

**Arcsine laws and interval partitions
derived from a stable subordinator**

by

Jim Pitman and Marc Yor

**Technical Report No. 189
November 1990**

Research supported by NSF grant DMS-88-01808

**Department of Statistics
University of California
Berkeley, California 94720**

Arcsine laws and interval partitions derived from a stable subordinator

by Jim Pitman and Marc Yor.

November 28, 1990

Abstract. Lévy discovered that the fraction of time a standard one-dimensional Brownian motion B spends positive before time T has arcsine distribution, both for T a fixed time when $B_T \neq 0$ almost surely, and for T an inverse local time, when $B_T = 0$ almost surely. This identity in distribution is extended from the fraction of time spent positive to a large collection of functionals derived from the lengths and signs of excursions of B away from 0. Similar identities in distribution are associated with any process whose zero set is the range of a stable subordinator, for instance a Bessel process of dimension d for $0 < d < 2$.

Keywords. Brownian motion, Bessel process, local time, occupation time, excursions, Poisson point process, Palm distribution, length/size biased sampling

1. Introduction.

Let $(B_t, t \geq 0)$ be a Brownian motion on the line, starting at 0. Let $\Gamma_+(t)$ be the time B spends above 0 up to time t :

$$\Gamma_+(t) = \int_0^t 1(B_s > 0) ds.$$

Lévy (1939) showed that for each $t > 0$ the variable $\Gamma_+(t)/t$ has the arcsine law:

$$(1.a) \quad P(\Gamma_+(t)/t \in du) = \frac{du}{\pi \sqrt{u(1-u)}}, \quad \text{for } 0 < u < 1.$$

On the way to this result, Lévy showed that the same arcsine distribution is obtained if the fraction of time spent positive, $\Gamma_+(t)/t$, is considered at the random time $t = T_s$, where $(T_s, s \geq 0)$ is the inverse of the continuous local time process $(S_t, t \geq 0)$ which Lévy associated with the random set of zeros of B . We denote this identity in distribution by

$$(1.b) \quad \frac{\Gamma_+(t)}{t} \stackrel{d}{=} \frac{\Gamma_+(T_s)}{T_s}$$

As Lévy noted in the third paragraph of page 326 of his 1939 paper, it is remarkable that the same law for the fraction of time spent positive should appear both at a fixed time t when $B_t \neq 0$ almost surely, and for the random time $t = T_s$ when $B_t = 0$ almost surely. Our aim is to expose as best we can what lies behind the identity of laws (1.b). This enables us to extend the identity to a large collection of functionals derived

from the lengths and signs of excursions of B away from 0. We find similar identities for a process whose zero set is the range of a stable subordinator. These results are closely related to the multidimensional arcsine laws of Barlow, Pitman and Yor (1989b) referred to below as BPYb.

We start by recalling how Lévy derived the arcsine distribution of $\Gamma_+(T_s)/T_s$. He first showed that $(T_s, s \geq 0)$ is an increasing process with independent increments, or *subordinator*, which is stable with index $\alpha = 1/2$. There is a one to one correspondence between jumps of this subordinator and maximal open intervals during which the Brownian motion is away from zero. From the Poisson character of the jumps of the subordinator, and the fact that each of these jumps contributes to time positive with probability $1/2$, independently of all others, Lévy argued that

$$(1.c) \quad (\Gamma_+(T_s), T_s) \stackrel{d}{=} (T_{s/2}, T_s),$$

so that

$$(1.d) \quad \frac{\Gamma_+(T_s)}{T_s} \stackrel{d}{=} \frac{T_{s/2}}{T_{s/2} + T_{s/2}^*}$$

where $T_{s/2}^* = T_s - T_{s/2}$ is independent of $T_{s/2}$ with the same stable distribution with index $1/2$. Lévy showed the distribution of the right hand variable in (1.d) is arcsine by a two dimensional integration. This can also be seen more simply as in paragraph (1.1) of BPYb.

Using (1.c) and the formula

$$(1.e) \quad P(T_s \in dt)/dt = \frac{s}{\sqrt{2\pi}} t^{-3/2} e^{-s^2/2t}, \quad s > 0, t > 0,$$

Lévy gave an explicit formula for the density of the conditional distribution

$$(1.f) \quad P(\Gamma_+(T_s) \in \cdot \mid T_s = t) = P(\Gamma_+(t) \in \cdot \mid B_t = 0, S_t = s).$$

By integration with respect to the conditional distribution of S_t given $B_t = 0$, he then obtained his famous result for the time spent positive by a Brownian bridge:

$$(1.g) \quad \text{the conditional distribution of } \frac{\Gamma_+(t)}{t} \text{ given } B_t = 0 \text{ is uniform on } [0, 1].$$

Finally, Lévy pointed out that the unconditional distribution of $\Gamma_+(t)/t$ could be found by conditioning on the time G_t of the last zero before time t . Lévy showed that G_t/t too has arcsine distribution and that

$$(1.h) \quad \text{given } G_t = g \text{ the pre-}g \text{ process } (B_u, 0 \leq u \leq g) \text{ is a Brownian bridge of length } g.$$

He then derived the arcsine law (1.a) for $\Gamma_+(t)/t$ by applying (1.g) to the bridge in

(1.h), and using the fact that independently of what happens before g , on the final zero free interval of length $1 - g$, the Brownian path is equally likely to go positive or negative.

The most important ingredient in these results of Lévy is the fact that the random zero set Z of the Brownian motion B is the closure of the range of a stable subordinator of index $\alpha = 1/2$. Somewhat more generally, suppose B is replaced by a diffusion on the line whose zero set is the closed range of a subordinator $(T_s, s \geq 0)$, which is *stable with index α* , for some $0 < \alpha < 1$, meaning that for every $a > 0$,

$$(1.i) \quad (T(s), s \geq 0) \stackrel{d}{=} (aT(s/a^\alpha), s \geq 0).$$

(We often use notation like $T(s)$ instead of T_s for typographical convenience.) A diffusion B with such an inverse local time process at zero can be constructed as in BPYa using a Bessel process of dimension $d = 2(1 - \alpha)$ for the modulus, and coin tossing for the signs. Lévy's formulae (1.c) and (1.d) adapt at once to this setting to describe the law of $\Gamma_+(T_s)/T_s$. And as shown by a Laplace transform calculation in BPYb, the identity in law (1.b) between $\Gamma_+(t)/t$ and $\Gamma_+(T_s)/T_s$ continues to hold, even though Lévy's approach sketched above is blocked in this setting by the lack of any analogue of the explicit formula (1.e) for $\alpha \neq 1/2$.

Puzzled by why (1.b) should hold so generally, we were led to analyze the *interval partition* generated by the zeros up to time t . By this we mean the collection of lengths of maximal intervals comprising $Z^c \cap (0, t)$, without regard to the order in which these lengths appear. We describe the interval partition by the sequence of ranked lengths of intervals, as in the following theorem:

Theorem 1.1. *Fix α with $0 < \alpha < 1$. Suppose Z is the closure of the range of a stable subordinator (T_s) of index α . Let $V(t)$ be the infinite sequence of lengths of the maximal open subintervals of $Z^c \cap (0, t)$, arranged in descending order:*

$$(1.j) \quad V(t) = (V_1(t), V_2(t), V_3(t), \dots)$$

where $V_1(t) \geq V_2(t) \geq V_3(t) \geq \dots$. Then for $t > 0$ and $s > 0$,

$$(1.k) \quad \frac{V(t)}{t} \stackrel{d}{=} \frac{V(T_s)}{T_s}.$$

To illustrate this result in the Brownian case, $\alpha = 1/2$, Z is the zero set of the Brownian motion, and each interval of Z^c is given a sign by a fair coin tossing process independent of interval length. So the i th longest interval of $Z^c \cap (0, t)$, that is $V_i(t)$, contributes with probability $1/2$ to the sum of positive interval lengths $\Gamma_+(t)$, independently as i varies and independently of $V(t)$. The same can be said at the random time

T_s instead of the fixed time t . So Lévy's identity (1.b) follows at once from (1.k).

Notice that for a fixed time t , one of the lengths appearing in the sequence $V(t)$ represents the age A_t of the excursion in progress at time t :

$$(1.m) \quad A_t = t - G_t, \quad \text{where} \quad G_t = \sup \{Z \cap (0, t)\}$$

is the last time in Z before t . More precisely, $A_t = V_{\#(t)}(t)$, where $\#(t) - 1$ is the number of excursions completed before time t with lifetime longer than A_t . As a complement to Theorem 1.1 we show:

Theorem 1.2. *Conditionally given $V(t)$, the length of the final interval is picked by length biased sampling. In symbols:*

$$P(\#(t) = j \mid V(t)) = V_j(t)/t, \quad j = 1, 2, \dots$$

In contrast to a fixed time t , the random time T_s falls in Z almost surely, so $G_{T(s)} = T_s$ and $A_{T(s)} = 0$ almost surely. This means that every length appearing in $V(T_s)$ represents the lifetime of a complete excursion. And $\#(T_s)$ is undefined. So despite the equality in distribution of $V(t)/t$ and $V(T_s)/T_s$ it is nonsense to substitute the random time T_s instead of the fixed t in Theorem 1.2. But suppose V_s^* is the length of the complete excursion containing a random time point T_s^* , which given the subordinator is distributed uniformly on the interval $(0, T_s)$. Then, according to Theorems 1.1 and 1.2, the joint distribution of $V(t)/t$ and A_t/t is the same as that of $V(T_s)/T_s$ and V_s^*/T_s .

Kingman (1975, Section 7), considered the distribution of $V(T_s)/T_s$ for a more general subordinator. We conjecture that the conclusion of either of Theorems 1.1 or 1.2 is characteristic of the stable case. This motivates us to reformulate these results in Section 7 in a way which makes sense for more general subordinators. In particular, in case (T_s) is a gamma process, the law of $V(T_s)/T_s$ is the well known Poisson-Dirichlet distribution on the infinite simplex with parameter s . This distribution turns up in a number of different contexts, for example in the asymptotics for the ranked cycle lengths of a random permutation (Shepp and Lloyd (1966), Vershik and Schmidt (1977 and 78)), as the distribution of the ranked sizes of atoms in the Dirichlet process prior in Bayesian statistics (Blackwell and McQueen (1973), Ferguson (1973)), and in limiting models for the abundances of genes in population genetics and species in mathematical ecology (see e.g. Patil and Taillie (1977), Hoppe (1987), Joyce and Tavaré (1987)). Explicit but rather intractable formulae for features of the Poisson-Dirichlet distribution such as the joint density of the first n components can be found in Watterson (1976), Ewens (1988), and some of the other references above. Perman

(1990) gives extensions of these formulae for a class of subordinators including both the stable and gamma cases in terms of the solution of an integral equation. In the stable case the one-dimensional density of the largest component $V_1(T_s)/T_s$ is found quite simply on the interval $(1/2, 1]$, then determined recursively by the integral equation on the intervals $(1/3, 1/2]$, $(1/4, 1/3]$, and so on. But explicit integration seems possible only in the case $\alpha = 1/2$ and even then only as far as the interval $(1/3, 1/2]$. See also Gettoor (1979) and Knight (1985) for results on the distribution of $V_1(G_t)$ at fixed times t for a general subordinator.

While explicit description of the common finite dimensional distributions of $V(T_s)/T_s$ and $V(t)/t$ in the stable case turns out to be unmanageable, Theorem 1.1 hints that there may be a simple description for the size biased permutation of the values of $V(T_s)/T_s$, analogous to the following well known result in the Poisson-Dirichlet case: (Patil and Taillie (1977, Theorem 2.3), Donnelly and Joyce (1989, Theorem 5)):

If T_s has gamma(s) distribution and Y_1, Y_2, \dots represents a size biased permutation of $V(T_s)/T_s$, then

$$(1.n) \quad Y_i = \xi_i \prod_{j=1}^{i-1} (1 - \xi_j) \quad , i \geq 1$$

where the ξ_i are independent with identical beta($1, s$) distribution.

Indeed, Perman (1990, Corollary 3.19) has applied Theorems 1.1 and 1.2, to derive the following remarkable analog in the stable case:

If T_s has stable (α) distribution, then the size-biased permutation of $V(T_s)/T_s$ is described by (1.n) for independent and non-identically distributed ξ_i , where ξ_i has beta($1-\alpha, i\alpha$) distribution.

Due to Theorem 1.1, (Y_i) constructed as in (1.n) from such beta variables also describes the length biased permutation of the unit interval partition $V(t)/t$ associated with the stable (α) subordinator, for any $t > 0$. These and other results concerning size biased sampling of the jumps of a subordinator are studied in Perman, Pitman and Yor (1990).

Thinking now of Z as the zero set of an underlying Brownian motion or Bessel process, Theorems 1.1 and 1.2 invite interpretation at the level of some kind of path transformation. In the case of classical identities in distribution related to arc sine laws for random walks, this is achieved by the Sparre-Andersen transformations described by Feller (1971 Section XII.8). Karatzas and Shreve (1987) offer an analogous transformation in the Brownian case. Our set up is a little different, as we have

moved away from processes with independent increments to include the Bessel case with zero set defined by a stable subordinator. But by a combination of Brownian scaling and a rearrangement of intervals in the zero set we come up with the following:

Theorem 1.3. *Let $(B_t, t \geq 0)$ be a Brownian motion on \mathbb{R} , or a Bessel process of dimension d on \mathbb{R}^+ , for $0 < d < 2$, starting at $B_0 = 0$. Let $T_1 = \inf\{t : S_t = 1\}$ where $(S_t, t \geq 0)$ is the local time process of B at 0. Let $(X_t, 0 \leq t \leq 1)$ be derived from $(B_u, 0 \leq u \leq T_1)$ by Brownian scaling:*

$$X_t = \frac{1}{\sqrt{T_1}} B(t T_1), \quad 0 \leq t \leq 1.$$

Let U be distributed uniformly on $[0, 1]$, independent of (X_t) . Let (G, D) be the maximal interval free of zeros of X which contains U , so $G < U < D$ and $X_G = X_D = 0$. Let $H = 1 - (D - G)$, and let Y be the process with lifetime H obtained by excising the excursion of X away from zero on the interval (G, D) , and closing up the gap:

$$\begin{aligned} Y_t &= X_t, & 0 \leq t \leq G \\ &= X_{D+t-G}, & G \leq t \leq H. \end{aligned}$$

Then

$$(Y_t, 0 \leq t \leq H) \stackrel{d}{=} (B_t, 0 \leq t \leq G_1)$$

where $G_1 = \sup\{t : t \leq 1, B_t = 0\}$.

The process $(X_t, 0 \leq t \leq 1)$ we call a *pseudo-bridge*. According to Theorem 5.2 below, which extends the result of Biane, Le Gall and Yor (1987) from the Brownian case to the Bessel case, the law of the pseudo-bridge has density $\text{constant}/S_1$ relative to the that of the standard bridge $(B_t, 0 \leq t \leq 1 \mid B_1 = 0)$, where S_1 is the bridge local time at zero up to time 1.

We note in passing that that G_1 has beta($\alpha, 1 - \alpha$) law, where $\alpha = 1 - d/2$, (see e.g. Dynkin (1961), BPYb), and that given G_1 , $(B_t, 0 \leq t \leq G_1)$ is a Brownian or Bessel bridge of length G_1 , as in (1.h).

Theorem 1.4 contains Theorems 1.1 and 1.2 implicitly. To see this, just note that $V(1)$ and $1 - G_1$ are measurable functions of $(B_t, 0 \leq t \leq G_1)$, hence have the same joint law as the corresponding functions of $(Y_t, 0 \leq t \leq H)$, which are $V(T_s)/T_s$ and $D - G$ by examination of the construction. Spelled out even further,

$$(Y_t, 0 \leq t \leq H ; 1 - H ; V(T_1)/T_1 ; T_1^{-1/\alpha})$$

$$\stackrel{d}{=} (B_t, 0 \leq t \leq G_1; 1 - G_1; V(1); S_1).$$

In the Brownian case, two variables that can be added to the first list are the sign of X on (G, D) (i.e. the sign of B on (GT_1, DT_1)), and $\Gamma_+(T_1)/T_1$, where $\Gamma_+(t)$ is the time spent positive by B up to time t . Here the sign is independent of $(Y_t, 0 \leq t \leq H)$, positive with probability $1/2$, and by construction

$$\frac{\Gamma_+(T_1)}{T_1} = \int_0^H 1(Y_t > 0) dt + (1 - H) 1(+)$$

where $1(+)$ is the indicator of a positive sign of X on (G, D) . The corresponding variables in the second list are the sign of B on $(G_1, 1)$, and $\Gamma_+(1)$, for similar reasons.

One would naturally like to extend the definition of Y beyond time H and up to time 1 to get a Brownian motion for $0 \leq t \leq 1$. This is done in Theorem 6.3 by continuing Y after time H with a signed Brownian (or Bessel) meander of length $1 - H$, which given H is independent of Y on $[0, H]$. This meander could be created by some transformation of the signed excursion of length $1 - H$ which was excised from X to make Y , or by further randomization. But we do not see any particular way of doing this with interesting implications.

The rest of the paper is organized in sections as follows:

2. Reformulation and proof of Theorems 1.1 and 1.2. Theorem 1.1 is reformulated as Theorem 2.1. Theorem 2.1 is then proved using Lemma 2.2, which describes in particular the joint distribution of X and N^- , where N is a point process, and N^- is obtained from N by deletion of a point X picked from the points of N in an arbitrary probabilistic way.

3. Analysis at an independent exponential time. This section offers an alternative derivation of Theorems 1.1 and 1.2. A substitute for Lemma 2.2 is provided by Theorem 3.1. This gives a remarkable description of the joint distribution of X and N^- , where N^- is derived from a Poisson process N by removal of a single point X picked from N by h -biased sampling for some positive function h : a random variable Y is created such that conditionally given Y , X and N^- are independent, and N^- is still Poisson.

4. Application to occupation times. Here we recover the result of BPYb concerning the joint distribution of the occupation times in sectors of Walsh's singular diffusion in the plane.

5. Results in terms of bridges.

6. Rearrangements. Here consider the conditional distribution of the whole process given the interval partition it creates on $(0, t)$. This leads to the derivation of Theorem 1.3.

7. More general subordinators. Here we reformulate some of the results in the stable case in a way that makes sense for more general subordinators.

2. Reformulation and proof of Theorems 1.1 and 1.2

We assume throughout this section that $(T_s, s \geq 0)$ is a stable subordinator with index α . As is well known, the Laplace transform of T_s must then be of the form

$$(2.a) \quad E[e^{-\theta T(s)}] = \exp(-c s \theta^\alpha) = \exp\left(-s \int_0^\infty (1 - e^{-\theta x}) \Lambda(dx)\right), \quad \theta > 0$$

where $c > 0$ is a scale constant, and Λ is the Lévy measure on $(0, \infty)$ defined by

$$(2.b) \quad \Lambda(dx) = \frac{c \alpha}{\Gamma(1-\alpha)} \frac{dx}{x^{\alpha+1}}, \quad 0 < x < \infty.$$

Think of the range of the subordinator as a subset of a *real time* scale $t \geq 0$ parameterized by a *local time* scale $s \geq 0$. Assume $(T_s, s \geq 0)$ is right continuous, and let $(S(t), t \geq 0)$ be the continuous local time process, inverse to the subordinator. So

$$S(t) = \inf\{s : T(s) > t\}; \quad T(s) = \inf\{t : S(t) > s\}.$$

The counting process associated with the local times and jump lengths of the subordinator is a Poisson process, homogeneous in the local time, with intensity measure for jump lengths Λ per unit local time.

We start with an elementary remark, namely that the distribution of $S(t)$ for any fixed t can be derived from that of T_s for any fixed s . Indeed, for any $t > 0$ and $x > 0$

$$\begin{aligned} P(S(t)/t^\alpha < x) &= P(S(t) < t^\alpha x) \\ &= P(T(t^\alpha x) > t) && \text{by the inverse relation} \\ &= P(T(t^\alpha x/a^\alpha) > t/a) && \text{for any } a > 0, \text{ by (1.i)} \\ &= P(T(s) > (s/x)^{1/\alpha}) && \text{by choice of } a = t(x/s)^{1/\alpha} \\ &= P(s/T_s^\alpha < x). \end{aligned}$$

To summarize, for $t > 0$ and $s > 0$,

$$(2.c) \quad \frac{S(t)}{t^\alpha} \stackrel{d}{=} \frac{s}{T(s)^\alpha}.$$

In particular, we note that to get the usual normalization of local time in the *Brownian case*, so as to make $S(t) \stackrel{d}{=} |B_t|$, the parameters are $\alpha = 1/2$ and $c = \sqrt{2}$. Then (2.c) gives $T(s) \stackrel{d}{=} s^2/B_1^2$, hence (1.e), as noted by Lévy (1939).

Consider now the sequence of ranked interval lengths $V(t)$ as in Theorem 1.1. It is well known that

$$(2.d) \quad S(t) = \lim_{\varepsilon \rightarrow 0} \frac{\#\{i : V_i(t) > \varepsilon\}}{\Lambda(\varepsilon, \infty)}, \quad t > 0, \quad \text{where}$$

$$(2.f) \quad \Lambda(\epsilon, \infty) = \frac{c \epsilon^{-\alpha}}{\Gamma(1-\alpha)}$$

is the Lévy measure for jumps lengths greater than ϵ , and the convergence in (2.d) is uniform on bounded t -intervals almost surely. This formula shows that the local time $S(t)$ may be regarded a measurable function of $V(t)$, say

$$(2.g) \quad S(t) = \text{Loc} [V(t)].$$

For any strictly positive, fixed or random normalizing factor X , (2.d) and (2.f) give

$$(2.h) \quad \text{Loc} \left[\frac{V(t)}{X} \right] = \frac{S(t)}{X^\alpha}$$

simultaneously for all $t \geq 0$ a.s., hence also for random times as well as fixed ones.

Now fix $t > 0$ and $s > 0$, and consider the assertion of Theorem 1.1, namely that

$$(2.i) \quad \frac{V(t)}{t} \stackrel{d}{=} \frac{V(T_s)}{T_s}.$$

Applying Loc to both sides shows that (2.i) implies (2.c), and indeed

$$(2.j) \quad \left(\frac{V(t)}{t}, \frac{S(t)}{t^\alpha} \right) \stackrel{d}{=} \left(\frac{V(T_s)}{T_s}, \frac{s}{T_s^\alpha} \right).$$

This in turn is equivalent by the same kind of trick to

$$(2.k) \quad \frac{V(t)}{S(t)^{1/\alpha}} \stackrel{d}{=} \frac{V(T_s)}{s^{1/\alpha}}.$$

So the random scaling can be transferred from one side to the other. Since for each s the sequence $V(T_s)/s^{1/\alpha}$ just consists of the descending sequence of points of a Poisson point process governed by Λ (abbreviated $\text{PPP}(\Lambda)$), and since the terms of $V(t)$ are also descending, Theorem 1.1 can be reformulated as follows:

Theorem 2.1. *For subsets B of $(0, \infty)$ let*

$$N_t(B) = \#\{i : \frac{V_i(t)}{S_t^{1/\alpha}} \in B\}.$$

Then for each fixed $t > 0$, the counting process N_t is a PPP (Λ) .

Let $X_t = A_t/S_t^{1/\alpha}$ be the point of N_t corresponding to the final interval in $V(t)$. In view of (2.g), Theorem 1.2 amounts to the assertion

(2.m) X_t is a point of N_t picked at random by size biased sampling:

$$P(X_t \in dx | N_t) = x N_t(dx) / \int y N_t(dy).$$

To summarize the above discussion: Theorems 1.1 and 1.2 admit an equivalent formulation as Theorem 2.1 and (2.m). The rest of this section offers a proof of the latter results by application of the following general lemma. An alternative approach is offered in the following section.

Lemma 2.2. (Poisson sampling) *Let $(\Omega^*, \mathcal{F}^*)$ be the space of counting measures on a measurable space (S, \mathcal{S}) . Let Q be the law on $(\Omega^*, \mathcal{F}^*)$ of a PPP (μ) over (S, \mathcal{S}) . Let $f(x, n)$ be a non-negative jointly measurable function of a point $x \in S$ and a counting measure $n \in \Omega^*$. Let X be a random point in S and N^- a random counting measure over S defined jointly on some basic probability space (Ω, \mathcal{F}, P) . Let $N = N^- + \delta_X$ be the point process obtained from N^- by addition of a single extra point at X . The following statements are equivalent:*

$$(2.n) \quad P(X \in dx, N^- \in dn) = f(x, n + \delta_x) \mu(dx) Q(dn)$$

$$(2.o) \quad P(X \in dx, N \in dn) = f(x, n) n(dx) Q(dn)$$

Proof. Let Q_x denote the distribution of $n + \delta_x$ when n is a PPP (μ) with law Q . By a change of variable, (2.n) amounts to

$$(2.p) \quad P(X \in dx, N \in dn) = f(x, n) \mu(dx) Q_x(dn).$$

But it is well known that $(Q_x, x \in S)$ serves as a family of *Palm distributions* for Q , meaning that there is the identity of measures on $(\Omega^* \times S, \mathcal{F}^* \otimes \mathcal{S})$

$$(2.q) \quad \mu(dx) Q_x(dn) = n(dx) Q(dn).$$

See for instance Daley and Vere-Jones (1988), Section 12.1. \square

We note that (2.o) displays the most general possible joint distribution of X and N where N has a law absolutely continuous with respect to Q , and X is a point of N . The proof shows how to formulate the lemma for a more general distribution Q on the space of counting measures $(\Omega^*, \mathcal{F}^*)$ instead of the law of a PPP. Simply replace Q in (2.n) by the Q_x distribution of $n - \delta_x$, where $(Q_x, x \in S)$ serves as a family of Palm distributions for Q relative to μ , as in (2.q). The existence of such a disintegration of the law Q of a point process is known in great generality (see e.g. Kallenberg (1983)).

Proof of Theorem 2.1 and (2.m). Let Q be the law of a PPP (Λ) . Theorem 2.1 and (2.m) combined amount to

$$(2.r) \quad P(X_i \in dx, N_i \in dn) = \frac{x}{\int y n(dy)} n(dx) Q(dn),$$

which, by Lemma 2.2, is equivalent to

$$(2.s) \quad P(X_t \in dx, N_t^- \in dn) = \frac{x}{x + \int y n(dy)} \Lambda(dx) Q(dn),$$

where $N_t^- = N_t - \delta_{X_t}$ is the point process N_t with the point at X_t removed. Thus N_t^- has a point at $V_n(t)/S_t^{1/\alpha}$ for each complete excursion interval $V_n(t)$. Now it is clear in principle that both X_t and N_t^- are functions of the point process M on $(0, \infty)^2$ defined by the jumps of the subordinator $(T_s, s \geq 0)$:

$$M(B) = \#\{s : (s, T_s - T_{s-}) \in B\}, \quad B \subset (0, \infty)^2.$$

And M is a PPP($ds \wedge (dx)$). We proceed to verify (2.s) by a change of variables from the distribution of M . Let $Y_t = D_t - G_t$ be the length of the jump interval covering t . Given $M = m$, (S_t, Y_t) is a.s. the unique point (s, y) of M such that $T_{s-} < t$ and $T_{s-} + y \geq t$, where $T_{s-} = T_{s-}(m) = \int x 1(s' < s) m(ds' dx)$. That is to say,

$$P(S_t \in ds, Y_t \in dy, M \in dm) = 1(T_{s-} < t, T_{s-} + y \geq t) m(ds dy) P(M \in dm).$$

Let M^- be the M with the point (S_t, Y_t) removed. By Lemma 2.2

$$(2.t) \quad P(S_t \in ds, Y_t \in dy, M^- \in dm) = 1(T_{s-} < t, T_{s-} + y \geq t) ds \wedge (dy) P(M \in dm),$$

where $T_{s-} = T_{s-}(m)$. Now for a fixed $s > 0$, let K_s be the point process on $(0, \infty)$ obtained by first restricting M to $(0, s) \times (0, \infty)$, then projecting onto the second coordinate and scaling via the transformation

$$(2.u) \quad (s', x) \rightarrow x/s^{1/\alpha}.$$

That is to say, for subsets C of $(0, \infty)$,

$$K_s(C) = \#\{s' : s' < s, (T_{s'} - T_{s'-})/s^{1/\alpha} \in C\}.$$

For each fixed s , the point process K_s is a PPP(Λ), with law denoted by Q . Now $T_{s-} = T_{s-}(m) = s^{1/\alpha} T(k_s)$, where k_s is the image of m under the scaling transformation (2.u), and $T(n) = \int x n(dx)$. Since N_t^- as in (2.s) is just $N_t^- = K_{S_t}$, a change of variables from (2.t) shows that

$$(2.v) \quad P(S_t \in ds, Y_t \in dy, N_t^- \in dn) = 1(s^{1/\alpha} T < t, s^{1/\alpha} T + y \geq t) ds \wedge (dy) Q(dn)$$

where $T = T(n)$. Since $X_t = A_t/S_t^{1/\alpha} = t/S_t^{1/\alpha} - T(N_t^-)$, we get

$$P(S_t \in ds, Y_t \in dy, X_t \in dx, N_t^- \in dn) = 1\left(\frac{t}{s^{1/\alpha}} - T \in dx, x \leq \frac{y}{s^{1/\alpha}}\right) ds \wedge (dy) Q(dn),$$

$$\frac{P(X_t \in dx, N_t^- \in dn)}{Q(dn)} = \iint ds \wedge (dy) 1\left(\frac{t}{s^{1/\alpha}} - T \in dx, x \leq \frac{y}{s^{1/\alpha}}\right)$$

$$= \int ds 1\left(\frac{t}{s^{1/\alpha}} - T \in dx\right) \Lambda(s^{1/\alpha} x, \infty)$$

$$\begin{aligned}
 &= |dx| \left| \frac{ds}{dx} \right| \Lambda(s^{1/\alpha}x, \infty), \quad \text{where} \quad \frac{t}{s^{1/\alpha}} - T = x, \quad s = \left(\frac{t}{x+T}\right)^\alpha \\
 &\qquad\qquad\qquad s^{1/\alpha}x = \frac{tx}{x+T}; \quad \frac{ds}{dx} = -\alpha \frac{t^\alpha}{(x+T)^{1+\alpha}} \\
 &= |dx| \alpha \frac{t^\alpha}{(x+T)^{\alpha+1}} \frac{c}{\Gamma(1-\alpha)} \left(\frac{xt}{x+T}\right)^{-\alpha}, \\
 &= \Lambda(dx) \frac{x}{x+T} \quad \text{by (2.b).} \quad \square
 \end{aligned}$$

3. Analysis at an Independent Exponential Time.

This section presents an alternative approach to Theorems 1.1 and 1.2, based on analysis of the intervals up to an independent exponential time. As in Section 2, we establish Theorems 1.1 and 1.2 as reformulated in Theorem 2.1 and (2.m). The argument brings out some remarkable general features of size-biased sampling from a Poisson point process, which are presented in Theorem 3.1.

Let T be an exponentially distributed random variable with rate 1 independent of the stable subordinator (T_s) . Let N_T be the point process whose points in descending order form the normalized sequence of ranked interval lengths $V_T/S_T^{1/\alpha}$. Let $X_T = A_T/S_T^{1/\alpha}$, which is one of the points of N_T . Since it is clear a priori by scaling invariance that N_T is independent of T , it suffices to show that

(3.a) N_T is a PPP(Λ);

(3.b) given T and N_T , the point X_T is picked from N_T by size-biased sampling.

For the rest of the argument, we write simply S for S_T , and v for $1/\alpha$. Each length in the sequence V_T represents either the length of one of the infinite number of jumps of the subordinator occurring at some local time strictly before S , or the age A_T of the excursion in progress at time T . The counting process N_T associated with V_T/S^v is therefore the sum of two processes:

(3.c) the process N_T^- counting the lengths of jumps of the subordinator strictly before local time S , with each jump normalized by S^v ;

(3.d) the process counting a single point at $X_T = A_T/S^v$.

It will now be argued that these two counting processes are conditionally independent given S , with just the right conditional distributions given S to make (3.a) and (3.b) hold.

Think of T as the time of the first mark in a Poisson process of marks with rate 1 on the real time scale, independent of the subordinator. Associate with each jump of the subordinator the real times of any marks that fall in that jump interval, measured from the start of the jump (i.e. the beginning of the corresponding excursion). A big Poisson process, homogeneous in the local time coordinate, is then obtained by counting the local times of jumps of the subordinator, jointly with jump lengths embellished by their mark times if any. And S is the local time of the first marked jump in this big Poisson process. This idea has been used in a number of contexts. See e.g. Greenwood and Pitman (1980), Rogers (1981), and Rogers-Williams (1987, section VI.53) for a careful formulation. The independent decomposition of this Poisson point process into points with marks and points without implies the following formulae, where

we assume for simplicity (and without loss of generality) that the constant c in (2.a) is $c = 1$:

(3.f) S is exponential with rate $\int (1 - e^{-x}) \Lambda(dx) = 1$

(3.g) the local times and lengths of the unmarked jumps appear according to a homogeneous Poisson point process with intensity measure $e^{-x} \Lambda(dx)$ per unit local time, independently of S ;

(3.h) independently of S and all unmarked jumps, the age A_T , which is the time of the first mark measured from the beginning in the first marked jump interval, has the gamma $(1 - \alpha, 1)$ distribution:

$$P(A_T \in dx) = e^{-x} dx \Lambda(x, \infty) = \frac{dx}{\Gamma(1 - \alpha)} e^{-x} x^{-\alpha}, \quad x > 0.$$

Now condition on $S = s$, and normalize all the lengths by s^ν , where $\nu = 1/\alpha$. From (3.g) by the linear change of variable $x = y/s^\nu$, the process of normalized completed jumps is Poisson with mean measure

$$(3.i) \quad s^\nu \exp(-s^\nu x) \frac{\alpha}{\Gamma(1 - \alpha)} \frac{s dx}{(xs^\nu)^{\alpha+1}} = \exp(-s^\nu x) \Lambda(dx), \quad x > 0.$$

And independently of this Poisson counting process given $S = s$, the normalized age $X_T = A_T / s^\nu$ has gamma $(1 - \alpha, s^\nu)$ distribution: for $x > 0$,

$$(3.j) \mathcal{P}(X_T \in dx \mid S = s) = \frac{dx}{\Gamma(1 - \alpha)} \exp(-s^\nu x) x^{-\alpha} (s^\nu)^{1-\alpha} = \nu s^{\nu-1} x \exp(-s^\nu x) \Lambda(dx).$$

Now (2.a), (3.f), (3.i) and (3.j) show that conditions (d) (e) and (f) of the following theorem are satisfied with

$$N = N_T, \quad Y = S^\nu, \quad X = X_T, \quad h(x) = x, \quad \mu = \Lambda,$$

The desired conclusions (3.a) and (3.b) then follow at once from conditions (a) and (c) of the theorem.

Theorem 3.1. *Let $(\Omega^*, \mathcal{F}^*)$ be the space of counting measures on a measurable space (S, \mathcal{S}) , let μ be a σ -finite measure on S . Let h be a non-negative measurable function defined on S such that*

$$(3.k) \quad \text{if } N \text{ is a PPP}(\mu) \text{ then } P(0 < Nh < \infty) = 1,$$

where $Nh = \int h dN$. Suppose Y is a non-negative random variable, X a random point in S , and N^- a point process on S , all defined on some basic probability space (Ω, \mathcal{F}, P) . Let $N = N^- + \delta_X$ be the point process obtained by addition to N^- of a single extra point at X . The following statements (a), (b) and (c) combined are equivalent to (d), (e) and (f) combined:

$$\left\{ \begin{array}{l} \text{(a) } N \text{ is a PPP}(\mu); \\ \text{(b) } Y = T / Nh \text{ where } T \text{ has standard exponential distribution independently of } N; \\ \text{(c) given } N \text{ and } Y, X \text{ is picked from } N \text{ by } h\text{-biased sampling.} \end{array} \right.$$

$$\left\{ \begin{array}{l} \text{(d) } P(Y > y) = \exp[-\int (1 - e^{-h(x)y}) \mu(dx)]; \\ \text{(e) given } Y = y, X \text{ and } N^- \text{ are conditionally independent, with} \\ \quad P(X \in dx | Y = y) = q(y)^{-1} h(x) e^{-h(x)y} \mu(dx) \\ \quad \text{for a constant of integration } q(y); \\ \text{(f) given } Y = y, N^- \text{ is a PPP}[e^{-h(x)y} \mu(dx)]. \end{array} \right.$$

Proof. Since the joint law of either of the triples (N, X, Y) or (N^-, X, Y) determines that of the other, it suffices to establish that (d), (e) and (f) hold for any particular triple (N, X, Y) satisfying (a), (b) and (c). This is conveniently done by supposing that N is of the form

$$N(B) = \sum_i 1(X_i \in B)$$

for some sequence of S -valued random variables (X_i) , and that given (X_i) there is a sequence of independent variables (Y_i) such that Y_i is exponential with rate $h(X_i)$. Let $Y = \min_i Y_i = Y_I$ say, and let $X = X_I$. Then (a) holds by construction, and (b) and (c) follow from standard properties of independent exponential variables. It is well known that if the points of a Poisson process on S are assigned marks independently according to a transition probability function from S to some other space, the result is a Poisson process on the product space. Thus the process N^* on $S \times (0, \infty)$ defined by

$$N^*(\cdot) = \sum_i 1[(X_i, Y_i) \in \cdot]$$

is a PPP $[\mu(dx) q(x, y) dy]$, where $q(x, y) = h(x) e^{-h(x)y}$ is the conditional density of Y_i given $X_i = x$. Now (X_I, Y_I) is picked from N^* by the deterministic rule of taking the point with minimum y -value, and N^- is the projection of $N^* - \delta_{(X_I, Y_I)}$ onto the x -axis.

So (d), (e) and (f) follow immediately from standard Poisson features of N^* . \square

We continue this section with a series of remarks concerning Theorem 3.1.

Remark 3.2 Assuming that (a) holds, the right side of (e) may be recognized as the Laplace transform of Nh :

$$(3.m) \quad P(Y > y) = E \exp[-y Nh] = \exp[-F(y)],$$

say. Then the constant of integration $q(y)$ in (d) turns out to be simply the derivative of F at y :

$$(3.n) \quad q(y) = F'(y).$$

Remark 3.3 Hidden in Theorem 3.1 is the following consequence of Lemma 2.2, noted already in a special case in (2.s):

if N^- is obtained by deleting a point X from N , a PPP(μ), by h -biased sampling, then

$$(3.p) \quad P(X \in dx, N^- \in dn) = \frac{h(x)}{nh + h(x)} \mu(dx) Q(dn),$$

where Q is the P distribution of N . Suppose Y is constructed as in condition (b) of the theorem, so $Y = T/Nh$ for a standard exponential variable T independent of N and X . Then (a), (b) and (c) hold, and it is interesting to see how (3.p) emerges from (d) (e) and (f). Let Q_y denote the law of a PPP (μ_y) where $\mu_y(dx) = e^{-h(x)y} \mu(dx)$. So Q_y is the law of N^- given $Y = y$ as prescribed in (f). It is easy to check that

$$Q_y(dn) = \frac{e^{-(nh)y}}{\phi(y)} Q(dn) \quad \text{where} \quad nh = \int n(dx) h(x), \text{ and}$$

$$\phi(y) \stackrel{\text{def}}{=} \int Q(dn) e^{-(nh)y} = P(Y > y)$$

by (3.m). Now (3.m) and (3.n) make

$$P(Y \in dy) = -\phi'(y) dy = \phi(y) q(y) dy,$$

So conditions (d), (e) and (f) of Theorem 3.1 imply

$$\begin{aligned} P(X \in dx, N^- \in dn) &= \int P(X \in dx, N^- \in dn \mid Y = y) P(Y \in dy) \\ &= \left[\int q(y)^{-1} h(x) e^{-h(x)y} \frac{e^{-(nh)y}}{\phi(y)} \phi(y) q(y) dy \right] Q(dn) \mu(dx) \end{aligned}$$

which reduces to (3.p).

Remark 3.4. Suppose again that N is a PRM (μ) and X is picked from N by h -biased sampling. According to Theorem 3.1, provided that the space (Ω, \mathcal{F}, P) on which N and X are defined allows room for a continuously distributed random variable T independent of both N and X , a random variable Y can be created from N and T such

that conditionally on Y , N^- is a Poisson process independent of X . In particular, this implies that the law of N^- is a mixture of Poisson laws, a fact that does not seem obvious a priori, and is even well hidden in formula (3.p). In common terminology, N^- is a *Cox process*.

Remark 3.5 The part of the assumption in (3.k) that $Nh > 0$ a.s. rules out the case when μ is a finite measure. But this assumption was made only for simplicity of exposition. It is clear from the proof that the result still holds without assuming $Nh > 0$ a.s. provided the various conditions are modified to allow the possibility of the event $(Y = \infty)$, in which case X is undefined and N^- is identically zero. In particular, in case S is a single point, μ and h are identified with positive real numbers, and N with a non-negative integer valued random variable, Theorem 3.1 reduces to the following elementary fact: *N has Poisson(μ) distribution on $\{0, 1, 2, \dots\}$ iff the distribution of $(N - 1)$ restricted to the event $N \geq 1$ is the mixture over λ of Poisson (λ) distributions with respect to $d\lambda e^{\mu-\lambda} 1(\lambda \leq \mu)$.* (This fact is obvious by conditioning on the time of the last arrival before time μ in a Poisson process with rate 1 on $(0, \infty)$.)

Remark 3.6. Say (U, V, W) is Markov if U and W are conditionally independent given V . A curious feature of the joint law described in Theorem 3.1 is that (Y, N, X) is Markov, while $(N - \delta_X, Y, X)$ is also Markov. A third equivalent description of the Markov triple (Y, N, X) is obtained by combining (d) (the law of Y) and (c) (the law of X given Y and N) with the following formula for the law of N given Y :

$$(3.q) \quad P(N \in dn \mid Y = y) = c(y)^{-1} (nh) e^{-y(nh)} Q(dn)$$

where Q is the law of a PPP(μ), and $c(y)$ is a constant of integration which can be calculated variously as

$$c(y) = \int (nh) e^{-y(nh)} Q(dn) = -\phi'(y) = \phi(y) q(y) = P(Y \in dy) / dy.$$

This follows easily from (e) and the following proposition applied conditionally given $Y = y$, with $e^{-h(x)y} \mu(dx)$ instead of $\mu(dx)$, and $h(x)/q(y)$ instead of $h(x)$.

Proposition 3.7. Suppose under P that N^- is a PPP (μ) with distribution Q , and $N = N^- + \delta_X$ where X is independent of N^- with law

$$P(X \in dx) = h(x) \mu(dx)$$

for some density h . Then N has law

$$P(N \in dn) = (nh) Q(dn),$$

and X is picked from N by h -biased sampling.

Proof. Apply the Poisson sampling Lemma 2.2. \square

To conclude this section we mention two natural extensions of Theorem 3.1 which can be obtained by minor variations of the same argument. One is a description of the joint distribution of N and $m \geq 1$ points of N picked by repeated h -biased sampling, which is given by Perman, Pitman and Yor (1991). The other, spelled out as Theorem 3.2 below, is obtained by using a more general transition density $q(x, y)$, $x \in S$, $y \geq 0$, instead of $q(x, y) = h(x)e^{-h(x)y}$. We do not know of any interesting applications of this result, but it seems nonetheless worth recording:

Theorem 3.2 *Let $q(x, y)$, $x \in S$, $y \geq 0$, be a transition probability density from S to $(0, \infty)$. Let*

$$F(x, y) = \int_0^y q(x, u) du$$

$$G(x, y) = 1 - F(x, y) = \int_y^\infty q(x, u) du.$$

For $y > 0$ define

$$q(y) = \int q(x, y) \mu(dx)$$

$$F(y) = \int_0^y q(u) du = \int F(x, y) \mu(dx).$$

Assume $F(y) < \infty$ for all $y > 0$. Suppose defined on (Ω, \mathcal{F}, P) a triple (N, X, Y) such that

- (a) N is a PRM (μ) ;
- (b)
$$P(Y > y | N) = \prod_i G(X_i, Y)$$

$$= \exp \int \log G(x, y) N(dx)$$
- (c)
$$P(X \in dx | N, Y = y) = \frac{h(x, y) N(dx)}{\int h(x, y) N(dx)}$$

where $h(x, y) = q(x, y)/G(x, y)$ is the hazard rate at y associated with the density $q(x, \cdot)$. Let $N^- = N - \delta_X$. Then

- (d)
$$P(Y > y) = \exp(-F(y))$$
- (e) Given $Y = y$, X and N^- are conditionally independent, with

$$P(X \in dx | Y = y) = \frac{q(x, y)}{q(y)} \mu(dx),$$

- (f) Given $Y = y$, N^- is a PRM $G(x, y) \mu(dx)$.

Conversely, if (N^-, X, Y) satisfy (d), (e) and (f), $N = N^- + \delta_X$, then (a), (b), (c) hold for (N, X, Y)

Notes. (i) taking expectations in (b) gives

$$(d') \quad P(Y > y) = P \exp - \int \log(G(x, y)) N(dx)$$

which can be used as a substitute for (d).

(ii) Integrating out y gives the formula

$$P(X \in dx, N^- \in dn) = \mu(dx) P(dn) \int q(x, y) \exp \left[\int \log G(x', y) n(dx') \right] dy$$

4. Application to Occupation times.

Barlow Pitman and Yor (1989b) (= BPYb) found the joint distribution of the occupation times in sectors of Walsh's singular diffusion in the plane. This result, reformulated here slightly more generally as Theorem 4.1, is an extremely close relative of Theorem 1.1. Indeed, it was Theorem 4.1 that first led us to Theorem 1.1, and either result can easily be derived from the other.

Walsh (1978) defined a diffusion process X in the plane as follows. Assume X starts at the origin, that the radial part of X is a reflecting Brownian motion on $[0, \infty)$, and that each excursion of the radial part away from zero corresponds to an excursion of X within a ray emanating from the origin at an angle chosen independently according to some given distribution F on $[0, 2\pi)$. See BPYa for a more careful description of this process. In particular, in case F puts probability p on angle 0 and probability $1 - p$ on angle π , the process X can be identified as a *skew Brownian motion* such that

$$(4.a) \quad P(X_t > 0) = p ; P(X_t < 0) = 1 - p.$$

Skew Brownian motion, introduced by Itô and McKean (1965), has been studied by Walsh (1978), Harrison and Shepp (1981) and Brooks and Chacon (1983). It appears naturally in certain limit theorems for real diffusions considered by Rosenkrantz (1975), Le Gall (1984) and Franchi (1989). BPYb considered also the process X in the plane defined as above, but whose radial part is a Bessel process of dimension $d \in (0, 2)$ instead of a reflecting Brownian motion. As pointed out by Molchanov and Ostrowski (1969), (see also BPYb), the inverse local time at zero of this Bessel process is stable subordinator with index $\alpha = 1 - d/2$.

Walsh's construction involves a random labelling of the jump intervals of the subordinator, which we can describe as follows. Suppose that every jump interval $I = (T_{s-}, T_s)$ of the subordinator (T_s) is labelled by a random variable Θ_I . In Walsh's construction, the Θ_I are understood to represent angles, but any measurable space of labels is now allowed. Assume these Θ_I are mutually independent as I varies, independent also of the lengths of all intervals, and that the Θ_I have some common distribution F . To be pedantic, this means the following: for some exhaustive listing $(I(n), n = 1, 2, \dots)$ of the maximal open intervals of Z^c , where the definition of the endpoints of $I(n)$ may depend measurably on the subordinator (T_s) whose closed range is Z , random variables $\Theta_{I(n)}$ are independent and identically distributed according to F , independently also of (T_s) . This defines Θ_I for every jump interval I . And for any other such (T_s) -dependent listing of the intervals, exhaustive or not, the sequence of Θ values will have the same probabilistic properties.

Define a process $(\Theta_t, t \in Z^c)$ by

$$\Theta_t = \Theta_I \text{ if } t \in I,$$

and consider for a fixed or random time T the random occupation measure $\Gamma(\cdot, T)$ defined on measurable subsets D of Θ values by

$$(4.b) \quad \Gamma(D, T) = \int_0^T 1_D(\Theta_s) ds.$$

Put another way,

$$(4.c) \quad \Gamma(D, T) = \sum_n V_n(T) 1[\Theta_n(T) \in D]$$

where $\Theta_n(T)$ is the label of the n th longest subinterval of $Z^c \cap (0, T)$, whose length is $V_n(T)$. Provided T depends only on the subordinator, these labels are independent with common distribution F .

Theorem 4.1. (BPYb) *For disjoint measurable sets of labels $D_i, i = 1, \dots, n$, and $s, t > 0$, the random variables*

$$(4.d) \quad \frac{\Gamma(D_i, t)}{S(t)^{1/\alpha}}, \quad i = 1, \dots, n,$$

are independent with the same joint distribution as

$$(4.e) \quad \frac{\Gamma(D_i, T_s)}{s^{1/\alpha}}, \quad i = 1, \dots, n,$$

Here $\Gamma(D_i, T_s)/s^{1/\alpha}$ has the same stable distribution with index α as $T(F(D_i))$. This, and the independence of $\Gamma(D_i, T_s)$ as i varies, are immediate consequences of the Poisson character of the counting process on $[0, \infty) \times \text{labels}$ which counts the local times and labels associated with jump intervals of the subordinator. As an immediate consequence of Theorem 4.1, *the random variables*

$$(4.f) \quad \frac{\Gamma(D_i, t)}{t}, \quad i = 1, \dots, n; \quad \frac{S(t)}{t^{1/\alpha}}$$

have the same joint distribution as

$$(4.g) \quad \frac{\Gamma(D_i, T_s)}{T_s}, \quad i = 1, \dots, n; \quad \frac{s}{T_s^{1/\alpha}}$$

See BPYb for further description of this joint law in case $\alpha = 1/2$.

Derivation of Theorem 4.1 from Theorem 2.1. When assigned independent labels according to distribution F , the points of the Poisson process N_t in Theorem 2.1 that are assigned labels in D_i form independent Poisson processes with mean measures

$F(D_i)\Lambda$, $i = 1, \dots, n$. Since adding the values of points with labels D_i gives $\frac{\Gamma(D_i, t)}{S(t)^{1/\alpha}}$, these must be independent stable (α) random variables, as asserted by Theorem 4.1. \square

As Theorem 4.1 was proved in BPYb by a different method (computation of the joint Laplace transform via excursion theory), we offer also the following argument:

Derivation of Theorem 2.1 from Theorem 4.1. Take the distribution of labels to be uniform on $[0, 1]$. Fix $t > 0$. Theorem 4.1 implies that

$$(4.h) \quad \left(\frac{\Gamma([0, a], t)}{S(t)^{1/\alpha}}, 0 \leq a \leq 1 \right)$$

is a process with stationary independent increments, identical in law to $(T_a, 0 \leq a \leq 1)$. So the counting process M_t defined on subsets B of $(0, \infty)$ by

$$(4.i) \quad M_t(B) = \# \left\{ a: \frac{\Gamma(\{a\}, t)}{S(t)^{1/\alpha}} \in B \right\}$$

is Poisson with mean measure Λ . This process counts the sizes of the atoms of the random measure $\Gamma(\cdot, t)/S(t)^{1/\alpha}$, which for each n puts an atom of size $V_n(t)/S(t)^{1/\alpha}$ at a label chosen uniformly at random from $[0, 1]$, independently as n varies. Since there is probability zero that any two labels are equal, the counting process N_t in Theorem 2.1 and M_t in (4.i) are identical almost surely. \square

In view of the argument preceding Theorem 2.1, it is now plain that Theorems 1.1, 2.1 and 4.1 are equivalent in the sense that any one of these results can easily be derived from any of the others. But to build up any of these results from scratch seems to require some non-trivial calculation.

This section would not be complete without mentioning the analog in this setting of Theorem 1.2. Recall from (4.c) that $\Theta_n(t)$ is the label of the n th longest subinterval of $Z^c \cap (0, t)$, whose length is $V_n(t)$.

Theorem 4.2. *Let G_t be the σ -field generated by the random variables $\Theta_n(t)$ and $V_n(t)$ for $n = 1, 2, \dots$. Then*

$$(4.j) \quad P(\Theta_t \in \cdot \mid G_t) = \Gamma(\cdot, t)/t.$$

Proof. Due to the independence of labels and interval lengths, Theorem 1.2 implies that given $\Theta_n(t)$ and $V_n(t)$ for $n = 1, 2, \dots$, $\Theta_t = \Theta_n(t)$ with probability $V_n(t)$. So the conclusion is immediate from formula (4.c). \square

On the way to his arcsin law (1.a) for Brownian motion, Lévy (1939, formula (51)), found that

$$(4.k) \quad P(\Gamma_+(t)/t \in du \mid B_t > 0) = \frac{2}{\pi} \sqrt{\frac{u}{1-u}} du, \quad 0 < u < 1.$$

In combination with the arcsin law (1.a), this amounts to

$$(4.m) \quad P(B_t > 0 \mid \Gamma_+(t)/t = u) = u \quad 0 < u < 1,$$

which is a special case of Theorem 4.2. According to Theorem 4.2, formula (4.m) must hold also in the case of a skew Brownian motion B , regardless of what probability p there is for positive excursions, and even for a "skew Bessel process".

5. Results in Terms of Bridges

Let P_{0t} be the law of a *bridge of length t* , obtained as a conditional distribution given $t \in Z$, where Z is the closed range of a subordinator $(T_s, s \geq 0)$. If we assume for simplicity that Z is the zero set of a Brownian motion Bessel process $(B_t, t \geq 0)$ set up on the canonical path space,

$$P_{0t}(\cdot) = P(\cdot \mid B_t = 0)$$

governs a Brownian or Bessel bridge of length t . And there are nice versions of conditional distributions so the basic *switching identity*

$$(5.a) \quad P(\cdot \mid T_s = t) = P_{0t}(\cdot \mid S_t = s)$$

holds exactly for every $s > 0$ and $t > 0$. In the Brownian case this is the basis of Lévy's argument (1.f). See also Karatzas-Shreve (1987b, Section 4) Kallenberg (1981 Lemma 4.1). The switching identity allows us to reformulate Theorem 1.1 in terms of bridges:

Theorem 5.1. *For a stable subordinator (T_s) , the conditional distribution of $V(t)$ given S_t is the same for the bridge distribution P_{0t} as for the original distribution P .*

Proof of Theorem 5.1. Compute as follows:

$$\begin{aligned} P_{0t} \left[\frac{V(t)}{t} \in \cdot \mid S_t = s \right] &= P \left[\frac{V(T_s)}{T_s} \in \cdot \mid T_s = t \right] \text{ by the switching identity (5.a)} \\ &= P \left[\frac{V(T_s)}{T_s} \in \cdot \mid \frac{s}{T_s^\alpha} = \frac{s}{t^\alpha} \right] \\ &= P \left[\frac{V(t)}{t} \in \cdot \mid \frac{S_t}{t^\alpha} = \frac{s}{t^\alpha} \right] \text{ by (2.r)} \\ &= P \left[\frac{V(t)}{t} \in \cdot \mid S_t = s \right]. \quad \square \end{aligned}$$

As we give a different approach to an extension of Theorem 5.1 in Section 7, we record the following argument, which shows that Theorem 5.1 joins Theorems 1.1, 2.1 and 4.1 in a circle of "equivalent" results, any one of which can be derived quickly from any of the others:

Derivation of Theorem 1.1 from Theorem 5.1. Assuming the conclusion of Theorem 5.1, the quantities in the first and last lines of the preceding proof must be equal, hence so too must the second and third. Take $s/t^\alpha = x$ say, and use scaling, to deduce that

$$P \left[\frac{V(T_s)}{T_s} \in \cdot \mid \frac{s}{T_s^\alpha} = x \right] = P(V(1) \in \cdot \mid S_1 = x),$$

But as remarked in (2.c), $\frac{s}{T_s^\alpha} \stackrel{d}{=} S_1$, hence the conclusion of Theorem 1.1 for $t = 1$, then any t by scaling. \square

The present discussion is closely related to the work of Biane, Le Gall and Yor (1987) (abbreviated BLY), who considered the *pseudo-bridge*

$$(5.b) \quad B_u^* = \frac{1}{\sqrt{T_1}} B(u T_1), \quad 0 \leq u \leq 1,$$

in case B is Brownian motion and (T_s) its inverse local time at zero. Here is a straightforward generalization of their result:

Theorem 5.2. (due to BLY in the Brownian case $\delta = 1$).

Let (T_s) be the subordinator inverse of the local time at zero (S_t) of (B_t) , a Brownian motion, or a Bessel process of dimension δ , for $0 < \delta < 2$. Then the law P_{01}^ of the pseudo-bridge (5.b) is mutually absolutely continuous with respect to the law P_{01} of the corresponding bridge of length 1 with density*

$$(5.c) \quad \frac{dP_{01}^*}{dP_{01}} = \frac{1}{c \alpha \Gamma(\alpha) S_1}.$$

where $\alpha = 1 - \delta/2$ and c is the scale constant determined by the normalization of the local time, as in (2.a) and (2.b).

Proof. According to Molchanov and Ostrovski (1969) (or see BPYb), the subordinator (T_s) derived from a Bessel process of dimension δ is stable with index $\alpha = 1 - \delta/2$. And the BLY computation in the Brownian case $\alpha = 1/2$ extends easily to a general $0 < \alpha < 1$. \square

Theorem 5.3. *Let G_t be the σ -field generated by $V(t)$. Then P_{0t} and P are mutually absolutely continuous, with*

$$\frac{dP_{0t}}{dP} \Big|_{G_t} = c \alpha \Gamma(\alpha) \frac{S_t}{t^\alpha}.$$

Proof. By scaling, it suffices to consider the case $t = 1$. By construction of the pseudo-bridge, $V(T_1)/T_1$ is the vector of interval lengths generated by the pseudo-bridge B^* . The identity $V(T_1)/T_1 \stackrel{d}{=} V(1)$ of Theorem 1.1 therefore amounts to

$$(5.e) \quad P_{01}^* = P \text{ on } G_1,$$

so the result follows immediately from Theorem 5.2. \square

It follows at once from Theorem 5.3 and the factorization criterion for sufficient statistics that P_{0_t} and P must have the same conditional distribution for $V(t)$ given S_t . Thus Theorem 5.3 provides a strengthening of Theorem 5.1. Equally, Theorem 5.3 can be derived from Theorem 5.1 without reference to the pseudo-bridge by simply by checking that the formula $S_t / E(S_t)$ gives the correct one-dimensional density on $\sigma(S_t)$. For a way to do this, see formula (7.b).

In the setting of Section 4, where intervals are assigned independent and identically distributed labels, it is immediate from the independence of the subordinator and the labelling process that Theorem 5.3 still holds with G_t enlarged to include the labels generated up to time t , as in Theorem 4.2. And Theorem 5.1 can be similarly extended. In the Brownian case, this shows for example that for $\Gamma_+(t)$ the time spent positive up to t ,

$$(5.f) \quad P(\Gamma_+(t) \in \cdot \mid S_t = s, B_t = 0) = P(\Gamma_+(t) \in \cdot \mid S_t = s)$$

Lévy hints at this identity in the paragraph mentioned below (1.b), and he could have computed either conditional distribution by his methods. Explicit formulae for joint densities of occupation and local times of Brownian motion can be found in Perkins (1982 Theorem 10), and Karatzas and Shreve (1987a Section 6.3).

6. Rearrangements. Our concern in this section is how to describe precisely the relation between what happens up to fixed times t and up to inverse local times T_s if we keep track of the whole zero set, or what is the same information, the order in which the intervals in $V(t)$ or $V(T_s)$ appear as they are laid down by the process. The whole discussion can be conducted for a Bessel process, or just for a stable subordinator. But to be definite we shall assume for simplicity of exposition that we are dealing with the Brownian case $\alpha = 1/2$.

Consider various laws on $C[0, 1]$ with coordinates $(X_t, 0 \leq t \leq 1)$ and $S_1 =$ Brownian local time at 0 up to time 1.

Q = Law of standard Brownian motion started at 0

$Q^s = Q(\cdot | S_1 = s)$

Q_0 = Law of standard Brownian bridge = $Q(\cdot | X_1 = 0)$

$Q_0^s = Q_0(\cdot | S_1 = s) = Q(\cdot | X_1 = 0, S_1 = s)$.

Also, for a distribution μ on $(0, \infty)$, let

$$Q^\mu = \int Q^s \mu(ds),$$

and define Q_0^μ similarly in terms of the Q_0^s . Let

$$v(ds) = \frac{2}{\sqrt{2\pi}} e^{-s^2/2} ds \quad (\text{the } Q \text{ distribution of } S_1)$$

$$v_0(ds) = s e^{-s^2/2} ds \quad (\text{the } Q_0 \text{ distribution of } S_1)$$

Then $Q^v = Q$, $Q_0^{v_0} = Q_0$. According to Theorem 5.2, Q_0^v is the law of the BLY pseudo-bridge (5.b).

On $C[0, 1]$ define $V = (V_1, V_2, \dots)$ to be the vector of interval lengths. Under any of the above laws,

$$V_1 > V_2 > \dots \text{ and } \sum_n V_n = 1 \text{ a.s.}$$

Then we have the following basic result:

Proposition 6.0. *For every probability distribution μ on $(0, \infty)$, the distribution of V is the same under Q^μ as under Q_0^μ .*

Proof. The result for $\mu = \delta_s$ is Theorem 5.1, and the result for general μ follows by mixing on s . \square

Let (L_n, R_n) be the interval which contributes length $V_n = R_n - L_n$. So

$$(6.a) \quad \{t : 0 < t < 1, X_t \neq 0\} = \bigcup_n (L_n, R_n).$$

The stochastic structure of the zero set of X is thus defined by the joint distribution of any two of the random sequences L , R and V . It seems most convenient to consider V and R . Inasmuch as the law of V is the same under both Q^μ and Q_0^μ , whereas the structure of the whole zero set is obviously different for Q^μ and Q_0^μ , the difference lies in the conditional distribution of R given V .

The law of R given V is simplest to describe under Q_0^μ , and is the same no matter what the distribution of local time μ . Informally, for any n , the intervals of length V_1, \dots, V_n are equally likely to appear in any of $n!$ possible orderings. These orderings have an obvious consistency property as n varies. And once these orderings are determined, the value of R_n can be calculated as the sum of V_i over all i such that V_i is put before V_n .

This can be expressed more neatly by introducing the random variables on $C[0, 1]$

$$(6.b) \quad U_n = S(R_n)/S(1)$$

where $(S(t), t \geq 0)$ is the local time process. Here $U_n < U_m$ means the interval of length V_n is placed to the left of the interval of length R_m . Equivalently: $R_n < R_m$. And

$$(6.c) \quad R_m = \sum_n V_n 1(U_n \leq U_m), \quad n = 1, 2, \dots$$

This formula makes sense and holds both Q a.s. and Q_0^μ a.s. for any μ . In view of this formula, to specify the joint law of V and R under any of these probabilities, it suffices to specify the joint law of V and $U = (U_1, U_2, \dots)$.

Proposition 6.1. *Under Q_0^μ , for any μ ,*

$$U_1, U_2, \dots$$

are i.i.d. uniform $[0, 1]$, independent of V .

Proof. First, it is enough to consider the case $\mu = \delta_s$ for $0 < s < \infty$. But since $Q_0^s(\cdot) = Q_0^s(\cdot | S_1 = s)$, and S_1 is a function of V , the result is immediate from the presentation of Q_0^s as the law of the BLY pseudo-bridge (5.b). and standard Itô excursion theory. (Alternatively: note that under Q_0^s the process $(T_w, 0 \leq w \leq s)$ is a subordinator conditioned on $T_s = 1$, which is a process with exchangeable increments. And $U_n = S_n/s$, where S_n is the local time at which the n th largest jump of this process

occurs. Now use the canonical representation of processes with exchangeable increments due to Kallenberg (1973).

The corresponding situation under Q is only a little more complicated. Under Q there is almost surely a last interval

$$(G_1, 1) = (L_N, R_N)$$

for a random index N . This N is revealed in the U sequence as the unique n : $U_n = 1$. And according to Theorem 1.2,

$$Q(N = n | V) = V_n, \quad n = 1, 2, \dots$$

Now to be given V and N is the same as to be given $V(G_1)$ and G_1 . And we know that under Q the process $(X_t, 0 \leq t \leq G_1)$ given G_1 is just a Brownian bridge of length G_1 . So essentially we are being given G_1 and the interval lengths for this bridge which come from the sequence V by deleting the term V_N . By the previous proposition for bridges we therefore obtain the following, first for $\mu = v$, the Q distribution of S_1 , so $Q^v = Q$, then for general μ by first conditioning on S_1 , then mixing:

Proposition 6.2. *Under Q^μ , for any μ , there is a unique random index N such that $U_N = 1$;*

$$Q^\mu(N = n | V) = V_n, \quad n = 1, 2, \dots,$$

and the U sequence with the N th term deleted,

$$U_1, U_2, \dots, U_{N-1}, U_{N+1}, U_{N+2}, \dots$$

is i.i.d. uniform $[0, 1]$, independent of V and N .

Informally, Propositions 6.0, 6.1 and 6.2 tell us how the intervals forming

$$\{t : 0 < t < 1, X_t \neq 0\}$$

are created under Q^μ and Q_0^μ . According to Proposition 6.0, the ranked lengths are created with the same joint distribution in the two cases. According to Proposition 6.1, in the case of Q_0^μ these ranked lengths appear in a *purely random order*, the longest equally likely to be before or after the second longest, the 1st and 2nd and 3rd longest equally likely to be in any of the $3!$ possible orders, and so on. (The formula for R in terms of V and the relative local times U , and the iid property of U , is just a slick formalization of this idea).

According to Proposition 6.2, in case of Q^μ the only difference is that one of the lengths in V is picked out by length biased sampling to be the last interval $(G_1, 1)$, then the remaining intervals are laid down to the left of $(G_1, 1)$, in purely random order,

relative to each other.

Comparison of these two prescriptions suggests a natural way of creating the interval structure under Q^μ from that under Q_0^μ by a simple random shuffling scheme:

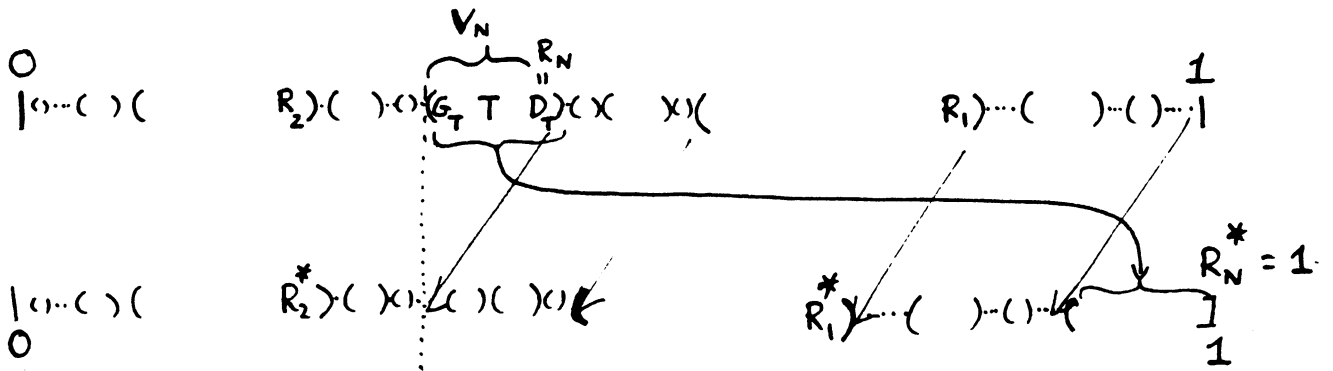
On a suitable probability space (Ω, \mathcal{F}, P) , let X have distribution Q_0^μ , and let T be a uniform $[0, 1]$ r.v. indpt of X . Let (G_T, D_T) denote the zero free interval of X that covers T , so

$$(G_T, D_T) = (L_N, R_N)$$

where N is the index at which $D_T - G_T$ appears in the terms of $V = V(X)$. Then by construction

$$P(N = n | V) = V_n, \quad n = 1, 2, \dots$$

Now from the interval structure (V, R) of X create a new interval structure (V, R^*) with the same lengths as follows: exchange the interval (G_T, D_T) with the pattern of intervals from D_T to 1, as in the following figure:



Thus

$$\begin{aligned} R_n^* &= R_n \text{ if } R_n < R_N \text{ (i.e. } U_n < U_N) \\ &= 1 \text{ if } R_n = R_N \text{ (i.e. } n = N, U_n = U_N) \\ &= R_n^* - V_N \text{ if } R_n > R_N \text{ (i.e. } U_n > U_N). \end{aligned}$$

These are the right ends of the zero-free intervals defined by the process

$$\begin{aligned} X^*(t) &= X(t), \quad 0 \leq t < G_T \\ &= X(D_T + t - G_T), \quad G_T \leq t < 1 - (D_T - G_T) \\ &= X(G_T + t - [1 - (D_T - G_T)]), \quad 1 - (D_T - G_T) \leq t \leq 1. \end{aligned}$$

Clearly, $V(X^*) = V(X)$. Now it is clear that the variables U_1, U_2, \dots defined earlier on the canonical space via local times can be composed with X^* to create U_1^*, U_2^*, \dots , by using e.g. the formula for local time in terms of number of intervals $> \varepsilon$. And it is

clear by construction that the P joint distribution of $V(X^*)$ and $U(X^*)$ is identical to the Q^μ joint distribution of V and U . Consequently, the structure of intervals forming $\{t : X_t \neq 0\}$ under Q^μ is identical to the structure of intervals forming $\{t : X_t^* \neq 0\}$ under P , where X^* is the random rearrangement of X with distribution Q_0^μ , as above. To lift this identity to the level of processes, introduce the Brownian scaling notation

$$X_{[a,b]}(u) = \frac{1}{\sqrt{b-a}} X_{a+u(b-a)}, 0 \leq u \leq 1,$$

for any process X and random times $a < b$. Let SSE and SSM stand for standard signed excursion and signed meander on $[0,1]$. Let Z be the σ -field generated by the zero set, meaning

$$Z = \sigma(V, R) = \sigma(L, R) = \sigma(V, U)$$

(where the last of these equalities is true only Q^μ and Q_0^μ a.s. for any μ , because the definition of U involves local times). Standard excursion theory tells us that for any μ :

Under Q_0^μ given Z , the processes $X_n(\cdot) = X_{[L_n, R_n]}(\cdot)$, $n = 1, 2, \dots$ are i.i.d. SSE's.

Under Q_μ given Z , if N denotes the index with $R_N = 1$,

$$X_N(\cdot) = X_{[G_1, 1]}(\cdot)$$

is a SSM, and independently the sequence of processes $X_n(\cdot)$, $n = 1, 2, \dots$, with the N th term deleted, is an i.i.d. sequence of SSE's.

Also, because under Q_0^μ or Q^μ the whole process X can be reconstructed in a measurable manner a.s. from (V, R) and the sequence of standardised fragments $X_1(\cdot), X_2(\cdot), \dots$, to show that a process X has one of these laws it suffices to check that X has continuous paths and the right laws for (V, R) and the $X_n(\cdot)$, are described above. We therefore obtain the following conclusion:

Theorem 6.3. *Let X have distribution Q_0^μ , let T independent of X be $U[0,1]$. Let (G_T, D_T) be the zero free interval for X covering T . Let M be a SSM indpt of both X and T . Define a new process X^{**} by*

$$X_t^{**} = \begin{cases} X(t), & 0 \leq t \leq G_T \\ X(D_T + t - G_T), & G_T \leq t \leq 1 - (D_T - G_T) \\ \sqrt{D_T - G_T} M\left(\frac{1}{D_T - G_T} [t + (D_T - G_T) - 1]\right), & 1 - (D_T - G_T) \leq t \leq 1 \end{cases}$$

*Then the law of X^{**} is Q^μ .*

Note. X^{**} is identical to X^* defined previously on the interval $[0, G_T + 1 - D_T]$. So we learn that

$$(X_t^*, 0 \leq t \leq G_T + 1 - D_T; P) \stackrel{d}{=} (X_t, 0 \leq t \leq G_1; Q^\mu)$$

This is Theorem 1.3.

7. More General Subordinators.

This section presents some extensions for more general subordinators of the previous results in the stable case. To go beyond the stable case it seems essential to formulate things in a way that does not involve scaling. Previous results of this kind are Theorems 1.2, 5.1 and 5.2. The main result of this section, Theorem 7.1 below, combines all three of these results in a more general setting.

We first introduce some notation for various functions associated with a subordinator.

To facilitate comparison with previous results, we use the notation " $\stackrel{*}{=}$ " below to denote "equals in the stable case". Suppose given a subordinator (T_s) with no drift component and Lévy measure Λ , so for $\theta > 0$

$$E [e^{-\theta T(s)}] = \exp \left(-s \int_0^{\infty} (1 - e^{-\theta x}) \Lambda(dx) \right) \stackrel{*}{=} \exp(-c s \theta^\alpha).$$

Suppose Λ has a density

$$\rho(x) = \frac{\Lambda(dx)}{dx} \stackrel{*}{=} \frac{c \alpha}{\Gamma(1-\alpha)} \frac{1}{x^{\alpha+1}}, \quad 0 < x < \infty.$$

Assume ρ is strictly positive except perhaps on an interval of the form $[a, \infty)$ for some $a > 0$. The functions

$$\bar{\Lambda}(x) = \Lambda(x, \infty) \stackrel{*}{=} \frac{cx^{-\alpha}}{\Gamma(1-\alpha)}$$

$$h(x) = \bar{\Lambda}(x)/\rho(x) \stackrel{*}{=} x/\alpha$$

will play an important role. So too will the density functions

$$p(t, s) = P(S_t \in ds)/ds$$

$$\theta(s, t) = P(T_s \in dt)/ds$$

In principle, either of these functions can be calculated from the other, due to the inverse relation between (S_t) and (T_s) . Neither function assumes a simple form in the stable case, except if $\alpha = 1/2$. But their ratio is simple in the stable case, whatever α , due to (2.c):

$$\Phi(t, s) = \frac{\theta(s, t)}{p(t, s)} \stackrel{*}{=} \alpha \frac{s}{t}.$$

Bridge distributions $(P_{0t}, t \geq 0)$ can be identified as in Kallenberg (1981) as the family of Palm distributions on the basic probability space (Ω, \mathcal{F}, P) associated with the local time random measure dS_t :

$$P_{\alpha}(F) = \frac{E(dS_t; F)}{E(dS_t)}, \quad F \in \mathcal{F}, \quad t > 0.$$

As suggested by Kallenberg (1981, Lemma 4.1), the basic switching identity

$$(7.a) \quad P(\cdot \mid T_s = t) = P_{0t}(\cdot \mid S_t = s)$$

holds in this setting modulo null sets. We shall assume for simplicity that good versions of the conditional distributions exist so that (7.a) holds identically in s and t . Then (7.a) implies

$$(7.b) \quad P_{0t}(S_t \in ds)/ds = \frac{\theta(s, t)}{m(t)} \stackrel{*}{=} \frac{c \alpha \Gamma(\alpha) s}{t^{\alpha}} p(t, s)$$

where

$$(7.c) \quad m(t) = \int_0^{\infty} \theta(s, t) ds = \frac{E(dS_t)}{dt} \stackrel{*}{=} \frac{t^{\alpha-1}}{c \Gamma(\alpha)}.$$

Theorem 7.1

Let G_t be the σ -field generated by the sequence of lengths

$$V_{1t} > V_{2t} > \dots$$

of maximal intervals in the random set $Z \{c \cap (0, t)\}$, where Z is the closure of the range of a subordinator of (T_s) as above. Let $\sup Z \cap (0, t) = t - A_t$, and let $H_t = \sum_n h(V_{nt})$. Then for $n = 1, 2, \dots$

$$(7.d) \quad P(A_t = V_{nt} \mid G_t) = h(V_{nt})/H_t.$$

The unconditional law P , and the bridge law P_{0t} are mutually absolutely continuous on G_t , with

$$(7.e) \quad \left. \frac{dP}{dP_{0t}} \right|_{G_t} = m(t) H_t / S_t.$$

Similarly, for the laws P_t^s and P_{0t}^s obtained by further conditioning on $S_t = s$,

$$(7.f) \quad \left. \frac{dP_t^s}{dP_{0t}^s} \right|_{G_t} = \Phi(t, s) H_t / s$$

and (7.d) holds also with P_t^s in lieu of P .

Proof. Part (7.f) together with (7.d) for P_t^s instead of P follow immediately from Lemmas 7.2 and 7.3 below. To derive (7.e) from (7.f), compute as follows: for $G \in G_t$

$$\begin{aligned}
 P(G; S_t \in ds) &= P_t^s(G) p(t, s) ds \\
 &= P_{0,t}^s(GH_t) \Phi(t, s) s^{-1} p(t, s) ds \quad \text{by (7.f)} \\
 &= P_{0,t}^s(GH_t) \theta(s, t) s^{-1} ds \\
 &= P_{0,t}(GH_t, S_t \in ds) m(t) s^{-1} ds \quad \text{by (7.b)}. \quad \square
 \end{aligned}$$

Lemma. 7.2 *Let N_t be the point process on $(0, \infty)$ which counts the interval lengths up to time $G_t = 1 - A_t$:*

$$N_t(\cdot) = \#\{n : V_{n,t} \in \cdot\} - 1(A_t \in \cdot).$$

Let Q_t^s be the law of a Poisson point process N with intensity $s \Lambda(dx)$ over $x \in (0, \infty)$, conditioned on $\int x N(dx) = t$, Then

$$(7.g) \quad P_t^s(N_t \in dn) = Q_t^s(dn).$$

$$(7.h) \quad P_t^s(A_t \in da, N_t \in dn) = da \bar{\Lambda}(a) \frac{\theta(s, t-a)}{p(t, s)} Q_{t-a}^s(dn).$$

Proof. Formula (7.g) follows at once from the switching identity (7.a) and the Poisson character of the jumps of the subordinator (T_s) . To derive (7.h), we recall the well known facts that

$$P(A_t \in da) = m(t-a) da \bar{\Lambda}(a),$$

and that under P given $A_t = a$, the process on $[0, t-a]$ is distributed as under the bridge law $P_{0,t-a}$. See e.g. Kallenberg [1981]. This enables us to compute

$$\begin{aligned}
 P(A_t \in da, N_t \in dn, S_t \in ds) &= m(t-a) da \bar{\Lambda}(a) P_{0,t-a}(N_{t-a} \in dn, S_{t-a} \in ds) \\
 &= m(t-a) da \bar{\Lambda}(a) P_{0,t-a}(S_{t-a} \in ds) Q_{t-a}^s(dn)
 \end{aligned}$$

where Q_u^s appears as the $P_{0,u}^s$ distribution of N_u , by (7.g). This expression and (7.b) yield (7.h). \square

Lemma 7.3. *Let M be a PRM (μ) , Σ a function such that $M\Sigma$ has probability density $\theta(t)$, $t > 0$, and let (Q_t) be a regular conditional distribution for M given $M\Sigma = t$, $t > 0$. Suppose defined on (Ω, \mathcal{F}, P) a point process N and a random variable A such that*

$$P(A \in da, N \in dn) = g(a) \mu(da) Q_{t-\Sigma(a)}(dn).$$

Let $N^+ = N + \delta_A$. Then $P(N^+ \in dn) = (nh_t) Q_t(dn)$, where

$$h_t(a) = g(a) \theta(t) / \theta(t - \Sigma(a)),$$

and A is picked from N by h_t -biased sampling:

$$P(A \in da | N) = h_t(a)N(da)/Nh_t.$$

Proof. Conditioning a point process n to have sum $n\Sigma = t$, and a point at a , is the same as conditioning n to have a point at a , and $n - \delta_a$ to have sum $t - \Sigma(a)$. Let Q denote the law of a PRM (μ) . From the description (2.q) of the Palm distributions for Q , it follows that for n with distribution $Q_t = Q(\cdot | n\Sigma = t)$, $Q_{t-\Sigma(a)}$ serves as the Palm distribution of $n - \delta_a$ given n with a point at a . The mean measure for Q_t is μ_t where

$$\begin{aligned}\mu_t(da) &= Q_t n(da) = Q[n(da) | n\Sigma = t] \\ &= \frac{Q[n(da)1(n\Sigma \in dt)]}{Q(n\Sigma \in dt)} \\ &= \mu(da)\theta(t - \Sigma(a))/\theta(t).\end{aligned}$$

Thus the assumptions on N and A make

$$P(A \in da, N \in dn) = h_t(a)\mu_t(da)Q_{t-\Sigma(a)}(dn).$$

By the above description of the Palm measures associated with Q_t and the remark below Lemma 2.2, this implies

$$P(A \in da, N^+ \in dn) = h_t(a)Q_t(dn)n(dx), \quad \square$$

Corollary 7.4 Let $\delta(t) = t\rho(t)/\bar{\Lambda}(t)$. Then for $G_t \in \mathbf{G}_t$

$$P_t^s[G_t \delta(A_t)] = P_{0t}^s(G_t)\Phi(t, s)t/s, = P_{0t}^s(G_t)P_t^s[\delta(A_t)].$$

Proof. The first equality is immediate from (7.f) and the fact that $\sum_n V_{nt} = t$ a.s.. The second equality follows by two applications of the first one.

Remarks. In the stable case, $h(t) = t/\alpha$. Hence also $H_t = t/\alpha$, and (7.e) shows that $[dP/dP_{0t}] = m(t)t/(\alpha S_t)$ on \mathbf{G}_t , which is the conclusion of Theorem 5.2. Consequently, in the stable case

$$(7.i) \quad P(G_t | S_t) = P_{0t}(G_t | S_t),$$

as asserted by Theorem 5.1. The above corollary offers one kind of extension of this identity to the general case. Another, immediate from (7.e) is

$$(7.j) \quad P(G_t | S_t) = \frac{P_{0t}(G_t H_t | S_t)}{P_{0t}(H_t | S_t)} \quad \text{where } H_t = \sum_n h(V_{nt}).$$

References:

M. Barlow, J. Pitman, M. Yor (1989a) On Walsh's Brownian motions. *Séminaire de probabilités XXIII*, Lecture Notes in Math. 1372, Springer, 275-293

M. Barlow, J. Pitman, M. Yor (1989b) Une extension multidimensionnelle de la loi de l'arc sinus, *Séminaire de probabilités XXIII*, Lecture Notes in Math. 1372, Springer, 294-314.

P. Biane, J.F. Le Gall, M. Yor (1987) Un processus qui ressemble au pont Brownien, *Séminaire de probabilités XXI*, Lecture Notes in Math. 1247, Springer, 270-275.

N. H. Bingham, R. A. Doney (1988) On higher-dimensional analogues of the arc-sine law. *Journal of Applied Probability* **25** , 120-131

D. Blackwell, J.B. MacQueen (1973). Ferguson distributions via Pòlya urn schemes, *Ann. Statist.*, **1**, 353-355

J.K. Brooks, R.V. Chacon (1983). Diffusions as a limit of stretched Brownian motions. *Adv. in Math.*, vol. 49, No. 2, 109-122.

D. J. Daley, D. Vere-Jones (1988). An introduction to the theory of point processes. Springer.

P. Donnelly, P. Joyce (1989). Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stochastic Processes and their Applications* **31**, 89-103.

E.B.Dynkin (1961) Some limit theorems for sums of independent random variables with infinite mathematical expectations.

IMS-AMS Selected Trans. in Math. Stat. and Prob. **1** , 171-189.

W.J. Ewens (1988). Population genetics theory - the past and the future, in: S.Lessard, ed., *Mathematical and Statistical Problems in Evolution* University of Montreal Press, Montreal.

T.S. Ferguson (1973) A Bayesian analysis of some nonparametric problems, *Ann. Stat.*, **1**, 209-230.

W. Feller (1971) An introduction to probability theory and its applications, Vol 2, 2nd ed. Wiley, New York.

J. Franchi (1989) Produits semi-directs de diffusions réelles at lois asymptotiques. *Adv. Appl. Prob.* **21**, 756-769.

R.K. Gettoor (1979) Excursions of a Markov process. *Ann. Prob.* **7** 244-266.

P. Greenwood, J. Pitman (1980). Fluctuation identities for Lévy processes and splitting at the maximum
Adv. Appl. Prob. **12**, 893-902.

J.M. Harrison, L.A. Shepp (1981). On skew Brownian motion. *Ann. Prob.* **9**, 309-313.

F.M. Hoppe (1987) The sampling theory of neutral alleles and an urn model in population genetics, *J. Math. Biol.*, **25**, 123-159.

K. Itô, H.P. McKean (1965). Diffusion processes and their sample paths. Springer.

P. Joyce, S. Tavaré (1987) Cycles, permutations and the structure of the Yule process with immigration. *Stochastic Processes and their Applications* **25**, 309-314.

O. Kallenberg (1983)*Random measures*, 3rd ed. Akademie Verlag, Berlin and Academic Press, London

O. Kallenberg (1981). Splitting at backward times in regenerative sets *Annals of Probability* **9** , 781-799.

I. Karatzas, S.E. Shreve (1987a). Brownian motion and stochastic calculus. Springer.

I. Karatzas, S.E. Shreve (1987b). A decomposition of the Brownian path. *Statistics and Probability Letters* **5** 87-93

J. F. C. Kingman (1975). Random discrete distributions, *J. Roy. Statist. Soc.*, **37**, 1-22.

F.B. Knight (1985). On the duration of the longest excursion. *Seminar on stochastic processes, 1985*, ed by Çinlar, Chung, Gettoor, 117-148.

J.F. Le Gall (1984). One-dimensional stochastic differential equations involving the local times of the unknown process. In: 'Stochastic Analysis and Applications' (eds: A. Truman, D. Williams) Lect. Notes in Math. 1095. Springer.

P. Lévy (1939). Sur certains processus stochastiques homogènes. *Compositio Mathematica* 7,283-339

S.A. Molchanov and E. Ostrovski (1969). Symmetric stable processes as traces of degenerate diffusion processes, J. Prob. Appl., 14, No. 1, 128-131.

G.P. Patil and C. Taillie (1977). Diversity as a concept and its implications for random communities, *Proc. 41st. I.S.I. New Delhi.*, 497-515.

E. Perkins (1982). Local time is a semi-martingale Z. *Wahrscheinlichkeitstheorie verw. Gebiete* 60,79-117

M. Perman (1990) Random discrete distributions derived from subordinators. Ph.D. thesis, Dept Statistics, U.C. Berkeley.

M. Perman, J. Pitman and M. Yor (1990). Size-biased sampling of Poisson point processes and excursions. Paper in preparation. Dept Statistics, U.C. Berkeley.

J. Pitman, M. Yor (1986). Asymptotic laws of planar Brownian motion. *Ann. Prob.* 14, 733-779.

D. Revuz and M. Yor (1990). *Continuous Martingales and Brownian motion*, to be published by Springer

L.C.G. Rogers (1981). Williams characterisation of the Brownian excursion law: proof and applications. *Séminaire de probabilités XV*, Lecture Notes in Math. 850, Springer, 227-250.

L.C.G. Rogers, D. Williams (1987). Diffusions, Markov processes and Martingales. Vol 2: Itô Calculus. J. Wiley.

W. Rosenkrantz (1975). Limit theorems for solutions to a class of stochastic differential equations. *Indiana Math. J.* 24, 613-625.

L.A. Shepp, S.P. Lloyd (1966). Ordered cycle lengths in a random permutation.

Trans. Amer. Math Soc., **121**, 340-357.

A.M. Vershik, A.A Schmidt (1977). Limit measures arising in the asymptotic theory of symmetric groups I., *Theory of Probability and its Applications*, **22**, No.1, 70-85.

A.M. Vershik, A.A Schmidt (1978). Limit measures arising in the asymptotic theory of symmetric groups II., *Theory of Probability and its Applications*, **23**, No.1, 36-49.

J.B. Walsh (1978). A diffusion with discontinuous local time. *Astérisque* **52-53**, 37-45.

G.A. Watterson (1976). The stationary distribution of the infinitely many neutral alleles diffusion model, *Journal of Applied Probability*, Vol.13, 639-651.