Adv. Appl. Prob. **39,** 1036–1053 (2007)
Printed in Northern Ireland

TRANSFORMATIONS OF GALTON–WATSON PROCESSES AND LINEAR FRACTIONAL REPRODUCTION

F. C. KLEBANER,* *Monash University*
U. RÖSLER,** *Christian-Albrechts-Universität zu Kiel*
S. SAGITOV,*** *Chalmers University of Technology*

Abstract

By establishing general relationships between branching transformations (Harris–Sevastyanov, Lamperti–Ney, time reversals, and Asmussen–Sigman) and Markov chain transforms (Doob's h-transform, time reversal, and the cone dual), we discover a deeper connection between these transformations with harmonic functions and invariant measures for the process itself and its space–time process. We give a classification of the duals into Doob's h-transform, pathwise time reversal, and cone reversal. Explicit results are obtained for the linear fractional offspring distribution. Remarkably, for this case, all reversals turn out to be a Galton–Watson process with a dual reproduction law and eternal particle or some kind of immigration. In particular, we generalize a result of Klebaner and Sagitov (2002) in which only a geometric offspring distribution was considered. A new graphical representation in terms of an associated simple random walk on $\mathbb{N}^2$ allows for illuminating picture proofs of our main results concerning transformations of the linear fractional Galton–Watson process.

Keywords: Galton–Watson process; Markov chain transform; time reversal; Doob's transform; cone dual; linear fractional generating function; Q-process

2000 Mathematics Subject Classification: Primary 60J80

1. Introduction

A famous model for population growth is the Galton–Watson process (GWP). In genealogy and population genetics it is important to be able to look back in time, and this leads to the consideration of time reversals of this process. Various time reversals of the classical GWP have been given in the literature. The purpose of this paper is to provide a synoptic survey of these and to establish new time reversals. More precisely, the first part of this paper achieves this for the general GWP, while the second part deals with the linear fractional case, where the results are more explicit. Linear fractional distributions are in fact modified geometric distributions; this fact is more suitable for our purposes and we use it throughout the paper. Since the

Received 22 February 2005; revision received 30 June 2007.
* Postal address: School of Mathematical Sciences, Monash University, Clayton, VIC 3800, Australia.
Email address: fima.klebaner@sci.monash.edu.au
Supported by the Australian Research Council grant DP0451657.
** Postal address: Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Ludewig Meyn Strasse 4, 24098 Kiel, Germany. Email address: roesler@math.uni-kiel.de
*** Postal address: School of Mathematical Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden. Email address: serik@math.chalmers.se
Partially supported by the Bank of Sweden Tercentenary Foundation.

GWP is a Markov chain, we review Markov chain (MC) transforms and establish new relationships, which become important for imbedding branching transformations into more general MC transforms. We consider the following branching transformations: Harris–Sevastyanov, Lamperti–Ney, and time reversals, and we show that they are in fact well-known MC transforms: Doob's h-transform, time reversal, and the cone dual. Specifically, we establish that Doob's h-transform is the Harris–Sevastyanov transformation, Doob's h-transform with the space–time harmonic function is the Lamperti–Ney transformation, Doob's h-transform for the space–time process is also a Lyons–Pemantle–Peres (LPP) size-biased GWP process, the Lamperti–Ney transformation is a composition of a Harris–Sevastyanov and the LPP transformations, and the Asmussen–Sigman transformation is the cone dual. For the case of modified geometric offspring distributions, the results are more explicit and show that branching transformations result in branching processes in the same class with possibly eternal particles, and that the time reversal of a GWP with a modified geometric offspring distribution is a Lamperti–Ney process. This generalizes the earlier results of [7], where only a geometric offspring distribution was considered. Construction of the time-reversed process allows us to answer some questions about the original branching population. For example, the question of the age of the population can be answered by considering the hitting time of 1 by the time-reversed process; see [7].

We close this section with the definitions of the GWP, its variants that allow immigration (GWPI), and processes with an eternal particle (GWPE).

In the GWP the number Z_n of individuals in the nth generation is given recursively by $Z_0 = 1$ and, for $n > 0$,

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i},$$

where $\xi_{n,i}$ denotes the offspring of the ith individual in the nth generation; see, e.g. [5, pp. 1–2]. All offspring random variables $\xi_{n,i}$ are independent and identically distributed (i.i.d.). In the GWPI $(Y_n)_n$ the $(n+1)$th generation consists of the offspring of the nth generation plus an immigration variable C_n, i.e.

$$Y_{n+1} = \sum_{i=1}^{Y_n} X_{n,i} + C_n,$$

where the C_ns are i.i.d. and $\mathrm{P}(C_n = 0) = 0$. Alternatively, we may think of one eternal particle with offspring C_n at time n and all other particles with offspring like ξ. Now we should count the eternal particle, $\hat{Y}_n = Y_n + 1$, and obtain the representation

$$\hat{Y}_{n+1} - 1 = \sum_{i=1}^{\hat{Y}_n - 1} X_{n,i} + C_n, \tag{1.1}$$

which we use throughout the paper and refer to as the GWPE.

The paper is organized as follows. In the rest of the introduction we review MC transforms, branching transformations, and time reversals and give the concept of the cone dual. Section 2 contains results on the linear fractional or modified geometric distributions and GWPs with such offspring distributions. Section 3 contains the main results on branching transformations as MC transforms with explicit results for linear fractional processes.

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

1.1. Doob's h-transform

Let $X = (X_n)_{n\in\mathbb{N}_0}$ be a time-homogeneous MC on a countable state space S, and let all states communicate (P_x (there exists an n, $X_n = y) > 0$ for all $x, y \in S$). Let $h\colon S \to \mathbb{R}$ be a strictly positive harmonic function ($\mathrm{P}\,h = h$) for the transition probability kernel $\mathrm{P} = (p(x, y))_{x,y\in S}$. Doob's h-transform [11, p. 327] is an MC with transition probabilities

$$q(x, y) := \frac{h(y)}{h(x)} p(x, y), \tag{1.2}$$

and the semigroup $q^n(y, x) = (h(y)/h(x))p^n(x, y)$.

Doob's h-transform is a measure transform. Let $M = (M_n)_n$ be a positive martingale of expectation 1. Let P denote the measure of the MC, and let P_n denote the restriction of P to the σ-field, $\mathcal{A}_n$, up to time n. Then $Q_n := M_n \mathrm{P}_n$ defines a consistent family of probability measures with a projective limit Q. If the martingale is regular (L^1 convergent) with limit M_∞ then $Q = M_\infty \mathrm{P}$.

In our case, $(h(X_n))_n$ is a positive martingale and the corresponding measure transform is the MC with transitions given in (1.2). If A is a tail event then $h(x) := \mathrm{P}_x(A)$ is a bounded harmonic function and Doob's h-transformed process X^h is the original process conditioned to end in the terminal set A.

For a corresponding intuitive description in the general case we need Martin boundary theory. Let ν denote a Radon measure on the state space. Let $\mathcal{P}$ denote the set of positive potentials (measurable functions g on the state space satisfying $\lim_{n\to\infty} \mathrm{P}^n\, g = 0$ pointwise) integrable with respect to ν, and let $\overline{\mathcal{P}}$ denote the closure with respect to pointwise convergence. Let $\mathcal{H}$ denote the set of ν-integrable harmonic functions ($\mathrm{P}\,h = h \geq 0$). Then $\mathcal{P}$, $\mathcal{H}$, and $\overline{\mathcal{P}}$ are convex, positive cones with a unique integral representation. There exists a bijection of functions g in the convex cone C, and there exist measures μ on the set $E_C \subset C$ of extremals of the cone endowed with the cone topology (here Martin topology of pointwise convergence). The bijection $g \leftrightarrow \mu_g$ can be given by

$$g = \int_{E_C} e\mu_g(\mathrm{d}e).$$

Furthermore, $\overline{\mathcal{P}} = \mathcal{P} \oplus \mathcal{H}$ with a unique representation $\overline{\mathcal{P}} \ni \overline{g} = g + h$ and a unique decomposition $\mu_{\overline{g}} = \mu_g + \mu_h$ with respect to the disjoint union $E_{\overline{\mathcal{P}}} = E_{\mathcal{P}} \cup E_{\mathcal{H}}$.

Martin boundary theory provides a tool for finding the extremals, $E_{\mathcal{H}}$. The extremals for the potentials are the Greens kernel, $G(\cdot, y) = \sum_{n\geq 0} p^n(\cdot, y)$ for $y \in S$. (For simplicity we assume finiteness.) Form the Martin kernel,

$$K(x, y) := \frac{G(x, y)}{G(\nu, y)},$$

where $G(\nu, y) = \int G(s, y)\nu(\mathrm{d}s)$, and consider all limits of the Martin functions $K(\cdot, y)$ with respect to pointwise convergence. The set of Martin limits contains the extremals $\mathcal{H}$ of harmonic functions. In most cases Martin limits are extremals. Finally, we note that the standard normalization is by a point measure ν on some state x_0 and that the Martin kernel becomes $K(x, y) := G(x, y)/G(x_0, y)$.

Returning to the interpretation, Doob's h-transform is the MC, X^h, such that X^h_n converges weakly in Martin topology to the representing measure μ_h. If h is extremal then X^h converges

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

to a single point in the Martin boundary. The intuitive description is that the paths of the transformed process are Martin sequences for the extremal function.

We also use Doob's h-transform for the space–time process defined as follows. Let $(P^{m,n})_{m\le n\in\mathbb{N}_0}$ denote a semigroup of a time-dependent MC on S. The space–time transition probability kernel R is a stochastic matrix on $S\times\mathbb{Z}$ defined by

$$r((x,s),(y,t)) = \mathbf{1}_{\{s+1=t\}}\, p^{s,t}(x,y).$$

The semigroup corresponding to R is given by $r^n((x,s),(y,t)) = \mathbf{1}_{\{s+n=t\}}\, p^{s,t}(x,y)$. The Greens kernel becomes $G((x,s),(y,t)) = \mathbf{1}_{\{s\le t\}}\, p^{s,t}(x,y)$. A space–time harmonic function (parabolic) is a positive function h on space–time $S\times\mathbb{Z}$ satisfying $Rh = h$. Doob's h-transform for the space–time setting becomes

$$q((x,s),(y,t)) = \mathbf{1}_{\{s\le t\}}\,\frac{h(y,t)}{h(x,s)}\,p^{t-s}(x,y).$$

Of special interest for a time-homogeneous MC are factorizing parabolic functions of the special form

$$h(x,s) = g(x)\mathrm{e}^{\lambda s}.$$

Here h is space–time harmonic if and only if the function g on S is a right eigenfunction of P with eigenvalue $\mathrm{e}^{-\lambda}$, i.e. $Pg = \mathrm{e}^{-\lambda}g$. Doob's h-transform becomes

$$q((x,s),(y,s+1)) = \frac{g(y)\mathrm{e}^{\lambda}}{g(x)}\,p(x,y). \tag{1.3}$$

This is a measure transform corresponding to the martingale $h(X_n,n)$. For the GWP, for example, the function $h(x,n) = x/m^n$ is space–time harmonic.

Doob's h-transform is a pathwise transform. This means that the probability of every path $(i_0,i_1,\dots,i_n)$ is changed by a factor depending on $i_0,i_1,\dots,i_n$ only. In the case of a GWP it depends on i_0,i_n only; more formally, the probability of a cylindrical set $\{\omega \mid X_j(\omega) = i_j,\ j = 0,1,\dots,n\}$ is changed by the factor $h(i_n)/h(i_0)$.

1.2. The Harris–Sevastyanov transformation

The Harris–Sevastyanov transformation, Z_n^*, of a supercritical GWP, $(Z_n)_{n\ge0}$, with offspring generating function f satisfying $f(0) > 0$ is the subcritical GWP with generating function $f^*(s) := f(sq)/q$, where $q < 1$ denotes the extinction probability of $(Z_n)_{n\ge0}$ and, thus, the smallest fixed point of f in $[0,1]$ (see [5, pp. 47–53]). The transformation is obtained by conditioning $(Z_n)_{n\ge0}$ upon ultimate extinction. Therefore, it is a Doob's h-transform with the harmonic function being the extinction probability starting with x individuals, $h(x) = \mathrm{P}_x(\lim_n Z_n = 0) = q^x$. We shall see that the Harris–Sevastyanov transformation of a supercritical GWP with a modified geometric offspring distribution $\mathrm{Geom}(u_0,u)$ is a subcritical GWP again with a modified geometric distribution $\mathrm{Geom}(1-u,1-u_0)$.

1.3. The Lamperti–Ney transformation

The Lamperti–Ney transformation, also known as the Q transform, introduced in [8] (see [5, p. 58]) is a Markov chain on $\mathbb{N}$ with transition matrix $Q = (q(i,j))$ defined by

$$q(i,j) := \frac{jq^{j-i}p(i,j)}{if'(q)}, \tag{1.4}$$

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

where it is assumed that $q > 0$. It is easy to see that the function

$$h(x, s) = xq^{x-1} f'(q)^{-s}$$

is space–time harmonic. This follows from the following formal calculation, which can be justified:

$$\mathrm{E}_x(Z_1 s^{Z_1-1}) = \mathrm{E}_x\left(\frac{\mathrm{d}}{\mathrm{d}s} s^{Z_1}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\,\mathrm{E}_x(s^{Z_1}) = \frac{\mathrm{d}}{\mathrm{d}s} f^x(s) = x f^{x-1}(s) f'(s).$$

Taking $s = q$ leads to h. Now it is clear, from (1.3), that the Q-transform, (1.4), is Doob's h-transform of the space–time process with the above h.

1.4. Time reversals

Given an MC and the knowledge of the starting measure μ_{-n} at time $-n$, we may reverse the process by conditioning on specific events today, like entering a specific state and looking back in history, $(Y_m) = (X_{-m})$. Any distributional limit of Y_m as $-n$ tends to $-\infty$ is called a time reversal. We present the time reversals for the GWP, which lead to quasi-invariant measures [1]. The Esty reversal is the limit for conditioning on just entering the state 0 at time 0 and starting in the state 1 when $-n$ tends to $-\infty$. The Esty reversal is included in the standard approach.

1.4.1. *Time reversal based on an invariant measure.* Let μ denote an invariant Radon measure ($\mu \mathrm{P} = \mu$) for a time-homogeneous MC. Note that μ is a Radon measure if μ is σ-finite and if, for any measurable set A of finite measure and for $\varepsilon > 0$, there exists a compact set $K \subset A$ such that $\mu(A \setminus K) < \varepsilon$. (For simplicity, take a finite measure and assume that $\mu(x) > 0$ for all $x \in S$.) The time reversal of an MC with respect to μ is an MC with transition probability kernel

$$q(y, x) := \frac{\mu(x) p(x, y)}{\mu(y)},$$

and semigroup $q^n(y, x) = (\mu(x) p^n(x, y))/\mu(y)$.

If μ is a stationary probability measure and $X = (X_n)_{n\in\mathbb{Z}}$ is an MC to μ on the whole time scale $\mathbb{Z}$, then reversing the time $Y_n := X_{-n}$ will be sufficient. If μ is only stationary, but not finite, the same interpretation holds although the construction of $X = (X_n)_{n\in\mathbb{Z}}$ is somewhat different, leaving probability spaces.

Note the analogy with harmonic function interchanging the first and second coordinate (see (1.2)). This is more than a formal similarity. The function $x \mapsto 1/\mu(x)$ is a harmonic function for the Q-kernel and vice versa. This provides a one-to-one correspondence between the invariant measures of P and the harmonic functions of Q. Martin boundary theory, as outlined above, is applicable.

For a GWP, since 0 is absorbing, the time reversal has to be carried out for the space–time process and consequently via the quasi-stationary measures [1], [5, p. 67]. Note that there are no invariant Radon measures besides multiples of δ_0 (where δ_0 is the unit mass at 0).

Let τ denote the extinction time for a subcritical GWP. The general approach to the time reversal $Y_n := Z_{\tau-n}$ for $n \in \mathbb{N}$ was considered in [1] via quasi-stationary measures. The set of quasi-invariant measures is a convex cone and has a unique integral representation over the extremals. Therefore, a characterization of the extremals, including Martin topology (here pointwise convergence), suffices for the knowledge of all quasi-invariant measures by integration. Alsmeyer and Roesler [1] gave a complete description of all extremal quasi-invariant Radon measures including Martin topology without any $L \log L$ condition. The Martin

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

boundary stated in [3] and [5, pp. 69–73] is stated correctly under only the $L \log L$ condition, basically following an announcement of Spitzer [12], [2]. Note the construction in [1] of the quasi-invariant measures as pointwise limits of the Martin (measures) functions $G(x, \cdot)/G(x, y_0)$ now for sequences in x.

1.4.2. *The Esty time reversal.* The Esty reversal is obtained by conditioning a GWP in negative time upon entering state 0 (extinction) at time 0 when starting at state 1 at time $-n$ and letting n tend to ∞. (Esty [6] had the additional technical condition $\mathrm{P}(\xi = 1) > 0$, which we think can, in his setting, be replaced by the smallest additive group generated by the support of ξ as integers.) Provided the additional condition $\mathrm{P}(\xi = 1) > 0$ holds, the following limit exists (monotone ratio lemma [5, p. 12]):

$$a_j = \lim_{N\to\infty} \frac{\mathrm{P}(Z_N = j \mid Z_0 = 1)}{\mathrm{P}(Z_N = 1 \mid Z_0 = 1)}, \qquad j \geq 1,$$

and the Esty reversal is a time-homogeneous MC with transition probabilities

$$\mathrm{P}(X_{n+1} = i \mid X_n = j) = \frac{a_j}{a_i b} p(i, j), \tag{1.5}$$

$b = f'(q)$, $n \in \mathbb{N}_0$, and $i, j \in \mathbb{N}$. The vector $(a_j)_{j\in\mathbb{N}}$ is a quasi-stationary measure [5, p. 12] for the GWP

$$\sum_{i\in\mathbb{N}} a_i \,\mathrm{P}(Z_{n+1} = j \mid Z_n = i) = b a_j, \qquad j \in \mathbb{N}.$$

A closer look shows that the limit

$$\lim_{N\to\infty} \frac{\mathrm{P}(Z_N = j \mid Z_0 = i)}{\mathrm{P}(Z_N = 1 \mid Z_0 = 1)}$$

is independent of i and is equal to a_j. This allows us to view the Esty reversal as a time reversal based on an invariant measure for the space–time MC on $\mathbb{N}$. It is a remarkable result that, in the linear fractional case (modified geometric offspring distribution), the Esty reversal is the same as the Lamperti–Ney process, shown in Theorem 3.2, below. This generalizes the earlier results of [7], where only a geometric offspring distribution was considered.

1.5. The cone dual

Another transform via stochastic monotonicity was introduced by Asmussen and Sigman [4]. We call this the cone dual, since we will imbed this dual in a more general approach using cones with unique integral representations. (Uwe Roesler has learnt this from Hans Foellmer (private communication).) We are not aware of any reference to the following. Let C be a cone of functions on some nice topological space S with a unique integral representation, i.e. for every $c \in C$, there exists a unique Radon measure μ_c on the set E of extremals such that $c(\cdot) = \int_E e(\cdot)\mu_c(\mathrm{d}e)$. Let P be a probability kernel on S. We will use $P(x, f)$ for the integral $\int f(y)P(x, \mathrm{d}y)$, and we will use P as a map of functions on S to itself. The important assumption for this duality to work is

$$PC \subset C.$$

For every extremal $e \in E$, there exists a unique probability measure μ_e on E such that

$$P(\cdot, e) = \int_E f(\cdot)\mu_e(\mathrm{d}f).$$

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

By the same argument, there exists a unique $\mu_e^n \in E$ such that

$$P^n(x, e) = \int_E f(x)\mu_e^n(\mathrm{d}f).$$

The family $(e, \cdot) \mapsto \mu_e^n(\cdot)$, $n \in \mathbb{N}$, is a transition kernel semigroup, as the following calculation shows:

$$\begin{aligned}\int f(x)\mu_e^{s+t}(\mathrm{d}f) &= P^{s+t}(x, e)\\ &= \iint e(z)P^t(y, \mathrm{d}z)P^s(x, \mathrm{d}y)\\ &= \int P^t(y, e)P^s(x, \mathrm{d}y)\\ &= \iint g(y)\mu_e^t(\mathrm{d}g)P^s(x, \mathrm{d}y)\\ &= \int P^s(x, g)\mu_e^t(\mathrm{d}g)\\ &= \iint h(x)\mu_g^s(\mathrm{d}h)\mu_e^t(\mathrm{d}g).\end{aligned}$$

The uniqueness of the representation provides the semigroup property $\mu^{s+t} = \mu^s \star \mu^t$. We call the MC with transition $Q(e, \cdot) = \mu_e(\cdot)$ the cone dual of an MC to P relative to the cone C. We specialize this setting to stochastic monotone matrices on $\mathbb{N}_0$. An MC is called stochastically monotone if $\mathrm{P}(X_{n+1} \geq y \mid X_n = x)$ is monotone increasing in x for every y. Consider the cone C of positive and increasing functions. This cone has a unique integral representation over the extremals $e_y = \mathbf{1}_{[y,\infty)}$, $y \in \mathbb{N}_0$. The property $PC \subset C$ follows by stochastic monotonicity. Thus, we obtain a cone dual with respect to C via

$$Q(e, \{e_z \mid z \in [0, x]\}) = P(x, e).$$

The cone dual corresponds to a transition matrix on the set E of extremals. We now identify the extremal function $e = e_y$ with the representing integer $y = \varphi(e_y)$. Then the cone dual $\overline{Q}(y, A) := Q(e_y, \varphi^{-1}(A))$ is a transition kernel on the integers

$$\overline{Q}(y, [0, x]) = P(x, [y, \infty)).$$

This leads to the process defined by the transition probabilities

$$\overline{q}(y, x) = P(x, [y, \infty)) - P(x - 1, [y, \infty)),$$

as in [4].

2. Linear fractional or modified geometric reproduction

Let $p = (p_0, p_1, \dots)$ denote a probability vector not concentrated at a point. Let $\xi_{n,i}$, $n, i \in \mathbb{N}$, be independent random variables (RVs) with distribution $p = (\mathrm{P}(\xi = j))_j$. The associated GWP, $(Z_n)_{n\geq 0}$, with

$$Z_{n+1} := \sum_{i=1}^{Z_n} \xi_{n,i},$$

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

is a time-homogeneous MC with transitions

$$p(i, j) = \mathrm{P}(S_i = j),$$

where S_i is the ith partial sum of i.i.d. copies of ξ.

Let f denote the probability generating function of the offspring distribution p:

$$f(s) = \mathrm{E}(s^{\xi}) = \sum_{i=0}^{\infty} p_i s^i.$$

For our purpose, we take a positive s and allow infinite values of the generating function. The generating function of Z_n, given $Z_0 = i$, is

$$\mathrm{E}(s^{Z_n} \mid Z_0 = i) = f_n^i(s),$$

where f_n is the nth iterate of f. These generating functions are all strictly increasing whenever finite. These have at most two fixed points, one of which is always 1.

The extinction probability q,

$$q := \lim_{n\to\infty} \mathrm{P}(Z_n = 0 \mid Z_0 = 1),$$

is the smallest fixed point of the generating function f. Depending on the offspring mean $m = \mathrm{E}(\xi) = f'(1)$ (the left-hand derivative), the extinction probability q is either equal to 1 if $m \le 1$ or strictly less than 1 if $m > 1$. The derivative $b := f'(q)$ is strictly larger than 1 if $q < 1$ and less than or equal to 1 for $q = 1$.

A geometric distribution $\mathrm{Geom}(u)$, $u \in (0, 1)$ is a distribution p on $\mathbb{N}_0$ given by

$$p_n = vu^n, \qquad n \in \mathbb{N}_0,\ v = 1 - u.$$

A modified geometric distribution $\mathrm{Geom}(u_0, u)$, $u_0, u \in (0, 1)$, or a zero-modified geometric distribution is a distribution on $\mathbb{N}_0$ given by

$$p_0 = v_0, \qquad p_n = u_0 v u^{n-1}, \quad n \in \mathbb{N}, \tag{2.1}$$

where $v_0 = 1 - u_0$; see, e.g. [10].

It is easy to see the following.

Proposition 2.1. *The* $\mathrm{Geom}(u_0, u)$ *distribution is obtained as a product of independent RVs* B *and* ξ*, where* B *is Bernoulli distributed with parameter* u_0 *and* ξ *is geometrically distributed with parameter* u*. Then* $B(\xi + 1)$ *has a modified geometric distribution,* $\mathrm{Geom}(u_0, u)$.

It follows that the expectation of a $\mathrm{Geom}(u_0, u)$ distribution is $m = u_0/v$. Note that $\mathrm{Geom}(u, u) = \mathrm{Geom}(u)$ is a solution of the stochastic fixed point equation

$$\xi \overset{\mathrm{D}}{=} B(\xi + 1),$$

where '$\overset{\mathrm{D}}{=}$' denotes equality in distribution, B and ξ are independent, and B is Bernoulli distributed with parameter u.

Often a $\mathrm{Geom}(u_0, u)$ distribution is referred to as a linear fractional distribution owing to the form of its generating function,

$$f(s) = v_0 + \frac{u_0 v s}{1 - us} = \frac{v_0 + (u_0 - u)s}{1 - us} = \frac{v_0 + (v - v_0)s}{1 - us} \quad \text{for } s < \frac{1}{u},$$

and $f(s) = \infty$ for $s \ge 1/u$.

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

The key feature of the modified geometric distribution $\mathrm{Geom}(u_0, u)$ is the following explicit formula [5, p. 7] for the iterations:

$$f_n(s) = 1 - m^n v_n + \frac{m^n v_n^2 s}{1 - u_n s},$$

$$v_n = \begin{cases} \dfrac{u - v_0}{um^n - v_0} & \text{if } m \neq 1, \\ \dfrac{v}{v + nu} & \text{if } m = 1, \end{cases} \qquad u_n = 1 - v_n \quad \text{for } n \geq 1. \tag{2.2}$$

In particular, $f_1 = f$ and $u_1 = u$. Thus, given $Z_0 = 1$, the size of the nth generation Z_n has the zero-modified geometric distribution $\mathrm{Geom}(m^n v_n, u_n)$ with the following counterpart of (2.1):

$$\mathrm{P}(Z_n = 0 \mid Z_0 = 1) = 1 - m^n v_n, \qquad \mathrm{P}(Z_n = i \mid Z_0 = 1) = m^n v_n^2 u_n^{i-1}. \tag{2.3}$$

More generally, in the case of $i \geq 1$ initial particles we have

$$\mathrm{P}(Z_n = 0 \mid Z_0 = i) = (1 - m^n v_n)^i,$$

$$\mathrm{P}(Z_n = j \mid Z_0 = i) = \sum_{l=1}^{j} \binom{j-1}{l-1} \binom{i}{l} v_n^{2l} u_n^{j-l} (1 - m^n v_n)^{i-l} m^{nl},$$

$i, j \in \mathbb{N}$. These formulae follow from Proposition 2.2, below.

Proposition 2.2. *If S_i denotes the sum of i independent RVs with a common* $\mathrm{Geom}(u_0, u)$ *distribution then*

$$\mathrm{P}(S_i = 0) = u_0^i, \qquad \mathrm{P}(S_i = j) = \sum_{l=1}^{i} \binom{j-1}{l-1} \binom{i}{l} u^l v^{j-l} u_0^{i-l} v_0^l, \tag{2.4}$$

$i, j \in \mathbb{N}$.

Here and elsewhere we set $\binom{j}{l} = 0$ for $j < l$.

Proof of Proposition 2.2. To verify (2.4) we use the representation $S_i = \sum_{l=1}^{i} B_l(\xi_l + 1)$, where B_l and ξ_l are independent RVs, the Bs have a Bernoulli distribution, $\mathrm{Ber}(u_0)$, and the ξs have a geometric distribution, $\mathrm{Geom}(u)$, by Proposition 2.1. The number of strictly positive summands in S_i is a binomially distributed random variable $L \sim \mathrm{Bin}(i, u_0)$ with $\mathrm{P}(L = l) = \binom{i}{l} v_0^{j-l} u_0^l$, $0 \leq l \leq i$. Given $L = l$, the distribution of $S_i - l \sim \mathrm{NB}(l, p)$ is a negative binomial with

$$\mathrm{P}(S_i - L = k \mid L = l) = \binom{k + l - 1}{k} v^l u^k, \qquad k \in \mathbb{N}_0.$$

Applying the law of total probabilities we arrive at (2.4).

The extinction probability of the GWP with $\mathrm{Geom}(u_0, u)$ reproduction equals $q = \min(v_0/u, 1)$, in accordance with the first part of (2.3). Note that q is the smaller of the two fixed points v_0/u and 1 of the generating function f. The parameter $b = f'(q)$ computes to

$$b = \begin{cases} \dfrac{v}{u_0} & \text{if } m > 1, \\ \dfrac{u_0}{v} & \text{if } m \leq 1. \end{cases}$$

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

Let us point out the mathematical clue in order to establish the formulae [5]. Consider the set of generating functions of the form $f(s) = (as + b)/(cs + d)$, where $a, b, c, d > 0$ and $a + b = c + d$. Identify this generating function with a matrix

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Observe that the composition of these generating functions corresponds to the multiplication of the associated matrices. In our setting of a modified geometric distribution, $\mathrm{Geom}(u_0, u)$, the associated matrix is

$$\boldsymbol{A} = \begin{pmatrix} v - v_0 & v_0 \\ -u & 1 \end{pmatrix}.$$

Firstly we consider the case in which $u \neq v_0$. The matrix has eigenvalues v and u_0 with the eigenvectors $(1, 1)$ and $(1, u/v_0)$, respectively. A diagonal form is given by

$$\boldsymbol{\mathcal{O}}^{-1}\boldsymbol{A}\boldsymbol{\mathcal{O}} = \boldsymbol{D},$$

where

$$\boldsymbol{\mathcal{O}} = \begin{pmatrix} 1 & 1 \\ 1 & u/v_0 \end{pmatrix}, \qquad \boldsymbol{D} = \begin{pmatrix} v & 0 \\ 0 & u_0 \end{pmatrix}.$$

The inverse of $\boldsymbol{\mathcal{O}}$ is

$$\boldsymbol{\mathcal{O}}^{-1} = \frac{1}{u - v_0} \begin{pmatrix} u & -v_0 \\ -v_0 & v_0 \end{pmatrix}.$$

The powers $\boldsymbol{A}^n$, given by

$$\boldsymbol{A}^n = \frac{1}{u - v_0} \begin{pmatrix} v^n u - u_0 v_0 & u_0^n v_0 - v_0 v^n \\ v^n u - u u_0^n & u u_0^n - v_0 v^n \end{pmatrix},$$

correspond to the iterates f_n.

The argument for the case in which $u = v_0$ is similar. The matrix has the eigenvalue v with multiplicity 2. The eigenvector is $(1, 1)$ (up to multiplication), and $y = (1, 1 + 1/u)$ satisfies $\boldsymbol{A}y = vy + (1, 1)$. The powers $\boldsymbol{A}^n$ are given by

$$\boldsymbol{A}^n = v^{n-1} \begin{pmatrix} v - nu & nu \\ -nu & v + nu \end{pmatrix}.$$

3. Branching transformations as MC transforms

3.1. The Harris–Sevastyanov transformation and Doob's h-transform

Here we give a more general version of the Harris–Sevastyanov transformation, which also applies to subcritical processes, and show that in fact it is a Doob's h-transform with an appropriate function h.

Definition 3.1. Let $r > 0$ be a fixed point of the offspring distribution generating function f. The Harris–Sevastyanov transformation [5, pp. 47–53] is the GWP with the transformed offspring generating function

$$f^*(s) = \frac{f(rs)}{r}.$$

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

The iterates of f^* are

$$f_n^*(s) = \frac{f_n(rs)}{r}.$$

If $r = 1$, the Harris–Sevastyanov transform is the process itself. The Harris–Sevastyanov transform is symmetric in the sense that $f = f^{**}$, where $f^{**}(s) = f^*(r^*s)/r^*$ is the Harris–Sevastyanov transform of f^* with the fixed point $r^* := 1/r$.

If the generating function f has exactly two fixed points q and r, $0 < q < r$, then the Harris–Sevastyanov transform f^* has exactly two fixed points q^* and r^*, $q^* = 1 < 1/q = r^*$, for the case in which $q < r = 1$, and exactly two fixed points q^* and r^*, $q^* = 1/r < 1 = r^*$, for the case in which $q = 1 < r$. If the GWP is supercritical or subcritical then the Harris–Sevastyanov transform is respectively subcritical or supercritical and vice versa. (The generating function f, defined on the positive reals and allowing for the value of ∞, is strictly convex and strictly increasing whenever finite. This function has at most one fixed point, $r \in [1, \infty)$, besides the extinction probability $q \in [0, 1]$. For the case in which $m > 1$, there are two fixed points q and r, $0 \le q < r = 1$, and, for the case in which $m = 1$, there is only the fixed point, $r = 1$. So, we set $r = 1$, and for the case in which $m < 1$, the smallest fixed point is $q = 1$ and there may be another one, $r > 1$, depending on the finiteness range of f.)

Theorem 3.1. *Let $r > 0$ be a fixed point of f and h be the function $h(n) := r^n$. Then h is a harmonic function for P, and Doob's h-transform is again a GWP with generating function*

$$f^*(s) = \frac{f(rs)}{r}.$$

Proof. The function h is harmonic for P, since

$$\sum_{j \in \mathbb{N}_0} p(i, j)h(j) = \mathrm{E}(r^{Z_1} \mid Z_0 = i) = f^i(r) = r^i = h(i) \quad \text{for } i \in \mathbb{N},$$

and similarly for $i = 0$.

The generating function for the probability vector $q^n(i, j)$, $i, j \in \mathbb{N}_0$, for Doob's h-transform (see the introduction) is

$$q^n(i, j) = \frac{1}{r^i} p^n(i, j) r^j,$$

and

$$\sum_j q^n(i, j)s^j = \left(\frac{f_n(sr)}{r}\right)^i.$$

By the uniqueness of generating functions, Doob's h-transform is the Harris–Sevastyanov process to r.

Next we specialize the offspring distribution to a modified Geom(u_0, u) geometric distribution. The generating function has the two fixed points $r = 1$ and $r = v_0/u$. The generating function f^* of the Harris–Sevastyanov transform with $r = v_0/u$ is

$$f^*(s) = \frac{f(v_0 s/u)}{v_0/q} = \cdots = \frac{u + (u_0 - u)s}{1 - v_0 s} = u + \frac{u_0 v s}{1 - v_0 s}.$$

Thus, we obtain the following result.

Proposition 3.1. *The Harris–Sevastyanov transform of a GWP with the modified geometric distribution* Geom(u_0, u) *is a GWP with the modified geometric distribution* Geom(v, v_0).

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

3.2. Immigration and Doob's *h*-transform for the space–time process

Let Y_0, C_n, and $\xi_{n,i}$, $n \in \mathbb{N}_0, i \in \mathbb{N}$, denote independent RVs with values in the positive integers. The ξ RVs are identically distributed, and the C RVs are identically distributed. Let g denote the generating function of the immigration C. Taking the view of an eternal particle producing the immigration and counting this particle, we arrive at the generating function of the GWPE $\hat{Y}_n$ (see (1.1)),

$$\mathrm{E}(s^{\hat{Y}_{n+1}} \mid \hat{Y}_n = i) = sg(s)f^{i-1}(s).$$

Let Z_n denote a GWP with offspring distributed as ξ. Since Z_n/m^n is a martingale, we consider Doob's h-transform, $(X)_{n\in\mathbb{N}_0}$, on the state space $\mathbb{N}$ with transition matrix $Q = (q(i, j))_{i,j\in\mathbb{N}}$ and $h(i) = i$, where

$$q(i, j) = p(i, j)\frac{j}{mi}, \qquad i, j \in \mathbb{N}.$$

Note that 0 is not in the state space. Since $\mathrm{E}(Z_1 \mid Z_0 = i) = mi$, it is clear that Q is a transition matrix. Apparently, this transform appears in [9] as a size-biased GWP.

The corresponding generating function for the transformed process $(X_n)_n$ is given by

$$\begin{aligned}\mathrm{E}(s^{X_{n+1}} \mid X_n = i) &= \sum_{j\in\mathbb{N}} q(i, j)s^j \\ &= \frac{1}{im}\sum_{j=1}^{\infty} p(i, j)js^j \\ &= \frac{s}{im}\sum_{j=1}^{\infty} p(i, j)js^{j-1} \\ &= \frac{s}{im}(f^i(s))' \\ &= \frac{s}{m}f'(s)f^{i-1}(s).\end{aligned}$$

Thus, $(X_n)_n$ is a GWPE with the same offspring distribution and eternal particle offspring generating function $g(s) = (s/m)f'(s)$.

Proposition 3.2. *The Lamperti–Ney [8] transform is a composition of a Harris–Sevastyanov transform with the fixed point $r > 0$ and the LPP transform. It is an MC on $\mathbb{N}$ with transition matrix $Q = (q(i, j))_{i,j\in\mathbb{N}}$,*

$$q(i, j) = p(i, j)\frac{jr^j}{cir^i}, \qquad i, j \in \mathbb{N}, \tag{3.1}$$

where $c = f'(r)$. It is a GWPE with offspring generating function $(s/f'(r))f'(sr)(f(sr)/r)^{i-1}$ and eternal particle offspring generating function $g(s) = (s/b)f'(sq)$.

The special case in which $r = 1$ is the LPP transform.

Using (1.3) and the fact that the function $h\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$,

$$h(i, n) := im^{-n}, \tag{3.2}$$

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

is space–time harmonic, we can see that the LPP transform is a Doob's h-transform for the space–time process.

Note that $h(i,\cdot) > 0$ for $i > 0$ and that $h(0,\cdot) = 0$. Therefore, we take the space–time process on the state space $\mathbb{N} \times \mathbb{N}_0$. We use the notation given in the introduction. Doob's h-transform has the space–time (time-homogeneous) transition probabilities

$$q(i,s,j,t) = \mathbf{1}_{\{s+1=t\}}\, p(i,j)\frac{j}{mi}. \tag{3.3}$$

Doob's h-transform is the space–time process of the LPP process as MC with transitions on $\mathbb{N}$ given in (3.3). (Note that the space–time harmonic function, (3.2), factors.) A pathwise interpretation as a measure transform is obvious, the transform depending on space and time.

Consecutive Doob's h-transforms are a Doob's h-transform for the product of harmonic functions. We summarize our findings in Proposition 3.3, below, for a modified geometric distribution.

Proposition 3.3. *Let $(Z_n)_n$ be a GWP with a modified* $\mathrm{Geom}(u_0, u)$ *offspring distribution. Then the LPP process is a GWPE with the same offspring distribution and eternal particle offspring generating function*

$$\frac{s}{m} f'(s) = s\frac{v}{1-us}\frac{v}{1-us},$$

which is the convolution of a point measure at 1 *with two geometric* $\mathrm{Geom}(u)$ *distributions.*

The Lamperti–Ney process, $(X_n)_n$, given in (3.1), is Doob's h-transform with the space–time harmonic function

$$h(i,n) = \frac{ir^{i-1}}{c^n},$$

where $c = f'(r)$. The process $(X_n)_n$ is a GWPE process with offspring generating function $f(sr)/r$ and eternal particle offspring distribution $\delta_1 \star \mathrm{Geom}(ur) \star \mathrm{Geom}(ur)$.

3.3. The Esty reversal as a time reversal

Theorem 3.2, below, provides the Esty reversal for a modified geometric offspring distribution.

Theorem 3.2. *The Esty reversal of a subcritical or critical GWP with a* $\mathrm{Geom}(u_0, u)$ *offspring distribution, $u_0 \le 1-u$, is a GWPE with the same offspring distribution. The eternal offspring distribution is a convolution of the point measure at* 1 *and two geometric distributions with parameter u. In the supercritical case the Esty reversal of a GWP with a* $\mathrm{Geom}(u_0, u)$ *offspring distribution is a GWPE with the dual offspring distribution* $\mathrm{Geom}(v, v_0)$. *The eternal offspring distribution is a convolution of the point measure at* 1 *and two geometric distributions with parameter v_0.*

Using Proposition 3.2 we obtain the following result.

Corollary 3.1. *The Esty reversal of a GWP with a* $\mathrm{Geom}(u_0, u)$ *distribution, $u_0 < 1-u$, is a Lamperti–Ney process with* $\mathrm{Geom}(v_0, v)$ *reproduction.*

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

Proof. The proof relies on a result of independent interest (Proposition 3.4, below).

Proposition 3.4. *Let S_n denote the nth partial sum of i.i.d. RVs with a* Geom(u_0, u) *distribution. Then*

$$\sum_{i=0}^{\infty} r^i \,\mathrm{P}(S_i = j) = r u_0 v \frac{1}{1 - r v_0} \frac{1}{1 - r v_0} \hat{f}^{j-1}(r) \quad \text{for } j \in \mathbb{N}, \tag{3.4}$$

where $\hat{f}$ is the generating function of a Geom(v, v_0) *distribution.*

Proof. Let f be the generating function of a Geom(u_0, u) distribution and $\hat{f}$ be the generating function of a Geom(v, v_0) distribution. After some manipulation of the power series, we obtain

$$\begin{aligned}
\sum_{i,j \in \mathbb{N}_0} r^i \,\mathrm{P}(S_i = j) s^j &= \sum_i r^i f^i(s) \\
&= \frac{1}{1 - r f(s)} \\
&= \frac{1}{1 - r(v_0 + s(v - v_0))/(1 - su)} \\
&= \frac{1 - su}{1 - ru - s(v_0 + t(v - v_0))} \\
&= \frac{1}{1 - r v_0} \frac{1 - su}{1 - s(u + r(u_0 - u))/(1 - r v_0)} \\
&= \frac{1}{1 - r v_0} (1 - su) \sum_{j \in \mathbb{N}_0} s^j \hat{f}^j(r) \\
&= \frac{1}{1 - r v_0} \left(1 + \sum_{j \in \mathbb{N}} s^j \hat{f}^j(r) (\hat{f}(r) - u) \right).
\end{aligned}$$

A coefficient comparison yields

$$\sum_{i \in \mathbb{N}_0} r^i \,\mathrm{P}(S_i = j) = \frac{1}{1 - r v_0} (\hat{f}(r) - u) \hat{f}^{j-1}(r)$$

and therefore the partial claim.

Proof of Theorem 3.2. Using (2.2) and (2.3), we obtain

$$a_j = \lim_n u_n^{j-1} = \left(\frac{u}{v_0} \right)^{j-1} \quad \text{for } j \in \mathbb{N}.$$

We will show that

$$\sum_{i \in \mathbb{N}_0} q(j, i) r^i = r \frac{v}{1 - ru} \frac{v}{1 - ru} f^{j-1}(r) \quad \text{for } j \in \mathbb{N},$$

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

where f is the generating function of a $\text{Geom}(u_0, u)$ distribution. Using the previous lemma, we obtain

$$\begin{aligned}\sum_{i\in\mathbb{N}_0} q(j,i)r^i &= \frac{1}{b}\sum_i \left(\frac{ru}{v_0}\right)^i p(i,j)\\ &= \frac{1}{b}\frac{ru}{v_0}u_0 v \frac{1}{1-ru}\frac{1}{1-ru}\hat{f}^{j-1}\left(\frac{ru}{v_0}\right)\\ &= r\frac{v}{1-ru}\frac{v}{1-ru}f^{j-1}(r).\end{aligned}$$

Note that the theorem also applies in reverse as there exists a one-to-one correspondence between the invariant distributions of an MC and the harmonic functions of the transformed process.

Since we have a pathwise transform, an interpretation based upon looking back where we came from is obvious for a stationary probability measure. The same intuition works for stationary measures and stationary space–time measures.

3.4. Uniform prior time reversal of a GWP

Let $(Z_n)_n$ be a GWP with a modified $\text{Geom}(u_0, u)$ distribution. Letting $t = 1$ in (3.4) yields $\sum_{i\geq 1} \text{P}(S_i = i) = v/u_0$. This provides a quasi-invariant measure, the uniform distribution on the integers. Thus, we have found another space–time invariant measure μ,

$$\mu(i,n) = \left(\frac{v}{u_0}\right)^n.$$

The space–time time reversal with respect to μ is a space–time process of the time-homogeneous MC with transitions

$$q(j,i) = \frac{u_0}{v}p(i,j), \qquad i,j \in \mathbb{N}.$$

This corresponds (see Proposition 3.4) to a GWPE with a modified geometric offspring distribution, $\text{Geom}(v, v_0)$. The eternal offspring distribution is a convolution of the point measure at 1 and two geometric distributions with parameter v_0.

Taking this view, the time reversal process could be called the reversed chain with the uniform prior as the quasi-stationary distribution.

Proposition 3.5. *The uniform prior time reversal of the GWP with* $\text{Geom}(u_0, u)$ *reproduction is a GWPE,* $(Y_n)_n$*, with the* $\text{Geom}(v, v_0)$ *reproduction and a* $\delta_1 \star \text{Geom}(v_0) \star \text{Geom}(v_0)$ *eternal offspring distribution.*

3.5. The Asmussen–Sigman transformation and the cone dual

Asmussen and Sigman introduced a dual GWP, $(V_n)_{n\in\mathbb{N}_0}$, in [4] using the following formula for the transition probabilities:

$$\text{P}(V_{n+1} = i \mid V_n = j) = \text{P}(S_{i+1} > j \geq S_i), \qquad i,j \in \mathbb{N}_0. \tag{3.5}$$

Proposition 3.6. *The cone dual to the GWP with a* $\text{Geom}(u_0, u)$ *reproduction is a GWPI process with a* $\text{Geom}(v, v_0)$ *reproduction and* $\text{Geom}(v_0)$ *immigration.*

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

Proof. Using (3.5) and

$$\sum_{i=0}^{\infty} r^i \,\mathrm{P}(S_i > j) = \sum_{k=j+1}^{\infty} \sum_{i=0}^{\infty} r^i \,\mathrm{P}(S_i = k)$$
$$= \frac{v_0 r}{(1-r)(1-u_0 r)} \left(v + \frac{u v_0 r}{1-u_0 r} \right)^j,$$

we obtain

$$\sum_{i=0}^{\infty} r^i \,\mathrm{P}(V_{n+1} = i \mid V_n = j) = \frac{1-r}{s} \sum_{i=0}^{\infty} r^i \,\mathrm{P}(S_i > j) = \frac{v_0}{1-u_0 r} \left(v + \frac{u v_0 r}{1-u_0 r} \right)^j.$$

4. A picture proof of Theorem 3.2

In this section we present a graphical representation of the one-step transitions in the linear fractional GWP using simple random walks on an $\mathbb{N}^2$ grid. This representation leads to illuminating picture proofs of several statements made in this paper concerning time reversal and duality, including Theorem 3.2, which is somewhat counterintuitive in view of its corollary. We visualize this as follows. Let S_n denote a sum of i.i.d. RVs with a Geom(u_0, u) distribution. Define recursively $\tau_1 = 1$ and $\tau_{n+1} = S_{\tau_n}$ for $n \in \mathbb{N}$. Then the sequence $(Z_n)_n$ has the same distribution as $(\tau_n - \tau_{n-1})_n$, letting $\tau_0 = 0$. Draw the points (n, S_n) on the x–y plane. Connect two points (n, y) and $(n+1, z)$ by first drawing a line from (n, y) to (n, z) and then drawing a line from (n, z) to $(n+1, z)$. Assign v_0 to the line $[(n, z), (n+1, z)]$, u_0 to the line $[(n, y), (n, y+1)]$, and u to every remaining line $[(n, \cdot), (n, \cdot + 1)]$. This picture represents the path S_n. The probability of the path is the product of the letters written on the side. For the reverse process, we have to take the same path, but replace the letter u_0 by v and the letter u by v_0.

Panel 1 of Figure 1 gives an example of a one-step transition for a linear fractional GWP with the Geom(u_0, u) reproduction law. Here the generation size changes from seven to eight after a certain random walk on the $\mathbb{N}^2$ grid (which will be called a Geom(u_0, u)-walk) makes seven horizontal jumps. As the picture indicates, there are two types of horizontal jumps in a Geom(u_0, u)-walk. A u_0-jump corresponds to a particle in generation n that has zero offspring and a u-jump corresponds to a particle with at least one offspring. Therefore, each horizontal u-jump is preceded by several vertical jumps depicting the offspring of the successful particle in question. Since the offspring number, conditional on being positive, is one plus a Geom(u) RV, every streak of vertical arrows in panel 1 starts with a v_0-arrow and is followed by a Geom(u) number of v-arrows. The number of horizontal u-arrows is equal to the number of vertical v_0-arrows and corresponds to the number l of successful particles in generation n. Note that every streak of horizontal u_0-arrows has a Geom(v_0) distribution.

The labels attached to the jumps give the probabilities, so that the product of the labels along a trajectory with $l = 3$ vertical arrows gives the term $u^3 v^5 u_0^4 v_0^3$ (see (2.4)). The binomial coefficients in (2.4) reflect the fact that $\binom{7}{2}\binom{7}{3} = 735$ trajectories with the given number $l = 3$ may lead to the same transition from $i = 7$ to $j = 8$.

Panel 2 of Figure 1 is the inverse random walk picture representing a naive attempt to obtain a reverse transition from eight to seven particles. There is no straightforward way to meaningfully interpret this panel even though the new trajectory looks like a dual Geom(v, v_0)-walk. The problem with this panel is that it is impossible to interpret the beginning and the end of the trajectory in a way consistent with the previous panel interpretation.

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

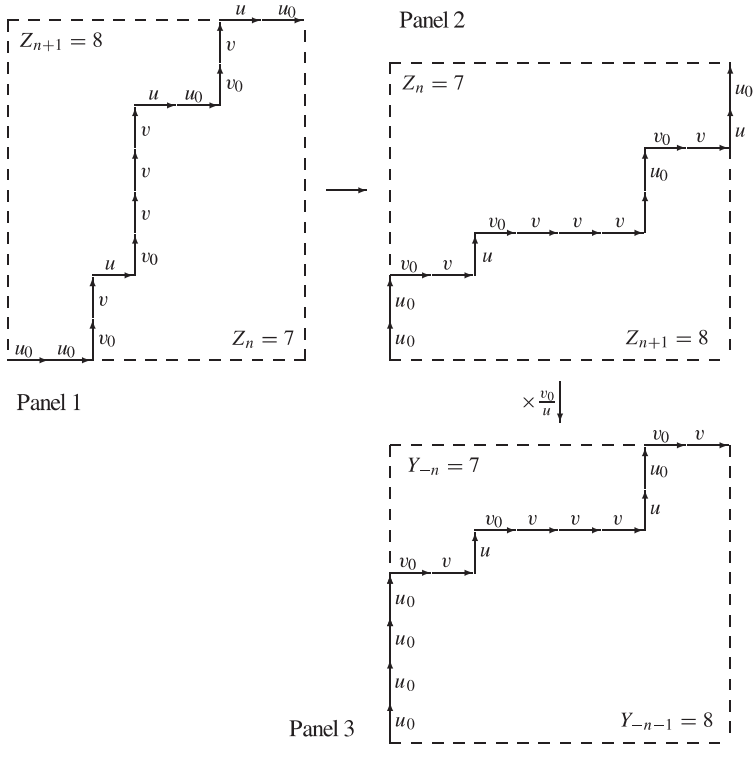

FIGURE 1: Picture proof of Theorem 3.2 in the supercritical case.

Panel 3 of Figure 1 shows how the arrows in panel 2 can be rearranged leading to a meaningful illustration of Theorem 3.2 in the supercritical case. Recall that, according to (1.5), Esty's time reversal is based on a transformation of the original GWP transitional probability by the factor $a_j/a_i b$, which in the supercritical case equals v_0/u for any pair (i, j). It follows that we have to add a v_0-arrow and remove a u-arrow to arrive at a dual random walk trajectory describing the transition in the time-reversed process.

The change of measure shown in panel 3 corresponds to the following rearrangements of the arrows in panel 2:

1. remove the last u-arrow,
2. place an additional vertical v_0-arrow to the beginning of the path,
3. move the last u_0-arrow(s) to the beginning of the path below the just-added v_0-arrow.

In this way we obtain the correct transition probability for the reversed process and obtain its size-biased process interpretation, which comes next.

The first horizontal v_0-arrow describes the eternal particle among eight particles in the current generation of the reverse process. Its offspring is depicted by the initial stretch of vertical arrows with all of them except one being u_0-arrows. The only vertical v_0-arrow describes the eternal particle among seven particles in the next generation of the reverse process. It separates two streaks of u_0-arrows, each corresponding to a Geom(v_0) random number of particles stemming from the eternal particle. The rest of the arrows have the same interpretation as in panel 1

https://doi.org/10.1239/aap/1198177238 Published online by Cambridge University Press

with the dual parameters: a v-arrow represents a childless particle and a horizontal v_0-arrow, when preceded by one plus a Geom(u) number of vertical arrows, represents the offspring of a successful particle.

References

[1] Alsmeyer, G. and Rösler, U. (2002). Asexual versus promiscuous bisexual Galton–Watson processes: the extinction probability ratio. *Ann. Appl. Prob.* **12,** 125–142.

[2] Alsmeyer, G. and Rösler, U. (2006). The Martin entrance boundary of the Galton–Watson process. *Ann. Inst. H. Poincaré Prob. Statist.* **42,** 591–606.

[3] Asmussen, S. and Hering, H. (1983). *Branching Processes*. Birkhäuser, Boston, MA.

[4] Asmussen, S. and Sigman, K. (1996). Monotone stochastic recursions and their duals. *Prob. Eng. Inf. Sci.* **10,** 1–20.

[5] Athreya, K. B. and Ney, P. E. (1972). *Branching Processes*. Springer, New York.

[6] Esty, W. W. (1975). The reverse Galton–Watson process. *J. Appl. Prob.* **12,** 574–580.

[7] Klebaner, F. C. and Sagitov, S. (2002). The age of a Galton–Watson population with a geometric offspring distribution. *J. Appl. Prob.* **39,** 816–828.

[8] Lamperti, J. and Ney, P. (1968). Conditioned branching processes and their limiting diffusions. *Theory Prob. Appl.* **13,** 128–139.

[9] Lyons, R., Pemantle, R. and Peres, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Prob.* **23,** 1125–1138.

[10] Mitov, K. V., Pakes, A. G. and Yanev, G. P. (2003). Extremes of geometric variables with applications to branching processes. *Statist. Prob. Lett.* **65,** 379–388.

[11] Rogers, L. C. G. and Williams, D. (1994). *Diffusions, Markov Processes, and Martingales*, Vol. 1, *Foundations*, 2nd edn. John Wiley, Chichester.

[12] Spitzer, F. (1967). Two explicit Martin boundary constructions. In *1967 Symp. Prob. Methods Anal.* (Lecture Notes Math. **31**), Springer, Berlin, pp. 296–298.