assumption that $X(0) = x$ is natural because the local time at x does not grow until the first hitting time of $\{x\}$. However, in stating theorems and proving results we have also assumed that $\{x\}$ is recurrent; that is, that Z is unbounded with probability one. This assumption is unnecessary for the validity of our constructions. In this section we examine the changes needed when $\{x\}$ is transient; that is, when Z is bounded with probability one.

The local time A still exists, but as a function of t it will be bounded. The right continuous inverse is still a subordinator T provided it is understood that T will jump to $+\infty$ at a finite (random) time. Accordingly, μ_0 will have a finite atom at $+\infty$. The formula (2.5) is still correct provided, of course, that $1 - e^{-\theta\infty} = 1$ for $\theta > 0$.

Excursions of infinite duration now play a role, but we still require our functions f to be finite-valued in which case the corresponding μ's have no atom at $+\infty$. The function f_0, which assigns to each excursion its duration, is an exception; μ_0 does have an atom at $+\infty$.

Itô's measure ν does not in general have as natural an interpretation as it does under the assumption that $\{x\}$ is recurrent. Recall that, under this assumption,

$$\#\{(\tau_W, V(W)) \in D: V(W) \text{ defined by (2.6)}\}$$

is Poisson with mean $(\lambda \times \nu)(D)$. In general this cannot be the case since there are no excursions corresponding to τ's larger than $A(\infty)$. However, we can think of there being such excursions which, since they will correspond to "infinite t", will not affect X at all. Here is an indication of how to make this idea precise. For C a measurable subset of $\mathscr{V}$, let $\nu(C) = \lim_{\sigma \downarrow 0} E(\rho)/\sigma$ where ρ is the number of excursions W for the process X which actually occur and satisfy

$$\tau_W \leqq \sigma, \; V(W) \in C.$$

Use this measure ν on excursion space to define spurious excursions W for which $\tau_W > A(\infty)$, and proceed initially to include both actual and spurious excursions in the analysis.

Now μ corresponding to a function f can still be defined via (2.8). In Lemma 2.1 the definition of G_m needs to be altered to the following:

$$G_m(\tau) = \sum_{\tau_W \leqq \tau} \frac{f_m(W)}{b_m},$$

where the spurious excursions just introduced are included. *However*, the definition (2.10) of F_m requires *no* alterations, and it is F_m which is involved in the major theorems of Sects. 3 and 4.

We now claim that Lemma 2.2 and all the results in Sects. 3 and 4 are valid without the assumption that $\{x\}$ is recurrent. The proofs are valid with only minor changes except for the proof of Lemma 3.4 and the argument that (A) $\Rightarrow$(D) in the cases of a.s. convergence and convergence-in-probability in the proof of Lemma 2.2. For (A)$\Rightarrow$(D) in the transient case two steps are needed – the first a proof that (A) implies

$$G_m(\tau)\,\mathbf{1}_{\{\omega:\ A(t) \geqq \tau \text{ for some } t\}} \to \tau\,\mathbf{1}_{\{\omega:\ A(t) \geqq \tau \text{ for some } t\}}$$

and the second a proof, using the fact that G_m is a subordinator for which $A(\infty)$ is a stopping time, that $G_m(\tau) \to \tau$. In (3.18) the expression $T((k+1)/p) - T(k/p)$ may, in the transient case, be the meaningless expression $\infty - \infty$ in which case $\exp(-\theta T(k/p)) = 0$; so, the argument leading from (3.17) to (3.19) remains valid. The equality between (3.19) and (3.20) fails in the transient case: a correction is needed to account for the possibility that $T(j/p)$ may equal $+\infty$ if $j \neq 0$. That correction is just what is needed in (3.22); for, since $A(\infty)$ may be less than 1, (3.22), as it stands, is not valid in the transient case.

In the transient case $\varphi(0+) > 0$, so, for an appropriate version of the local time, $\varphi(0+) = 1$. For this version we can let $\theta \downarrow 0$ in Lemma 3.4 to obtain

$$\int_{(0,\infty)} r\mu(dr) = E\Big(\sum_{W^- < \infty} f(W)\Big). \tag{5.1}$$

We have introduced $W^- < \infty$ in (5.1) as a reminder that only actual excursions (as opposed to the spurious ones introduced earlier in this section) are included. Since ν assigns finite, positive measure (in the transient case) to the set of excursions having infinite duration, there is no loss of generality in assuming that set to have ν-measure one.

Lemma 5.1. *Suppose that*

$$\nu\{V \in \mathscr{V} : V^+ = \infty\} = 1.$$

Let $\mathscr{E}$ be a measurable subset of $\mathscr{V}$ such that $V \in \mathscr{E} \Rightarrow V^+ < \infty$. Then

$$\nu(\mathscr{E}) = \frac{p(\mathscr{E})}{1 - p(\mathscr{E})}, \tag{5.2}$$

where $p(\mathscr{E})$ equals the probability that at least one excursion belonging to $\mathscr{E}$ occurs during finite time in the t-scale.

Proof. Let $\mathscr{D}=\{V\in\mathscr{V}: V^+=\infty\}$. Then $p(\mathscr{E})$ equals the probability that the first (in the τ-scale, say) excursion belonging to $\mathscr{D}\cup\mathscr{E}$ belongs to $\mathscr{E}$. So, since $\mathscr{D}\cap\mathscr{E}=\varnothing$,

$$p(\mathscr{E})=\frac{\nu(\mathscr{E})}{\nu(\mathscr{D}\cup\mathscr{E})}=\frac{\nu(\mathscr{E})}{1+\nu(\mathscr{E})},$$

from which (5.2) follows. □

6. Counting Constructions

The a.s. portion of the following result is due to Maisonneuve ([1974, Theoreme X. 4], [1980]).

Corollary 6.1. *Let ν be the measure on the space $\mathscr{V}$ of excursions described in Sect. 2. Let $(\mathscr{E}_\varepsilon\subset\mathscr{V}: \varepsilon>0)$ have the property that $\infty>\nu(\mathscr{E}_\varepsilon)\to\infty$ as $\varepsilon\downarrow 0$ and let*

$$N_\varepsilon(t)=\#\{W: V(W)\in\mathscr{E}_\varepsilon, W^-\leqq t\}. \tag{6.1}$$

Then, as $\varepsilon\downarrow 0$,

$$\frac{N_\varepsilon(t)}{\nu(\mathscr{E}_\varepsilon)}\to A(t) \tag{6.2}$$

in L^2 for each t. If, in addition, $\mathscr{E}_\varepsilon\uparrow$ as $\varepsilon\downarrow 0$, then (6.2) *also holds almost surely.*

Proof. Let f_ε denote the indicator function of $\mathscr{E}_\varepsilon$. Apply Corollary 3.7 to obtain L^2-convergence. Apply Theorem 4.2 to obtain a.s. convergence. □

The idea of obtaining corollaries such as Corollary 6.1 by examining Itô's [1970] Poisson point process of excursions also appears in [Greenwood and Pitman, 1980].

Even though the next corollary is not "a counting construction result" we include it here because it has an easy proof based on Corollary 6.1.

Corollary 6.2. *Let f be a measurable function from $\mathscr{V}$ to $\mathbb{R}^+$ and let μ be defined as in* (2.8), *with $\mu\{\infty\}>0$ being permitted. Suppose that $\int(r\wedge 1)\,\mu(dr)<\infty$ and $\mu(0,1]=\infty$. Then, as $\varepsilon\downarrow 0$,*

$$\frac{\sum_{W^-\leqq t}\varepsilon\wedge f(W)}{\int_0^\varepsilon \mu(r,\infty]\,dr}\to A(t) \text{ in } L^2 \text{ and a.s.} \tag{6.3}$$

for each t.

Proof. Let $\mathscr{E}_\varepsilon=\{V\in\mathscr{V}: f(V)>\varepsilon\}$. Our hypotheses on μ imply that we may use Corollary 6.1 to deduce that, as $r\downarrow 0$,

$$\frac{N_\varepsilon(t)}{\mu(\varepsilon,\infty]}\to A(t) \text{ in } L^2 \text{ and a.s.} \tag{6.4}$$

The finiteness of $\int(r\wedge 1)\,\mu(dr)$ implies that the denominator in (6.3) is finite. Thus, we integrate the denominator and numerator in (6.4) to obtain

$$\frac{\int_0^{\varepsilon} N_r(t)\,dr}{\int_0^{\varepsilon} \mu(r,\infty]\,dr} \to A(t) \text{ in } L^2 \text{ and a.s.} \tag{6.5}$$

as $\varepsilon\downarrow 0$. Integrate by parts in the numerator at (6.5) to obtain (6.3). □

The next two examples are applications of Corollary 6.1. In each of them, $\mathscr{E}_\varepsilon$ of Corollary 6.1 is the set of excursions which hit a certain set of real numbers.

Example 6.3. A process X is called *strictly stable of index* α if X has stationary, independent increments, and if, for any $a>0$, the process $t\mapsto a^{-1}X(a^\alpha t)$ is stochastically equivalent to X, a property that is called the *scaling property*. Necessarily, $0<\alpha\leqq 2$. We consider a one-dimensional strictly stable process X of index $\alpha>1$ starting at 0. It is known that (2.1) and (2.2) hold. For $\alpha\leqq 1$, (2.1) is false.

Let $\mathscr{E}_\varepsilon$ denote the set of excursions that hit the set $(-\infty,-\varepsilon]\cup[\varepsilon,\infty)$. If $\nu(\mathscr{E}_\varepsilon)=+\infty$ the number of excursions belonging to $\mathscr{E}_\varepsilon$ and occuring by the (random) time $T(1)$ would be Poisson distributed with mean $+\infty$; that is, it would equal $+\infty$ a.s. in contradiction of the fact that processes with stationary, independent increments have discontinuities of only the first kind. Hence, $\nu(\mathscr{E}_\varepsilon)<\infty$ for each $\varepsilon>0$ and, therefore, Corollary 6.1 is applicable: $N_\varepsilon(t)/\nu(\mathscr{E}_\varepsilon)\to A(t)$ a.s. and in L^2. We will use Lemma 3.4 to show $\nu(\mathscr{E}_\varepsilon)=c\varepsilon^{-(\alpha-1)}$.

For each $W\in\mathscr{W}$ and each $r\in(0,\infty)$ let W_r denote the excursion defined by $(W_r)^-=r^\alpha W^-$ and $W_r(t)=rW(r^{-\alpha}t)$. By the scaling property, the family $\{W\}$ is stochastically equivalent to $\{W_r\}$. By Lemma 3.4,

$$\begin{aligned}\nu(\mathscr{E}_\varepsilon)&=\varphi(\theta)\,E\big(\textstyle\sum e^{-\theta W^-}\,\mathbf{1}_{\mathscr{E}_\varepsilon}(W)\big)\\&=\varphi(\theta)\,E\big(\textstyle\sum e^{-\theta(W_\varepsilon)^-}\,\mathbf{1}_{\mathscr{E}_\varepsilon}(W_\varepsilon)\big)\\&=\varphi(\theta)\,E\big(\textstyle\sum e^{-\theta\varepsilon^\alpha W^-}\,\mathbf{1}_{\mathscr{E}_1}(W)\big).\end{aligned}$$

Set $\theta=\varepsilon^{-\alpha}$ and use $\varphi(\theta)=\theta^{(\alpha-1)/\alpha}$ [Stone, 1963] to obtain $\nu(\mathscr{E}_\varepsilon)=c\varepsilon^{-(\alpha-1)}$. □

Example 6.4. Let X be an asymmetric Cauchy process specified by

$$\log E(e^{iuX(t)})=-t|u|\,(1+ih\,\mathrm{sgn}\,(u)\log|u|), \tag{6.6}$$

where $h\in[-\pi/2,0)\cup(0,\pi/2]$ is a constant. (We omit $h=0$ for it corresponds to the (symmetric) Cauchy process which does not satisfy (2.1).) Since X has infinitely many jumps in every time intervals, (2.2) holds. Since the real part of (6.6) is integrable for u in a neighborhood of 0, X is transient. Port and Stone [1969] showed that (2.1) holds. Let $\mathscr{E}_\varepsilon$ denote the set of excursions from 0 that hit $(-\infty,-\varepsilon]$ and $\mathscr{E}_\varepsilon^*=\{V\in\mathscr{E}_\varepsilon\colon V^+<\infty\}$. Pruitt and Taylor [1982] analyze the local structure of the process X and their Lemma 4.2 gives, in the notation of our Lemma 5.1, that, as $\varepsilon\downarrow 0$,

$$1-p(\mathscr{E}_\varepsilon^*)\sim c/\log(1/\varepsilon)$$

for some positive, finite constant c. (In [Pruitt and Taylor, 1982], the value of c is given in terms of h, but this value is unimportant and is changed by the normalization of Lemma 5.1.) By Lemma 5.1, $v(\mathscr{E}_\varepsilon^*)\sim c^{-1}\log(1/\varepsilon)$. Since $\mathscr{E}_\varepsilon$ has at most one more member than $\mathscr{E}_\varepsilon^*$, $v(\mathscr{E}_\varepsilon)\sim c^{-1}\log(1/\varepsilon)$. Now Corollary 6.1 gives:

$$\frac{N_\varepsilon(t)}{\log(1/\varepsilon)}\to A(t) \text{ a.s. and in } L^2 \text{ as } \varepsilon\downarrow 0 \tag{6.7}$$

where $N_\varepsilon(t)$ is the number of excursions that begin by time t and hit $(-\infty, -\varepsilon]$. □

The next result is in some ways more general and in other ways less general than a result of Getoor [1976]. For $\theta>0$ and x the point at which we want to study the local time A let

$$G(\theta,y)=E^x(e^{-\theta\inf\{t:X(t)=y\}}),$$
$$H(\theta,y)=E^y(e^{-\theta\inf\{t:X(t)=x\}}).$$

Corollary 6.5. *Suppose $(x_m: m=1,2,\dots)$ is a sequence such that each $x_m\neq x$ and, for some $\theta>0$, $G_m(\theta,x_m)\to 1$ as $m\to\infty$. Then, as $m\to\infty$,*

$$\frac{1-G(\theta,x_m)H(\theta,x_m)}{\varphi(\theta)}N(x_m,x,t)\to A(t) \tag{6.8}$$

in L^2 for each t, where $N(x_m,x,t)$ is the number of passages from x_m to x by time t.

Remark. In one sense, $\varphi(\theta)$ is superfluous in the corollary since A is only determined up to a multiplicative constant. We have inserted it so that the same version of local time is obtained for different values of θ. If $G(\theta,x_m)\to 1$ for some $\theta>0$, it does so for all $\theta>0$. In the transient case $0+$ may be inserted for θ.

Proof. Since $G(\theta,x_m)\to 1$ and (2.1) holds, $H(\theta,x_m)\to 1$ [Blumenthal and Getoor, 1968, V Exercise (3.36)]. Let X_θ denote the process obtained by terminating X using an independent random time with an exponential distribution having mean $1/\theta$. Clearly, we only need prove the corollary for X_θ.

Let $\mathscr{E}_m$ denote the set of excursions V which hit $\{x_m\}$ and satisfy $V^+<\infty$. Since X_θ is a strong Markov process, $G(\theta,x_m)H(\theta,x_m)$ equals the probability that X_θ has at least one excursion belonging to $\mathscr{E}_m$. Thus, from Lemma 5.1, we obtain

$$v_\theta(\mathscr{E}_m)\sim c_\theta[1-G(\theta,x_m)H(\theta,x_m)]^{-1}$$

as $m\to\infty$, where v_θ is the measure on $\mathscr{V}$ induced by X_θ and c_θ is a constant that frees us from the normalizing hypothesis in Lemma 5.1; accordingly,

$$c_\theta=v_\theta\{V\in\mathscr{V}: V^+=\infty\}.$$

The L^2 portion of Corollary 6.1 is now applicable despite the fact that the numerator there may differ by 1 from the number of passages by X_θ from x_m to x by time t, so it only remains to show that we get the same version of the local time for different θ's by choosing $c_\theta = \varphi(\theta)$.

One way of fixing one version of the local time of X independent of which value of θ is used in the preceding argument is to regard ν_θ as being induced by ν (rather than by X_θ in which case there is an arbitrary constant for each θ): ν_θ is obtained by shifting some of the measure ν to excursions having infinite duration. Let $\mathscr{D} = \{V \in \mathscr{V}: V^+ = \infty\}$. The probability that for X_θ there exists an excursion W with $V(W) \in \mathscr{D}$ and $\tau_W \in (0, 1]$ (see (2.6) and (2.7)) is the probability that the local time for X_θ never exceeds 1 which is the probability that the local time A of X does not exceed 1 when it is evaluated at the independent exponentially distributed random variable S_θ having mean $1/\theta$. Thus, it equals

$$P\{T(1) > S_\theta\} = \int_{(0,\infty]} [1 - e^{-\theta t}] P\{T(1) \in dt\} = 1 - e^{-\varphi(\theta)}.$$

Since this probability is also the probability that a Poisson random variable with mean $\nu_\theta(\mathscr{D})$ is non-zero (The spurious excursions introduced in Sect. 5 are playing a role), we obtain $\nu_\theta(\mathscr{D}) = \varphi(\theta)$ as desired. □

In the following example we give a result for diffusions that Itô and McKean [1965, 6.5a] obtained by using a time substitution in Brownian motion. Even for Brownian motion this result has attracted recent attention ([Getoor, 1976, p. 2], [Chung and Durrett, 1976], [Williams, 1977] and [Maisonneuve, 1981]).

Example 6.6. Let ρ, satisfying $\rho(0) = 0$, be a scale of a non-singular diffusion X on an interval containing $[0, h]$ for some $h > 0$. As is usual we allow a variety of behavior at the endpoints of the interval except that we do require (2.1) and (2.2) for $x = 0$ even if 0 is the left endpoint.

Suppose $x_m \downarrow x = 0$ as $m \to \infty$. We note that Corollary 6.5 and, because of the continuity of X, the a.s. portion of Corollary 6.1 are applicable so that (6.8) holds almost surely. We now analyze $[1 - G(\theta, x_m) H(\theta, x_m)]/\varphi(\theta)$.

Let G denote the generator of X and let g_θ and h_θ denote, respectively, increasing and decreasing solutions of $Gf = \theta f$ chosen so that the Wronskian $g'_\theta h_\theta - g_\theta h'_\theta = 1$ where (′) denotes right derivative with respect to ρ. Itô and McKean [1965, 4.10] give:

$$G(\theta, y) = g_\theta(0)/g_\theta(y),$$
$$H(\theta, y) = h_\theta(y)/h_\theta(0).$$

Hence,

$$\begin{aligned}\frac{1 - G(\theta, x_m) H(\theta, x_m)}{\varphi(\theta)} &= \frac{g_\theta(x_m) h_\theta(0) - g_\theta(0) h_\theta(x_m)}{\varphi(\theta) g_\theta(x_m) h_\theta(0)} \\ &= \frac{[g(x_m) - g(0)] h_\theta(0) - g_\theta(0) [h_\theta(x_m) - h_\theta(0)]}{\rho(x_m) g_\theta(x_m) h_\theta(0)} \cdot \frac{\rho(x_m)}{\varphi(\theta)} \\ &\sim \frac{\rho(x_m)}{g_\theta(0) h_\theta(0) \varphi(\theta)}.\end{aligned}$$

So $\varphi(\theta)=c[g_\theta(0)\,h_\theta(0)]^{-1}$ in agreement with [Itô and McKean, 1965, 6.2] and

$$\rho(x_m)\,N(x_m, x, t)\to A(t)$$

a.s. and in L^2. □

For a discontinuous X, the a.s. part of Corollary 6.1 may not be applicable in the context of Corollary 6.5. The following sufficient condition for a.s. convergence follows easily from Corollary 4.5.

Corollary 6.7. *Under the hypotheses of Corollary* 6.5, *the convergence at* (6.7) *takes place a.s. if*

$$\sum_{m=1}^{\infty} \exp(-\gamma[1-G(\theta, x_m)\,H(\theta, x_m)]^{1/2})<\infty.$$

for some $\theta>0$ *and every* $\gamma>0$.

The sufficient condition for a.s. convergence that one can obtain from Corollary 6.7 in case X is a Lévy process is weaker than Getoor's [1976] in our context (which is more restrictive than is his). For comparison of our Corollaries 6.5 and 6.7 with his paper, his (1.9), (1.10), (2.11), and (2.22) are relevant.

7. Intrinsic Constructions

In this section we consider functions that depend on the excursion $V\in\mathscr{V}$ only through its duration V^+. In other words we are interested in obtaining $A(t)$ as a limit of approximations that depend only on the strong Markov set being studied rather than on some other aspect of a Markov process whose zero set, say, happens to be this strong Markov set.

We shall restrict our attention to possible limit theorems involving three one-parameter families of random functions:

$$N_\varepsilon(t)=\#\{W\colon W^-\leqq t,\ W^+-W^->\varepsilon\}, \tag{7.1}$$

$$R_\varepsilon(t)=\sum_{W^-\leqq t} \varepsilon\wedge(W^+-W^-), \tag{7.2}$$

$$S_\varepsilon(t)=\sum_{\substack{W^-\leqq t\\ W^+-W^-\leqq\varepsilon}} (W^+-W^-) \tag{7.3}$$

– each defined for $\varepsilon>0$.

Itô and McKean [1965, 6.3] proved the following result by a method to which some of our methods are similar.

Corollary 7.1. *For each* t, *as* $\varepsilon\downarrow 0$,

$$\frac{N_\varepsilon(t)}{\mu_0(\varepsilon,\infty]}\to A(t) \text{ in } L^2 \text{ and a.s.,}$$

where N_ε *is defined at* (7.1).

Proof. This corollary is a special case of Corollary 6.1. □

Even though S_ε is studied more often than is R_ε we consider the simpler R_ε first, and obtain a result of Kingman [1973]. As a reminder to the reader that for some theorems, but not all, $W^+ \leqq t$ can replace $W^- \leqq t$, we depart from our usual format and include both possibilities in our statement.

Corollary 7.2. *For each t, as $\varepsilon \downarrow 0$,*

$$\frac{R_\varepsilon(t)}{\int_0^\varepsilon \mu_0(s, \infty]\, ds} \to A(t) \text{ in } L^2 \text{ and a.s.,} \tag{7.4}$$

where R_ε is defined at (7.2) *and*

$$\frac{\sum_{W^+ \leqq t} [\varepsilon \wedge (W^+ - W^-)]}{\int_0^\varepsilon \mu_0(s, \infty]\, ds} \to A(t) \text{ in } L^2 \text{ and a.s.} \tag{7.5}$$

Proof. From Corollary 6.2, (7.4) follows immediately. The difference between the numerators on the left sides of (7.4) and (7.5) is no larger than ε. The equivalence of (7.4) and (7.5) follows from

$$\frac{\varepsilon}{\int_0^\varepsilon \mu_0(s, \infty]\, ds} \leqq \frac{\varepsilon}{\varepsilon \mu_0(\varepsilon, \infty]} \to 0. \quad \square \tag{7.6}$$

One way of constructing a standard Markov process X whose zero set is, except for a countable set, a particular strong Markov set Z is to let

$$X(t) = \inf\{s > t : s \in Z\} - t.$$

The numerator on the left side of (7.4) differs by no more ε from the occupation time of $[0, \varepsilon]$ by X before time t, provided that the Lebesgue measure of Z equals 0. Noting (7.6), we can conclude, in case the drift $\kappa = 0$ (see (2.5)):

$$\frac{\int_0^t \mathbf{1}_{\{u : X(u) \leqq \varepsilon\}}}{\int_0^\varepsilon \mu_0(s, \infty]\, ds} \to A(t)$$

in L^2 and a.s. as $\varepsilon \downarrow 0$.

We turn to the study of S_ε, defined at (7.3). What is interesting is the variety of possibilities, depending on μ_0, for the behavior of $S_\varepsilon(t)/b_\varepsilon$ as $\varepsilon \downarrow 0$. Accordingly, we look at some examples in which $f_\varepsilon(V) = V^+$ if $V^+ \leqq \varepsilon$ and $= 0$ if $V^+ > \varepsilon$ and μ_ε denotes the corresponding measure defined via (2.8). Clearly, $\mu_\varepsilon(B) = \mu_0(B \cap [0, \varepsilon])$. In the following example convergence in probability fails.

Example 7.3. Let $\mu_0(s, \infty] = 0 \vee \log(1/s)$. For $\varepsilon < 1$,

$$\frac{1}{b} \int_{(0,b]} r\mu_\varepsilon(dr) = \frac{1}{b} \int_0^{b \wedge \varepsilon} r(1/r)\,dr = \frac{b \wedge \varepsilon}{b}$$

so, a_ε, defined via (3.9), equals ε. We try $t = \frac{1}{2}$ in (3.10):

$$\mu_\varepsilon(\tfrac{1}{2}\varepsilon, \infty) = \log(2/\varepsilon) - \log(1/\varepsilon) = \log 2 \not\to 0.$$

By Proposition 3.3 we conclude that, as $\varepsilon \downarrow 0$, $S_\varepsilon(t)/b_\varepsilon$ does not converge to $A(t)$ in probability whatever be the numbers b_ε. If X is an asymmetric Cauchy process, then [Hawkes, 1970, Lemma 2.1]

$$-\log E(\exp(-\theta T(1))) \sim c \log \lambda$$

as $\lambda \to \infty$. By a Tauberian theorem

$$\mu_0(s, \infty] \sim \log(1/s)$$

as $s \downarrow 0$. Accordingly, the argument just given shows that $S_\varepsilon(t)/b_\varepsilon$ does not converge to $A(t)$ in probability for any choice of b_ε. □

The next example is at the opposite extreme from the preceding one and contains as a special case one of Lévy's results for Brownian motion.

Example 7.4. For $0 < \beta < 1$, let $\mu_0(s, \infty) = s^{-\beta}$, $s > 0$. By Theorem 3.5, as $\varepsilon \downarrow 0$, $S_\varepsilon(t)/\varepsilon^{1-\beta} \to A(t)$ in L^2. By Theorem 4.1 and Lemma 2.3, as $\varepsilon \downarrow 0$, $S_\varepsilon(t)/\varepsilon^{1-\beta} \to A(t)$ a.s. If X is a strictly stable process. (See Example 6.3) with index $\alpha > 1$, then [Fristedt, 1974, Example 5.11] μ_0 for its zero set is given by $\mu_0(s, \infty) = s^{-\beta}$ with $\beta = (\alpha - 1)/\alpha$. Setting $\alpha = 2$ (Brownian motion), we obtain (1.4). □

In the next example L^2 convergence holds even though a.s. convergence fails.

Example 7.5. Let μ_0 be supported by

$$\{2^{-k}/\log k : k = 2, 3, 4, \ldots\}$$

with

$$\mu_0\{2^{-k}/\log k\} = \log k.$$

We have

$$\int_{(0,\varepsilon]} r\mu_0(dr) = 2^{-j(\varepsilon)+1}$$

and, as $\varepsilon \downarrow 0$,

$$\int_{(0,\varepsilon]} r^2 \mu_0(dr) \sim 3^{-1} 4^{-j(\varepsilon)+1}/\log(j(\varepsilon)),$$

where the integer $j(\varepsilon)$ satisfies

$$2^{-j(\varepsilon)} \leqq \varepsilon < 2^{-j(\varepsilon)+1}.$$

By Theorem 3.5, $2^{j(\varepsilon)-1}S_\varepsilon(t)\to A(t)$ in L^2. In order to prove that a.s. convergence fails we set $\varepsilon_n=2^{-n}/\log n$, $b(\varepsilon_n)=2^{-n+1}$, $\rho=\frac{1}{2}$, and $\delta=1$ in Corollary 4.7 and examine

$$P\left\{\left|\frac{\varepsilon_n \#\{W\colon W^-\leqq T(1),\, W^+-W^-=\varepsilon_n\}}{2^{-n}}-1\right|>1\right\}$$

$$=P\{\#\{W\colon W^-\leqq T(1),\, W^+-W^-=\varepsilon_n\}>2\log n\}$$

$$=\sum_{i>2\log n}\frac{(\log n)^i e^{-\log n}}{i!}\geqq\frac{(\log n)^{2\log n}}{n\Gamma(2(1+\log n))}$$

(Γ = gamma function) which by Striling's Formula, is asymptotic to $cn^{-(\log 4-1)}(\log n)^{-5/2}$, the general term of a divergent series. So, by Corollary 4.7, a.s. convergence does fail. □

In the next example convergence in probability holds with the normalizing function

$$\int_{(0,\varepsilon]} r\mu_0(dr)$$

even though both L^2 convergence and a.s. convergence fail.

Example 7.6. Let μ_0 be supported by

$$\{k^{-1}2^{-k}\colon k=2,3,\ldots\}\cup\{k^{1/2}2^{-k}\colon k=2,3,\ldots\}$$

with

$$\mu_0\{k^{-1}2^{-k}\}=k, \qquad \mu_0\{k^{1/2}2^{-k}\}=k^{-1}.$$

As $\varepsilon\downarrow 0$

$$\int_{(0,\varepsilon]} r\mu_0(dr)\sim 2^{1-j(\varepsilon)},$$

where the integer $j(\varepsilon)$ satisfies

$$j(\varepsilon)^{-1}2^{-j(\varepsilon)}\leqq\varepsilon<[j(\varepsilon)-1]^{-1}2^{1-j(\varepsilon)};$$

and

$$\int_{(0,\varepsilon]} r^2\mu_0(dr)\sim\begin{cases} j(\varepsilon)^{-1}4^{1-j(\varepsilon)}/3 & \text{if } \varepsilon<j(\varepsilon)^{1/2}2^{-j(\varepsilon)} \\ 4^{1-j(\varepsilon)}/3 & \text{if } \varepsilon\geqq j(\varepsilon)^{1/2}2^{-j(\varepsilon)}. \end{cases}$$

For any fixed $\beta>0$, $\beta 2^{1-j(\varepsilon)}>j(\varepsilon)^{-1}2^{-j(\varepsilon)}$ for ε sufficiently small. So, as $\varepsilon\downarrow 0$,

$$2^{j(\varepsilon)-1}\int_{(\beta 2^{1-j(\varepsilon)},\varepsilon]} r\mu_0(dr)\leqq 2^{-1}j(\varepsilon)^{-1/2}\to 0.$$

Corollary 3.2 and Theorem 3.5 now apply:

$$2^{j(\varepsilon)-1}S_\varepsilon(t)\to A(t) \tag{7.7}$$

in probability but *not* in L^2.

To consider (7.7) in the almost-sure mode we prepare to use Corollary 4.7 by letting $\delta=1$, $\rho>\frac{1}{2}$, and

$$\varepsilon_n = n^{1/2}\, 2^{-n}.$$

The series at (4.12) can be written as

$$\sum_n P\{|n^{-1} U_n + n^{1/2} V_n - 1| > 1\}, \tag{7.8}$$

where U_n and V_n are independent Poisson random variables with means n and n^{-1}, respectively. The series at (7.8) dominates

$$\sum_n P\{|n^{-1} U_n - 1| < 1\}\, P\{V_n > 0\}$$

which dominates

$$c \sum_{n=1}^{\infty} P\{V_n > 0\} = c \sum_{n=1}^{\infty} [1 - \exp(n^{-1})]$$

which diverges by comparison with $\sum n^{-1}$. Corollary 4.7 thus implies that (7.7) does *not* hold a.s. □

In the following two examples a.s. convergence holds but L^2 convergence fails. In the first

$$\int_{(0,\varepsilon]} r\mu_0(dr) \tag{7.9}$$

is a correct normalizing function. In the second (7.9) is finite but not a correct normalizing function.

Example 7.7. Let μ_0 be supported by

$$\{k^{-2}\, 2^{-k}: k=4,5,\ldots\} \cup \{k 2^{-k}: k=4,5,\ldots\}$$

with $\mu_0\{k^{-2} 2^{-k}\} = k^2$ and $\mu_0\{k 2^{-k}\} = k^{-2}$ and define the integer $j(\varepsilon)$ via

$$j(\varepsilon)^{-2}\, 2^{-j(\varepsilon)} \leqq \varepsilon < [j(\varepsilon) - 1]^{-2}\, 2^{1-j(\varepsilon)}.$$

Calculations similar to those in Example 7.6 show that

$$2^{j(\varepsilon)-1} S_\varepsilon(t) \to A(t) \tag{7.10}$$

in probability but *not* in L^2, and that

$$2^{1-j(\varepsilon)} \sim \int_{(0,\varepsilon]} r\mu_0(dr).$$

To consider (7.10) in the almost-sure mode we prepare to use Corollary 4.9 by fixing $\delta>0$ and setting

$$\varepsilon_n = n^{-2}\, 2^{-n}. \tag{7.11}$$

If we define $b_\varepsilon = 2^{1-j(\varepsilon)}$, then b_ε is constant in $(\varepsilon_{n+1}, \varepsilon_n)$ so that

$$b_{(\varepsilon_{n+1})+} = b_{\varepsilon_n -}.$$

Even though the sequence in (7.11) does not depend on ρ we have still satisfied (4.15). To successfully apply Corollary 4.9 it remains to check (4.16) and (4.17), as (4.18) is trivially satisfied. The left side of (4.16) equals

$$\sum_n P\{|n^{-2}U_n+nV_n-1|>\delta\} \tag{7.12}$$

where U_n and V_n are independent Poisson random variables with means n^2 and n^{-2}, respectively. The sum at (7.12) is bounded by

$$\sum_n P\{|U_n-n^2|>\delta n^2\}+\sum_n P\{V_n>0\},$$

which, by Chebyshev, is bounded by

$$\sum_n \frac{n^2}{(\delta n^2)^2}+\sum_n [1-\exp(n^{-2})]<\infty.$$

We omit the verification of (4.17); it is similar to the preceding argument but easier since V_n is not involved. By Corollary 4.9, (7.10) holds in the almost-sure mode. □

Example 7.8. Let μ_0 be absolutely continuous with support $\bigcup_{k=3}^{\infty}(I_k\cup J_k)$, where $\{I_k, J_k\colon k=3,4,\ldots\}$ is a disjoint family of closed intervals such that the right endpoints of I_k and J_k are $k^{-2}2^{-k}$ and $k^3 2^{-k}$, respectively; and let $\mu_0(I_k)=k^2$ and $\mu_0(J_k)=k^{-2}$ with μ_0 uniform on each I_k and on each J_k. The intervals I_k and J_k should be thought of as short. Indeed, many of the following assertions are only valid if, as $k\to\infty$, the lengths of I_k and J_k go to 0 sufficiently fast. However we will not explicitly say this each time.

From Proposition 3.3 it is straightforward to show that if any normalizing b_ε will yield convergence of $b_\varepsilon^{-1}S_\varepsilon(t)$ as $\varepsilon\downarrow 0$, then the continuous function $\varepsilon\mapsto b_\varepsilon$ that satisfies

$$b_\varepsilon=2^{1-k} \quad \text{if } \varepsilon=k^{-2}2^{-k},$$

is linear on each I_k, and is constant on intervals in the complement of $\bigcup I_k$ is such a normalizing function.

For $\varepsilon=k^3 2^{-k}$,

$$\int_{(0,\varepsilon]} r\mu_0(dr)\sim k2^{1-k}$$

and, so, the "natural" normalizing function given at (7.9) will not give convergence; rather

$$\liminf b_\varepsilon\Big/\int_{(0,\varepsilon]} r\mu_0(dr)=0 \quad (\varepsilon\downarrow 0).$$

Checking condition (3.24) we see that we can *not* obtain convergence in the L^2 mode.

We fix $\delta>0$ and $\rho<1$ and let ε_n be defined as in Theorem 4.6. A necessary and sufficient condition for $\varepsilon_n\in I_k$ is $2^{-(k+1)}<\rho^n\leqq 2^{-k}$ or, equivalently,

$$(k+1)\left(\frac{\log 2}{\log(1/\rho)}\right)>n\geqq k\left(\frac{\log 2}{\log(1/\rho)}\right). \tag{7.13}$$

We want to apply Lemma 4.3 to the n^{th} term in (4.12).

We first suppose that $\{\varepsilon_n, \varepsilon_{n+1}\} \subset I_k$ so that

$$(k+1)\left(\frac{\log 2}{\log(1/\rho)}\right) - 1 > n \geqq k\left(\frac{\log 2}{\log(1/\rho)}\right). \tag{7.14}$$

We may take 2^{-k} for M in Lemma 4.3. The numbers b, c, and a of that lemma are, respectively, close to

$$\rho^n(1-\rho),\ k^{-2}2^{-k}\rho^n(1-\rho), \text{ and } k2^{k/2}\rho^{n/2}(1-\rho)^{1/2}.$$

From the second inequality in (7.14),

$$k2^{k/2}\rho^{n/2}(1-\rho)^{1/2}M \leqq k(1-\rho)^{1/2}2^{-k} \to 0.$$

Accordingly, the first of the bounds in Lemma 4.3 is valid, and, absorbing the approximations into the unspecified constant c in the exponent, we obtain

$$P\left\{\left|\frac{\sum_{W^- \leqq T(1)} [f_{\varepsilon_n}(W) - f_{\varepsilon_{n+1}}(W)]}{b_{\varepsilon_n} - b_{\varepsilon_{n+1}}} - 1\right| > \delta\right\}$$
$$\leqq 2\exp[-c\delta^2 k^2 2^k \rho^n(1-\rho)]$$

which, by the first inequality in (7.14), is less than

$$2\exp[-c\delta^2 k^2] < 2\exp[-c\delta^2 n^2],$$

the general term of a convergent series.

Next we suppose that $\varepsilon_n \in I_j$ and $\varepsilon_{n+1} \in I_k$ for some $k > j$. From (7.13) there is, since ρ is fixed, a fixed bound K on $k-j$. The expression we want to consider is

$$P\{|\sum_{W^- \leqq T(1)} [f_{\varepsilon_n}(W) - f_{\varepsilon_{n+1}}(W)] - [b_{\varepsilon_n} - b_{\varepsilon_{n+1}}]| > \delta\rho^n(1-\rho)\}$$

which can be bounded by the sum of at most K terms of the kind considered in the previous paragraph plus

$$\sum_{i=j+1}^{k} P(\#\{W\colon W^- \leqq T(1), W^+ - W^- \in J_i\} > 0)$$
$$= \sum_{i=j+1}^{k} [1 - \exp(i^{-2})] \leqq \sum_{i=j+1}^{k} i^{-2}.$$

As we sum over n we pick up no more than $c\sum i^{-2}$ which is finite.

By Theorem 4.6,

$$b_\varepsilon^{-1} S_\varepsilon(t) \to A(t) \text{ a.s.} \quad \square$$

The existence of atoms in Example 7.7 did not play a crucial role. We could have spread them out as we did in Example 7.8. Generally, if there are atoms and $b_\varepsilon^{-1} S_\varepsilon(t) \to A(t)$ either in probability or in L^2, then the atoms can be spread out and b_ε adjusted so that the convergence is not lost. If there are atoms, and convergence in some particular mode is not possible for any choice of $\varepsilon \mapsto b_\varepsilon$, then such convergence will still be impossible if the atoms are spread out slightly. In the following example there are atoms and there is a.s. convergence. However, if the atoms are spread out (see Example 7.10), then a.s. convergence is not possible.

Example 7.9. Let μ_0 be supported by

$$\{k^{-2} 2^{-2^k}: k=1,2,3\ldots\}$$

with

$$\mu_0\{k^{-2} 2^{-2^k}\}=k^2.$$

Let

$$b_\varepsilon = 2^{-2^k}, 2^{-2^k} \leqq \varepsilon < 2^{-2^{k-1}}.$$

In order to apply Corollary 4.9 we fix $\delta>0$ and set

$$\varepsilon_n = 2^{-2^n}.$$

Conditions (4.15) and (4.18) are trivially satisfied and conditions (4.16) and (4.17) are identical – namely, that

$$\sum_n P\{|n^{-2} U_n - 1| > \delta\} < \infty \tag{7.15}$$

where U_n is a Poisson random variable having mean n^2. Chebshev implies (7.15) and so

$$b_\varepsilon^{-1} S_\varepsilon(t) \to A(t) \text{ a.s.} \quad \square$$

Example 7.10. Let μ_0 be absolutely continuous with support $\bigcup_{k=1}^{\infty} I_k$, where $\{I_k,: k=1,2,\ldots\}$ is a disjoint family of closed intervals such that the right endpoint of I_k is

$$k^{-2} 2^{-2^k};$$

and let $\mu_0(I_k)=k^2$ with μ_0 uniform on each I_k. As in Example 7.8 some of the following steps depend on $|I_k| \to 0$ sufficiently fast.

By Proposition 3.3, we see that if $b_\varepsilon^{-1} S_\varepsilon(t) \to A(t)$ a.s. for some function $\varepsilon \mapsto b_\varepsilon$, then we will have convergence with:

$$b_\varepsilon = 2^{-2^k} \quad \text{if } \varepsilon = k^{-2} 2^{-2^k},$$

and $\varepsilon \mapsto b_\varepsilon$ continuous, linear on each I_k, and constant on intervals in the complement of $\bigcup I_k$. For this b_ε, $\delta=\frac{1}{2}$ and fixed $\rho<1$, let ε_n be defined as in Theorem 4.6. For each k consider the n such that $\varepsilon_{n+2} \notin I_k$, $\varepsilon_{n+1} \in I_k$ and, therefore, for k sufficiently large $\varepsilon_n \notin I_k$:

$$2^{k+1}\left(\frac{\log 2}{\log(1/\rho)}\right) - 1 > n \geqq 2^{k+1}\left(\frac{\log 2}{\log(1/\rho)}\right) - 2. \tag{7.16}$$

The n^{th} term in (4.12) is approximately

$$\begin{aligned} &P\{|k^{-2} 2^{-2^k} \rho^{-n} (1-\rho)^{-1} U_n - 1| > \tfrac{1}{2}\}, \\ &\geqq P\{U_n = 0\} \end{aligned} \tag{7.17}$$

where U_n is a Poisson random variable with mean

$$k^2 2^{2^k} \rho^n (1-\rho)$$

which, by the second inequality in (7.16), is dominated by

$$k^2 2^{2^k} \rho^{-2} 2^{-2^{k+1}} (1-\rho)$$

which approaches 0 as $k \to \infty$. Hence, the quantity at (7.17) approaches 1 as $n \to \infty$ through the sequence satisfying (7.16) for some k. Thus, (4.12) fails and $b_\varepsilon^{-1} S_\varepsilon(t)$ does not converge to $A(t)$ a.s.

In this example it is not essential to have a uniform distribution on each I_k: a similar argument will work provided the distribution of μ_0 is continuous on each I_k. □

Erickson [1981] obtains, in case μ_0 is non-atomic, a good sufficient condition on μ_0 for

$$\frac{S_\varepsilon(t)}{\int_{(0,\varepsilon]} r \mu_0(dr)} \to A(t) \text{ a.s.} \tag{7.18}$$

We can also obtain a sufficient condition by using Theorem 4.6 in conjunction with Lemma 4.3. Because our sufficient conditions can be satisfied for a μ_0 with atoms they are clearly not equivalent to those of Erickson. It would be interesting to obtain necessary and sufficient conditions for (7.18). The examples of this section show that a set of conditions which covers all relevant cases may be quite complicated.

8. A Theorem of Knight

It is well known that, for Brownian motion $Y(t)$, the result of reflecting in the maximum

$$X(t) = \sup\{Y(s): 0 \leqq s \leqq t\} - Y(t) \tag{8.1}$$

is a process $X(t)$ which behaves like $|Y(t)|$ not only away from the zeros of $X(t)$ but also on the zero set. For a symmetric stable process of index α $(0<\alpha<2)$, the local behavior at zero is different from that at other points. However, Knight [1971] shows that the local time at zero can be obtained as the limit of suitably normalized occupation times of $[0, \varepsilon]$. It is clear that this result fits into the framework of the present paper. We can generalize the result to include all strictly stable $Y(t)$. We exclude the monotone $Y(t)$ from the statement because the corresponding zero set for $X(t)$ is trivial.

Corollary 8.1. *Suppose Y is a strictly stable process in $\mathbb{R}^1$ with characteristic function*

$$E e^{i\theta Y(t)} = e^{-c[1 - i\beta \operatorname{sgn}(\theta)]|\theta|^\alpha}$$

where $c>0$, $0<\alpha \leqq 2$ and

$$|\beta| \begin{cases} <|\tan(\pi\alpha/2)| & \text{if } \alpha \leqq 1 \\ \leqq |\tan(\pi\alpha/2)| & \text{if } \alpha > 1. \end{cases}$$

Then the process $X(t)$ *defined by* (8.1) *has a local time* $A(t)$ *at* 0 *and, as* $\varepsilon \downarrow 0$,

$$\frac{\int_0^t \mathbf{1}_{[0,\varepsilon)}(X(s))\,ds}{\varepsilon^{\alpha/2-(1/\pi)\arctan\beta}} \xrightarrow[\text{a.s.}]{L^2} A(t). \tag{8.2}$$

Remark. If $\beta=0$, Y is symmetric and this theorem is due to Knight. If $0<\alpha<1$, the exponent in (8.2) can take all values in $(0,\alpha)$. Positive values of β correspond to an asymmetric process in which positive jumps are more likely than negative jumps. Ths makes Y "stick" near its maximum so the numerator in (8.2) is larger and the exponent of ε has to be smaller than for $\beta=0$. For $\alpha=1$, the processes Y included in the theorem are usually called symmetric Cauchy with a linear drift corresponding to the value of β. A positive drift corresponds to $\beta>0$ and a smaller exponent in (8.2). For $1<\alpha<2$ we can again see how the lack of symmetry in the stable process Y leads to exponents in range $[\alpha-1,1]$. For $\alpha=2$ we must have Brownian motion and our theorem is just a restatement of (1.1).

Proof. For all the processes Y included,

$$\inf\{s>0\colon Y(s)>0\}=0=\inf\{s>0\colon Y(s)<0\} \quad \text{a.s.}$$

so the process X obviously satisfies conditions (2.1) and (2.2); thus $A(t)$ exists. We can apply our framework to the excursions of X. For $\varepsilon>0$ put

$$f_\varepsilon(V)=\int_0^{V^+} \mathbf{1}_{[0,\varepsilon)}(V(s))\,ds$$

and let μ_ε be the measure corresponding to f_ε via (2.8). We will show that there exist finite positive constants c_1 and c_2 such that

$$\int s\,\mu_\varepsilon(ds)=c_1\,\varepsilon^{\alpha/2-(1/\pi)\arctan\beta}, \tag{8.3}$$

and

$$\int s^2\,\mu_\varepsilon(ds)=c_2\,\varepsilon^{3\alpha/2-(1/\pi)\arctan\beta}. \tag{8.4}$$

Since $|\arctan(\pm\tan\pi\alpha/2)|<\pi\alpha/2$ for $\alpha>1$, straightforward uses of (8.3) and (8.4) in Theorems 3.5 and 4.1 and Lemma 2.3 give the desired convergence. It remains to prove (8.3) and (8.4).

The process X inherits the scaling property (see Example 6.3) from Y. We use the scaling property, Lemma 3.4, $\theta=\varepsilon^{-\alpha}$, and the known [Fristedt, 1974, Example 9.13] formula

$$\varphi(\theta)=\theta^{\frac{1}{2}+\frac{1}{\pi\alpha}\arctan\beta}$$

to obtain

$$\begin{aligned}\int_{(0,\infty)} r\,\mu_\varepsilon(dr)&=\theta^{\frac{1}{2}+\frac{1}{\pi\alpha}\arctan\beta}\,E\Big(\int_0^\infty e^{-\theta t}\,d\sum_{W^-\leqq t} f_\varepsilon(W)\Big)\\&=\varepsilon^{-\frac{\alpha}{2}-\frac{1}{\pi}\arctan\beta}\,E\Big(\int_0^\infty e^{-\varepsilon^{-\alpha}t}\,d\sum_{W^-\leqq t\varepsilon^{-\alpha}} f_\varepsilon(\varepsilon W(\varepsilon^{-\alpha}\cdot))\Big)\\&=\varepsilon^{\frac{\alpha}{2}-\frac{1}{\pi}\arctan\beta}\,E\Big(\int_0^\infty e^{-s}\,d\sum_{W^-\leqq s} f_1(W)\Big).\end{aligned}$$

A similar argument shows

$$\int_0^\infty r^2\,\mu_\varepsilon(dr)=\varepsilon^{\frac{3\alpha}{2}-\frac{1}{\pi}\arctan\beta}\,E\left(\int_0^\infty e^{-s}\,d\sum_{W^-\leqq s} f_1^2(W)\right).$$

It remains to prove that $\int r\,\mu_\varepsilon(dr)$ and $\int r^2\,\mu_\varepsilon(dr)$ are finite. That is the case but we omit the argument since it is similar to and slightly easier than an analogous argument in Sect. 9. □

9. Residual Area for a Strictly Stable Process

Our final result gives a construction which is new even for Brownian motion. The arguments works for any strictly stable process (defined in Example 6.3) so we state it in that context.

Corollary 9.1. *Suppose X is a strictly stable process in $\mathbb{R}$ with index $\alpha>1$. Then, as $\varepsilon\downarrow 0$,*

$$\frac{\varepsilon t-\int_0^t(\varepsilon\wedge|X(s)|)\,ds}{\varepsilon^2}\to A(t) \tag{9.1}$$

a.s. and in L^2, where A is the local time at 0.

Remark. The quantity $\varepsilon t-\int_0^t \varepsilon\wedge|X(s)|\,ds$ is the area above the graph of $|X|$ below ε.

Proof. The total area can be calculated excursion by excursion. Let

$$f_\varepsilon(V)=\int_0^{V^+}[\varepsilon-(\varepsilon\wedge|V(s)|)]\,ds.$$

We use Lemma 3.4, the scaling property, $\varphi(\theta)=\theta^{(\alpha-1)/\alpha}$, and $\theta=\varepsilon^{-\alpha}$ as in Example 6.3 to obtain

$$\begin{aligned}\int_{(0,\infty)} r\,\mu_\varepsilon(dr)&=\theta^{(\alpha-1)/\alpha}E\left(\int_0^\infty e^{-\theta t}\,d\sum_{W^-\leqq t} f_\varepsilon(W)\right)\\&=\varepsilon^{-\alpha+1}E\left(\int_0^\infty e^{-t\varepsilon^{-\alpha}}\,d\sum_{W^-\leqq t\varepsilon^{-\alpha}} f_\varepsilon(\varepsilon W(\varepsilon^{-\alpha}\cdot))\right)\\&=\varepsilon^2 E\left(\int_0^\infty e^{-s}\,d\sum_{W^-\leqq s} f_1(W)\right).\end{aligned} \tag{9.2}$$

By using f_ε^2 in lieu of f_ε we obtain

$$\int_{(0,\infty)} r^2\,\mu_\varepsilon(dr)=\varepsilon^{3+\alpha}E\left(\int_0^\infty e^{-s}\,d\sum_{W^-\leqq s} f_1^2(W)\right). \tag{9.3}$$

We will next show that (9.2) is finite. Integration by parts, correct even if the integral is infinite, gives

$$\int_0^\infty e^{-s} d \sum_{W^- \leqq s} f_1(W) = \int_0^\infty e^{-s} \sum_{W^- \leqq s} f_1(W)\, ds$$

$$\leqq \int_0^\infty e^{-s}\left[s + \int_s^{\inf\{t>s: X(t)=0\}} \mathbf{1}_{[-1,1]}(X(u))\, du\right] ds$$

$$\leqq 1 + \int_0^\infty e^{-s}(K_s+1)\, ds$$

where K_s, conceivably $+\infty$, is the supermum of integers k such that

$$\int_s^{\inf\{t>s: X(t)=0\}} \mathbf{1}_{[-1,1]}(X(u))\, du > k.$$

Since X is recurrent, for any k we can define

$$\mathrm{t}_k = \inf\left\{u: \int_s^u \mathbf{1}_{[-1,1]}(X(u))\, du = k\right\}$$

By quasi-left-continuity $|X(\mathrm{t}_k -)| \leqq 1$. By restarting X at time t_k we see that, conditioned on $K_s \geqq k$, the probability that $K_s \geqq k+q$ is no larger than

$$1 - \inf_{|x| \leqq 1} P^x\{X(u)=0 \text{ for some } u<q\}$$

which, by the scaling property (see Example 6.3), is equal to

$$1 - P^{(-1)}\{X(u)=0 \text{ for some } u<q\} \wedge P^1\{X(u)=0 \text{ for some } u<q\},$$

which is less than 1 for some q (actually, for all q). Hence, $E(K_s)$ is bounded by a finite number independent of s and, thus, (9.2) is finite. A similar argument shows (9.3) to be finite.

By Theorem 3.5, Theorem 4.1, and Lemma 2.3,

$$\frac{\sum_{W^- \leqq t} f_\varepsilon(W)}{\varepsilon^2} \to A(t) \tag{9.4}$$

a.s. and in L^2. To obtain (9.1) from (9.4) notice that the numerator in (9.1) is no larger than the numerator in (9.4) but is at least as large as the numerator one obtains by replacing t by s_γ, $\gamma > 0$, in (9.4), where

$$s_\gamma = \sup\{u: A(u) = A(t) - \gamma\}. \quad \square$$

Question. Is Corollary 9.1 true for every Lévy process satisfying (2.1) and (2.2)? We note that the normalizing function ε^2 does not depend on the index of the strictly stable process.

10. Concluding Remarks

In this paper we have produced a general method which includes as special cases all the constructions (1.1) to (1.4) first considered for Brownian motion.

Provided the construction can be made by "adding up" the behavior on separate excursions we have general theorems which will apply, although it is not always easy to check the truth of the hypotheses. However, there are cases where the method does not apply because the construction cannot be divided according to excursions. This is the case for the result due to Getoor [1976] for downcrossings from x_m to y_m, both different from x. However if one knows enough about the process one can sometimes obtain this type of result from the result based on excursions. This is done in [Pruitt and Taylor, 1982] to obtain a construction of local time for the asymmetric Cauchy process different from that given in Example 6.4.

References

Blumenthal, R.M., Getoor, R.K.: Markov Processes and Potential Theory. New York-London: Academic Press 1968

Chung, K.L., Durrett, R.: Downcrossings and local time. Z. Wahrscheinlichkeitstheorie Verw. Gebiete **35**, 147–149 (1976)

Erickson, K.B.: A limit theorem for renewal sequences with an application to local time. Z. Wahrscheinlichkeitstheorie Verw. Gebiete **57**, 535–558 (1981)

Fristedt, B.: Sample functions of stochastic processes with stationary, independent increments. Advances in Probability, Volume 3 edited by Ney, P. and Port, S. pp. 241–396. New York: Marcel Dekker 1974

Getoor, R.K.: Another limit theorem for local time. Z. Wahrscheinlichkeitstheorie Verw. Gebiete **34**, 1–10 (1976)

Getoor, P.K., Millar, P.W.: Some limit theorems for local time. Compositio Math. **25**, 123–134 (1972)

Greenwood, P., Ptiman, J.: Construction of local time and Poisson point processes from nested arrays. J. London Math. Soc. (2) **22**, 182–192 (1980)

Hall, W.J.: On Wald's equations in continuous time. J. Appl. Probab. **7**, 59–68 (1970)

Hawkes, J.: Measure function properties of the asymmetric Cauchy process. Mathematika **17**, 68–78 (1970)

Hoffmann-Jørgensen, J.: Markov sets. Math. Scand. **24**, 145–166 (1969)

Horowitz, J.: Semilinear Markov processes, subordinators, and renewal theory. Z. Wahrscheinlichkeitstheorie Verw. Gebiete **24**, 167–194 (1972)

Itô, K.: Poisson point processes attached to Markov processes. Proceedings of the Sixth Berkeley Symposium, 225–239 (1970)

Itô, K., McKean, H.P.: Diffusion Processes and Their Sample Paths. Berlin-Heidelberg-New York: Springer 1965

Kingman, J.F.C.: An intrinsic description of local time. J. London Math. Soc. (2) **6**, 725–731 (1973)

Knight, F.: The local time at zero of the reflected symmetric stable process. Z. Wahrscheinlichkeitstheorie Verw. Gebiete **19**, 180–190 (1971)

Lévy, P.: Processus Stochastiques et Mouvement Brownien. Paris: Gauthier-Villars 1948

Loève, M.: Probability Theory, Second Edition. Princeton-Toronto-New York-London: D. Van Nostrand Company 1960

Maisonneuve, B.: Systèms régénératifs. Astérisque **15**, 1–124 (1974)

Maisonneuve, B.: Temps local et dénombrements d'excursions. Z. Wahrscheinlichkeitstheorie Verw. Gebiete **52**, 109–113 (1980)

Maisonneuve, B.: On Lévy's downcrossing theorem and various extensions. Z. Wahrscheinlichkeitstheorie Verw. Gebiete. Séminaire de Probabilités XV, pp. 191–205. Lecture Notes in Mathematics Vol. 850. Berlin-Heidelberg-New York: Springer 1981

Port, S.C., Stone, C.J.: The asymmetric Cauchy process on the line. Ann. Math. Statist. **40**, 137–143 (1969)

Pruitt, W.E., Taylor, S.J.: The local structure of the sample paths of asymmetric Cauchy processes. To appear

Stone, C.J.: The set of zeros of a semi-stable process. Illinois J. Math. **7**, 631–637 (1963)

Taylor, S.J.: Sample path properties of processes with stationary independent increments. Stochastic Analysis edited by Kendall, D.G. and Harding, E.F. pp. 387–414. London-New York-Sydney-Toronto: John Wiley and Sons 1973

Williams, D.: Lévy's downcrossing theorem. Z. Wahrscheinlichkeitstheorie Verw. Gebiete **40**, 157–158 (1977)

Received January 25, 1982; in final form July 30, 1982