Elect. Comm. in Probab. **5** (2000) 77–84

ELECTRONIC COMMUNICATIONS in PROBABILITY

A CHARACTERIZATION OF THE SET OF FIXED POINTS OF THE QUICKSORT TRANSFORMATION

JAMES ALLEN FILL [1]
Department of Mathematical Sciences, The Johns Hopkins University
34th and Charles Streets, Baltimore, MD 21218-2682 USA
email: `jimfill@jhu.edu`

SVANTE JANSON
Department of Mathematics, Uppsala University
P. O. Box 480, S-751 06 Uppsala, SWEDEN
email: `svante.janson@math.uu.se`

submitted May 19, 2000 *Final version accepted* May 26, 2000

AMS 1991 Subject classification: Primary 68W40; secondary 60E05, 60E10, 68P10.
`Quicksort`, fixed point, characteristic function, smoothing transformation, domain of attraction, coupling, integral equation.

Abstract

The limiting distribution μ of the normalized number of key comparisons required by the `Quicksort` *sorting algorithm is known to be the unique fixed point of a certain distributional transformation T—unique, that is, subject to the constraints of zero mean and finite variance. We show that a distribution is a fixed point of T if and only if it is the convolution of μ with a Cauchy distribution of arbitrary center and scale. In particular, therefore, μ is the unique fixed point of T having zero mean.*

1 Introduction, motivation, and summary

Let $\mathcal{M}$ denote the class of all probability measures (distributions) on the real line. This paper concerns the transformation T defined on $\mathcal{M}$ by letting $T\nu$ be the distribution of

$$UZ + (1-U)Z^* + g(U),$$

where U, Z, and Z^* are independent, with $Z \sim \nu$, $Z^* \sim \nu$, and $U \sim \text{unif}(0,1)$, and where

$$g(u) := 2u \ln u + 2(1-u)\ln(1-u) + 1. \tag{1.1}$$

Of course, T can be regarded as a transformation on the class of characteristic functions ψ of elements of $\mathcal{M}$. With this interpretation, T takes the form

$$(T\psi)(t) = \mathbf{E}\big[\psi(Ut)\,\psi((1-U)t)\,\exp[itg(U)]\big] = \int_{u=0}^{1} \psi(ut)\,\psi((1-u)t)\,e^{itg(u)}\,du, \quad t \in \mathbf{R}.$$

It is well known [7] that (i) among distributions with zero mean and finite variance, T has a unique fixed point, call it μ; and (ii) if C_n denotes the random number of key comparisons

[1] Research supported by NSF grant DMS–9803780, and by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.

required by the algorithm `Quicksort` to sort a file of n records, then the distribution of $(C_n - \mathbf{E}C_n)/n$ converges weakly to μ.

There are other fixed points. For example, it has been noted frequently in the literature that the location family generated by μ is a family of fixed points. But there are many more fixed points, as we now describe. Define the Cauchy(m, σ) distribution (where $m \in \mathbf{R}$ and $\sigma \geq 0$) to be the distribution of $m + \sigma C$, where C has the standard Cauchy distribution with density $x \mapsto [\pi(1+x^2)]^{-1}$, $x \in \mathbf{R}$; equivalently, Cauchy(m, σ) is the distribution with characteristic function $e^{imt-\sigma|t|}$. [In particular, the Cauchy$(m, 0)$ distribution is unit mass at m.] Now let $\mathcal{F}$ denote the class of all fixed points of T, and let $\mathcal{C}$ denote the class of convolutions of μ with a Cauchy distribution. Using characteristic functions it is easy to check that $\mathcal{C} \subseteq \mathcal{F}$, and that all of the distributions in $\mathcal{C}$ are distinct. In this paper we will prove that, conversely, $\mathcal{F} \subseteq \mathcal{C}$, and thereby establish the following main result.

Theorem 1.1. *The class $\mathcal{F}$ equals $\mathcal{C}$. That is, a measure ν is a fixed point of the* `Quicksort` *transformation T if and only if it is the convolution of the limiting* `Quicksort` *measure μ with a Cauchy distribution of arbitrary center m and scale σ. In particular, $\mathcal{F}$ is in one-to-one correspondence with the set $\{(m, \sigma) : m \in \mathbf{R},\ \sigma \geq 0\}$.*

The following corollary is immediate and strengthens Rösler's [7] characterization of μ as the unique element of $\mathcal{F}$ having zero mean and finite variance.

Corollary 1.2. *The limiting* `Quicksort` *measure μ is the unique fixed point of the* `Quicksort` *transformation T having finite expectation equal to* 0. $\square$

The present paper can be motivated in two ways. First, the authors are writing a series of papers refining and extending Rösler's [7] probabilistic analysis of `Quicksort`. No closed-form expressions are known for any of the standard functionals (e.g., characteristic function, distribution function, density function) associated with μ; information to be obtained about μ must be read from the fixed-point identity it satisfies. We were curious as to what extent additional known information about μ, such as the fact that it has everywhere finite moment generating function, must be brought to bear. As one example, it is believed that the continuous Lebesgue density f (treated in [3]) for μ decays at least exponentially quickly to 0 at $\pm\infty$, cf. [5]. But we now know from Theorem 1.1 that there can be no proof for this conjecture that solely makes use of the information that $\mu \in \mathcal{F}$.

Second, we view the present paper as a pilot study of fixed points for a *class* of distributional transformations on the line. In the more general setting, we would be given (the joint distribution of) a sequence $(A_i : i \geq 0)$ of random variables and would define a transformation T on $\mathcal{M}$ by letting $T\nu$ be the distribution of $A_0 + \sum_{i=1}^{\infty} A_i Z_i$, where $Z_1, Z_2, \ldots$ are independent random variables with distribution ν. [To ensure well-definedness, one might (for example) require that (almost surely) $A_i \neq 0$ for only finitely many values of i.] For probability measures ν on $[0, \infty)$, rather than on $\mathbf{R}$, and with the additional restrictions that $A_0 = 0$ and $A_i \geq 0$ for all $i \geq 1$, such transformations are called *generalized smoothing transformations*. These have been thoroughly studied by Durrett and Liggett [2], Guivarc'h [4], and Liu [6], and by other authors; consult the three papers we have cited here for further bibliographic references. Generalized smoothing transformations have applications to interacting particle systems, branching processes and branching random walk, random set constructions, and statistical turbulence theory. The arguments used to characterize the set of fixed points for generalized smoothing transformations make heavy use of Laplace transforms; unfortunately, these arguments do not carry over readily to distributions on the line. Other authors (see, e.g., [8] [9] [10]) have treated fixed points of transformations of measures ν on the whole line as discussed above, but not without finiteness conditions on the moments of ν.

We now outline our proof of Theorem 1.1. Let ψ be the characteristic function of a given $\nu \in \mathcal{F}$, and let $r(t) := \psi(t) - 1$, $t \in \mathbf{R}$. In Section 2 we establish and solve (in a certain sense) an integral equation satisfied by r. In Section 3 we then use the method of successive substitutions to derive asymptotic information about $r(t)$ as $t \downarrow 0$, showing first that $r(t) = O(t^{2/3})$, next that $r(t) = \beta t + o(t)$ for some $\beta = -\sigma + im \in \mathbf{C}$ with $\sigma \geq 0$, and finally that $r(t) = \beta t + O(t^2)$. In

Section 4 we use this information to argue that there exist random variables $Z_1 \sim \nu$, $Z_2 \sim \mu$, and $C \sim \text{Cauchy}(m, \sigma)$ such that $Z_1 = Z_2 + C$. We finish the proof by showing that one can take Z_2 and C to be independent, whence $\nu \in \mathcal{C}$.

Remark 1.3. Neither the proof in [7] of uniqueness of the fixed point μ among distributions with zero mean and finite variance nor any of the proofs in the present paper rely on any properties of the function g beyond zero mean and finite variance.

2 An integral equation

Let ψ denote the characteristic function of a given $\nu \in \mathcal{F}$. Since $\psi(-t) \equiv \overline{\psi(t)}$, we shall only need to consider $\psi(t)$ for $t \geq 0$. For notational convenience, define

$$r(t) := \psi(t) - 1, \quad t \geq 0.$$

Rearranging the fixed-point integral equation $(T\psi)(t) \equiv \psi(t)$, we obtain the following result.

Lemma 2.1. *The function r satisfies the integral equation*

$$r(t) = 2 \int_{u=0}^{1} r(ut)\, du + b(t), \quad t \geq 0,$$

where

$$b(t) := \int_{u=0}^{1} r(ut)\, r((1-u)t)\, du + it \int_{u=0}^{1} [\psi(ut)\, \psi((1-u)t) - 1]\, g(u)\, du + a(t) \tag{2.1}$$

with

$$\left| a(t) := \int_{u=0}^{1} \psi(ut)\, \psi((1-u)t)\, [e^{itg(u)} - 1 - itg(u)]\, du \right| \leq \tfrac{1}{2} \mathbf{E} g^2(U) t^2 = (\tfrac{7}{6} - \tfrac{1}{9}\pi^2) t^2. \tag{2.2}$$

□

Note that r and b are continuous on $[0, \infty)$, with $r(0) = 0 = b(0)$. Regarding b as "known", the integral equation in Lemma 2.1 is easily "solved" for r:

Proposition 2.2. *For some constant $c \in \mathbf{C}$, we have*

$$\frac{r(t)}{t} = c - 2 \int_{v=t}^{1} \frac{b(v)}{v^2}\, dv + \frac{b(t)}{t}, \quad t > 0.$$

Proof. Setting $h(t) := t[r(t) - b(t)]$, Lemma 2.1 implies

$$h(t) = 2 \int_{v=0}^{t} \left[\frac{h(v)}{v} + b(v) \right] dv, \quad t > 0.$$

Thus h is continuously differentiable on $(0, \infty)$ and satisfies the differential equation

$$h'(t) = \tfrac{2}{t} h(t) + 2b(t)$$

there. This is an easy differential equation to solve for h, and we find that

$$h(t) = ct^2 - 2t^2 \int_{v=t}^{1} \frac{b(v)}{v^2}\, dv, \quad t > 0,$$

for some $c \in \mathbf{C}$. After rearrangement, the proposition is proved. □

3 Behavior of r near 0

We now proceed in stages, using Proposition 2.2 as our basic tool, to get ever more information about the behavior of r (especially near 0).

Lemma 3.1. *Let $\psi \equiv 1 + r$ denote the characteristic function of a given $\nu \in \mathcal{F}$. Then there exists a constant $C < \infty$ such that*

$$|r(t)| \leq Ct^{2/3} \text{ for all } t \geq 0.$$

Proof. Let

$$M(t) := \max\{|r(s)| : 0 \leq s \leq t\} \leq 2, \quad t \geq 0.$$

From (2.1) and (2.2), we see immediately that, for $0 < t \leq 1$,

$$|b(t)| \leq M^2(t) + O(t).$$

Therefore, for $0 < t < 1$, Proposition 2.2 yields

$$|r(t)| \leq M^2(t) + 2t \int_{v=t}^{1} \frac{M^2(v)}{v^2}\, dv + \epsilon(t) = M^2(t) + 2 \int_{u=t}^{1} M^2(t/u)\, du + \epsilon(t),$$

where

$$\epsilon(t) = O\left(t \log\left(\tfrac{1}{t}\right)\right) + O(t) = O(t^{2/3}).$$

Consequently, again for $0 < t < 1$ (but then trivially for all $t \geq 0$),

$$M(t) \leq M^2(t) + 2 \int_{u=0}^{1} M^2(t/u)\, du + O(t^{2/3}).$$

Fix $0 < a < 1$; later in the proof we shall see that $a = 1/8$ suffices for our purposes. Since $M(t) \to 0$ as $t \to 0$, we can choose $t_0 > 0$ such that $M(t_0) \leq a$. Then, for $0 \leq t \leq t_0$,

$$\begin{aligned} M(t) &\leq M^2(t) + 2 \int_{u=0}^{t/t_0} M^2(t/u)\, du + 2 \int_{u=t/t_0}^{1} M^2(t/u)\, du + O(t^{2/3}) \\ &\leq aM(t) + 8\frac{t}{t_0} + 2a \int_{u=t/t_0}^{1} M(t/u)\, du + O(t^{2/3}) \end{aligned}$$

and thus

$$M(t) \leq \frac{2a}{1-a} \int_{u=0}^{1} M(t/u)\, du + O(t^{2/3}).$$

Since M is bounded, this is trivially true also for $t > t_0$. Summarizing, for some constant $\tilde{C} < \infty$ we have, with $U \sim \text{unif}(0, 1)$,

$$M(t) \leq \frac{2a}{1-a} \mathbf{E} M(t/U) + \tilde{C} t^{2/3}, \quad t \geq 0. \tag{3.1}$$

Now fix the value of a to be any number in $(0, 1/7)$, say $a = 1/8$. Then a straightforward induction [substituting (3.2) into (3.1) for the induction step] shows that for any nonnegative integer n we have, for all $t \geq 0$,

$$M(t) \leq \left(\frac{2a}{1-a}\right)^n \mathbf{E}\, M\left(\frac{t}{U_1 \ldots U_n}\right) + \frac{1-a}{1-7a} \tilde{C} t^{2/3}. \tag{3.2}$$

Recalling that M is bounded and letting $n \to \infty$, we obtain the desired conclusion, with $C := \frac{1-a}{1-7a} \tilde{C}$. ∎

Lemma 3.2. *Let $\psi \equiv 1 + r$ denote the characteristic function of a given $\nu \in \mathcal{F}$, and define b by* (2.1)*. Then*

$$r(t) = (c - 2J)t + o(t) \text{ as } t \downarrow 0,$$

where J is the absolutely convergent integral

$$J := \int_{v=0}^{1} \frac{b(v)}{v^2}\, dv. \tag{3.3}$$

Proof. Combining (2.1)–(2.2) and Lemma 3.1, we obtain

$$|b(t)| \le O(t^{4/3}) + O(t^{1+(2/3)}) + O(t^2) = O(t^{4/3}).$$

Thus the integral J converges absolutely, and from Proposition 2.2 we obtain the desired conclusion about r. □

Lemma 3.2 is all we will need in the next section, but the following refinement follows readily and takes us as far as we can go with the method of successive substitutions.

Corollary 3.3. *Let ψ denote the characteristic function of a given $\nu \in \mathcal{F}$. Then there exists a constant $\beta = im - \sigma \in \mathbf{C}$ with $\sigma \ge 0$ such that*

$$\psi(t) = 1 + \beta t + O(t^2) \text{ as } t \downarrow 0.$$

Proof. Combining (2.1)–(2.2) and Lemma 3.2, we readily obtain $b(t) = O(t^2)$. Therefore, by Proposition 2.2,

$$\psi(t) - 1 = r(t) = (c - 2J)t + 2t \int_{v=0}^{t} \frac{b(v)}{v^2}\, dv + b(t) = \beta t + O(t^2),$$

with $\beta = im - \sigma := c - 2J$. Since $|\psi(t)| \le 1$ for all t, we must have $\sigma \ge 0$. □

4 Proof of the main theorem

4.1 Further preliminaries

In Sections 4.1–4.2 we complete the proof of our main Theorem 1.1. To do this, we begin with a key result that any characteristic function with expansion as in Corollary 3.3 [more generally, we allow the remainder term there to be simply $o(t)$] is in the domain of attraction of (iterates of) the "homogeneous" analogue T_0 of T. (Here $\Longrightarrow$ denotes weak convergence of probability measures.)

Theorem 4.1. *Let ψ be any characteristic function satisfying*

$$\psi(t) = 1 + \beta t + o(t) = 1 + imt - \sigma t + o(t) \text{ as } t \downarrow 0 \tag{4.1}$$

for some $\beta = im - \sigma \in \mathbf{C}$, with $m \in \mathbf{R}$ and $\sigma \ge 0$. Let ν be the corresponding probability measure. Then

$$T_0^n \nu \Longrightarrow \text{Cauchy}(m, \sigma),$$

where T_0 is the homogeneous analogue of the `Quicksort` *transformation T mapping distributions as follows (in obvious notation):*

$$T_0 : Z \mapsto UZ + (1 - U)Z^*. \tag{4.2}$$

Proof. Let $Z_1, Z_2, \dots; U_1, U_2, \dots$ be independent random variables, with every $Z_i \sim \nu$ and every $U_j \sim \text{unif}(0,1)$. Then, using the definition of T_0 repeatedly,

$$W_n := \sum_{i=1}^{2^n} V_i^{(n)} Z_i \sim T_0^n \nu, \qquad n \geq 0,$$

where we define the random variables $V_i^{(n)}$ as follows. Using U_1 in the obvious fashion, split the unit interval into intervals of lengths U_1 and $1-U_1$. Now using U_2 and U_3, split the first interval into subintervals of lengths U_1U_2 and $U_1(1-U_2)$ and the second interval into subintervals of lengths $(1-U_1)U_3$ and $(1-U_1)(1-U_3)$. Continue in this way (using $U_1, \dots, U_{2^n-1}$) until the unit interval has been divided overall into 2^n subintervals. Call their lengths, from left to right, $V_1^{(n)}, \dots, V_{2^n}^{(n)}$.

Let $L_n := \max(V_1^{(n)}, \dots, V_{2^n}^{(n)})$. We show that L_n converges in probability to 0 as $n \to \infty$. Luckily, the complicated dependence structure of the variables $V_i^{(n)}$ does not come into play; the only observation we need is that that each $V_i^{(n)}$ marginally has the same distribution as $U_1 \cdots U_n$. Indeed, abbreviate $V_1^{(n)}$ as V_n; briefly put, we derive a Chernoff's bound for $\ln(1/V_n)$ and then simply use subadditivity. To spell things out, let $x > 0$ be fixed and let $t \geq 0$. Then

$$\mathbf{P}(V_n \geq e^{-x}) \leq e^{tx}\mathbf{E}V_n^t = e^{tx}\prod_{j=1}^{n}\mathbf{E}U_j^t = e^{tx}(1+t)^{-n} = \exp[-(n\ln(1+t) - xt)].$$

Choosing the optimal $t = \frac{n}{x} - 1$ (valid for $n \geq x$), this yields

$$\mathbf{P}(V_n \geq e^{-x}) \leq \exp[-(n\ln(n/x) - n + x)] = \exp[-(n(\ln n - \ln(ex)) + x)]$$

and thus

$$\mathbf{P}(L_n \geq e^{-x}) \leq 2^n \exp[-(n(\ln n - \ln(ex)) + x)] = \exp[-(n(\ln n - \ln(2ex)) + x)] \to 0$$

as $n \to \infty$.

Since L_n converges in probability to 0, we can therefore choose $\epsilon_n \to 0$ so that $\mathbf{P}(L_n > \epsilon_n) \to 0$. To prove the theorem, it then suffices to prove

$$\widetilde{W}_n := \mathbf{1}(L_n \leq \epsilon_n) W_n \implies \text{Cauchy}(m, \sigma).$$

For this, we note that the characteristic function ϕ_n of $\widetilde{W}_n$ is given for $t \in \mathbf{R}$ by

$$\phi_n(t) = \mathbf{P}(L_n > \epsilon_n) + \mathbf{E}\left[\mathbf{1}(L_n \leq \epsilon_n)\prod_{i=1}^{2^n}\psi(V_i^{(n)}t)\right]. \tag{4.3}$$

We will show that $\phi_n(t)$ converges to $e^{\beta t} = e^{imt - \sigma t}$ for each fixed $t \geq 0$, and [since, further, $\phi_n(-t) \equiv \overline{\phi_n(t)}$] this will complete the proof of the lemma.

Indeed, we need only consider the second term in (4.3). For that, the calculus estimates outlined in the proof of the lemma preceding Theorem 7.1.2 in [1] demonstrate that, when $L_n \leq \epsilon_n$,

$$\prod_{i=1}^{2^n}\psi(V_i^{(n)}t) = (1 + D_n)e^{\beta t}$$

for complex random variables D_n (depending on our fixed choice of $t \geq 0$) satisfying $|D_n| \leq \delta_n$ for a deterministic sequence δ_n [$\equiv \delta(\epsilon_n t)$] $\to 0$ [with $\delta(s) \to 0$ as $s \to 0$]. [Leaving out the error estimates, the argument is

$$\begin{aligned}\log\left[\prod_{i=1}^{2^n}\psi(V_i^{(n)}t)\right] &\approx \sum_{i=1}^{2^n}\left(\psi(V_i^{(n)}t) - 1\right) \\ &\approx \sum_{i=1}^{2^n}\beta V_i^{(n)}t = \beta t. \qquad]\end{aligned}$$

It now follows easily that $\phi_n(t) \to e^{\beta t}$, as desired. $\square$

Both the next lemma and its immediate corollary (Lemma 4.3) will be used in our proof of Theorem 1.1.

Lemma 4.2. *Let $\nu_i \in \mathcal{F}$, $i = 1, 2$. Suppose that (Z_1, Z_2) is a coupling of ν_1 and ν_2 such that the characteristic function of $Z_1 - Z_2$ satisfies* (4.1). *Then there exists a coupling $(\tilde{Z}_1, \tilde{Z}_2)$ of ν_1 and ν_2 such that $\tilde{Z}_1 - \tilde{Z}_2 \sim \mathrm{Cauchy}(m, \sigma)$.*

Proof. Extend T to a transformation T_2 on the class $\mathcal{M}_2$ of probability measures on $\mathbf{R}^2$ by mapping the distribution $\xi \in \mathcal{M}_2$ of (X, Y) to the distribution $T_2\xi$ of

$$(UX + (1-U)X^* + g(U),\ UY + (1-U)Y^* + g(U)),$$

where U, (X, Y), and (X^*, Y^*) are independent, with $(X, Y) \sim \xi$, $(X^*, Y^*) \sim \xi$, and $U \sim \mathrm{unif}(0,1)$, and where g is given by (1.1). (Note that we use the *same* uniform U for the Ys as for the Xs!) Of course, T_2 maps the marginal distributions $\xi_1(\cdot) = \xi(\cdot \times \mathbf{R})$ of X and $\xi_2(\cdot) = \xi(\mathbf{R} \times \cdot)$ of Y into $T\xi_1$ and $T\xi_2$, respectively; more importantly for our purposes, it maps the distribution, call it $\hat{\xi}$, of $X - Y$ into the distribution $T_0\hat{\xi}$, with T_0 defined at (4.2).
Now let $\nu \in \mathcal{M}_2$ have marginals ν_i, $i = 1, 2$. Then $(T_2^n \nu)_{n \geq 1}$ has constant marginals (ν_1, ν_2) as n varies and so is a tight sequence. We then can find a weakly convergent subsequence, say,

$$T_2^{n_k} \nu \implies \nu^\infty \in \mathcal{M}_2;$$

of course, the limit ν^∞ again has marginals ν_i, $i = 1, 2$. Moreover,

$$T_0^{n_k} \hat{\nu} = \widehat{T_2^{n_k} \nu} \implies \widehat{\nu^\infty}.$$

But, by supposition, the characteristic function of $\hat{\nu}$ satisfies (4.1), so Theorem 4.1 implies that $\widehat{\nu^\infty}$ is $\mathrm{Cauchy}(m, \sigma)$. Thus $\nu^\infty \in \mathcal{M}_2$ supplies the desired coupling. $\square$

Lemma 4.3. *Let $\nu_i \in \mathcal{F}$, $i = 1, 2$. Suppose that (Z_1, Z_2) is a coupling of ν_1 and ν_2 such that $Z_1 - Z_2$ has zero mean and finite variance. Then $\nu_1 = \nu_2$.* $\square$

4.2 The proof

We now complete the proof of Theorem 1.1.

Proof. As discussed in Section 1, it is simple to check that $\mathcal{C} \subseteq \mathcal{F}$ (and that the elements of $\mathcal{C}$ are all distinct).
Conversely, given $\nu \in \mathcal{F}$, let $Z_1 \sim \nu_1 := \nu$ and $Z_2 \sim \nu_2 := \mu$ be independent random variables (on some probability space); recall that μ is the limiting `Quicksort` measure, with zero mean and finite variance. Write ψ_i, $i = 1, 2$, for the characteristic functions corresponding respectively to ν_i, $i = 1, 2$. By Lemma 3.2 (or see Corollary 3.3), ψ_1 satisfies (4.1) [for some (m, σ)]. Of course, ψ_2 satisfies (4.1) with β taken to be 0, so the characteristic function $t \mapsto \psi_1(t)\psi_2(-t)$ of $Z_1 - Z_2$ satisfies (4.1) for the same (m, σ) as for ψ_1. Applying Lemma 4.2, there exists a coupling $(\tilde{Z}_1, \tilde{Z}_2)$ of ν_1 and ν_2 such that $C := \tilde{Z}_1 - \tilde{Z}_2 \sim \mathrm{Cauchy}(m, \sigma)$. Without loss of generality (by building a suitable product space), we may assume the existence of a random variable $Y \sim \mu$ on the same probability space as $\tilde{Z}_1$ and $\tilde{Z}_2$ such that Y and C are independent.
We know that the distribution ν_1 of $\tilde{Z}_1 = \tilde{Z}_2 + C$ is a fixed point of T. But so is the distribution $\nu_1' \in \mathcal{C}$ of $Z := Y + C$. By Lemma 4.3 applied to $(\tilde{Z}_1, Z)$, $\nu = \nu_1 = \nu_1' \in \mathcal{C}$, as desired. $\square$

Acknowledgment. We thank Persi Diaconis for pointing us to the literature on the smoothing transformation.

References

[1] Chung, K. L. *A Course in Probability Theory.* Second edition. Academic Press, New York, 1974.

[2] Durrett, R. and Liggett, T. M. Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete* **64** (1983), 275–301.

[3] Fill, J. A. and Janson, S. Smoothness and decay properties of the limiting Quicksort density function. Refereed article, to apppear in a book edited by D. Gardy and A. Mokkadem, published in 2000 by Birkhäuser, and based on the Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities (University of Versailles–St. Quentin, Versailles, France, September, 2000). Preprint available from `http://www.mts.jhu.edu/~fill/` or `http://www.math.uu.se/~svante/`.

[4] Guivarc'h, Y. Sur une extension de la notion de loi semi-stable. *Ann. Inst. H. Poincaré Probab. Statist.* **26** (1990), 261–285.

[5] Knessl, C. and Szpankowski, W. Quicksort algorithm again revisited. *Discrete Math. Theor. Comput. Sci.* **3** (1999), 43–64.

[6] Liu, Q. Fixed points of a generalized smoothing transformation and applications to the branching random walk. *Adv. in Appl. Probab.* **30** (1998), 85–112.

[7] Rösler, U. A limit theorem for 'Quicksort'. *RAIRO Inform. Théor. Appl.* **25** (1991), 85–100.

[8] Rösler, U. A fixed point theorem for distributions. *Stochastic Process. Appl.* **42** (1992), 195–214.

[9] Rösler, U. The analysis of stochastic divide and conquer algorithms. *Algorithmica* (1999), to appear.

[10] Rösler, U. and Rüschendorf, L. The contraction method for recursive algorithms. *Algorithmica* (1999), to appear.