The Annals of Probability
1986, Vol. 14, No. 3, 1024–1036

SPLITTING INTERVALS

BY MICHAEL D. BRENNAN[1] AND RICHARD DURRETT[2]

Utah State University and University of California, Los Angeles

In the processes under consideration, an interval of length L splits with probability (or exponential rate) proportional to L^α, $\alpha \in [-\infty, \infty]$, and when it splits, it splits into two intervals of length LV and $L(1-V)$ where V has d.f. F on $(0,1)$. When $\alpha = 1$ and $F(x) = x$, the split points are i.i.d. uniform on $(0,1)$ and when $\alpha = \infty$ (a longest interval is always split), the model is a splitting process invented by Kakutani. In both these cases, the empirical distribution of the split points converges almost surely to the uniform distribution on $(0,1)$. On the other hand, when $\alpha = 0$, the model is a binary cascade and the empirical distribution of the split points converges almost surely to a random, continuous, singular distribution. In this paper, we show what happens in the other cases. Can the reader guess at what point the character of the limiting behavior changes?

1. Introduction. In this paper we consider a class of discrete and continuous time stochastic processes in which the unit interval undergoes random subdivision at successive points $\{X_i,\ i \geq 1\}$. The problem is to determine the limiting behavior, in the sense of weak convergence of probability measures, of the empirical distribution function

$$n^{-1}N_n(\cdot) = n^{-1}\sum_{i=1}^{n} I_{\{X_i \leq \cdot\}}.$$

In discrete time the processes are as follows: $\{V_n,\ n \geq 0\}$ are i.i.d. with distribution function F on $(0,1)$. At time $n=0$, there is one interval, $[0,1]$. At time $n=1$, $[0,1]$ splits into a left interval of length V_1 and a right interval of length $1-V_1$. At time n, there are $n+1$ intervals with lengths $L_1, L_2, \dots, L_{n+1}$. One interval is chosen from the $n+1$ intervals according to the probability law which assigns mass proportional to L_i^α to the ith interval, $\alpha \in [-\infty, \infty]$. (For $\alpha = \infty$ a longest interval is always chosen and for $\alpha = -\infty$ a shortest interval is always chosen.) The chosen interval, say of length L, splits at time $n+1$ into a left interval of length LV_{n+1} and a right interval of length $L(1-V_{n+1})$.

Special cases of this model have been previously studied. For $\alpha = 1$ and $F(x) = x$, the $X_1, X_2, \dots$ are i.i.d. uniform on $(0,1)$ and $n^{-1}N_n(x) \to x$ a.s. for each $x \in (0,1)$ by the classical strong law of large numbers. $\alpha = \infty$ is a stochastic version of a deterministic scheme for splitting $[0,1]$ invented by Kakutani (1975).

Received December 1980; revised October 1985.

[1]The work of this author was initiated at the University of Southern California and partially supported by NSF grant MCS81-02730.

[2]The work of this author was partially supported by NSF grants MCS 77-02121, MCS 80-02732, and MCS 83-00836 and an Alfred P. Sloan fellowship. Now at Cornell University.

AMS 1980 *subject classifications*. Primary 60F15; secondary 60K99.

Key words and phrases. Random subdivision, splitting process, empirical distribution function, uniform distribution, branching random walk, renewal equation.

In this case and indeed whenever $\alpha > 1$, there is a greater tendency to split longer intervals than when $\alpha = 1$, so it is natural to conjecture that $n^{-1}N_n(\cdot)$ converges to the uniform distribution for $\alpha = \infty$ and for $1 < \alpha < \infty$ by "interpolation." This conjecture is correct and has been proved by Lootgieter (1977) and van Zwet (1978) for $\alpha = \infty$ and $F(x) = x$, Lootgieter (1978) for $\alpha = \infty$ and F with a Lebesgue component, and by Lootgieter (1981) for $1 \le \alpha < \infty$ and F arbitrary. The case $\alpha = 0$ is radically different. Here $n^{-1}N_n(\cdot)$ converges weakly to a random, strictly increasing, continuous distribution function which depends on F and is always singular with respect to the uniform distribution. These random measures were first constructed by Dubins and Freedman (1967) and have been studied in detail by Peyriere (1979) and Graf, Mauldin, and Williams (1986).

The case $\alpha = -\infty$ is trivial. For $n \ge 1$, $|X_{n+1} - X_n| < 2^{-n-1}$ and thus $\{X_n\}$ converges to a limit X_∞ and $n^{-1}N_n(\cdot)$ converges weakly to the point mass at X_∞.

At this point the reader might like to test his or her intuition by guessing the limit behavior for $\alpha \in (-\infty, 0)$ or $(0,1)$. $\alpha \in (0,1)$ is the most intriguing case. Why is $\alpha = 0$ singular? Roughly because small intervals are as likely to split as large intervals so when a subinterval gets a split point it increases its probability of getting later split points and the split points "pile up." To a lesser degree this is true for $0 < \alpha < 1$; since $x^\alpha + (1-x)^\alpha > 1$ for $0 < x < 1$ and $\alpha < 1$, when an interval splits it increases its probability of splitting later. Is this effect strong enough to produce singular limit measures for $0 < \alpha < 1$?

Our first step in the solution of this problem is to reformulate the splitting processes in continuous time. The continuous time process is as follows: At $t = 0$, there is the interval $[0,1]$. After creation, an interval of length L waits mean $L^{-\alpha}$ exponential time and then splits into a left interval of length LV and a right interval of length $L(1-V)$. V has distribution function F on $(0,1)$ as above. Both V and the waiting time to split are independent of the past and the other present intervals.

Let $Z(t)$ = the number of intervals at time t and $N(t,x)$ = the number of intervals contained in $[0,x]$ at time t. Since a version of the discrete time splitting process is embedded in the continuous time splitting process at the split times $T_n = \inf\{t \ge 0\colon Z(t) = n+1\}$, the original problem is solved by determining

$$\lim_{t\to\infty} \frac{N(t,\cdot)}{Z(t)}$$

for each α and F.

There is a nice way to view the continuous time processes which makes the correct answer intuitively plausible. If an interval of length L is associated with a particle at $-\log L$, then the system becomes a Markovian branching random walk where a particle at x waits a mean $e^{\alpha x}$ exponential time and then splits into two particles at $x - \log V$ and $x - \log(1-V)$. From this viewpoint, the behavior for $\alpha < 0$ is easy to see. The time between splits decreases exponentially fast, so the number of intervals (or particles) explodes in finite time and at the explosion time only one nested sequence of intervals has split infinitely many times. It follows that the split points $\{X_n\}$ for this sequence of intervals converge almost surely to a random limit X_∞ and $n^{-1}N_n(\cdot)$ almost surely converges weakly to the point mass at X_∞. Details are given in Section 3.

For $\alpha > 0$ as particles move away from the origin their waiting times to split grow exponentially large. This keeps the particles close together, i.e., the intervals about the same length, and suggests that the limit will be uniform. Unlike $\alpha < 0$, this argument is far from a proof and requires quite a bit of work.

A key to our analysis for $\alpha > 0$ is to observe that the splitting processes on $[0, X_1]$ and $[X_1, 1]$ are independent (this is not true in discrete time); moreover, after scaling they are replicas of the original process on $[0,1]$. More precisely,

$$(1.1)\qquad Z(t) = \begin{cases} 1 & \text{for } t < T_1, \\ Z_1(X_1^\alpha(t - T_1)) + Z_2((1 - X_1)^\alpha(t - T_1)) & \text{for } t \ge T_1, \end{cases}$$

where Z_1 and Z_2 are independent with the same distribution as Z and independent of T_1 and X_1. Consequently,

$$(1.2)\qquad \lim_{t\to\infty} \frac{N(t, X_1)}{Z(t)} = \lim_{t\to\infty} \frac{Z_1(X_1^\alpha t)}{Z_1(X_1^\alpha t) + Z_2((1 - X_1)^\alpha t)}.$$

To understand how to attack (1.2) for $\alpha > 0$, we first analyze (1.2) for $\alpha = 0$ which is easy. When $\alpha = 0$, all intervals split at rate one so $Z(t)$ is a binary Yule process. Consulting Ross (1983), we find $Z(t)$ has a geometric distribution with mean e^t and $e^{-t}Z(t) \to W$ a.s. where W is a mean one exponential. Substituting these observations into (1.2), gives

$$(1.3)\qquad \begin{aligned} \lim_{t\to\infty} \frac{N(t, X_1)}{Z(t)} &= \lim_{t\to\infty} \frac{e^{-t}Z_1(t)}{e^{-t}Z_1(t) + e^{-t}Z_2(t)} \\ &= W_1(W_1 + W_2)^{-1} = U \quad \text{a.s.} \end{aligned}$$

W_1 and W_2 are independent copies of W, hence U is uniform on $(0,1)$ (and independent of X_1). By considering other split points, it is easy to see that $N(t, \cdot)/Z(t)$ almost surely converges weakly to a random distribution function which is strictly increasing and continuous. To see that this measure is Lebesgue-singular is a little more difficult. It turns out this construction is a special case of a general random construction by Dubins and Freedman (1967) which almost surely produces Lebesgue-singular measures on $(0,1)$. Here Dubins and Freedman's μ is $dF \times dx$ on the unit square. Other results about the limit measure and the $\alpha = 0$ discrete time splitting process are found in Peyriere (1979) and Graf, Mauldin, and Williams (1986).

Returning to $\alpha > 0$ and motivated by computation (1.3), we next analyze the behavior of $Z(t)/m(t)$ for $m(t) = EZ(t)$. Taking expectations in (1.1), we obtain

$$m(t) = e^{-t} + 2\int_0^t e^{-s}\int_0^1 m(x^\alpha(t - s))\, d\bar{F}(x)\, ds,$$

or in differential form

$$(1.4)\qquad \begin{aligned} m'(t) &= -m(t) + 2\int_0^1 m(x^\alpha t)\, d\bar{F}(x), \\ m(0) &= 1. \end{aligned}$$

$\bar{F}(x) = (F(x) + 1 - F(1 - x -))/2$ is the symmetrization of F about $\frac{1}{2}$. Since only $\bar{F}$ enters into (1.4), we can and will, without loss, assume F is symmetric about $\frac{1}{2}$.

(1.4) is easy to solve when $1/\alpha$ is an integer and then $m(t)$ is a polynomial of degree $1/\alpha$. For example, when $F(x) = x$ and $\alpha = 1, \frac{1}{2}$, and $\frac{1}{3}$, $m(t) = 1 + t$, $1 + t + t^2/6$, and $1 + t + t^2/4 + t^3/60$. The polynomials do not tell us much about $m(t)$ for other α's, but they were one clue that led us to guess $m(t)$ grows like $t^{1/\alpha}$. This guess is correct and in Section 2 we analyze (1.4) to show

$$m(t) \sim Kt^{1/\alpha} \tag{1.5}$$

as $t \to \infty$. $K > 0$ is a constant whose value depends on α and F. A similar computation for $v(t) = \operatorname{Var} Z(t)$ gives

$$\begin{aligned} v'(t) &= -v(t) + 2\int_0^1 v(x^\alpha t)\, dF(x) - 2m(t)m'(t) \\ &\quad + \int_0^1 f(x, t)h(x, t)\, dF(x), \\ v(0) &= 0 \end{aligned} \tag{1.6}$$

for $f(x, t) = m(x^\alpha t) + m((1 - x)^\alpha t) + m(t)$ and $h(x, t) = m(x^\alpha t) + m((1 - x)^\alpha t) - m(t)$. In Section 2 we show (1.6) implies

$$v(t) \le Ct^{2/\alpha - \theta}, \tag{1.7}$$

where $C > 0$ and $\theta > 0$ are constants depending on α and F. Here we run into some technical difficulties. We are unable to get good enough estimates of $h(x, t)$ in (1.6) to prove (1.7) without additional assumptions on F. To estimate $h(x, t)$, which we do in Section 2, we use a coupling constructed by Ney (1981) to estimate the rate of convergence in the classical renewal theorem (i.e., the rate of convergence of the expected number of renewals in $(t, t + h]$ to $\mu^{-1}h$ as $t \to \infty$). To use Ney's coupling, we must assume that the distribution function $\hat{G}$, defined by $\hat{G}(y) = 1 - \int_0^{e^{-y}-} 2x\, dF(x)$, has a nonzero Lebesgue component for some n-fold convolution. In particular, this hypothesis is satisfied if F has a nonzero Lebesgue component.

In Section 2 we show that (1.5) and (1.7) imply

$$Z(t) \sim Kt^{1/\alpha} \quad \text{a.s.} \tag{1.8}$$

as $t \to \infty$. Substituting (1.8) into (1.2), we obtain

$$\frac{N(t, X_1)}{Z(t)} = \frac{Z_1(X_1^\alpha t)}{Z(t)} \sim \frac{K(X_1^\alpha t)^{1/\alpha}}{Kt^{1/\alpha}} = X_1 \quad \text{a.s.} \tag{1.9}$$

as $t \to \infty$. Clearly (1.9) remains true with X_1 replaced by any split point and with X_1 replaced by x by density of the split points.

Combining our results with the earlier results mentioned above, we obtain our main result.

THEOREM 1.1. *Let the distribution function $\hat{G}(y) = \int_0^{e^{-y}} 2x\, dF(x)$ have a Lebesgue component for some n-fold convolution. For the discrete or continuous* *time splitting process with splitting distribution F and parameter α, there are three possibilities depending on α:*

(1) *For $\alpha \in (0, \infty]$, $n^{-1}N_n(\cdot)$ almost surely converges weakly to the uniform distribution.*
(2) *For $\alpha = 0$, $n^{-1}N_n(\cdot)$ almost surely converges weakly to a random distribution function which is strictly increasing, continuous, and singular with respect to the uniform distribution.*
(3) *For $\alpha \in [-\infty, 0)$, $n^{-1}N_n(\cdot)$ almost surely converges weakly to a random point mass.*

We have proved our contribution to this theorem except for the following details: The analysis of the integral–differential equations (1.4) and (1.6) for $m(t)$ and $v(t)$ leading to asymptotic estimate (1.8) is in Section 2. The explosion argument for $\alpha < 0$ is in Section 3. We only use the hypothesis on $\hat{G}$ for $0 < \alpha < \infty$ and $\alpha \neq 1/n$ for some integer n.

2. Asymptotic estimates for $Z(t)$, $0 < \alpha < \infty$. In this section we prove the asymptotic formulas

$$m(t) \sim Kt^{1/\alpha}, \tag{2.1}$$

$$v(t) \leq Ct^{2/\alpha - \theta}, \tag{2.2}$$

$$Z(t) \sim Kt^{1/\alpha} \quad \text{a.s.} \tag{2.3}$$

as $t \to \infty$. K, C, and θ are positive constants depending on α and F. (Here and below C denotes a positive constant whose value is unimportant and may change from line to line.)

First we recall that $m(t)$ satisfies the integral–differential equation

$$m'(t) = -m(t) + 2\int_0^1 m(x^\alpha t)\, dF(x),$$
$$m(0) = 1.$$

It will be convenient to rewrite this equation as

$$\begin{aligned} m'(t) &= -m(t) + \int_0^1 x^{-1} m(x^\alpha t)\, d\hat{F}(x), \\ m(0) &= 1, \end{aligned} \tag{2.4}$$

where $\hat{F}(x) = \int_0^x 2u\, dF(u)$ is a distribution function by the symmetry of F.

The first result of this section establishes the polynomial solutions to (2.4) when $1/\alpha$ is a positive integer.

PROPOSITION 2.1. *$m(t)$ has a power series representation $m(t) = 1 + t + \sum_{k=2}^{\infty} a_k t^k$ which is absolutely convergent for $0 \leq t < \infty$. If $1/\alpha = n$, the series terminates and $m(t)$ is a polynomial of degree n with positive coefficients. If $1/\alpha$ is not an integer, set $p =$ the smallest integer larger than $1/\alpha$, then $a_k > 0$ for $k \leq p$, $a_{p+1} < 0$, and the series alternates thereafter.*

PROOF. It is an easy computation to check that (2.4) uniquely determines the a_k's by the formula $a_0 = 1$ and $a_{k+1} = a_k(k+1)^{-1}\int_0^1 (x^{k\alpha-1} - 1)\,d\hat{F}(x)$ and that the resulting series is absolutely convergent and satisfies (2.4). □

Except when $1/\alpha$ is an integer, the power series for $m(t)$ tells us little about $m(t)$ for large t. To determine the behavior of $m(t)$ as $t \to \infty$, we will make repeated use of renewal theory and the renewal theorem. Chapter 11 of Feller (1971) contains all the renewal theory we use. The renewal theorem is on page 363.

We need a technical result.

LEMMA 2.2. *Let $X_1, X_2, \dots$ be an i.i.d. sequence of positive random variables with distribution function G having mean μ and satisfying $\int_0^\infty e^x\,dG(x) = M < \infty$. Then for $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$, and $N_t = \inf\{k > 0\colon S_k > t\}$, it follows that* $\sup_{0 \le t < \infty} E(\exp(S_{N_t} - t)) < \infty$.

PROOF. Setting $R(t) = E(\exp(S_{N_t} - t))$, we have

$$R(t) = \sum_{n=0}^{\infty} \int_0^t \int_{t-y}^{\infty} e^{x+y-t}\,dG(x)\,P(S_n \in dy)$$
$$\le M \sum_{n=0}^{\infty} P(S_n \le t) < \infty$$

and R satisfies the renewal equation

$$R(t) = \int_{t+}^{\infty} e^{y-t}\,dG(y) + \int_0^t R(t-y)\,dG(y),$$

so

$$\lim_{t\to\infty} R(t) = \mu^{-1}\int_0^\infty e^y(1 - G(y))\,dy = \mu^{-1}(M - 1). \square$$

Returning to the analysis of (2.4), for $t > 0$ and $-\infty < u < \infty$, we make the substitutions $t = e^{\alpha u}$, $g(u) = e^{-u}m(e^{\alpha u})$, $h(u) = e^{-u}m'(e^{\alpha u})$, and $y = -\log x$ in (2.4) and obtain

$$g(u) = \int_0^\infty g(u-y)\,d\hat{G}(y) - h(u), \tag{2.5}$$

where $\hat{G}(y) = 1 - \hat{F}(e^{-y} -)$. $\hat{G}$ has mean $\hat{\mu} = \int_0^1 -2x\log x\,dF(x) < \infty$.

We let $Y_1, Y_2, \dots$ be an i.i.d. sequence with distribution function $\hat{G}$, $S_0 = 0$, and $S_n = \sum_{i=1}^n Y_i$. By iterating (2.5) $n-1$ times, we obtain

$$g(u) = E\left(g(u - S_n) - \sum_{k=0}^{n-1} h(u - S_k)\right). \tag{2.6}$$

For $-\infty < T \le u$, we set $N(u) = \inf\{k \ge 1\colon u - S_k < T\}$ and $N(u,k) = \min\{N(u), k\}$. It is a straightforward exercise in manipulation of stopping times to check that (2.6) remains true when n is replaced by $N(u,k)$. For $y \le u$,

$g(y) \le m(e^{\alpha u})e^{-y}$. Consequently,

$$\begin{aligned} g(u - S_{N(u,k)}) &\le m(e^{\alpha u})\exp(S_{N(u,k)} - u) \\ &\le m(e^{\alpha u})\exp(S_{N(u)} - u) \end{aligned}$$

and

$$E(\exp(S_{N(u)} - u)) = e^{-T}E(\exp(S_{N(u)} - (u - T))) < \infty$$

by Lemma 2.2. So replacing n by $N(u, k)$ in (2.6) and letting $k \to \infty$, we obtain by dominated convergence on the first term and monotone convergence on the second term

$$\begin{aligned} g(u) &= E(g(u - S_{N(u)})) - E\left(\sum_{k=0}^{N(u)-1} h(u - S_k)\right) \\ &= J(u) - H(u). \end{aligned} \tag{2.7}$$

$J(u) < \infty$ and $J(\cdot)$ satisfies the renewal equation

$$J(u) = \int_{(u-T)+}^{\infty} g(u - y)\, d\hat{G}(y) + \int_0^{u-T} J(u - y)\, d\hat{G}(y)$$

for $u \ge T$. By the renewal theorem, we obtain

$$\begin{aligned} \lim_{u \to \infty} J(u) &= \hat{\mu}^{-1}\int_T^{\infty}\int_{(s-T)+}^{\infty} g(s - y)\, d\hat{G}(y)\, ds \\ &= \hat{\mu}^{-1}\int_0^{\infty} g(T - s)(1 - \hat{G}(s))\, ds < \infty. \end{aligned} \tag{2.8}$$

This integral is finite since $g(T - s) \sim Ce^s$ as $s \to \infty$ and $\int_0^\infty e^s(1 - \hat{G}(s))\, ds = 1$.

We next establish bounds on the growth of $m(t)$ and $m'(t)$.

PROPOSITION 2.3.

$$m(t) = O(t^{1/\alpha}) \quad \textit{and} \quad m'(t) = O(t^{1/\alpha - 1}) \quad \textit{as } t \to \infty.$$

PROOF. $H(u) \ge 0$ so $m(t) = O(t^{1/\alpha})$ follows from (2.7) and (2.8). If $\{L_i: 1 \le i \le Z(t)\}$ are the interval lengths at time t and $L(t) = \sum_{i=1}^{Z(t)} L_i^\alpha$, then $m'(t) = EL(t)$. If $0 < \alpha < 1$, $m'(t)$ is increasing so $m'(t)t \le \int_t^{2t} m'(s)\, ds \le m(2t)$. If $1 \le \alpha$, then $m'(t)$ is nonincreasing so $m'(t)t \le m(t)$. □

By Proposition 2.3, $h(u) = O(e^{-\alpha u})$ as $u \to \infty$; so returning to (2.7), we see that $H(u)$ satisfies the renewal equation

$$H(u) = h(u) + \int_0^{u-T} H(u - y)\, d\hat{G}(y)$$

for $u \ge T$, and by the renewal theorem

$$\lim_{u \to \infty} H(u) = \hat{\mu}^{-1}\int_T^{\infty} h(s)\, ds < \infty.$$

We have established

$$\lim_{u \to \infty} g(u) = \hat{\mu}^{-1}\int_0^{\infty} g(T - s)(1 - \hat{G}(s))\, ds - \hat{\mu}^{-1}\int_T^{\infty} h(s)\, ds, \tag{2.9}$$

so (2.1) is proved once it is shown that the right side of (2.9) which is evidently constant in T is not zero. The next proposition is needed to show this. It relates the values of $m(t)$ for fixed t, fixed F, and varying α.

PROPOSITION 2.4. *Let $m(t,\alpha)$ be the solution of* (2.4), *then if $t>0$ and $\alpha_1<\alpha_2$, $m(t,\alpha_1)>m(t,\alpha_2)$.*

PROOF. We construct two splitting processes $\mathrm{SP}(\alpha_1)$ and $\mathrm{SP}(\alpha_2)$ on the same probability space such that $P(Z(t,\alpha_1)\ge Z(t,\alpha_2))=1$ and $P(Z(t,\alpha_1)>Z(t,\alpha_2))>0$ for $t>0$.

Both processes split at the same set of points $\{U_i\colon i\ge 1\}$ determined by splitting distribution function F. We think of these points as predetermined before $t=0$. U_1 is the split point of $[0,1]$, U_2 is the split point of $[0,U_1]$, U_3 of $[U_1,1]$, U_4 of $[0,U_2]$, U_5 of $[U_2,U_1]$, and so on.

The processes are coupled as follows: Both processes split at the same time at U_1. Thereafter, if I is an unsplit subinterval of length L of $\mathrm{SP}(\alpha_2)$, there are two possibilities: (1) I is also an unsplit subinterval of $\mathrm{SP}(\alpha_1)$. In this case, both processes split I simultaneously with exponential rate L^{α_2} and $\mathrm{SP}(\alpha_1)$ splits I alone with rate $L^{\alpha_1}-L^{\alpha_2}$. (2) I has already split in $\mathrm{SP}(\alpha_1)$. In this case, $\mathrm{SP}(\alpha_1)$ on the subintervals of I and $\mathrm{SP}(\alpha_2)$ on I run independently. □

Returning to the analysis of (2.9), we assume $1/\alpha$ is not an integer and set $q=$ the integer such that $1/(q+1)<\alpha<1/q$ if $0<\alpha<1$ and $q=0$ if $\alpha>1$. When $q\ge 1$, $m(t,1/q)$ is a polynomial of degree q with positive coefficients by Proposition 2.1. By Proposition 2.4, $m(t,\alpha)\ge m(t,1/q)>a_q t^q$. When $q=0$, $m(t)\ge 1$. In either case, $g(u)\ge C\exp((\alpha q-1)u)$. We have

$$
\begin{aligned}
\hat{\mu}^{-1}\int_0^\infty g(T-s)(1-\hat{G}(s))\,ds &\ge \hat{\mu}^{-1}\int_0^\infty C\exp((\alpha q-1)(T-s))(1-\hat{G}(s))\,ds\\
&\ge Ce^{(\alpha q-1)T}\hat{\mu}^{-1}\int_0^\infty(1-\hat{G}(s))\,ds \qquad (2.10)\\
&= Ce^{(\alpha q-1)T}.
\end{aligned}
$$

Substituting (2.10) into (2.9) and using Proposition 2.3 to bound the second integral of (2.9), we obtain

$$\lim_{u\to\infty} g(u)\ge Ce^{(\alpha q-1)T}-C_1e^{-\alpha T}>0$$

for T sufficiently large since $-\alpha<\alpha q-1<0$. The proof of (2.1) is complete.

We now turn our attention to proving (2.2). We recall that $v(t)=\operatorname{Var}Z(t)$ satisfies the integral–differential equation

$$
\begin{aligned}
v'(t) &= -v(t)+2\int_0^1 v(x^\alpha t)\,dF(x)-2m(t)m'(t)\\
&\quad+\int_0^1 f(x,t)h(x,t)\,dF(x), \qquad (2.11)\\
v(0) &= 0
\end{aligned}
$$

for

$$f(x,t) = m(x^\alpha t) + m((1-x)^\alpha t) + m(t)$$

and

$$h(x,t) = m(x^\alpha t) + m((1-x)^\alpha t) - m(t).$$

We need bounds for the last integral in (2.11). $f(x,t) \le 3m(t) \le C(1 + t^{1/\alpha})$, and we establish bounds on the growth of $h(x,t)$ next.

LEMMA 2.5.

$$\int_0^1 |h(x,t)|\, dF(x) \le C(1 + t^{1/\alpha - \theta})$$

for a $\theta > 0$ which depends on α and F.

PROOF. We first consider $0 < \alpha < 1$. Let $0 < z \le x < 1$ and $X_1, X_2, \dots$ be i.i.d. with distribution function $\hat{F}$. $X_0 = 1$, $W_i = X_0 X_1 \cdots X_i$, and $J = \inf\{i \ge 1: W_i < z\}$. Starting with

$$m(t) = \int_0^1 y^{-1} m(y^\alpha t)\, d\hat{F}(y) - m'(t), \tag{2.12}$$

we iterate this equation $J - 1$ times to obtain

$$m(t) = E\big(W_J^{-1} m(W_J^\alpha t)\big) - E\left(\sum_{i=0}^{J-1} W_i^{-1} m'(W_i^\alpha t)\right). \tag{2.13}$$

[As J is random, a more rigorous development of (2.13) is to recognize that it is a rearrangement of (2.7).] Since $m'(t)$ is increasing,

$$\begin{aligned} E\left(\sum_{i=0}^{J-1} W_i^{-1} m'(W_i^\alpha t)\right) &\le m'(t) E\left(\sum_{i=0}^{J-1} W_i^{-1}\right) \\ &= z^{-1} m'(t) E\left(\sum_{i=0}^{J-1} e^{-(-\log z - S_i)}\right) \\ &= z^{-1} m'(t) R(z), \end{aligned} \tag{2.14}$$

where $S_i = -\log W_i$. An argument similar to the proof of Lemma 2.2 shows $\sup_{0 < z \le 1} R(z) < \infty$ so we have verified

$$0 \le E\big(W_J^{-1} m(W_J^\alpha t)\big) - m(t) \le C z^{-1} m'(t). \tag{2.15}$$

Next, we repeat the above procedure with a new i.i.d. sequence $\tilde{X}_1, \tilde{X}_2, \dots$. Everything is as above except $\tilde{X}_0 = x$. Then (2.12) is replaced by

$$x^{-1} m(x^\alpha t) = \int_0^1 (yx)^{-1} m((yx)^\alpha t)\, d\hat{F}(y) - x^{-1} m'(x^\alpha t).$$

The right-hand term in (2.14) becomes $z^{-1} m'(t) R(z/x)$ and (2.15) remains true with W_J replaced by $\tilde{W}_{\tilde{J}}$ and $m(t)$ replaced by $x^{-1} m(x^\alpha t)$. C on the right side of (2.15) does not depend on x in this case.

Next we use a coupling of Ney (1981) to construct the sequences $X_0, X_1, X_2, \dots$ and $\tilde{X}_0, \tilde{X}_1, \tilde{X}_2, \dots$ so that for small z, W_J, and $\tilde{W}_{\tilde{J}}$ have high probability of being the same random variable. Ney constructed his coupling to estimate the rate of convergence in the classical renewal theorem so it is easier to apply his results by initially working with the sums $S_i = -\log W_i$ and $\tilde{S}_i = -\log \tilde{W}_i$. We set $K(r) = \inf\{n \geq 0: S_n > r\}$ with a similar definition for $\tilde{K}(r)$.

Ney constructs a joint realization of $\{S_i\}$ and $\{\tilde{S}_i\}$ on the same probability space with a random variable, τ, such that

$$S_{K(r)} I_{\{\tau \leq r\}} =_d \tilde{S}_{\tilde{K}(r)} I_{\{\tau \leq r\}} \tag{2.16}$$

and

$$P(\tau > r) \leq CP\left(\sum_{i=1}^{N} Z_i > r + \log x \right), \tag{2.17}$$

where N and $\{Z_i\}_{i=1}^{\infty}$ are independent. $\{Z_i\}$ are i.i.d. with $P(Z_i > r) \leq C\int_r^{\infty} 1 - \hat{G}(y)\,dy$ and $P(N \geq i) \leq C\delta^i$ for some $0 < \delta < 1$. [Equation (2.17) follows from Lemma 2.1 of Ney's paper.]

$$1 - \hat{G}(y) = \hat{F}(e^{-y}-) \leq \int_0^{e^{-y}} 2u\,dF(u) \leq 2e^{-y}.$$

Consequently, $P(Z_i > r) \leq Ce^{-r}$ and

$$P\left(\sum_{i=1}^{N} Z_i > r \right) \leq e^{-\gamma r} \sum_{i=1}^{\infty} \left(E(\exp(\gamma Z_1)) \right)^i P(N = i) \leq Ce^{-\gamma r}$$

for $0 < \gamma \leq 1$ sufficiently small.

We next set $r = -\log z$ in (2.17) and convert (2.16) to a statement about W_J and $\tilde{W}_{\tilde{J}}$. From our computations above

$$W_J I_{\{\tau \leq -\log z\}} =_d \tilde{W}_{\tilde{J}} I_{\{\tau \leq -\log z\}} \tag{2.18}$$

and

$$P(\tau > -\log z) \leq Cx^{-\gamma} z^{\gamma}. \tag{2.19}$$

Next,

$$\begin{aligned}
&\left| x^{-1} m(x^{\alpha} t) - m(t) \right| \\
&\quad \leq E\left(W_J^{-1} m(W_J^{\alpha} t) - m(t) \right) \\
&\qquad + E\left(\tilde{W}_{\tilde{J}}^{-1} m(\tilde{W}_{\tilde{J}}^{\alpha} t) - x^{-1} m(x^{\alpha} t) \right) \\
&\qquad + \left| E\left(W_J^{-1} m(W_J^{\alpha} t) - \tilde{W}_{\tilde{J}}^{-1} m(\tilde{W}_{\tilde{J}}^{\alpha} t) \right) \right| \\
&\quad \leq Cz^{-1} m'(t) + E\left(\left| W_J^{-1} m(W_J^{\alpha} t) - \tilde{W}_{\tilde{J}}^{-1} m(\tilde{W}_{\tilde{J}}^{\alpha} t) \right| I_{\{\tau > -\log z\}} \right) \\
&\quad \leq Cz^{-1} m'(t) + E\left(W_J^{-1} m(W_J^{\alpha} t) I_{\{\tau > -\log z\}} \right) \\
&\qquad + E\left(\tilde{W}_{\tilde{J}}^{-1} m(\tilde{W}_{\tilde{J}}^{\alpha} t) I_{\{\tau > -\log z\}} \right)
\end{aligned}$$

by (2.15) and (2.18). $m(t) \leq C(1 + t^{1/\alpha})$ and

$$E\left(W_J^{-1} \right) = E(\exp(S_J)) = z^{-1} E(\exp(S_J + \log z)) \leq Cz^{-1}$$

by Lemma 2.2, so

$$\begin{aligned}
E\left(W_J^{-1} m(W_J^{\alpha} t) I_{\{\tau > -\log z\}} \right) &\leq C\left(E\left(W_J^{-1} \right) + t^{1/\alpha} P(\tau > -\log z) \right) \\
&\leq C(z^{-1} + x^{-\gamma} z^{\gamma} t^{1/\alpha})
\end{aligned}$$

from (2.19), and a similar calculation holds for the $\tilde{W}$ term. We have

$$|x^{-1}m(x^{\alpha}t) - m(t)| \le C(z^{-1}m'(t) + x^{-\gamma}z^{\gamma}t^{1/\alpha}),$$

and finally replacing z by $t^{-\rho}$ for $0 < \rho < 1$ gives

$$|x^{-1}m(x^{\alpha}t) - m(t)| \le C(t^{\rho}m'(t) + x^{-\gamma}t^{-\rho\gamma+1/\alpha})$$

for $0 < t^{-\rho} \le x < 1$ and

$$\begin{aligned}
\int_0^1 |h(x,t)|\,dF(x) &\le \int_0^1 x|x^{-1}m(x^{\alpha}t) - m(t)| \\
&\qquad + (1-x)\left|(1-x)^{-1}m((1-x)^{\alpha}t) - m(t)\right| dF(x) \\
&= 2\int_0^1 x|x^{-1}m(x^{\alpha}t) - m(t)|\,dF(x) \\
&\le C\Big(\int_0^{t^{-\rho}} (1 + xt^{1/\alpha})\,dF(x) \\
&\qquad + \int_{t^{-\rho}}^1 x(t^{\rho}m'(t) + x^{-\gamma}t^{-\rho\gamma+1/\alpha})\,dF(x)\Big) \\
&\le C(1 + t^{\rho}m'(t) + t^{-\rho\gamma+1/\alpha}).
\end{aligned} \tag{2.20}$$

The lemma is now proved for $0 < \alpha < 1$ since $m'(t) \le C(1 + t^{1/\alpha-1})$. If $1 \le \alpha$, $m'(t)$ is nonincreasing and we bound $m'(W_i^{\alpha}t)$ by 1 in the argument leading to (2.14) and (2.15). (2.15) and (2.20) remain true with $m'(t)$ replaced by 1. The lemma follows from (2.20) by taking $0 < \rho < 1/\alpha$. □

Since we have made no attempt to compute the largest possible θ in Lemma 2.5, in the computations that follow we assume $0 < \theta\alpha < 1$. By Lemma 2.5 and our earlier bound on $f(x,t)$, we have

$$\int_0^1 |f(x,t)h(x,t)|\,dF(x) \le C(1 + t^{2/\alpha-\theta}).$$

We let $B(t) = C(1 + t^{2/\alpha-\theta})$. Since $m(t)m'(t) > 0$, by a simple comparison test $v(t) \le w(t)$ where $w(t)$ satisfies

$$\begin{aligned}
w'(t) &= -w(t) + \int_0^1 x^{-1}w(x^{\alpha}t)\,d\hat{F}(x) + B(t), \\
w(0) &> 0.
\end{aligned} \tag{2.21}$$

$w'(t) > 0$ as $w'(0) = w(0) + C > 0$, and if $\inf\{t\colon w'(t) = 0\} = s < \infty$, then

$$w''(s) = \int_0^1 x^{\alpha-1}w'(x^{\alpha}s)\,d\hat{F}(x) + B'(s) > 0,$$

which is a contradiction. Replacing $w'(t)$ by zero and making the substitutions $t = e^{\alpha u}$, $y = -\log x$, $g(u) = e^{-u}w(e^{\alpha u})$, and $h(u) = e^{-u}B(e^{\alpha u})$ in (2.21), we obtain

$$g(u) \le \int_0^{\infty} g(u-y)\,d\hat{G}(y) + h(u).$$

By the argument used to obtain (2.7), we have

$$g(u) \le E\big(g(u - S_{N(u)})\big) + E\left(\sum_{k=0}^{N(u)-1} h(u - S_k)\right)$$
$$= J(u) + H(u).$$

By the same renewal arguments used to analyze (2.7), $\lim_{u\to\infty} J(u)$ exists and is finite and $H(u) \sim Ce^{(1-\alpha\theta)u}$. This shows that $w(e^{\alpha u}) \le Ce^{(2-\alpha\theta)u}$ as $u \to \infty$, which in turn shows $v(t) \le Ct^{2/\alpha-\theta}$ as $t \to \infty$. This completes the proof of (2.2).

We are finally ready to show (2.3) which follows easily from (2.1) and (2.2). For any $\varepsilon > 0$,

$$P\left(\left|\frac{Z(t)}{m(t)} - 1\right| > \varepsilon\right) \le \frac{v(t)}{\varepsilon^2 m^2(t)} \le Ct^{-\theta}. \tag{2.22}$$

We let $\lambda > 0$ be such that $\lambda\theta > 1$. By the Borel–Cantelli lemma and (2.22), we have

$$\lim_{n\to\infty} \frac{Z(n^\lambda)}{m(n^\lambda)} = 1 \quad \text{a.s.} \tag{2.23}$$

If $n^\lambda \le t \le (n+1)^\lambda$, then

$$\frac{Z(n^\lambda)}{m\big((n+1)^\lambda\big)} \le \frac{Z(t)}{m(t)} \le \frac{Z\big((n+1)^\lambda\big)}{m(n^\lambda)}. \tag{2.24}$$

Since $m(t) \sim Kt^{1/\alpha}$, $\lim_{n\to\infty} m((n+1)^\lambda)/m(n^\lambda) = 1$. So by (2.23) and (2.24), we have

$$\lim_{t\to\infty} \frac{Z(t)}{m(t)} = 1 \quad \text{a.s.}$$

which together with (2.1) implies (2.3).

3. Explosion argument for $-\infty < \alpha < 0$. This section provides the details of the explosion argument sketched in Section 1 which leads to the conclusion $n^{-1}N_n(\cdot)$ almost surely converges weakly to a random point mass for $\alpha < 0$.

$T_n = \inf\{t \ge 0\colon Z(t) = n + 1\}$ are the split times and X_n the split point added to $[0,1]$ at time T_n. Set $T_\infty = \sup T_n = \sup\{t \ge 0\colon Z(t) < \infty\}$. For any subinterval I, determined by the split points, set $T_\infty(I) =$ the first time after the creation of I that the number of subintervals contained in I is infinite. The proof now proceeds in steps.

STEP 1. $P(T_\infty < \infty) = 1$ and T_∞ has a continuous distribution.

PROOF. Let $0 < S_1 < S_2 < \cdots$ be the successive split times of the leftmost interval.

$$S_n =_d \xi_0 + \sum_{k=1}^{n-1} W_k^{-\alpha}\xi_k,$$

where $\{\xi_k\}$ are mean one exponential, $W_k = \prod_{i=1}^{k} V_i$ where $\{V_i\}$ are i.i.d. with distribution function F on $(0,1)$, and $\{\xi_k\}$ and $\{V_i\}$ are independent.

$$T_\infty \le \lim_{n\to\infty} S_n = \xi_0 + \sum_{k=1}^{\infty} W_k^{-\alpha}\xi_k,$$

$$ET_\infty \le 1 + \sum_{k=1}^{\infty} \left(E(V_1^{-\alpha})\right)^k < \infty.$$

The second assertion is proved by noting that

$$T_\infty = T_1 + \min\{T_\infty([0, X_1]), T_\infty([X_1, 1])\}.$$

The summands are independent and T_1 has an exponential distribution so T_∞ has a continuous distribution. □

STEP 2. Construct a nested sequence of closed intervals as follows: $I_0 = [0,1]$ and if I_n splits into I_n' and I_n'', $I_{n+1} = I_n'$ if $T_\infty(I_n') < T_\infty(I_n'')$, and $= I_n''$ otherwise. This construction is well defined by step 1 since $L^\alpha(I)T_\infty(I) =_d T_\infty$ for $L(I) =$ the length of I. Let $\{X_\infty\} = \cap I_n$ which consists of only one point since $L(I_n) \to 0$.

STEP 3. Given $\varepsilon > 0$, we set m so that $L(I_m) < \varepsilon$. By the construction in step 2, only finitely many splits occur outside I_m before T_∞. Let T be the last such split time and N be the first split time after T. Then for $n \ge N$, X_n belongs to I_m and $|X_n - X_\infty| < \varepsilon$. This proves that $X_n \to X_\infty$ a.s. and $n^{-1}N_n(\cdot)$ almost surely converges weakly to the measure with mass one at X_∞.

REFERENCES

DUBINS, L. E. and FREEDMAN, D. A. (1967). Random distribution functions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2** 183–214. Univ. California Press.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

GRAF, S., MAULDIN, R. D. and WILLIAMS, S. (1986). Random homeomorphisms. *Adv. in Math.* To appear.

KAKUTANI, S. (1975). A problem of equidistribution on the unit interval [0, 1]. *Lecture Notes in Math.* **541** 369–376. Springer, Berlin.

LOOTGIETER, J. C. (1977). Sur la repartition des suites de Kakutani I. *Ann. Inst. H. Poincaré Sect. B* (*N.S.*) **13** 385–410.

LOOTGIETER, J. C. (1978). Sur la repartition des suites de Kakutani II. *Ann. Inst. H. Poincaré Sect. B* (*N.S.*) **14** 279–302.

LOOTGIETER, J. C. (1981). Sur une conjecture de L. E. Dubins. *Ann. Inst. H. Poincaré Sect. B* (*N.S.*) **17** 97–122.

NEY, P. (1981). A refinement of the coupling method in renewal theory. *Stochastic Process. Appl.* **11** 11–26.

PEYRIERE, J. (1979). A singular random measure generated by splitting [0, 1]. *Z. Wahrsch. verw. Gebiete* **47** 289–297.

ROSS, S. M. (1983). *Stochastic Processes*. Wiley, New York.

VAN ZWET, W. R. (1978). A proof of Kakutani's conjecture on random subdivision of longest intervals. *Ann. Probab.* **6** 133–137.

DEPARTMENT OF MATHEMATICS
UTAH STATE UNIVERSITY
LOGAN, UTAH 84322-4125

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853