Title
Some models for dependence in stochastic processes

Author
You, Zhiyi

Publication Date
2022

Peer reviewed|Thesis/dissertation

Some models for dependence in stochastic processes

by

Zhiyi You

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Statistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor James Pitman, Chair
Professor Alan Hammond
Associate Professor Adityanand Guntuboyina
Associate Professor Thomas Courtade

Spring 2022

Some models for dependence in stochastic processes

Abstract

Some models for dependence in stochastic processes

by

Zhiyi You

Doctor of Philosophy in Statistics

University of California, Berkeley

Professor James Pitman, Chair

This thesis is composed of five chapters, regarding several models for dependence in stochastic processes. We first discuss the class L of selfdecomposable laws, which is a subclass of the class of infinitely divisible laws and contains all stable laws. We show an example of selfdecomposable law whose selfdecomposability is related to path decomposition of planar Brownian motions. Then we introduce the family of self-similar additive processes, which is known to have a close relationship with the class L of selfdecomposable laws. The discussion is suggested by the scale invariant Poisson spacings theorem, which arose in various contexts including records, extremal processes and random permutations. We are able to show that the range of a self-similar gamma process is a scale invariant Poisson point process ($\theta x^{-1} dx$) and also conversely, this distribution of the range characterizes the gamma process among all self-similar additive processes. We then turn to a discussion of counting processes in discrete times. In particular, when the counting process is stationary 1-dependent, its distribution is determined by the bivariate probability generating function in terms of run probability generating functions. A probabilistic explanation is provided, alongside with comparison to other known encodings including the determinantal representation and a combinatorial enumeration formula. We also compare the bivariate generating function for 1-dependent sequences with similar generating functions derived from other dependence structures. Lastly, we discuss a positivity problem related to a bivariate probability generating function for renewal processes, allowing signed measures. Fascinating graphs and qualitative observational results are provided, as well as natural but challenging open problems to explain these facts.

To my friends, Xiaowei and Kun.

Contents

Acknowledgments

First and foremost, I am greatly indebted to my advisor Jim Pitman, for his guidance, patience, motivation and understanding. He has helped me generously not just in doing research, but also in being an independent researcher. I am extremely grateful for what he has offered. This thesis is adapted from joint works with Jim, including [89, 88].

Secondly, my gratitude goes to my wife, Xuwen Shi for her long-lasting love, support and company, physically and mentally. It is a great comfort and relief to have her backing me.

I sincerely thank Prof. Alan Hammond, who offered me many help especially in the first two years of my Ph.D. life. I would also like to thank Prof. Adityand Guntuboyina and Prof. Thomas Courtade, who kindly accepted to be part of my committee.

My thanks also goes to my friends and fellows: Futianyi Wang, Xiao Li, Chen Dan, Lihua Lei, Wenpin Tang, Yumeng Zhang, Yuchen Wang, Da Xu, Alexander Tsigler, Theodore Zhu, Feynman Liang, Ella Hiesmayr and Orhan Ocal. Their assistance, inspiration and support made my Ph.D. life a memorable experience.

Besides, I would like to thank the Ph.D. graduate advisor of our statistics department, La Shana Polaris, who has helped me too many times to tell.

Last but not least, my parents Zhongxiao You and Qiaoyan Xu deserve my sincerest gratitude. They are not only the ones that give birth to me but also the ones make me who I am today.

Chapter 1

Introduction

Generally speaking, a *stochastic process* is defined as a family of random variables, usually coming with particular dependence relationship. Stochastic processes are widely employed as mathematical models to describe phenomena that vary in a random manner but also follow some patterns. Examples of stochastic processes include the movement of a gas molecule, the traffic jams occurring in California, the change of stock prices, the global population growth, etc. Stochastic processes have applications in many disciplines such as biology, chemistry, ecology, neuroscience, physics, computer science, cryptography, telecommunications and finance.

The proposal of stochastic processes is also, in many cases, inspired by applications and the study of phenomena. One of the most well-known and important process is the Brownian motion, named after the Scottish botanist Robert Brown in studying pollen grains suspended in water. The second one is the Poisson process, named after the French mathematician Siméon Denis Poisson despite Poisson himself never having studied the process. A Poisson process is usually used to describe the number of events occurring during a certain period of time. The third one is the Markov chain, named after the Russian mathematician Andrey Markov in studying an extension of independent random sequences.

There are various ways to classify stochastic processes, for example, by the state space, the index set, or the dependence structure. A common way of classification is by the cardinality of the index set. Usually, the term *stochastic process* or *random funcion* is used when indexed by the integers or a subinterval of the real line, often interpreted as time. The process is said to be in *discrete time* if indexed by integers (or equivalently, a finite or countable number of elements), and in this case, the stochastic process can also be called a *random sequence*; otherwise, if indexed by a subinterval of the real line, the process is then said to be in *continuous time*. If the collection of variables is indexed by a higher dimensional space, then usually the term *random field* is used instead. In Chapters 2 and 3, the discussion is focused on processes in continuous time, including Brownian motions, Poisson processes and Lévy processes. While in Chapter 4, discrete-time counting/indicator processes are considered.

In this introductory chapter, we review some basic facts regarding stochastic processes that will be used in the other chapters. We also clarify some notations and conventions that we will use. All the facts and results mentioned in this chapter can be found quite easily in any popular graduate level probability theory textbook, see e.g. Kallenberg [63].

1.1 Definitions

Throughout this thesis, we use the notation $L := R$ to indicate that the left mathematical object L is set to be or defined as the right one R.

Stochastic process and indexing

Formally, a *stochastic process*, or simply a *process* is defined as a collection of *random variables* defined on a *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a sample space, $\mathcal{F}$ is a σ-algebra and $\mathbb{P}$ a probability measure. And the random variables in this process all take values in the same measurable space (S, Σ), where S is a mathematical space called the *state space* of the stochastic process and Σ is a σ-algebra to which S is measurable. Theses random variables are indexed by a set I, which, historically, had a meaning of time.

A stochastic process can be written in many ways, including $(X(t), t \in I)$, $\{X(t)\}_{t \in I}$, $(X(t))$, (X_t), or simple as $X(t)$ or X, or complex as $(X(t, \omega) : t \in I)$ to reflect the outcome $\omega \in \Omega$. To avoid ambiguity,

1. write $(X(t), t \in I)$ for a stochastic process with an uncountable set I of indices, where I is usually an interval of reals, such as the set $\mathbb{R}$ of all real numbers itself or the set $\mathbb{R}_+ := (0, \infty)$ of all positive reals, see Chapter 2 and Chapter 3;

2. write $(X_t, t \in I)$ for a stochastic process with a finite or countable set I of indices, such as the set of integers $\mathbb{Z} := \{\ldots, -1, 0, 1, 2, \ldots\}$ or its subset, including the set of (strictly) positive integers $\mathbb{Z}_+ := \{1, 2, 3, \ldots\}$, see Chapter 4;

3. and write $\{X_t : t \in I\}$ for a *simple point process* (which will be defined later in this section), where I is finite or countable and the order of the elements is not of our interest, see also Chapter 3 for more discussion on simple point processes.

This convention is especially helpful when a process in continuous time and one in discrete time are placed in the same probability space and we wish to use the same capital letter for some reason, for example in the case of studying the discrete skeleton of a continuous-time Markov jump process.

As long as there is no conflict in definitions, we may write $X := (X(t), t \in I)$, $X := (X_t, t \in I)$ or $X := \{X_t : t \in I\}$ so that the whole process is shorthanded as X. And sometimes when the indexing is not important, we may simply write X to denote the process of our interest.

Sample path and stochastic continuity

A *sample path* is a single outcome of a stochastic process, formed by taking a single outcome of each component random variable of the process. More specifically, consider a stochastic process $X := (X(t, \omega) : t \in I)$ with state space S, then for each $\omega \in \Omega$, the mapping

$$X(\cdot, \omega) \; : \; I \to S \tag{1.1}$$

is called a sample path of X.

We are particularly interested in studying sample paths when the process is in continuous time. And often times we require the process to be *stochastically continuous*, or *continuous in probability*, i.e. for a continuous-time process $(X(t), t \in I)$ and for $s, t \in I$ and for each $\epsilon > 0$,

$$\lim_{t \to s} \mathbb{P}(|X(t) - X(s)| \geq \epsilon) = 0. \tag{1.2}$$

Independent increments

An *increment* of a stochastic process $X := (X(t), t \in I)$ is the difference $X(t) - X(s)$ between two random variables $X(t)$ and $X(s)$ of the same stochastic process X. When the index set I is interpreted as time, an increment $X(t) - X(s)$ is how much the stochastic process changes over a certain time period $t - s$. We say the process X is with *independent increments* if for all indices $t_1 < t_2 < \cdots < t_n$ in I (usually $\mathbb{Z}, \mathbb{Z}_+$ or a subinterval of $\mathbb{R}$) such that

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \ldots, X_{t_n} - X_{t_{n-1}} \tag{1.3}$$

are independent.

Lévy processes, named after the French mathematician Paul Lévy, are important examples of stochastic processes with independent increments. A stochastic process $X := (X(t), t \in \mathbb{R}_+)$ is said to be *additive* if:

- X starts at zero:
$$X(0+) = 0; \tag{1.4}$$
- X has independent increments;
- X is stochastically continuous.

If an additive process X is also with *stationary increments*, i.e. for all $0 < s < t$,

$$X(t) - X(s) \stackrel{d}{=} X(t - s), \tag{1.5}$$

where $\stackrel{d}{=}$ reads equality in distribution, then call X a *Lévy process*. In particular, Brownian motion and Poisson process are both well-known examples of Lévy processes.

See, e.g. Kallenberg [63, Chapter 16] for more on independent increments and Lévy processes.

Finite-dimensional distributions

For a stochastic process $(X(t), t \in I)$, its *finite-dimensional distribution* for $t_1, t_2, \ldots, t_n \in I$ is defined as the joint distribution of $(X(t_1), X(t_2), \ldots, X(t_n))$.

With finite-dimensional distribution defined, we say two processes X and Y are *equal in distribution (in the sense of finite-dimensional distribution)* to each other, if they share the same finite-dimensional law.

Versions

For a stochastic process $X := (X(t), t \in I)$, we say another close related process $Y := (Y(t), t \in I)$ with the same index set I, state space and probability space is a *version* of X if for all $t \in I$,

$$\mathbb{P}(X(t) = Y(t)) = 1. \tag{1.6}$$

In other words, each random variable is modified only on a set of probability 0.

In this thesis, we will be working entirely in a setup where it is well-known that there exist versions with almost surely càdlàg path.

Random field and point process

The term *random field* is used when the index of the stochastic process has two or more dimensions, such as a $n(\geq 2)$-dimensional Euclidean space.

A *point process* X is a random collection of points located on some mathematical space S such as $\mathbb{R}^n$, interpreted as the *random counting measure*

$$N_X(B) = \#\{z \in \mathbb{Z} : X_z \in B\} \tag{1.7}$$

counting the numbers of points X_z in measurable subsets $B \subset S$, for some indexing of these points by the set of integers $\mathbb{Z}$. In the sense of index set, the term *random point field* is more appropriate, but we will stick to the term point process in this thesis. A point process is called *simple* if all points are distinct. See Sections 3.1 and 3.2 for more discussion.

See e.g. Kallenberg [63, Chapter 15], Daley and Vere-Jones [17, 18] and Khoshnevisan [66] for more on point processes and random fields.

1.2 Interpretations and encodings

Traditional ways of characterize a real-valued random variable X or the associated distribution include:

- using *cumulative distribution function* (*c.d.f.*)

$$F_X(t) := \mathbb{P}(X \leq t) \qquad (t \in \mathbb{R}). \tag{1.8}$$

- using either *probability density function* (*density*)

$$f_X(x) := \frac{\mathbb{P}(X \in dx)}{dx} = \frac{d}{dx}F_X(x) \qquad (x \in \mathbb{R}), \tag{1.9}$$

  or *probability mass function*

$$p_X(x) := \mathbb{P}(X = x) \qquad (x \in \mathbb{R}), \tag{1.10}$$

  whichever is applicable.

- using *characteristic function* (*ch.f.*)

$$\phi_X(t) := \mathbb{E}e^{itX} \qquad (t \in \mathbb{R}), \tag{1.11}$$

  or *Fourier transform*

$$\mathcal{F}_X(t) := \mathbb{E}e^{-itX} \qquad (t \in \mathbb{R}). \tag{1.12}$$

- using *moment-generating function* (*m.g.f.*)

$$M_X(t) := \mathbb{E}e^{tX} \qquad (t \in \mathbb{R}), \tag{1.13}$$

  or *Laplace transform*

$$\mathcal{L}_X(t) := \mathbb{E}e^{-tX} \qquad (t \geq 0), \tag{1.14}$$

  usually for X non-negative.

These characterizations provide us with both an interpretation of the distribution itself and an encoding that many powerful analytic and algebraic tools can be employed, including variable transform, Fourier transform, stochastic integral, and many others to be discussed in the following chapters.

If we further suppose the law of a real-valued random variable X is *infinitely divisible*, i.e. the law of X can be expressed as the probability distribution of the sum of an arbitrary number of *independent and identically distributed* (*i.i.d.*) random variables, then it has a Lévy-Khinchine representation given by its characteristic function:

$$\phi_X(u) := \mathbb{E}e^{iuX} = \exp\left\{ibu - \tfrac{1}{2}au^2 + \int_{\mathbb{R}-\{0\}}\left(e^{iux} - 1 - \frac{iux}{1+u^2}\right)\nu(dx)\right\}, \tag{1.15}$$

where $b \in \mathbb{R}$, $a \geq 0$ and ν is a σ-finite measure on $\mathbb{R}-\{0\}$ called the *Lévy measure*, satisfying the following integral condition

$$\int_{\mathbb{R}-\{0\}}(x^2 \wedge 1)\nu(dx) < \infty. \tag{1.16}$$

And it is well-known that each infinitely divisible law of X corresponds to a Lévy process $(X(t), t > 0)$ with $X(1) \overset{d}{=} X$ hence

$$\phi_{X(t)}(u) := \mathbb{E}e^{iuX(t)} = \mathbb{E}e^{iutX} =: \phi_X(ut). \qquad (t > 0, u \in \mathbb{R}) \tag{1.17}$$

Therefore, the law of this Lévy process is uniquely determined by the *Lévy-Khinchine triplet* (a, b, ν), where the term triplet suggest the three independent components:

1. a linear drift component b;
2. a Brownian component with variance a;
3. and a *Lévy jump process* determined by Lévy measure ν.

See e.g. Loève [73, Section 24] and Kallenberg [63, Chapter 7] for more on infinitely divisibility and Lévy-Khinchine representation.

Our discussion of selfdecomposable laws in Chapter 2 is based on the Lévy-Khinchine representation, while the Lévy measure plays an essential role in the discussion of self-similar additive processes in Chapter 3.

1.3 Organization

The rest of this thesis is organized as follows:

- In Chapter 2, we introduce the class L of selfdecomposable laws, as a proper subclass of the class of infinitely divisible laws whose Lévy density is in a special form [71, 65, 95]. We discuss the Talacko [102]-Zolotarev [108] distribution and prove its selfdecomposability, as well as interpret the distribution with a planar Brownian motion. We also show that the Talacko-Zolotarev distribution is not complete-selfdecomposable by considering an inverse Fourier transform.

- In Chapter 3, we place the scale invariant Poisson spacings theorem [4] in a broader context of self-similar additive processes. The self-similar additive processes are known to be closely related to the selfdecomposable laws [96]. We give the hold-jump description of a self-similar additive process from the Lévy density associated with the selfdecomposable law. We also provide a history note section to explain early works related or involved with self-similar processes. This Chapter is mainly adapted from [89].

- In Chapter 4, we introduce the bivariate probability generating functions as a generalization to the ordinary probability generating functions as an encoding of discrete-time counting processes. Under particular dependence structure, this encoding may uniquely determine the finite-dimensional distributions of the counting process. We are particularly interested in counting process associated with the stationary 1-dependent indicator process and give formulae of its bivariate probability generating function via its run generating functions. We also compare the formulae with alternative expressions from the theory of determinantal point processes and the combinatorial enumeration formula of sequences, as well as with the one of counting processes with other dependence structures. This Chapter is mainly adapted from [88].

- In Chapter 5, we discuss a positivity problem related to a renewal process and the associated Riordan array, by generalizing the renewal sequence to allow some of the coefficients to be negative while the Riordan array remains positive. We provide some necessary conditions bounding the negative coefficients. Additionally, we provide graphs about the coefficients of powers of polynomials which appear in a clear pattern, and describe some observational results.

Chapter 2

Selfdecomposable laws

In the probability theory and many mathematical statistic courses, the normal distribution and the central limit theorem are undoubtedly of great importance. Extensions of the class of normal distributions include the class $\mathscr{S}$ of stable laws and the whole class ID of all infinitely divisible laws. In this chapter we introduce the class of *selfdecomposable laws*, as a proper subclass of ID. In the 1930s', Lévy [71] generalized the notion of stable laws and introduced the class of *lois-limites*, while Khintchine [65] called it the *class L*. It is known to be identical to the class of selfdecomposable laws [95, Theorem 15.3]. Recently, selfdecomposable distributions had appeared in mathematical finance [10, 103] and in stochastic simulation [50].

We will also discuss the Talacko [102]-Zolotarev [108] distribution as an example and prove its selfdecomposability. Lastly, we mention the sequence of nested classes

$$ID \supset L := L_0 \supset L_1 \supset \cdots \supset L_\infty \supset \mathscr{S} \tag{2.1}$$

introduced by Urbanik [105], where L_∞ is the minimal class containing $\mathscr{S}$ closed under convolution and convergence. We prove that the Talacko-Zolotarev distribution is not of class L_∞.

2.1 Definitions

A random variable X is said to have the *class L property* if there exists a sequence of independent random variables $(Y_n, n \in \mathbb{Z}_+)$ and suitable scaling constants $b_n > 0$ and centering constants c_n such that the sums

$$X_n := b_n \sum_{k=1}^{n} Y_k + c_n, \qquad (n = 1, 2, \ldots) \tag{2.2}$$

converge in distribution to X and the triangular array of summands $(b_n Y_k, k = 1, 2, \ldots, n)$ is uniformly infinitesimal when $n \to \infty$. The distribution of X is said to be of *class L*. In other words, a distribution is of class L if it is the limit of a sequence of properly normalized partial sums of independent random variables.

The infinitesimality assumption is necessary, otherwise the property is trivial by setting $b_n \equiv 1$ and $Y_n \equiv 0$ for $n = 2, 3, \ldots$. If we further assume that $(Y_n, n \in \mathbb{Z}_+)$ is an i.i.d. sequence, then the limits of (2.2) give the class $\mathscr{S}$ of stable laws. On the other hand, the class ID of infinitely divisible laws coincides with the class of the limit distributions of uniformly infinitesimal triangular arrays with independent entries in each row. Hence

$$ID \supset L \supset \mathscr{S}. \tag{2.3}$$

The collection of laws of class L is an important generalization of the normal distributions, as (2.2) generalizes the central limit theorem in two ways: for one, the scaling constants are no longer restricted to be $n^{-1/2}$; for another, the summands can be chosen to have different orders of magnitude. Meanwhile, in Chapter 3 we will see how self-similar additive processes are related to the laws of class L.

A random variable X, or its distribution, is said to be *selfdecomposable* [101, 95] if for each $0 < a < 1$, there is the equality in distribution

$$X \stackrel{d}{=} aX + R_a, \tag{2.4}$$

for some random variable R_a independent of X. In terms of the characteristic function $\phi(u) := \mathbb{E}e^{iuX}$, this means for each $0 < a < 1$,

$$\phi(u) = \phi(au)\phi_a(u), \qquad u \in \mathbb{R}, \tag{2.5}$$

with ϕ_a a characteristic function. It is known [95, Theorem 15.3] that the class of selfdecomposable laws is identical to the class L introduced by Lévy [71] and Khintchine [65]. Lévy showed that such laws are infinitely divisible with a special structure of their Lévy measure. Specifically, from Sato and Yamazato [97], the Lévy-Khinchine representation of a selfdecomposable distribution of X is

$$\mathbb{E}e^{iuX} = \exp\left\{ibu - \tfrac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}-\{0\}} \left(e^{iux} - 1 - \frac{iux}{1+u^2}\right)\frac{k(x)}{x}dx\right\}, \tag{2.6}$$

where b is real, $\sigma^2 \geq 0$ and $k(x)$ is both

(a) non-negative, non-increasing on $\mathbb{R}_+$ and non-positive, non-increasing on $\mathbb{R}_-$;

(b) subject to the usual requirement for a Lévy density $k(x)/x$ that

$$\int_{\mathbb{R}-\{0\}} (x^2 \wedge 1)\frac{k(x)}{x}dx < \infty. \tag{2.7}$$

Hence, assuming the distribution of X is infinitely divisible, the distribution is selfdecomposable if and only its Lévy measure has a density of the form $k(x)/x$ where $k(x)$ satisfies condition (a) above.

When X is non-negative with no drift term, formula (2.6) can be simply re-written as:

$$\log \mathbb{E}e^{-tX} = \int_0^\infty \left(e^{-tx} - 1\right) \frac{k(x)}{x} dx, \tag{2.8}$$

where $k(\cdot)$

(a) is non-negative, non-increasing on $\mathbb{R}_+$;

(b) and satisfies the usual integrable property for non-negative Lévy density $k(x)/x$

$$\int_0^\infty (x \wedge 1) \frac{k(x)}{x} dx < \infty. \tag{2.9}$$

In spirit of Vervaat [106], Wolfe [107] and Jurek and Vervaat [61] characterized selfdecomposable laws in the manner of stochastic integrals with different proofs. Wolfe focused on the case of real-valued random variables, while Jurek and Vervaat proved results in the more general setting of random variables with values in a Banach space. For simplicity, we still restrict our discussion to real-valued random variables.

Theorem 2.1.1 ([61]) *X is selfdecomposable if and only if there exists a Lévy process $Y := (Y(r), r > 0)$ such that*

$$X \stackrel{d}{=} \int_0^\infty e^{-r} dY(r). \tag{2.10}$$

If so, the finite dimensional distributions of the process Y are uniquely determined, and

$$\mathbb{E} \log(1 + |Y(1)|) < \infty. \tag{2.11}$$

The Lévy process Y is called the *background driving Lévy process* (BDLP [60]) of X. Here, the stochastic integral (2.10) is understood as

- either a suitable limit as $t \to \infty$ of an integral $\int_0^t$ defined by integration by parts, as in Jurek and Vervaat [61];
- or a stochastic integral, since a Lévy process is a semi-martingale.

To summarize this section, a selfdecomposable law, or the corresponding random variable X, is characterized by each of the following:

- the limit of a sequence of properly normalized partial sums of independent random variables as in (2.2);

- the decomposition in distribution (2.4) for each $0 < a < 1$;
- the decomposition in characteristic function (2.5) for each $0 < a < 1$;
- the Lévy triple $[b, \sigma^2, x^{-1}k(x)]$ in (2.6) with $k(\cdot)$ satisfying the two conditions below (2.6);
- the BDLP in (2.10) with Y satisfying (2.11).

2.2 An example: the Talacko-Zolotarev distribution

The *Talacko-Zolotarev (T-Z) distribution* described in the proposition below was brought to notice independently in different contexts by Talacko [102] and Zolotarev [108].

Proposition 2.2.1 *Let C denote a standard Cauchy variable, $C_\alpha := -\cos(\alpha\pi)+\sin(\alpha\pi)C, \forall\alpha \in [0,1]$. For $\alpha \in (0,1]$, let S_α be a random variable with the conditional distribution of $\log C_\alpha$ given the event $(C_\alpha > 0)$, and the distribution of S_0 defined as the limit distribution of S_α as $\alpha \downarrow 0$. For each fixed $\alpha \in (0,1]$, we have the characteristic function*

$$\mathbb{E}e^{i\lambda S_\alpha} = \phi_\alpha(\lambda) = \frac{\sinh \alpha\pi\lambda}{\alpha \sinh \pi\lambda}, \qquad \forall\lambda \in \mathbb{R}. \tag{2.12}$$

This distribution of S_α is called here the Talacko-Zolotarev (T-Z) distribution. Talacko [102] regarded the family of S_α for $0 \le \alpha < 1$ as a one-parameter extension of the cases $\alpha = \frac{1}{2}$ and $\alpha = 0$ with

$$\phi_{\frac{1}{2}}(\lambda) = \frac{1}{\cosh \pi\lambda/2}, \qquad \phi_0(\lambda) = \frac{\pi\lambda}{\sinh \pi\lambda}. \tag{2.13}$$

These characteristic functions were found earlier by Lévy [72] in the study of the random area swept out by the path of a planar Brownian motion. He showed that S_0 and $S_{1/2}$ are infinitely divisible and had their Lévy measure computed. Inspired by this work, further studies were delivered to clarify the relations between various probability distributions derived from Brownian paths with hyperbolic type Fourier transforms. See e.g. Pitman and Yor [86] for these distributions and associated Lévy processes, and Pitman [83, section 3.1] for more on the T-Z distribution.

Jurek and Yor [62, Proposition 1] showed that S_0 and $S_{1/2}$ are selfdecomposable. In particular, observe that by definition there is an identity in characteristic functions

$$\phi_0(\lambda) = \phi_0(\alpha\lambda)\phi_\alpha(\lambda) \qquad (0 \le \alpha \le 1), \tag{2.14}$$

which immediately implies the selfdecomposability of S_0. We will prove that

Theorem 2.2.2 *For each $\alpha \in [0,1]$, S_α is selfdecomposable.*

The result is trivial when $\alpha = 1$ and is known for $\alpha = 0, \frac{1}{2}$. Therefore, Theorem 2.2.2 follows from the following two lemmas.

Lemma 2.2.3 *For $\alpha \in (0, \frac{1}{2})$, if $S_{2\alpha}$ is selfdecomposable, so is S_α.*

Lemma 2.2.4 *For all $\alpha \in (\frac{1}{2}, 1)$, S_α is selfdecomposable.*

Before we show the proofs, let $s(x) = s(x, \lambda) := \sinh(x\pi\lambda)$ and $c(x) = c(x, \lambda) := \cosh(x\pi\lambda)$. Hence, $\phi_\alpha(p\lambda) = \frac{s(p\alpha)}{\alpha s(p)}$ for simplicity.

Proof [Lemma 2.2.3] Observe that $S_\alpha \stackrel{d}{=} \frac{1}{2}S_{2\alpha} + S_{1/2}$ by

$$\phi_{2\alpha}(\tfrac{1}{2}\lambda)\phi_{1/2}(\lambda) = \frac{s(2\alpha \cdot \frac{1}{2})}{2\alpha s(\frac{1}{2})} \cdot \frac{s(\frac{1}{2})}{\frac{1}{2}s(1)} = \frac{s(\alpha)}{\alpha s(1)} = \phi_\alpha(\lambda), \tag{2.15}$$

where $S_{2\alpha}$ and $S_{1/2}$ are assumed to be independent. Then the result follows since both $S_{2\alpha}$ and $S_{1/2}$ are selfdecomposable. ■

Proof [Lemma 2.2.4] We will show by construction that

$$S_\alpha \stackrel{d}{=} \beta S_\alpha + T_{\alpha,\beta}, \qquad \forall \alpha \in (\tfrac{1}{2}, 1), \forall \beta \in [0, 1], \tag{2.16}$$

where

$$T_{\alpha,\beta} \stackrel{d}{=} U_{(1-\alpha)(1-\beta)}\left(\alpha\beta S_{\frac{1-\alpha}{\alpha}} + S_{\alpha(1-\beta)}\right) + (1 - U_{(1-\alpha)(1-\beta)})S_{\alpha+\beta-\alpha\beta} \tag{2.17}$$

is independent of S_α with U_p a Bernoulli(p) random variable and $S_{\frac{1-\alpha}{\alpha}}$, $S_{\alpha(1-\beta)}$ and $S_{\alpha+\beta-\alpha\beta}$ are assumed to be independent.

To check (2.16), it is equivalent to check $T_{\alpha,\beta}$ a random variable with characteristic function

$$\phi_{T_{\alpha,\beta}}(\lambda) = \frac{s(\alpha)s(\beta)}{s(1)s(\alpha\beta)}, \tag{2.18}$$

which is done by calculating the characteristic function $\phi_{RHS}(\cdot)$ of the right hand side of (2.17)

$$\phi_{RHS}(\lambda) =(1-\alpha)(1-\beta)\frac{s((1-\alpha)\beta)s(\alpha(1-\beta))}{(1-\alpha)(1-\beta)s(1)s(\alpha\beta)} + (\alpha+\beta-\alpha\beta)\frac{s((\alpha+\beta-\alpha\beta))}{(\alpha+\beta-\alpha\beta)s(1)} \tag{2.19}$$

$$=\frac{s((1-\alpha)\beta)s(\alpha(1-\beta))}{s(1)s(\alpha\beta)} + \frac{s((\alpha+\beta-\alpha\beta))}{s(1)} \tag{2.20}$$

$$=\frac{s(\beta-\alpha\beta)s(\alpha-\alpha\beta)) + s((\alpha+\beta-\alpha\beta))s(\alpha\beta)}{s(1)s(\alpha\beta)} \tag{2.21}$$

$$=\frac{s(\alpha)s(\beta)}{s(1)s(\alpha\beta)} = \phi_{T_{\alpha,\beta}}(\lambda). \tag{2.22}$$

■

2.3 A probabilistic explanation of the selfdecomposability of T-Z distribution

In this section, Theorem 2.2.2 is explained from the perspective of planar Brownian motions.

We start from recognizing the T-Z distributions as hitting distributions of a planar Brownian motion, which is discussed by Jurek and Yor [62]. Following the notations in Proposition 2.2.1, the Cauchy density of C_α is well known to be the hitting density of X_T on the real axis for a planar Brownian motion $(X_t + iY_t, t \geq 0)$ on the complex plane $\mathbb{C}$ started at the point $(-\cos\alpha\pi + i\sin\alpha\pi)$ on the unit semicircle in the upper half plane and stopped at the hitting time of real axis $T := \inf\{t : Y_t = 0\}$. In particular, $C = C_{1/2}$ with the standard Cauchy distribution has the starting point i.

Let

$$X_t + iY_t = R_t \exp(iW_t) \tag{2.23}$$

denote the usual representation of the planar Brownian motion in polar coordinates, starting from radial $R_0 = 1$ and angular $W_0 = (1-\alpha)\pi$. Then

$$C_\alpha \overset{d}{=} X_T = R_T 1(W_T = 0) - R_T 1(W_T = \pi). \tag{2.24}$$

According to conformal invariance of Brownian motion, the process $(\log R_t + iW_t, 0 \leq t \leq T)$ is a time-changed planar Brownian motion $(\Phi(u) - i\Theta(u), u \geq 0)$

$$\log R_t + iW_t = \Phi(U_t) - i\Theta(U_t), \tag{2.25}$$

where

$$U_t := \int_0^t \frac{ds}{R_s^2} \qquad \text{and} \qquad U_T = \inf\{u : \Theta(u) \in \{0, \pi\}\}. \tag{2.26}$$

The arguments above is summarized by the following proposition.

Proposition 2.3.1 *The distribution of S_α, as the conditional distribution of $\log C_\alpha$ given $C_\alpha > 0$ may also be represented as*

$$\mathbb{P}(S_\alpha \in \cdot) = \mathbb{P}_{(1-\alpha)\pi}(\Phi_T \in \cdot | \Theta_T = 0),$$

where $\mathbb{P}_\theta$ governs $(\Theta_t, t \geq 0)$ and $(\Phi_t, t \geq 0)$ two independent Brownian motions, started at $\Theta_0 = \theta \in (0, \pi)$ and $\Phi_0 = 0$, and $T := \inf\{t : \Theta_t \in \{0, \pi\}\}$.

With this proposition, we could easily translate the selfdecomposability of S_0, $S_{1/2}$ and more generally, every S_α for $\alpha \in (0, 1)$ with suitable scaling as displayed in the following figures.

Figures 2.1 and 2.2 are quite straightforward, by decomposition of the first hitting time of the corresponding vertical dashed line.

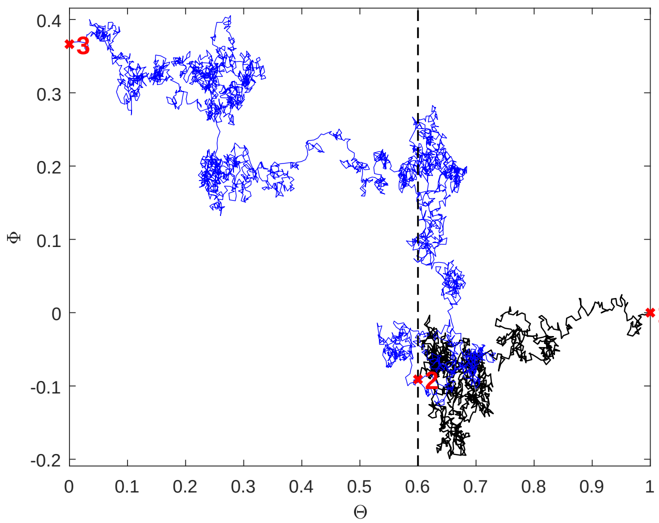


Figure 2.1: Decomposition for $\alpha = 0, \beta = 0.4$. The planar BM starts at point 1 and reaches point 3 through point 2. In distribution, the vertical difference of black part from 1 to 2 is a scaled copy of S_0 and the one of the blue part from 2 to 3 is a copy of $S_{1-\beta}$.

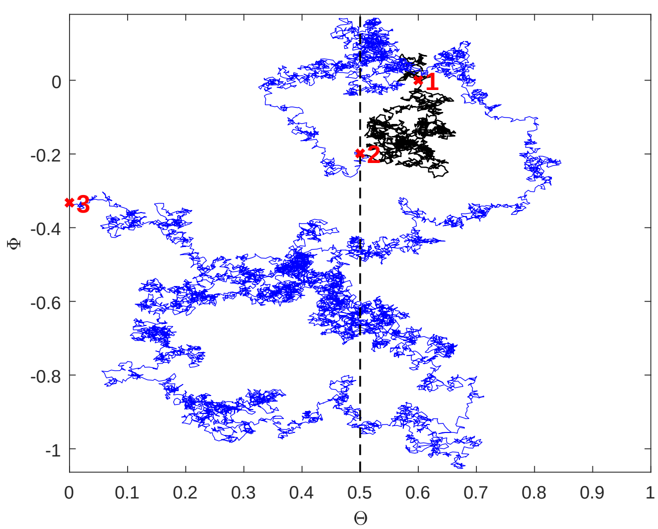


Figure 2.2: Illustration of Lemma 2.2.3 for $\alpha = 0.4$. Just like in figure 2.1, the path started at point 1 until hitting $\Theta = 0$ is cut into two parts. In distribution, the vertical differences of both parts are selfdecomposable, respectively. (Black: scaled $S_{2\alpha}$, Blue: $S_{1/2}$)

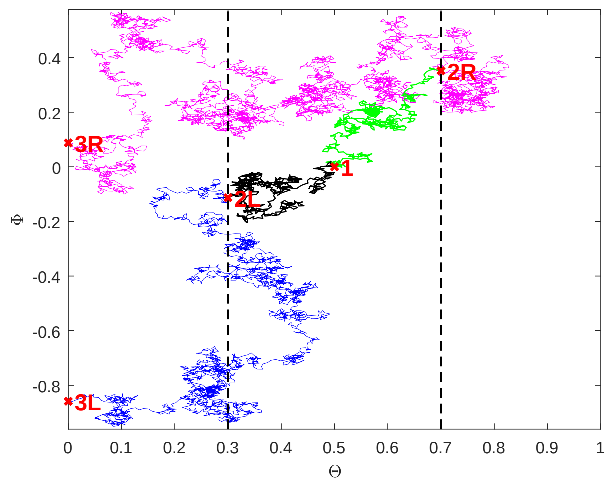


Figure 2.3: Decomposition for $\alpha = 0.5, \beta = 0.4$. In order to do the decomposition, we now need to consider two routes: one is $1 \to 2L \to 3L$; another is $1 \to 2R \to 3R$. Here L stands for the event that this BM hits the inner left bound $\Theta = (1-\alpha)(1-\beta) = 0.3$ before the right one $\Theta = 1 - \alpha + \alpha\beta = 0.7$, and vice versa. By symmetricity, both black part from 1 to $2L$ and green part from 1 to $2R$ are in distribution two scaled copies of $S_{0.5}$.

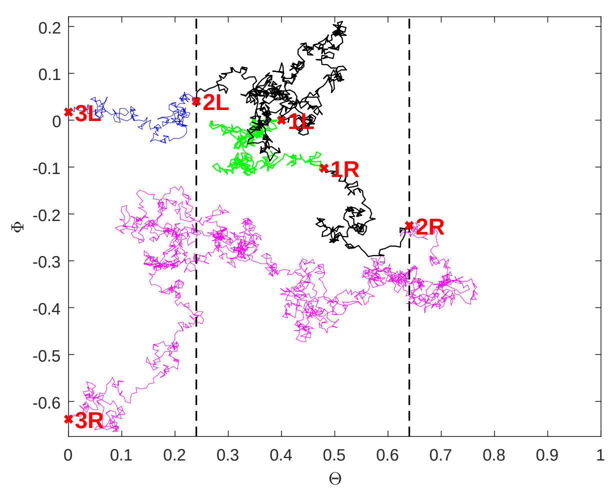


Figure 2.4: Illustration of Lemma 2.2.4 for $\alpha = 0.6, \beta = 0.4$. Here the two routes are $1L \to 2L \to 3L$ and $1L \to 1R \to 2R \to 3R$, where $1R$ is the hitting point on line $\Theta = 0.48$ given $R \cap (\Theta_T = 0)$, which makes the part from $1R$ to $2R$ in distribution a scaled copy of $S_{0.6}$.

To explain Figures 2.3 and 2.4, let L denote the event that the planar Brownian motion hits the inner left bound before the right one and R the event it hits the right one first. Easy conditional probability calculation shows that

$$\mathbb{P}(L \mid \Theta_T = 0) = \alpha + \beta - \alpha\beta. \tag{2.27}$$

Thus for figure 2.3 we have

$$T_{0.5,0.6} \overset{d}{=} 1_L S_{0.7} + 1_R S_{0.3}. \tag{2.28}$$

As for figure 2.4, we need to add in one extra term standing for the green part, which in fact equals to a scaled $S_{\frac{1-\alpha}{\alpha}}$ in distribution. In general, this extra term is

$$T_{\alpha,\beta} \overset{d}{=} 1_L S_{\alpha+\beta-\alpha\beta} + 1_R \left(\alpha\beta S_{\frac{1-\alpha}{\alpha}} + S_{\alpha(1-\beta)}\right) \tag{2.29}$$

which coincides with (2.17).

Another way to check the selfdecomposability is to treat S_α as an infinitely divisible random variable, and construct the associated Lévy process $(S_\alpha(t), t \geq 0)$ with characteristic function

$$\mathbb{E}e^{i\lambda S_\alpha(t)} = \phi_\alpha^t(\lambda) = \left(\frac{\sinh \alpha\pi\lambda}{\alpha \sinh \pi\lambda}\right)^t, \qquad \forall \lambda \in \mathbb{R}. \tag{2.30}$$

Let $k_\alpha(x)/x$ denote the Lévy density of S_α

$$\log\left(\frac{\sinh \alpha\pi\lambda}{\alpha \sinh \pi\lambda}\right) = \int_{-\infty}^{\infty} (e^{i\lambda x} - 1)\frac{k_\alpha(x)}{x} dx. \tag{2.31}$$

Here, we solve for ρ_α in spirit of [15, page 261] and [86, page 311]. Take derivative with respect to λ to both sides of (2.31)

$$\alpha\pi \coth \alpha\pi\lambda - \pi \coth \pi\lambda = \int_{-\infty}^{\infty} e^{i\lambda x}(ik_\alpha(x))dx, \tag{2.32}$$

then by inverting this Fourier transform we get

$$k_\alpha(x) = \coth\left(\frac{\pi x}{2}\right) - \coth\left(\frac{\pi x}{2\alpha}\right). \tag{2.33}$$

As discussed in Section 2.1, S_α is selfdecomposable if and only if $k_\alpha(x)$ is non-increasing for both $x > 0$ and $x < 0$, which is true by checking $k'_\alpha(x) < 0$.

2.4 Complete Selfdecomposability

We learn from Section 2.1 that the class L of selfdecomposable laws is a proper subclass of the class ID of infinitely divisible laws, and contains the class $\mathscr{S}$ of stable laws. Between $\mathscr{S}$ and

L, Urbanik [105] introduced some classes of limit distributions of sequences of independent random variables

$$ID \supset L := L_0 \supset L_1 \supset \cdots \supset L_\infty \supset \mathscr{S}. \tag{2.34}$$

Urbanik's original definition is a generalized version of Lévy 's definition of the class $L = L_0$, inductively through a triangular array of uniformly infinitesimal random variables, see Urbanik [105] and Jurek [59]. Here, we use a cleaner equivalent version, due to Urbanik [105, Proposition 1], by inductive use of (2.4).

Definition 2.4.1 *Fix $m \in \mathbb{Z}_+$, a random variable X is said to be* m-selfdecomposable *if for each* $0 < a < 1$

$$X \stackrel{d}{=} aX + R_c, \tag{2.35}$$

for some $(m-1)$-selfdecomposable random variable R_c independent of X.

Here, 0-selfdecomposable means selfdecomposable. Under this definition, Theorem 2.2.2 can be restated as '*S_0 is 1-selfdecomposable*'. As the limiting case,

Definition 2.4.2 *A random variable X is said to be* completely-selfdecomposable *if for each* $0 < a < 1$

$$X \stackrel{d}{=} aX + R_a, \tag{2.36}$$

for some completely-selfdecomposable random variable R_a independent of X.

The distribution of a m- (or completely-) selfdecomposable random variable is also said to be of *class L_m (L_∞)*.

It is also shown in Urbanik [105, Corollary 1] that the class L_∞ of completely-selfdecomposable laws is the smallest set containing all stable laws and closed under convolutions and convergence in law. See also [67, 80, 94, 57] for other related works on the class of L_m or L_∞.

To avoid the self-reference in the definition of completely-selfdecomposability above, the following alternative definition is employed by e.g. Roynette and Yor [93]:

Definition 2.4.3 *A random variable X_0 is said to be* completely-selfdecomposable *if for each $n \in \mathbb{Z}_+$ and for each $0 < a_i < 1, i \in \mathbb{Z}_+$,*

$$X_{i-1} \stackrel{d}{=} a_i X_{i-1} + X_i, \qquad i = 1, \cdots n, \tag{2.37}$$

where $X_1, \cdots, X_n$ is a sequence of independent random variables.

Proposition 2.4.4 *The T-Z random variable S_0 so defined as in Proposition 2.2.1 is not completely-selfdecomposable.*

Proof We show by contradiction. Suppose S_0 is completely-selfdecomposable, then for each $n \in \mathbb{Z}_+$ and each $0 < \gamma_i < 1, i \in \mathbb{Z}_+$, there exist random variables $S_{\gamma_1,\gamma_2}, \cdots S_{\gamma_1,\cdots,\gamma_n}$ such that

$$S_0 \stackrel{d}{=} \gamma_1 S_0 + S_{\gamma_1}, \tag{2.38}$$

$$S_{\gamma_1} \stackrel{d}{=} \gamma_2 S_{\gamma_1} + S_{\gamma_1,\gamma_2}, \tag{2.39}$$

$$\vdots$$

$$S_{\gamma_1,\cdots,\gamma_{n-1}} \stackrel{d}{=} \gamma_n S_{\gamma_1,\cdots,\gamma_{n-1}} + S_{\gamma_1,\cdots,\gamma_n}, \tag{2.40}$$

where these random variables are completely-selfdecomposable hence infinitely divisible. Suppose S_* has Lévy density $h_*(x)/x$, where all h-functions here are non-increasing on $\mathbb{R}_-$ and $\mathbb{R}_+$ by selfdecomposability.

To calculate these h-functions, notice that

- first,

$$\log(\phi_0(\lambda)) = \log(\pi\lambda) - \log(\sinh \pi\lambda), \tag{2.41}$$

$$\log(\phi_{\gamma_1}(\lambda)) = \log(\phi_0(\lambda)) - \log(\phi_0(\gamma_1\lambda)), \tag{2.42}$$

$$\log(\phi_{\gamma_1,\gamma_2}(\lambda)) = \log(\phi_{\gamma_1}(\lambda)) - \log(\phi_{\gamma_1}(\gamma_2\lambda)), \tag{2.43}$$

$$\vdots$$

- second,

$$\mathcal{F}^* \left(\frac{d}{d\lambda} \log(\sinh \alpha\pi\lambda) \right) = -\frac{\coth\left(\frac{\pi x}{2\alpha}\right)}{2x}, \tag{2.44}$$

  where $\mathcal{F}^*$ is inverse Fourier transform; and

- third, the linearity of operators $\frac{d}{d\lambda}$ and $\mathcal{F}^*$.

For short, set operator

$$\mathcal{H} := x\mathcal{F}^* \frac{d}{d\lambda}. \tag{2.45}$$

Then we get

$$h_*(x) = x \cdot \frac{h_*(x)}{x} = x\mathcal{F}^* \left(\frac{d}{d\lambda} \log(\phi(\lambda)) \right) = \mathcal{H}(\log \phi) \tag{2.46}$$

and

$$-\mathcal{H}(\log(\sinh \pi\lambda)) = \tfrac{1}{2} \coth(\tfrac{1}{2}\pi x). \tag{2.47}$$

Therefore,

$$\begin{aligned} h_0(x) =& \mathcal{H}(\log \phi_0) = \mathcal{H}(\log \pi\lambda) - \mathcal{H}(\log(\sinh \pi\lambda)) \\ =& -\mathcal{H}(\log(\sinh \pi\lambda)) = \frac{1}{2}\coth\left(\frac{\pi x}{2}\right); \end{aligned} \tag{2.48}$$

$$\begin{aligned} h_{\gamma_1}(x) =& \mathcal{H}(\log \phi_{\gamma_1}) = \mathcal{H}(\log(\phi_0(\lambda))) - \mathcal{H}(\log(\phi_0(\gamma_1\lambda))) \\ =& -\mathcal{H}(\log(\sinh \pi\lambda)) + \mathcal{H}(\log(\sinh \gamma_1\pi\lambda)) = \frac{1}{2}\left[\coth\left(\frac{\pi x}{2}\right) - \coth\left(\frac{\pi x}{2\gamma_1}\right)\right], \end{aligned} \tag{2.49}$$

$$h_{\gamma_1,\gamma_2}(x) = \frac{1}{2}\left[\coth\left(\frac{\pi x}{2}\right) - \coth\left(\frac{\pi x}{2\gamma_1}\right) - \coth\left(\frac{\pi x}{2\gamma_2}\right) + \coth\left(\frac{\pi x}{2\gamma_1\gamma_2}\right)\right], \tag{2.50}$$

$$\begin{aligned} h_{\gamma_1,\gamma_2,\gamma_3}(x) =& \frac{1}{2}\left[\coth\left(\frac{\pi x}{2}\right) - \coth\left(\frac{\pi x}{2\gamma_1}\right) - \coth\left(\frac{\pi x}{2\gamma_2}\right) - \coth\left(\frac{\pi x}{2\gamma_3}\right)\right. \\ & \left. + \coth\left(\frac{\pi x}{2\gamma_1\gamma_2}\right) + \coth\left(\frac{\pi x}{2\gamma_1\gamma_3}\right) + \coth\left(\frac{\pi x}{2\gamma_2\gamma_3}\right) - \coth\left(\frac{\pi x}{2\gamma_1\gamma_2\gamma_3}\right)\right], \end{aligned} \tag{2.51}$$

$$\vdots$$

An interesting fact is that all these functions above are symmetric with respect to γ_k's.

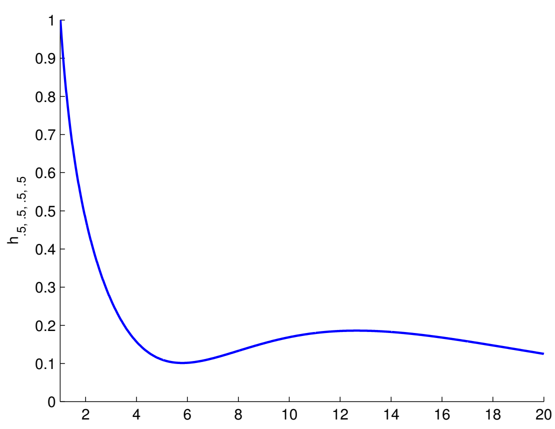


Figure 2.5: $h_{\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}}$

The counterexample appears when $n = 4$ and $\gamma_k = \frac{1}{2}, k = 1, 2, 3, 4$ since $h_{\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}}$ is no longer monotonically decreasing, as shown in figure 2.5. The desired result then follows. ∎

In fact, we proved that S_0 is not 4-selfdecomposable. However, it remains open if S_0 is either 2-selfdecomposable or 3-selfdecomposable.

Chapter 3

Self-similar additive processes

In this chapter, we discuss self-similar stochastic processes with independent increments, which are closely related to the selfdecomposable laws discussed in Chapter 2. It is shown that for a non-decreasing self-similar stochastic process T with independent increments, the range of T forms a Poisson point process with σ-finite intensity if and only if the one-dimensional distribution of $T(1)$ is of the gamma type. This follows from a general hold-jump description of such processes T, and implies the known result that the spacings between consecutive points of a scale invariant Poisson point process, with intensity $\theta x^{-1}dx$, are the points of another scale invariant Poisson point process with the same intensity.

3.1 The scale invariant Poisson spacings theorem

A point process $N_X(B) = \#\{z \in \mathbb{Z} : X_z \in B\}$, counting numbers of points in subintervals B of the positive half-line $\mathbb{R}_+ := (0, \infty)$, for some indexing of these points X_z by z in the set of integers $\mathbb{Z}$, is called *scale invariant* if for each $c > 0$ the point process N_{cX}, counting the scaled points $\{cX_z : z \in \mathbb{Z}\}$, has the same finite-dimensional distributions as N_X, as B varies over subintervals of $\mathbb{R}_+$. Assuming the point process N_X is *simple*, meaning the points X_z are all distinct, the point process N_X is then regarded as encoding the random countable set

$$range(X_z, \in \mathbb{Z}) := \{X_z, \in \mathbb{Z}\}. \tag{3.1}$$

So the identity in distribution of simple point processes N_X and N_{cX} may be indicated by the notation

$$range(X_z, z \in \mathbb{Z}) \stackrel{d}{=} range(cX_z, z \in \mathbb{Z}). \tag{3.2}$$

It is well known that a *Poisson point process* (*PPP*) on $\mathbb{R}_+$ is scale invariant if and only if its intensity measure is $\theta x^{-1}dx$ for some $\theta \geq 0$, when the process is called a $PPP(\theta x^{-1}dx)$, or a *scale invariant Poisson point process* with *rate* θ. So the parameter θ is the intensity of such a point process relative to the scale invariant measure $x^{-1}dx$ on the positive half-line, and for $0 < a < b < \infty$, the number of points in (a, b) has a Poisson distribution with mean $\theta \log(b/a)$.

The following result on scale invariant Poisson spacings arises in various contexts. The case $\theta = 1$ is contained in the theory of records and extremal processes, first developed by Dwass [23, 24, 25] and further studied by Resnick and Rubinovitch [91], and Shorrock [98]. The formulation for general $\theta > 0$ is due to Arratia [3, 2], who sketched a proof which was later detailed by Arratia, Barbour and Tavaré [4, Section 7].

Theorem 3.1.1 (Scale invariant Poisson spacings) *Fix $\theta > 0$. Let $(T_z, z \in \mathbb{Z})$, with $T_z < T_{z+1}$ for all $z \in \mathbb{Z}$, be an exhaustive listing of the points of a scale invariant PPP with rate θ. Then*

$$range(T_{z+1} - T_z, z \in \mathbb{Z}) \stackrel{d}{=} range(T_z, z \in \mathbb{Z}). \tag{3.3}$$

Less formally, the theorem states:

- *the spacings between consecutive points of a scale invariant Poisson point process on the positive half-line are the points of another scale invariant Poisson point process with the same rate.*

This chapter places Theorem 3.1.1 in the broader context of stochastic processes $T = (T(s), s > 0)$ which are *self-similar additive* (*SSA*), meaning that

- T is *self-similar*: for all $c > 0$,

$$(T(cs), s > 0) \stackrel{d}{=} (cT(s), s > 0); \tag{3.4}$$

- T is *additive*: meaning T has independent increments, and T is stochastically continuous with càdlàg paths starting at 0

$$\lim_{s \to 0+} T(s) = 0. \tag{3.5}$$

Such a process T is also called

- a *process of class* L [96], as the distribution of $T(1)$ is of class L, a subclass of infinitely divisible laws studied by Lévy [71], or
- a *Sato process* [10], and these processes were studied in depth by Sato [96, 95].

Here we focus on $range(T)$, the set of all values ever visited by T, regarded as a random subset of $\mathbb{R}_+$, for a non-negative, hence non-decreasing SSA process T with no drift component. As recalled in Section 3.3, it is known that for such an SSA process T, the Lévy-Khinchine representation of the Laplace transform of $T(1)$ has the special form

$$\log \mathbb{E} e^{-uT(1)} = \int_0^\infty \left(e^{-ux} - 1\right) \frac{k(x)}{x} dx, \tag{3.6}$$

for a uniquely determined right-continuous non-increasing *key function* $k : \mathbb{R}_+ \to \mathbb{R}_+$ subject to

$$\int_0^\infty (x \wedge 1) \frac{k(x)}{x} dx < \infty. \tag{3.7}$$

So the Lévy measure of $T(1)$ has a density relative to Lebesgue measure dx of the form $k(x)/x$ for such a key function $k(x)$.

Consider now the random countable set of jump times of T:

$$\{s > 0 : T(s) > T(s-)\}. \tag{3.8}$$

Let $k(0+) := \lim_{x \to 0+} k(x)$. It follows easily from self-similarity of T that

- either $k(0+) = \infty$ and the set of jump times is dense in $\mathbb{R}_+$;
- or $k(0+) < \infty$ and the jump times are the points of a scale invariant PPP with rate $\theta = k(0+)$, when we say the SSA process T has *(finite) rate* θ.

The identification of the rate $\theta = k(0+)$ is part of an explicit hold-jump construction of T with finite rate, provided later in Theorem 3.4.1. Taking the jump sizes into consideration, as well as the jump times, the Lévy-Itô representation of jumps, discussed in Section 3.4, shows that the random countable set of ordered pairs of jump times and jump sizes

$$\{(s, T(s) - T(s-)) : s > 0, T(s) > T(s-)\} \tag{3.9}$$

is the set of points of a scale invariant Poisson point process on $\mathbb{R}_+^2$. This leads to the following theorem:

Theorem 3.1.2 *For a SSA non-decreasing process T, with no drift component and key function $k(x)$, if $k(0+) = \theta$ is finite, then*

(I) $range(T) := \{T(s) : s > 0\}$ is a scale invariant point process on $\mathbb{R}_+$ with rate θ;

(II) the set of jump times of T is a scale invariant PPP on $\mathbb{R}_+$ with rate θ;

(III) the set of jump sizes of T is a scale invariant PPP on $\mathbb{R}_+$ with rate θ.

Here, (II) and (III) are just two coordinate projections of the self-similar Poisson point process (3.9) in the positive quadrant. Note that in (I) it is not asserted that $range(T)$ is Poisson point process, only that this point process is scale-invariant with mean intensity measure $\theta x^{-1} dx$. Beyond that, we know rather little about attributes of $range(T)$ as a point process on $\mathbb{R}_+$, such as its higher order factorial moments or Janossy measures, except in the special case when $T(1)$ is gamma distributed, and range(T) turns out to be Poissonian.

To provide a more detailed description of the range of a non-decreasing SSA process T with finite rate, let the scale invariant PPP of jump times of T be indexed by $\mathbb{Z}$ in an increasing way, say

$$0 < \cdots < S_{-1} < S_0 < S_1 < \cdots < \infty, \tag{3.10}$$

and define for each $z \in \mathbb{Z}$

$$T_z := T(S_z). \tag{3.11}$$

So the jump of T at time $s = S_z$ is from $T_{z-1} = T(S_z-)$ to $T_z = T(S_z)$. See Figure 3.1 for an illustration. Then the random countable range of T is

$$range(T(s), s > 0) = range(T_z, z \in \mathbb{Z}) \tag{3.12}$$

so the points in the range of T are also indexed by $\mathbb{Z}$ in increasing order

$$0 < \cdots < T_{-1} < T_0 < T_1 < \cdots < \infty, \tag{3.13}$$

as in the setup in for scale-invariant Poisson spacings in Theorem 3.1.1.

Here we are not specific about how S_0 is selected from the random set of jump times of T, since the choice is not important when considering the simple point process $\{(T_z, T_{z+1} - T_z) : z \in \mathbb{Z}\}$ as a random set on $\mathbb{R}_+^2$. It will be more convenient to define S_0 either so that $S_0 \leq 1 < S_1$ or so that $T_0 \leq 1 < T_1$.

Thus for a scale-invariant point process constructed as the range of a non-decreasing SSA process T with finite rate θ, comparing the point process of spacings between points, as on the left side of (3.3), and the points themselves on the right side of (3.3),

- the spacings between points are the jump sizes of T:

$$range(T_{z+1} - T_z, z \in \mathbb{Z}) = range(T(s) - T(s-) > 0 : s > 0) \tag{3.14}$$

  which forms a scale invariant PPP with rate θ, no matter what the key function $k(x)$ of T with $k(0+) = \theta$;

- the points themselves form the range of T:

$$range(T_z, z \in \mathbb{Z}) = range(T(s), s > 0) \tag{3.15}$$

  which is a scale invariant point process with rate θ, which might or might not be Poissonian, depending on the choice of the key function.

So to establish the spacings theorem 3.1.1, for a scale-invariant Poisson process with rate θ, it only remains to show that for a suitable choice of the key function $k(x)$ with $k(0+) = \theta$, the range of T is in fact Poissonian. That scale-invariant Poisson feature of range(T), with rate $\theta = 1$, was established by Dwass in the 1960s for the *SSA exponential process* T, with $k(x) = e^{-\lambda x}$ for some $\lambda > 0$, so $T(1) \sim \text{gamma}(1, \lambda)$ (or $\exp(\lambda)$) has the exponential distribution with rate λ, with tail probability

$$\mathbb{P}(T(1) > x) = e^{-\lambda x} 1(x > 0). \tag{3.16}$$

Our main point here is that the Poisson spacings theorem for a general rate $\theta > 0$ is a similar consequence of the following result, which we establish in Sections 3.5 and 3.6:

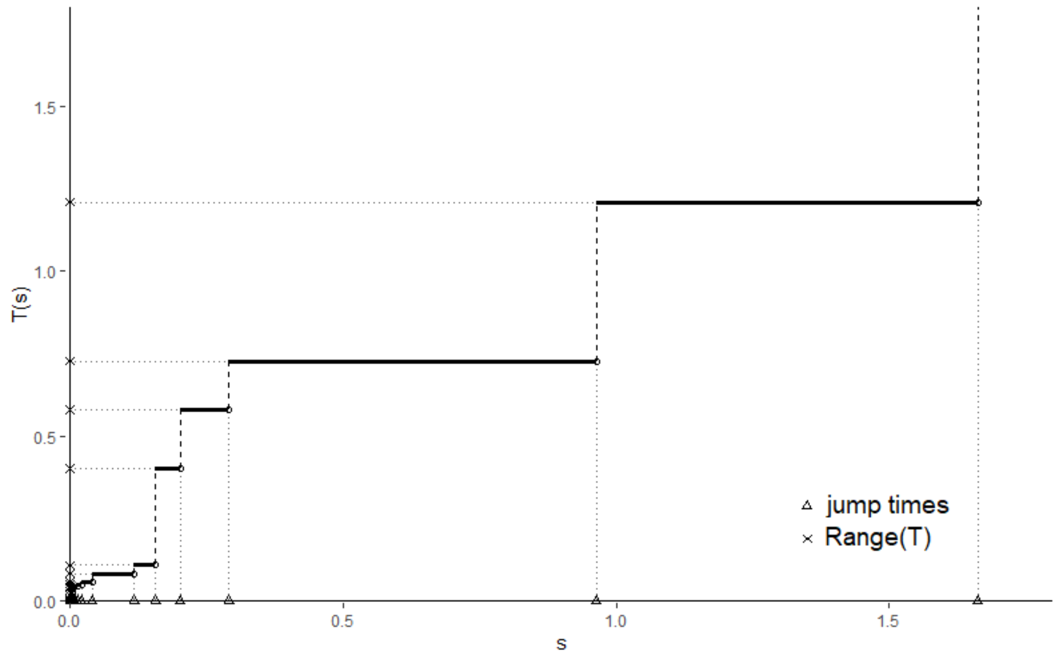


Figure 3.1: An SSA gamma process $(T(s), s > 0)$, its range, and jump times

Theorem 3.1.3 *Fix $\theta > 0$. For an SSA process $T := (T(s), s > 0)$,*

- *range(T) is a scale invariant PPP on $\mathbb{R}_+$ with rate θ,*

if and only if

- *the distribution of $T(1)$ is gamma(θ, λ) for some $\lambda > 0$, with probability density*

$$\frac{d}{dx}\mathbb{P}(T(1) \le x) = \frac{\lambda^\theta}{\Gamma(\theta)} x^{\theta-1} e^{-\lambda x} 1(x > 0). \tag{3.17}$$

Implicit in this result is a well known consequence of the Lévy-Khinchine formula:

the gamma(θ, λ) distribution of $T(1)$ has key function $k(x) = \theta e^{-\lambda x}$. (3.18)

However, some effort is required to pass from this fact to either of the implications of Theorem 3.1.3.

The following sections of this chapter is organized as follows:

- Section 3.2 collects for later use some easy results about a scale invariant PPP.
- Section 3.3 recalls basic facts about SSA processes and selfdecomposable laws.

- Section 3.4 characterizes an SSA non-decreasing process T with finite rate by its generic jump distribution, which can be read from the key function in the Lévy-Khinchine representation. The corresponding *hold-jump description* of T is the basis of proofs in Section 3.5.
- Section 3.5 offers proves the "if" part of Theorem 3.1.3, that the range of an SSA gamma process is a PPP, in two different ways, based on the hold-jump description provided in Section 3.4.
- Section 3.6 provides a general uniqueness theorem which includes the converse of Theorem 3.1.3: under a technical condition, the distribution of the range of a SSA non-decreasing process uniquely determines the distribution of the process itself, up to a scale factor.
- Section 3.7 investigates the conditional distribution of jump times given all jump magnitudes.
- Section 3.8 briefly discusses three different processes associated with a selfdecomposable law, then introduces a two-parameter process which provides a coupling of SSA processes with stationary independent increments in a second parameter.
- Section 3.9 raises an open question regarding the ratio of i.i.d. copies of a non-negative random variable, as an extension to results appeared previous sections.
- Section 3.10 offers some historical notes, including how the case $\theta = 1$ of Theorem 3.1.1 arises in the theory of extremal processes, and how the general case $\theta > 0$ is related to the Ewens sampling formula.

3.2 Scale invariant point processes

If a point process $X := \{X_z, z \in \mathbb{Z}\}$ on $\mathbb{R}_+$ is *scale invariant*, i.e.

$$range(X) \stackrel{d}{=} range(cX), \qquad \forall c > 0, \tag{3.19}$$

then X is a simple point process with intensity measure $\theta x^{-1}dx$ for some θ, with $0 < \theta \leq \infty$. Then we say X is a scale invariant point process on $\mathbb{R}_+$ with *rate* θ. We observe that

- the *inversion* $1/X := \{1/X_z, z \in \mathbb{Z}\}$ of a scale invariant point process X on $\mathbb{R}_+$ is a scale invariant point process on $\mathbb{R}_+$ with the same rate.

By considering $L := \log X$, this corresponds to a well known fact about stationary point processes on $\mathbb{R}$. Just as not all stationary point processes on the line are reversible, not all scale invariant point process are invariant under inversion. However, the distribution of a Poisson point process is entirely determined by its intensity measure, hence:

- if $X := \{X_z, z \in \mathbb{Z}\}$ is a scale invariant PPP then

$$range(1/X) \stackrel{d}{=} range(X). \tag{3.20}$$

The following lemma provides some useful characterizations of a scale invariant PPP. Recall that for U with uniform distribution on $[0,1]$, and $\theta > 0$, the distribution of $U^{1/\theta}$ is the beta$(\theta, 1)$ distribution on $[0,1]$ with probability density

$$\frac{d}{du}\mathbb{P}(U^{1/\theta} \leq u) = \theta u^{\theta-1} 1(0 < u < 1). \tag{3.21}$$

Lemma 3.2.1 *Fix $x > 0$ and $\theta > 0$. Suppose a scale invariant point process $X = \{X_z, z \in \mathbb{Z}\}$ is indexed in increasing manner with*

$$0 < \cdots < X_{-2} < X_{-1} < X_0 \leq x < X_1 < X_2 < \cdots < \infty. \tag{3.22}$$

Then the following three conditions are equivalent:

- *X is a scale invariant PPP with rate θ;*
- *x/X_1 and X_{z-1}/X_z for $z \geq 2$ are i.i.d. beta$(\theta, 1)$ random variables;*
- *X_0/x and X_{z-1}/X_z for $z < 0$ are i.i.d. beta$(\theta, 1)$ random variables.*

Proof Consider $L := \{L_z = \log(X_z), z \in \mathbb{Z}\}$. Then L is a stationary PPP with intensity measure $\theta d\ell$ on $\mathbb{R}$ if and only if X is a scale invariant PPP with rate θ. The statements of the lemma are just transformations of well known characterizations of a stationary PPP on the line. ■

3.3 Selfdecomposable laws and SSA processes

Following Sato [96, 95], we call a process $T := (T(s), s > 0)$ *self-similar with exponent H* if

$$(T(cs), s > 0) \stackrel{d}{=} (c^H T(s), s > 0) \qquad (c > 0), \tag{3.23}$$

in the sense of equality of finite-dimensional distributions. If T is also additive, as defined in Section 3.1, we say T is *H-self-similar additive (H-SSA)*. We omitted the prefix 'H-' when $H = 1$ in (3.4), because it has no impact on our discussion of $range(T)$ thanks to Lemma 3.3.2 below.

Easily from the definition, for an H-SSA process T, the one-dimensional distribution of $T(1)$ is *selfdecomposable*, discussed in Section 2.1. Sato [96] gave the following uniqueness theorem for the relationship between selfdecomposable laws and H-SSA processes.

Theorem 3.3.1 *For each H-SSA process $(T(s), s > 0)$, the marginal distribution of $T(s)$ is selfdecomposable. And for each selfdecomposable distribution μ and each $H > 0$, there exists an H-SSA process $(T(s), s > 0)$, unique in finite-dimensional distributions, such that $T(1)$ has distribution μ.*

We restrict the discussion to the case of $H = 1$ for most of the rest of this chapter, thanks to the following lemma.

Lemma 3.3.2 *Suppose $(T(s), s > 0)$ is an H-SSA non-decreasing process. Then the time change*

$$\tilde{T}(s) = T(s^{1/H}) \qquad (s > 0), \tag{3.24}$$

gives a 1-SSA non-decreasing process $(\tilde{T}(s), s \geq 0)$ with $range(\tilde{T}) = range(T)$.

We are primarily interested in the case of a 1-SSA process T that is non-decreasing and with no drift. Then (2.6) and (2.7) reduce to the formulas (3.6) and (3.7) for the Laplace transform of $T(1)$.

So the Lévy density of $T(s)$ at $x > 0$ is $k(x/s)/x$. As a result, the distribution of T is fully characterized by the *key function* $k(x)$, as detailed further in Section 3.4.

3.4 Hold-jump description

This section presents the relationship between the rate $\theta < \infty$ and the key function $k(x)$ of a 1-SSA non-decreasing process with finite rate θ. The hold-jump description after the jump over 1 will then be introduced as a framework to describe $range(T) \cap [1, \infty)$, as required to check whether $range(T)$ is Poisson using Lemma 3.2.1.

Rate and generic jump

Theorem 3.4.1 *Suppose $T := (T(s), s > 0)$ is a 1-SSA non-decreasing process with no drift, finite rate θ and key function $k(x)$. Then $k(0+) = \theta$ and*

$$T(s) = \sum_{z \in \mathbb{Z}} S_z J_z 1(S_z \leq s), \qquad (s > 0) \tag{3.25}$$

where

- *the jump times $(S_z, z \in \mathbb{Z})$ are the points of a scale invariant PPP with rate θ;*

Assuming also that these jump times are listed in an order depending only on their point process (3.8), for instance in an increasing order with S_0 the time of the first jump after time $s = 1$,

- *the* normalized jumps $(J_z, z \in \mathbb{Z})$ *form a sequence of i.i.d. copies of a positive random variable* J*, called the* generic jump *of* T*, with tail probability*

$$\mathbb{P}(J > x) = \frac{k(x)}{\theta}, \qquad (x \geq 0); \tag{3.26}$$

- *the sequence of jump times* $(S_z, z \in \mathbb{Z})$ *is independent of the sequence of normalized jumps* $(J_z, z \in \mathbb{Z})$*.*

Proof The self-similarity of T ensures that there are no jumps at fixed times. Thus, as a non-decreasing additive process, $T(s)$ has the following almost surely unique Lévy-Itô representation [63, Theorem 16.3]:

$$T(s) = a_s + \int_0^s \int_0^\infty x\eta(dy\ dx), \qquad (s > 0). \tag{3.27}$$

where a_s is a non-decreasing function with $a_0 = 0$ and $\eta(dy\ dx)$ is a Poisson point process on $(0, \infty)^2$ satisfying

$$\int_0^s \int_0^\infty (x \wedge 1)\mathbb{E}\eta(dy\ dx) < \infty. \qquad (s > 0) \tag{3.28}$$

In fact, η characterizes the jump structure of T

$$\eta(\cdot) = \sum_y 1\{(s, \Delta_y) \in \cdot\} \tag{3.29}$$

where the summation extends over all times $y > 0$ with $\Delta_y := T(y) - T(y-) > 0$. Hence

$$T(s) = a_s + \sum_y \Delta_y 1(y \leq s) \qquad (s > 0). \tag{3.30}$$

We call this η the *underlying Poisson point process* of the non-decreasing SSA process T. Taking into account self-similarity of T, it is easy to see that $a_s \equiv a_1 s$. If we further require T to be a pure jump process, then $a_s \equiv 0$. Moreover, as we know the jump times are from a scale invariant PPP with rate θ, the representation (3.25) follows immediately from self-similarity. To finish the proof, it remains to show (3.26).

Thanks again to self-similarity, note that $\{(S_z, \Delta_{S_z} = S_z J_z), z \in \mathbb{Z}\}$ is an exhaustive listing of points from the underlying Poisson point process η on $\mathbb{R}_+^2$ with intensity measure

$$\nu(ds\ dy) := \mathbb{E}\eta(ds\ dy) = \theta s^{-1} ds\ \mathbb{P}(sJ \in dy). \tag{3.31}$$

Hence, for each Borel set $B \subset \mathbb{R}_+$ and $a > 0$,

$$\int_0^a \theta s^{-1}\mathbb{P}(sJ \in B)ds = \int_{y\in\mathbb{R}_+} dF_J(y) \int_{s\in\mathbb{R}_+} \theta s^{-1} 1(sy \in B, s < a)ds \tag{3.32}$$

$$(\text{set } s = x/y) \qquad = \int_{y\in\mathbb{R}_+} dF_J(y) \int_{x\in\mathbb{R}_+} \theta x^{-1} 1(x \in B, y > x/a)dx \tag{3.33}$$

$$= \int_{x\in B} \theta x^{-1} dx \int_{y>x/a} dF_J(y) \tag{3.34}$$

$$= \int_{x\in B} \frac{\theta\ \mathbb{P}(J > x/a)}{x} dx. \tag{3.35}$$

But the Lévy measure of $T(1)$ is

$$\frac{k(x)}{x}dx = \int_{s\in(0,1]} \theta s^{-1}\mathbb{P}(sJ \in dx)ds, \tag{3.36}$$

which implies (3.26) by setting $a = 1$ in (3.35).

■

To illustrate Theorem 3.4.1, we give two examples.

Example 3.4.2 Fix $\theta, \lambda > 0$. A 1-SSA gamma process $(T(s), s > 0)$ with $T(1) \sim \text{gamma}(\theta, \lambda)$ has exponential generic jump $J \sim \exp(\lambda)$ and finite rate θ, thanks to the key function given in (3.18).

Example 3.4.3 This example is implicit in the discussion of the least concave majorant of a one-dimensional Brownian motion by Groeneboom [45] and Pitman and Ross [84]. Following the notations in [45], let ω denote a standard Brownian motion starting at the origin. For $a > 0$, let $\sigma(a)$ be the last time that maximum of $\omega(t) - at$ is attained, and $\tau(a) := \sigma(1/a), a > 0$ and $\tau(0) = 0$. The process τ is 2-SSA non-decreasing process with no drift. By Lemma 3.3.2 it can be reduced to fit our 1-SSA framework through the time change

$$T(a) := \tau(\sqrt{a}) = \sigma(1/\sqrt{a}). \tag{3.37}$$

Then T is a 1-SSA process, whose range is the random set of times of vertices of the least concave majorant of Brownian motion. The underlying Poisson point process driving T has intensity measure ν which can be read from Groeneboom's description of the process τ:

$$\nu(ds\ dx) = \frac{1}{s\sqrt{x}}\phi\left(\sqrt{\frac{x}{s}}\right)\frac{1}{2}s^{-\frac{1}{2}}dsdx = |k'|\left(\frac{x}{s}\right)\frac{ds\,dx}{s^2}, \tag{3.38}$$

with $\phi(\cdot)$ the standard normal probability density function, and

$$|k'|(x) = -\frac{dk(x)}{x} = \frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}}x^{-\frac{1}{2}}e^{-\frac{1}{2}x} \qquad (x > 0), \tag{3.39}$$

the absolute first derivative of the key function $k(x) = \mathbb{P}(J > x)/2$ for $J \sim \text{gamma}(\frac{1}{2}, \frac{1}{2})$, indicating that T has rate $\frac{1}{2}$. Therefore, the known results that

- *range*(T) is a scale invariant point process on $\mathbb{R}_+$ with rate $\frac{1}{2}$ [84, Corollary 9], and
- the set of jump times of T is a scale invariant PPP on $\mathbb{R}_+$ with rate $\frac{1}{2}$ [45, Theorem 2.1]

are the instances of parts I and II of Theorem 3.1.2 for this particular 1-SSA process T. Because the key function $k(x)$ of $T(1)$ is not just an exponential, the "only if" part of Theorem 3.1.3 shows that range (T), the set of times of vertices of the least concave majorant of Brownian motion, is not a Poisson point process. See Pitman and Ouaki [81] for a deeper study of Markovian structure in the concave majorant of Brownian motion.

Theorem 3.4.1 also yields the following *hold-jump description* of a 1-SSA non-decreasing process.

Corollary 3.4.4 *For each fixed time $s > 0$, and $t \geq 0$, the future of T after time s, conditional on $T(s) = t$, can be constructed by*

- 'Hold' - *at level t till the random time $H_s \overset{d}{=} s\beta^{-1}$ for $\beta \sim$ beta$(\theta, 1)$, i.e.*

$$\frac{d}{dx}\mathbb{P}(H_s \leq x) = \theta x^{-\theta-1} s^{\theta} 1(x > s); \tag{3.40}$$

- 'Jump' - *up by $H_s J$, where J is the generic jump and is independent of H_s, i.e.*

$$\mathbb{P}(T(H_s) - T(H_s-) > y \mid H_s = s') = \mathbb{P}(s'J > Y) = \mathbb{P}(J > Y/s'); \tag{3.41}$$

- *then repeat, conditioning on $T(s') = t'$ for $s' = H_s, t' = t + H_s J$.*

By setting a fixed starting time $S_1 = s$, (3.40) *and* (3.41) *specify a homogeneous pure jump-type Markov process $((S_n, T_n), n \in \mathbb{Z}_+)$ with state space $\mathbb{R}_+^2$, whose* entrance law *is given by $T_1 = T(s) \overset{d}{=} sT(1)$. Moreover, apart from this entrance law, the same description applies with the fixed time s replaced by any stopping time σ relative to the filtration of T, on the event $(\sigma > 0)$.*

The following corollary of Theorem 3.4.1 gives all possible distributions of generic jumps.

Corollary 3.4.5 *For each* 1*-SSA non-decreasing process T with finite rate θ, its generic jump J satisfies*

$$\mathbb{E}\log^+(J) < \infty. \tag{3.42}$$

Conversely, for each $\theta < \infty$ and each positive random variable J satisfying (3.42)*, there exists a unique* 1*-SSA process T with no drift, rate θ and generic jump J.*

Proof As per (3.26), each right-continuous, non-decreasing function $k(x)$ uniquely determines θ and J, and vice versa. So the only thing to check is the integrable condition of the Lévy density $k(x)/x$ in (3.7),

$$\int_0^\infty (x \wedge 1)\frac{k(x)}{x}dx = \theta\left[\int_0^1 \mathbb{P}(J > x)dx + \int_1^\infty \mathbb{P}(J > x)x^{-1}dx\right] < \infty, \tag{3.43}$$

where the integral is restricted on $\mathbb{R}_+$ since T is non-decreasing. The first term is finite no matter what distribution of J is, while

$$\int_1^\infty \mathbb{P}(J > x)x^{-1}dx = \int_1^\infty \mathbb{P}(\log J > \log x)d(\log x) \tag{3.44}$$

$$= \int_0^\infty \mathbb{P}(\log J > r)dr = \mathbb{E}\log^+ J, \tag{3.45}$$

which finishes the proof. ■

The convergence condition (3.42) appeared first in Vervaat [106, Theorem 1.6b], in the discussion of stochastic difference equations. See also Wolfe [107, Theorem 1].

The jump over 1

In considering whether or not $range(T)$ is a Poisson process, Lemma 3.2.1 shows it is sufficient to examine the ratios of adjacent points in $range(T) \cap (1,\infty)$ only. The hold-jump description provides us with sufficient information to calculate the ratios as long as we know the joint distribution of where and when the jump of T over 1 is made. This is given by the following lemma.

Lemma 3.4.6 *Consider the jump over level t of a 1-SSA non-decreasing process $(T(s), s > 0)$ with no drift and finite rate θ. Suppose the jump is made at time $S(t)$ from $G_t := T(S(t-)) \leq t$ to $D_t := T(S(t)) > t$. Then*

$$\mathbb{P}(S(t) \in ds, G_t \in dg, D_t - G_t > y) = \frac{\theta ds}{s}\mathbb{P}(T(s) \in dg)\mathbb{P}(sJ > y), \quad \forall 0 \leq g \leq t \leq g + y. \tag{3.46}$$

In particular, for $t = 1$ and $y = 1 - g$,

$$\mathbb{P}(S(1) \in ds, G_1 \in dg) = \frac{\theta ds}{s}\mathbb{P}(T(s) \in dg)\mathbb{P}(sJ > 1 - g), \quad \forall 0 \leq g \leq 1. \tag{3.47}$$

Proof Since T has independent increments and the set of jump times is not dense,

$$\begin{aligned}\mathbb{P}(S(t) \in ds, G_t \in dg, D_t - G_t > y) &= \mathbb{P}(T(s-) \in dg)\mathbb{P}(T(s+ds) - T(s-) > y) \\ &= \mathbb{P}(T(s-) \in dg) \int_{x \in (y,\infty)} \nu(ds, dx) \\ &= \mathbb{P}(T(s) \in dg) \int_{x \in (y,\infty)} \frac{\theta ds}{s} F_J\left(\frac{dx}{s}\right) \\ &= \frac{\theta ds}{s}\mathbb{P}(T(s) \in dg)\mathbb{P}(sJ > y), \quad \forall 0 < g \leq t \leq g+y. \end{aligned} \tag{3.48}$$

■

Now we give the following theorem as the *hold-jump description after the jump over* 1.

Theorem 3.4.7 *Suppose $T := (T(s), s > 0)$ is a 1-SSA non-decreasing process with no drift and finite rate θ. Let J denote its generic jump. Let S_1 be the time when T jumps over 1*

$$T(S_1-) \leq 1 < T(S_1), \tag{3.49}$$

and $S_1 < S_2 < \cdots$ be the times of successive jumps s of T with $T(s) > T(s-)$ and $T(s) > 1$, and define

$$T_0 := T(S_1-) \leq 1 < T_1 := T(S_1) < T_2 := T(S_2) < \cdots \tag{3.50}$$

so that $T_1, T_2, \ldots$ is the increasing sequence of values greater than 1 which are attained by T on the successive intervals $[S_1, S_2), [S_2, S_3), \ldots$. Then the joint distribution of the two sequences $(S_n, n \geq 1)$ and $(T_n, n \geq 0)$ is determined as follows:

$$\mathbb{P}(S_1 \in ds, T_0 \in dt) = \frac{\theta ds}{s}\mathbb{P}(T(s) \in dt)\mathbb{P}(sJ > 1-t), \tag{3.51}$$

$$S_n = S_1 \left(\prod_{i=1}^{n-1} \beta_i\right)^{-1} \qquad (n = 2, 3, \ldots), \tag{3.52}$$

$$T_n = T_{n-1} + S_n J_n \qquad (n = 1, 2, \ldots), \tag{3.53}$$

where $\beta_1, \beta_2, \ldots$ and $J_1, J_2, \ldots$ are independent random variables, with

- *$\beta_1, \beta_2, \ldots$ all with the beta$(\theta, 1)$ distribution* (3.21)*;*
- *$J_1, J_2, \ldots$ identically distributed copies of the generic jump J.*

Proof Formula (3.51) is just (3.47) with different notation, while (3.52) and (3.53) are due to the hold-jump description presented in Corollary 3.4.4. ■

Proof of Theorem 3.1.2

Suppose that T is a non-decreasing 1-SSA process with no drift and finite rate θ. Then Part (II) holds by definition, while (III) is observed by setting $a = \infty$ in (3.35). To see (I), set $g = t$ and $y = 0$ in (3.46) of Lemma 3.4.6. Suppose the rate of $range(T)$ is α, then for each Borel set $B \subset \mathbb{R}_+$,

$$\begin{aligned}
\int_B \frac{\alpha dt}{t} &= \int_{s\in(0,\infty)} \int_{t\in B} \mathbb{P}(S(1) \in ds, G_1 \in dt, D_1 - G_1 > 0) \\
&= \int_{s\in(0,\infty)} \frac{\theta ds}{s} \int_{t\in B} \mathbb{P}(T(s) \in dt)\mathbb{P}(sJ > 0) \\
&= \int_{s\in(0,\infty)} \frac{\theta ds}{s} \int_{t\in s^{-1}B} \mathbb{P}(T(1) \in dt) \\
&= \int_{s\in(0,\infty)} \int_{t\in s^{-1}B} \frac{\theta ds}{s} dF_{T(1)}(t) \\
&= \int_{t\in(0,\infty)} \int_{s\in t^{-1}B} \frac{\theta ds}{s} dF_{T(1)}(t) \\
(\text{set } r = st) &= \int_{t\in(0,\infty)} \int_{r\in B} \frac{\theta dr}{r} dF_{T(1)}(t) = \int_{r\in B} \frac{\theta dr}{r},
\end{aligned} \tag{3.54}$$

which shows $\alpha = \theta$.

3.5 Proof that the range of an SSA gamma process is a PPP

This is the "if" part of Theorem 3.1.3. We offer two different proofs.

First proof - through the hold-jump description after the jump over 1

This argument shows that the description of $range(T) \cap (1, \infty)$ implied by Theorem 3.4.7 matches that required by Lemma 3.2.1. We follow the notation of Theorem 3.4.7, which contains a complete description of the joint distribution of the first arrival times and levels (S_n, T_n) at all levels $t > 1$ that are ever attained by the path of T. By scaling, it is enough to consider the case $\lambda = 1$. From Example 3.4.2, the generic jump is exp(1) distributed: $\mathbb{P}(J > x) = e^{-x}$ for $x > 0$, so (3.51) becomes

$$\mathbb{P}(S_1 \in ds, T_0 \in dt) = \frac{\theta ds}{s} \frac{1}{s^\theta \Gamma(\theta)} t^{\theta-1} e^{-\frac{t}{s}} dt \; e^{-(\frac{1-t}{s})} = \frac{s^{-1-\theta} e^{-\frac{1}{s}} ds}{\Gamma(\theta)} \theta t^{\theta-1} dt. \tag{3.55}$$

That means T_0 and S_1 are independent, with T_0 distributed beta$(\theta, 1)$ and S_1 distributed as inverse gamma$(\theta, 1)$, meaning S_1^{-1} distributed as the gamma$(\theta, 1)$.

Hence by the memoryless property of exponential random variables, we may assume $T_0 = 1$ without hurting the joint distribution of $(S_n, n \geq 1)$ and $(T_n, n \geq 1)$. The joint distribution is now re-written as follows:

$$S_n = \left(\gamma_\theta \prod_{i=1}^{n-1} \beta_i\right)^{-1} \qquad (n = 1, 2, \ldots) \tag{3.56}$$

$$T_n = T_{n-1} + S_n \varepsilon_n \qquad (n = 1, 2, \ldots) \tag{3.57}$$

where $\gamma_\theta, \beta_1, \beta_2, \ldots$ and $\varepsilon_1, \varepsilon_2, \ldots$ are independent random variables, with

- γ_θ assigned the gamma$(\theta, 1)$ distribution;
- $\beta_1, \beta_2, \ldots$ all with the beta$(\theta, 1)$ distribution;
- $\varepsilon_1, \varepsilon_2, \ldots$ all with the exp(1) distribution.

All the ingredients needed for calculating $range(T) \cap (1, \infty)$ are in view. But the description of the levels T_n is tangled up with the description of the times S_n in such a way that it is not immediately obvious why the sequence of ratios T_{n-1}/T_n is also a sequence of i.i.d. copies of $beta(\theta, 1)$. However, the argument is completed by the following lemma.

Lemma 3.5.1 *Suppose that random variables $S_1 := 1/\gamma_\theta$ and $S_{n+1} := S_n/\beta_n$ for $n \geq 1$ are defined by the recursion* (3.56)*, along with $T_0 := 1 < T_1 < T_2 < \cdots$ by* (3.57)*, from independent random variables γ_θ, β_i and ε_i as above. Then for each $n = 1, 2, \ldots$, the $n+1$ ratios*

$$\frac{T_0}{T_1}, \ldots, \frac{T_{n-1}}{T_n}, \frac{T_n}{S_n} \tag{3.58}$$

are independent, with the first n consecutive T-ratios all distributed according to the common beta$(\theta, 1)$ distribution of all the β_i, and with the last of the $n+1$ ratios

$$\frac{T_n}{S_n} \stackrel{d}{=} \gamma_{\theta+1}, \tag{3.59}$$

the gamma$(\theta + 1, 1)$ distribution.

Proof

For $n = 1$, with $T_0 := 1$

$$T_1 = 1 + \frac{\varepsilon_1}{\gamma_\theta} = \frac{\gamma_\theta + \varepsilon_1}{\gamma_\theta} \tag{3.60}$$

and hence

$$\frac{T_0}{T_1} = \frac{1}{T_1} = \frac{\gamma_\theta}{\gamma_\theta + \varepsilon_1} \stackrel{d}{=} \beta_{\theta,1} \tag{3.61}$$

and this variable T_0/T_1 is independent of

$$\frac{T_1}{S_1} := T_1 \gamma_\theta = \gamma_\theta + \varepsilon_1 \stackrel{d}{=} \gamma_{\theta+1} \tag{3.62}$$

by the beta-gamma algebra mentioned in Lukacs [74]. The case of general $n = 1, 2, 3, \ldots$ now follows by induction on n, starting from this base case $n = 1$. Multiply the recursion (3.57) by $1/T_{n+1} = \beta_n/T_n$ to see that

$$\frac{T_{n+1}}{S_{n+1}} = \frac{T_n}{S_n}\beta_n + \varepsilon_{n+1} \stackrel{d}{=} \gamma_\theta + \varepsilon_1 \stackrel{d}{=} \gamma_{\theta+1} \tag{3.63}$$

because the gamma$(\theta + 1)$ distribution of T_n/S_n and the independence of this variable and $\beta_n \stackrel{d}{=} \beta_{\theta,1}$ makes their product $(T_n/S_n)\beta_n \stackrel{d}{=} \gamma_\theta$. Moreover

$$\frac{T_n}{T_{n+1}} = \frac{T_n/S_{n+1}}{T_{n+1}/S_{n+1}} = \frac{(T_n/S_n)\beta_n}{(T_n/S_n)\beta_n + \varepsilon_{n+1}} \stackrel{d}{=} \frac{\gamma_\theta}{\gamma_\theta + \varepsilon_1} \stackrel{d}{=} \beta_{\theta,1} \tag{3.64}$$

and this ratio T_n/T_{n+1} is independent of the ratio T_{n+1}/S_{n+1} in (3.63), again by beta-gamma algebra. Thus T_n/T_{n+1} and T_{n+1}/S_{n+1} are independent with the required distributions. By inductive assumption, T_n/S_n is independent of the n ratios T_{i-1}/T_i for $1 \le i \le n$. The two variables T_n/T_{n+1} and T_{n+1}/S_{n+1} are functions of T_n/S_n and two further independent variables β_n and ε_{n+1}. Hence the required independence of the $n + 2$ variables involved in (3.58) with $n + 1$ in place of n. ■

Second proof - to exploit symmetry

To see the symmetry, consider the *time-inversion* $\tilde{T} := (\tilde{T}(s) := T(s^{-1}), s > 0)$ of T. It is obvious that $\tilde{T}$ has a non-increasing staircase path, which is fully determined by its *corners*, i.e. points on the left end of each flat of the path.

To describe the corners, we restate, in a time-reversed manner, the *backward hold-jump description* of $\tilde{T}$. Conditional on $\tilde{T}(s) = t$, the *past* of $\tilde{T}$ before time s can be fully constructed by

- *'Hold'* - at level t till the random time $H_s \stackrel{d}{=} s\beta$ (this is going backward in time), where β has the common *beta*$(\theta, 1)$ distribution;
- *'Jump'* - up by $H_s^{-1}J$, where J is the generic jump and is independent of H_s;
- then repeat, conditioning on $\tilde{T}(s') = t'$ for $s' = H_s, t' = t + H_s^{-1}J$.

Lemma 3.5.2 *Suppose T is a 1-SSA gamma process with $T(1) \sim$ gamma$(\theta, 1)$ and $\tilde{T} := (\tilde{T}(s) := T(s^{-1}), s > 0)$ is the time-inversion of T. Then the set $\mathcal{C}$ of corners of the path of $\tilde{T}$ is symmetric about the bisectrix.*

Theorem 3.1.3 then follows since

$$range(T) = range(\tilde{T}) = \pi_t(\mathcal{C}) \stackrel{d}{=} \pi_s(\mathcal{C}), \tag{3.65}$$

where π_t and π_s are projections onto the t- and s-axes, and $\pi_s(\mathcal{C})$ is a scale invariant PPP on $\mathbb{R}_+$ with rate θ thanks the invariance under inversion (3.20).

Proof (Lemma 3.5.2)

Consider the joint density $p_m(a_1, \cdots, a_m)$ of the event that there are N consecutive points $a_i = (s_i, t_i) \in \mathcal{C}$ for $i = 1, 2, \cdots, N$ indexed decreasingly in s-coordinate

$$s_1 > s_2 > \cdots > s_N, \tag{3.66}$$

hence is indexed increasingly in t-coordinate. Knowing $J \sim \text{gamma}(1, 1)$,

$$p_N(a_1, \cdots, a_N) = \frac{s_1^\theta}{\Gamma(\theta)}(t_0)^{\theta-1}e^{-s_1 t_1} \cdot \frac{\theta}{s_1} \cdot \left[\prod_{n=1}^{N-1} \theta s_{i+1}^{\theta-1} s_i^{-\theta}\right] \cdot \left[\prod_{n=1}^{N-1} s_i e^{-s_i(t_{i+1}-t_i)}\right] \tag{3.67}$$

$$= \frac{\theta^N}{\Gamma(\theta)}(t_1 s_N)^{\theta-1} \exp\left\{\sum_{n=2}^{N-1} s_n t_n - \sum_{n=1}^{N-1} s_n t_{n+1}\right\}. \tag{3.68}$$

Everything in the first line is self-explanatory from the backward hold-jump description given above, except that the second term θ/s_1 accounts for a 'hold' with $\beta = 1$, meaning that there is an immediate jump at time s_1.

Expression (3.68) is invariant under the substitution

$$(\tilde{s}_1, \tilde{s}_2, \cdots, \tilde{s}_N) \leftrightarrow (\tilde{t}_N, \tilde{t}_{N-1}, \cdots, \tilde{t}_1), \tag{3.69}$$

which proved the symmetry as desired. ∎

This proof is inspired by Gnedin [37, Equation (5)] where a similar symmetry was shown for a different setup of corners constructed from a PPP on $\mathbb{R}_+^2$ with unit intensity. We also remark that the the proof of the Poisson spacing theorem by Arratia, Barbour and Tavaré [4, Lemma 7.1] is also done by checking the density of consecutive points. However, we manage to avoid the brutal integration in their proof by exploiting symmetry.

3.6 Uniqueness of the distribution of the range

In this section, we establish the following uniqueness theorem, from which the "if" part of Theorem 3.1.3 follows immediately.

Theorem 3.6.1 *Suppose T and $\tilde{T}$ are two 1-SSA non-decreasing processes with no drift and the same finite rate θ. Then*

$$\textit{range}(T) \stackrel{d}{=} \textit{range}(\tilde{T}) \qquad \textit{implies} \qquad T \stackrel{d}{=} c\tilde{T} \qquad \textit{for some} \qquad c > 0. \tag{3.70}$$

Proof Following (the results and also the notation of) Theorem 3.4.7, we may write S in terms of T and J in (3.52), then (3.53) becomes

$$\frac{T_n - T_{n-1}}{T_{n+1} - T_n} = \frac{\beta_n J_{n-1}}{J_n}, \qquad \forall n \geq 2, \tag{3.71}$$

where (still, as in Theorem 3.4.7) $1 < T_1 < T_2 < \cdots$ is an exhaustive ordered listing of points of T on $(1, \infty)$, β_n are i.i.d. Beta$(\theta, 1)$'s and J_n are i.i.d. copies of the generic jump J.

Similarly, for another 1-SSA process $\tilde{T}$ with rate θ whose range is equal in distribution as the one of T, we have

$$\frac{\tilde{T}_n - \tilde{T}_{n-1}}{\tilde{T}_{n+1} - \tilde{T}_n} = \frac{\tilde{\beta}_n \tilde{J}_{n-1}}{\tilde{J}_n}, \qquad \forall n \geq 2, \tag{3.72}$$

where every random variable is similarly defined as in (3.71) for $\tilde{T}$ instead of T.

By assumption, T and $\tilde{T}$ are both 1-SSA with the same finite rate θ. So it is sufficient to show $J \stackrel{d}{=} \tilde{J}$.

Note that $range(T) \stackrel{d}{=} range(\tilde{T})$ implies

$$(T_n, n \geq 1) \stackrel{d}{=} \left(\tilde{T}_n, n \geq 1\right). \tag{3.73}$$

Therefore, by (3.71) and (3.72),

$$\left(\frac{\beta_n J_{n-1}}{J_n}, n \geq 2\right) \stackrel{d}{=} \left(\frac{\tilde{\beta}_n \tilde{J}_{n-1}}{\tilde{J}_n}, n \geq 2\right). \tag{3.74}$$

By Lemma 3.6.2 below, (3.74) can be simplified to

$$\left(\frac{J_{n-1}}{J_n}, n \geq 2\right) \stackrel{d}{=} \left(\frac{\tilde{J}_{n-1}}{\tilde{J}_n}, n \geq 2\right) \tag{3.75}$$

which implies $J \stackrel{d}{=} c\tilde{J}$ thanks to Lemma 3.6.3. ■

This Lemma 3.6.2 is a multivariate extension of a simplified version of exercise 1.13.1 in Chaumont and Yor [12], since we only need the case when all coordinates are strictly positive.

Recall $\mathbb{R}_+ := (0, \infty)$. Suppose $Y = (Y_1, Y_2, \cdots, Y_n)$ is a $\mathbb{R}_+^n$-valued random variable. Let $\Phi_{\log(Y)}$ denote the characteristic function of $\log Y := (\log Y_1, \log Y_2, \cdots, \log Y_n)$

$$\Phi_{\log(Y)}(\lambda) := \mathbb{E} \exp\{i\lambda \cdot \log(Y)\}, \qquad \lambda \in \mathbb{R}^n, \tag{3.76}$$

where $\cdot$ is the usual inner product of vectors.

Lemma 3.6.2 (Multivariate simplifiable random variables) *If the non-zero set of the characteristic function of* $(\log Y)$, *i.e.* $\{\lambda : \Phi_{\log Y}(\lambda) \neq 0\}$, *is dense in* $\mathbb{R}^n$, *then* Y *is* multivariate simplifiable, *i.e. for all* $\mathbb{R}_+^n$*-valued random variables* X, Z *independent of* Y,

$$X \times Y \stackrel{d}{=} Z \times Y \qquad \textit{implies} \qquad X \stackrel{d}{=} Z, \tag{3.77}$$

where $\times$ *denotes the entry-wise product* $X \times Y := (X_1 Y_1, X_2 Y_2, \cdots, X_n Y_n)$.

In particular, for each $\theta > 0$, *if* $Y_1, Y_2, \cdots Y_n$ *are i.i.d. copies of* $\beta \sim$ *beta*$(\theta, 1)$, Y *is multivariate simplifiable.*

Proof

$$\Phi_{\log(X\times Y)}(\lambda) = \mathbb{E}\exp[i\lambda\cdot\log(X\times Y)] = \mathbb{E}\exp[i\lambda\cdot(\log X+\log Y)]$$
$$= \Phi_{\log X}(\lambda)\Phi_{\log Y}(\lambda), \qquad \forall\lambda\in\mathbb{R}^n. \quad (3.78)$$

and similarly $\Phi_{\log(Z\times Y)}(\lambda) = \Phi_{\log Z}(\lambda)\Phi_{\log Y}(\lambda)$. Hence, $X\times Y \stackrel{d}{=} Z\times Y$ implies

$$\Phi_{\log X}(\lambda)\Phi_{\log Y}(\lambda) = \Phi_{\log Z}(\lambda)\Phi_{\log Y}(\lambda), \qquad \lambda\in\mathbb{R}^n. \quad (3.79)$$

Then the cancellation of $\Phi_{\log Y}(\lambda)$ on a dense subset $\{\lambda : \Phi_{\log Y}(\lambda)\neq 0\}$ of $\mathbb{R}^n$ shows

$$\log X \stackrel{d}{=} \log Z \quad (3.80)$$

which easily implies $X\stackrel{d}{=} Z$.

Now it is left to show that $\{\lambda : \Phi_{\log Y}(\lambda)\neq 0\}$ is dense if $Y_1, Y_2, \cdots Y_n$ are i.i.d. copies of $\beta\sim$ beta$(\theta,1)$.

Observe that the characteristic function of $\log\beta$ is non-zero on $\mathbb{R}$

$$\phi_{\log\beta}(\lambda) = \mathbb{E}\exp[i\lambda\log(\beta)] = \mathbb{E}\beta_k^{i\lambda} = \int_0^1 x^{i\lambda}\,\theta x^{\theta-1}dx = \frac{\theta}{\theta+i\lambda}, \qquad \forall\lambda\in\mathbb{R}. \quad (3.81)$$

Hence the characteristic function of $\log Y$ is non-zero on $\mathbb{R}^n$. ■

Lemma 3.6.3 *Suppose X and Y are two positive random variables, with i.i.d. copies $X_n, Y_n, n = 1, 2, \cdots$, respectively. If for each n,*

$$\left(\frac{X_1}{X_2}, \frac{X_2}{X_3}, \cdots, \frac{X_{n-1}}{X_n}\right) \stackrel{d}{=} \left(\frac{Y_1}{Y_2}, \frac{Y_2}{Y_3}, \cdots, \frac{Y_{n-1}}{Y_n}\right), \quad (3.82)$$

then $Y\stackrel{d}{=} cX$ for some constant c.

Proof Identity (3.82) is equivalent to

$$\left(\frac{X_1}{X_2}, \frac{X_1}{X_3}, \cdots, \frac{X_1}{X_n}\right) \stackrel{d}{=} \left(\frac{Y_1}{Y_2}, \frac{Y_1}{Y_3}, \cdots, \frac{Y_1}{Y_n}\right), \quad (3.83)$$

which implies

$$\left(X_1\frac{Z_2}{X_2}, X_1\frac{Z_3}{X_3}, \cdots, X_1\frac{Z_n}{X_n}\right) \stackrel{d}{=} \left(Y_1\frac{Z_2}{Y_2}, Y_1\frac{Z_3}{Y_3}, \cdots, Y_1\frac{Z_n}{Y_n}\right), \quad (3.84)$$

where $Z_n, n = 2, 3, \cdots$ are i.i.d. copies of $Z\sim\exp(1)$ and are independent of the X- and Y-sequences.

The extra randomization Z is only to ensure that the cumulative distribution function of Z/X is continuous on $\mathbb{R}_+$. When $n := 2k \to \infty$, the $(k+1)$-th order statistic of the left sequence of (3.84) converges almost surely to $m_x X_1$ where $m_x > 0$ is the median of Z/X. Similarly, the $(k+1)$-th order statistic of the right sequence converges almost surely to $m_y Y_1$ where $m_y > 0$ is the median of Z/Y.

Hence

$$m_x X_1 \stackrel{d}{=} m_y Y_2, \tag{3.85}$$

which implies

$$X \stackrel{d}{=} cY \qquad \text{with} \qquad c = \frac{m_y}{m_x}. \tag{3.86}$$

■

3.7 Size-biased permutation

Suppose $T := (T(s), s > 0)$ is a 1-SSA non-decreasing process with finite intensity θ and generic jump J. We know that the set of pairs of jump times and jump magnitudes

$$range((s, T(s) - T(s-)) : s > 0, T(s) > T(s-)) \tag{3.87}$$

is a scale invariant PPP on $\mathbb{R}_+^2$ with intensity measure

$$\nu(ds\ dt) = \theta s^{-1} ds \mathbb{P}(sJ \in dt). \tag{3.88}$$

This is by no means symmetric. However, if we consider the time-inversion $\tilde{T} := (\tilde{T}(s) := T(s^{-1}), s > 0)$ which is employed once in Lemma 3.5.2, then

Theorem 3.7.1 *The set of pairs of jump times and jump magnitudes of* $\tilde{T}$

$$range((s^{-1}, T(s) - T(s-)) : s > 0, T(s) > T(s-)) \tag{3.89}$$

is a scale invariant PPP on $\mathbb{R}_+$ *with* symmetric *intensity measure*

$$\tilde{\nu}(ds\ dt) = \theta s^{-1} ds \mathbb{P}(s^{-1} J \in dt) \tag{3.90}$$

$$= \theta t^{-1} dt \mathbb{P}(t^{-1} J \in ds) \tag{3.91}$$

$$= \tilde{\nu}(dt\ ds). \tag{3.92}$$

Proof By easy measure transformation one obtain (3.90) from (3.88).

If J is continuous with density f_J, then (3.91) is also obvious by symmetry

$$\tilde{\nu}(ds\ dt) = \theta s^{-1} ds (s f_J(st) dt) = f_J(st) ds dt. \tag{3.93}$$

More generally, for any Borel sets $A, B \in \mathcal{B}$ (or simply intervals), expected number of points in $A \times B$, i.e. expected number of jumps of $\tilde{T}$ whose jump time is in A and jump magnitude is in B, is given by

$$\int_{s\in A} \theta s^{-1} ds \mathbb{P}(s^{-1}J \in B) = \int_0^\infty dF_J(y) \int_0^\infty \theta s^{-1} ds 1(s \in A, s^{-1}y \in B) \tag{3.94}$$

$$(\text{Put } t = s^{-1}y) \quad = \int_0^\infty dF_J(y) \int_0^\infty \theta t^{-1} dt 1(t^{-1}y \in A, t \in B) \tag{3.95}$$

$$= \int_{t\in B} \theta t^{-1} dt \mathbb{P}(t^{-1}J \in A). \tag{3.96}$$

Hence (3.91) follows. ■

Informally, (3.91) provides us with a way to generate jumps of $\tilde{T}$ given jump magnitudes:

- conditional on the set of jump magnitudes, a jump with magnitude t occurs at time $t^{-1}J$, conditionally independent of all other jumps.

Fix $s_0 > 0$. Consider the decomposition

$$\tilde{T}(s_0) = \sum_{i=1}^\infty \Delta_{(i)}, \tag{3.97}$$

where

$$\Delta_{(1)} \geq \Delta_{(2)} \geq \cdots \tag{3.98}$$

are the ranked valued of the first component of points (Δ, S) in a PPP on $(0, \infty) \times (s_0, \infty)$ with intensity measure

$$\tilde{\nu}(dt\ ds)1(s \geq s_0) = \left[\theta t^{-1}\mathbb{P}(t^{-1}J \geq s_0)dt\right] \mathbb{P}(t^{-1}J \in ds | t^{-1}J \geq s_0). \tag{3.99}$$

From the intensity measure above, we learn:

- $\Delta_{(1)} \geq \Delta_{(2)} \geq \cdots$ are ranked points from

$$PPP(\theta t^{-1}\mathbb{P}(t^{-1}J \geq s_0)dt) = PPP(t^{-1}k(ts_0)) \tag{3.100}$$

- The jump time S_i of $\Delta_{(i)}$ is given by

$$S_i = (J_i/\Delta_{(i)} \mid J_i/\Delta_{(i)} \geq s_0), \tag{3.101}$$

  where $(J_i, i = 1, 2, \cdots)$ is a sequence of i.i.d. copies of J independent of all $\Delta_{(i)}$'s.

Theorem 3.7.2 *Suppose T is 1-SSA non-decreasing with finite intensity. Set $S_{(1)} < S_{(2)} < \cdots$ the order statistics of S_i's as defined in* (3.101) *and $\Delta_1, \Delta_2, \cdots$ the corresponding Δ-values, then $(\Delta_1, \Delta_2, \cdots)$ is a size-biased permutation of $\Delta_{(i)}$ if and only if T is a 1-SSA gamma process.*

Proof (‘If’) For T a 1-SSA gamma process, without loss of generality set the generic jump $J \stackrel{d}{=} \varepsilon \sim \text{Exp}(1)$. Then by memoryless property, (3.101) becomes

$$S_i = \varepsilon_i/\Delta_{(i)} + s_0, \tag{3.102}$$

where ε_i is a sequence of i.i.d. copies of ε. Hence ordering by S-values is the same as ordering by $\varepsilon_i/\Delta_{(i)}$, and the desired result follows by [82, Lemma 4.4].

(‘Only if’) Since $(\Delta_1, \Delta_2, \cdots)$ is a size-biased permutation,

$$\mathbb{P}(S_1 < S_2 \mid \Delta_{(i)} = \delta_i, i = 1, 2, \cdots) =: \mathbb{P}_\delta(S_1 < S_2) = \frac{\delta_1}{\delta_1 + \delta_2}. \tag{3.103}$$

For simplicity we only prove for J continuous with differentiable tail probability $g(x) = \mathbb{P}(J > x) = k(x)/\theta$. Thus,

$$\frac{\delta_1}{\delta_1 + \delta_2} = \mathbb{P}\left(\frac{J_1}{\delta_1} < \frac{J_2}{\delta_2} \mid \frac{J_1}{\delta_1}, \frac{J_2}{\delta_2} \geq s_0\right) = \frac{\mathbb{P}(s_0 \leq \frac{J_1}{\delta_1} < \frac{J_2}{\delta_2})}{\mathbb{P}(\frac{J_1}{\delta_1}, \frac{J_2}{\delta_2} \geq s_0)}, \tag{3.104}$$

$$\delta_2 \mathbb{P}\left(s_0 \leq \frac{J_1}{\delta_1} < \frac{J_2}{\delta_2}\right) = \delta_1 \mathbb{P}\left(\frac{J_1}{\delta_1} > \frac{J_2}{\delta_2} \geq s_0\right), \tag{3.105}$$

$$\delta_2 \int_{\delta_1 s_0}^{\infty} g(dx) \int_{\delta_1^{-1}\delta_2 x}^{\infty} g(dy) = \delta_1 \int_{\delta_2 s_0}^{\infty} g(dx) \int_{\delta_2^{-1}\delta_1 x}^{\infty} g(dy). \tag{3.106}$$

Note that the above equations are true for all non-increasing sequences $\delta_1 \geq \delta_2 \geq \cdots$ and $s_0 > 0$, hence by differentiation with respect to s_0,

$$\delta_2 g(\delta_1 s_0) \frac{d}{ds_0} g(\delta_2 s_0) = \delta_1 g(\delta_2 s_0) \frac{d}{ds_0} g(\delta_1 s_0), \tag{3.107}$$

$$\frac{d}{ds_0} \ln g(\delta_2 s_0) = \frac{d}{ds_0} \ln g(\delta_1 s_0), \qquad \forall \delta_1 > \delta_2 > 0, s_0 > 0. \tag{3.108}$$

The solutions of the differentiation equation (3.108), satisfying the boundary conditions $g(0) = 1$ and $g(+\infty) = 0$ are given by

$$g(x) = e^{-\lambda x}, \qquad (x > 0) \tag{3.109}$$

for all fixed $\lambda > 0$, i.e. $J \sim \text{Exp}(\lambda)$. This implies that T is a 1-SSA gamma process. ■

3.8 Processes associated with selfdecomposable laws

In this section, we first introduce three different kinds of one-parameter processes associated with selfdecomposable laws, which are known in the previous study related to selfdecomposable laws. Then we introduce a two-parameter process with selfdecomposable margins that behaves differently along its two parameters.

One-parameter processes

Let X be a selfdecomposable random variable. According to the result of Sato [96] presented in Theorem 3.3.1,

- for each $H > 0$, there is a unique distribution of an H-SSA process $T^{(H)} := (T^{(H)}(s), s > 0)$ such that $T^{(H)}(1) \stackrel{d}{=} X$.

On the other hand, since X is infinitely divisible,

- there is a unique Lévy process $U := (U(s), s > 0)$ with $U(1) \stackrel{d}{=} X$.

See Sato [96, Section 4] for a comparison between these two processes, where it is mentioned that $T^{(H)} \stackrel{d}{=} U$, in the sense of equality in finite-dimensional distributions, if and only if X is a constant variable 0 or strictly stable with index $1/H$.

A third process associated with selfdecomposable X is known as the *background driving Lévy process* (*BDLP*) first discussed by Wolfe [107] and Jurek and Vervaat [61], and named by Jurek [60]. A Lévy process $Y := (Y(r), r > 0)$ is called the BDLP of X if

$$X \stackrel{d}{=} \int_0^\infty e^{-r} dY(r). \tag{3.110}$$

The relationship of $T^{(H)}$ and Y is given by

$$Y(r) = \int_{e^{-r}}^1 s^{-H} dT^{(H)}(s), \qquad r > 0, \tag{3.111}$$

and

$$T^{(H)}(s) = \int_{-\log(s)}^\infty e^{rH} dY(r), \qquad 0 < s < 1. \tag{3.112}$$

To illustrate the differences among these three processes, we observe if $k(x)/x$ is the Lévy density of X, then

- the Lévy density of $T^{(H)}(s) \stackrel{d}{=} T^{(1)}(s^H)$ is given by $k(s^{-H}x)/x$;
- the Lévy density of $U(s)$ is given by $sk(x)/x$;
- the Lévy density of $Y(r)$ is given by $rk(x)$.

The BDLP, denoted Y, can be easily extended for $s \leq 0$ by setting $Y(0) = 0$ and $(-Y(-s), s > 0)$ an independent copy of $(Y(s), s > 0)$, whence (3.111) holds for $r \in \mathbb{R}$ and (3.112) holds for $s > 0$. See Jeanblanc, Pitman and Yor [56, Theorem 1].

A two-parameter process

For simplicity, suppose the selfdecomposable random variable X is also non-negative with Lévy triple $(0, 0, k(x)x^{-1}dx)$ where $\theta := k(0+) < \infty$. So the associated 1-SSA process $(T(s), s > 0)$ with $T(1) \stackrel{d}{=} X$ is non-decreasing with no drift, finite rate θ and generic jump J whose distribution is determined by

$$\mathbb{P}(J > x) = k(x)/\theta. \tag{3.113}$$

Moreover, $T(s)$ has the following Lévy-Itô representation

$$T(s) = \int_0^s \int_0^\infty x\eta(dy\ dx) \qquad (s > 0), \tag{3.114}$$

where η is a Poisson point process on $\mathbb{R}_+^2$ satisfying (3.28), with intensity measure

$$\nu(dy\ dx) := \mathbb{E}\eta(dy\ dx) = \theta y^{-1}dy\mathbb{P}(sJ \in dx). \tag{3.115}$$

Based on η, define a Poisson point process on $\mathbb{R}_+^3$, say η^*, by setting its intensity measure $\nu^* := \mathbb{E}\eta^* = \nu \bigotimes \lambda$, where λ is the ordinary Lebesgue measure on $\mathbb{R}_+$. Now define a two-parameter process $(T(s,w), s > 0, w > 0)$ by setting

$$T(s,w) := \int_{z\in(0,w]} \int_{y\in(0,s]} \int_{x>0} x\eta^*(dz\ dy\ dx) \qquad (s, w > 0). \tag{3.116}$$

Then the following properties of this two-parameter process are easily checked:

- $T(s,w)$ is non-decreasing in both s and t;
- for each fixed positive integer w, $(T(s,w), s > 0)$ is a 1-SSA process obtained by adding w independent copies of T;
- for each fixed $w > 0$, $T_{(w)} := T(\cdot, w) = (T(s,w), s > 0)$ is a 1-SSA process with finite rate θw and generic jump J;
- for each fixed $s > 0$, $T^{(s)} := T(s, \cdot) = (T(s,\theta), \theta > 0)$ is a subordinator;
- the family $(T_{(w)}, w > 0)$ is coupled such that it is a “subordinator of 1-SSA processes”, i.e. for fixed $u > v > 0$, the increment $T_{(u)} - T_{(v)} \stackrel{d}{=} T_{(u-v)}$ is a 1-SSA non-decreasing process independent of $(T(s,w), s > 0, 0 < w < v)$;
- the family $(T^{(s)}, s > 0)$ is coupled such that it is an “1-SSA family of subordinators”, i.e. for fixed $s > 0$, $T^{(s)} \stackrel{d}{=} sT^{(1)}$ and for fixed $u > v > 0$, the increment $T^{(u)} - T^{(v)}$ is a subordinator independent of $(T(s,w), 0 < s < v, w > 0)$;
- for each finite interval I, when T is not identically 0, the jumps of $T^{(s)} := (T^{(s)}(w) = T(s,w), w > 0)$ as a subordinator are almost surely dense on I;

- however, for fixed $u > v > 0$, the number of jumps of $T^{(u)} - T^{(v)}$ is almost surely finite on I.

We only discussed above the case when T is non-decreasing and with finite rate . But it is easy to see that the bivariate process can be defined more generally. In particular, it is easy to add a deterministic or Brownian component. But the case when T has infinite rate and jumps of both signs would require more care.

3.9 An open problem on ratios of i.i.d. random variables

In Lemma 3.6.3, we have shown that if for each n the equality in distribution (3.82) of ratios of i.i.d. random variables holds then the distribution is uniquely determined up to a scaling factor. In fact, under some regularity condition, the identity (3.82) for $n = 3$

$$\left(\frac{X_1}{X_2}, \frac{X_2}{X_3}\right) \stackrel{d}{=} \left(\frac{Y_1}{Y_2}, \frac{Y_2}{Y_3}\right) \tag{3.117}$$

is sufficient, by writing the identity in joint distribution above in terms of characteristic functions $f(\lambda) := \phi_{\log X}(\lambda) = \mathbb{E}\exp[i\lambda\log(X)]$ and $g(\lambda) := \phi_{\log Y} = \exp[i\lambda\log(Y)]$

$$f(t_1)f(t_2)f(t_3) = g(t_1)g(t_2)g(t_3), \qquad \forall t_1 + t_2 + t_3 = 0, \tag{3.118}$$

which can be easily transformed into Cauchy's functional equation. This functional equation is first discovered and proved by Cauchy [11], and later been proved with various weaker conditions by many others, see e.g. Darboux [20].

On the other hand, we know it is insufficient to only have the identity (3.82) for $n = 2$, due to the following counterexample:

Example 3.9.1 ([85, Remark 7.6]) *Suppose $T := (T(s), s > 0)$ is a stable$(\frac{1}{2})$ subordinator. Let $X_1 = T(\frac{1}{2})$ and $X_2 = T(1) - T(\frac{1}{2}) \stackrel{d}{=} X_1$ since T is a Lévy process. Let $Y_1, Y_2 \stackrel{d}{=}$ gamma$(\frac{1}{2}, 1)$.*

Then

$$\frac{X_1}{X_1 + X_2} \stackrel{d}{=} \frac{Y_1}{Y_1 + Y_2} \sim \text{beta}(\tfrac{1}{2}, \tfrac{1}{2}) \tag{3.119}$$

hence

$$\frac{X_1}{X_2} \stackrel{d}{=} \frac{Y_1}{Y_2} \tag{3.120}$$

and X_1 is not equal in distribution to cY_1 for any constant c.

The second appearance of ratios of i.i.d. random variables in this Chapter is implicit in the proof to the 'only if' part of Theorem 3.7.2. To illustrate, consider the case $n = 2$ as the left equation of (3.104) in the proof, then take the limit $s_0 \to 0+$, i.e.

$$\frac{\delta_1}{\delta_1 + \delta_2} = \mathbb{P}\left(\frac{J_1}{\delta_1} < \frac{J_2}{\delta_2}\right) = \mathbb{P}\left(\frac{J_1}{J_2} < \frac{\delta_1}{\delta_2}\right), \qquad \forall \delta_1 \geq \delta_2 > 0. \tag{3.121}$$

In fact, since J_1 and J_2 are i.i.d., we may replace the restriction $\forall \delta_1 \geq \delta_2 > 0$ with $\forall \delta_1, \delta_2 > 0$. Note that for $\varepsilon_1, \varepsilon_2, \ldots$ i.i.d. copies of a standard exponential random variable ε,

$$\mathbb{P}\left(\frac{\varepsilon_1}{\varepsilon_2} < \frac{\delta_1}{\delta_2}\right) = \frac{\delta_1}{\delta_1 + \delta_2}, \qquad \forall \delta_1, \delta_2 > 0. \tag{3.122}$$

Hence, (3.121) is equivalent to

$$\frac{J_1}{J_2} \stackrel{d}{=} \frac{\varepsilon_1}{\varepsilon_2}. \tag{3.123}$$

And for general $n \geq 2$,

$$\left(\frac{J_1}{J_2}, \frac{J_2}{J_3}, \cdots, \frac{J_{n-1}}{J_n}\right) \stackrel{d}{=} \left(\frac{\varepsilon_1}{\varepsilon_2}, \frac{\varepsilon_2}{\varepsilon_3}, \cdots, \frac{\varepsilon_{n-1}}{\varepsilon_n}\right), \tag{3.124}$$

which is identical to (3.82) for $X \stackrel{d}{=} J$ and $Y \stackrel{d}{=} \varepsilon$, hence rounded back to a special case of Lemma 3.6.3 with one of the variables exponentially distributed.

The discussion above raises the question that, in Lemma 3.6.3, whether the condition when $n = 2$ alone might be sufficient when $Y \stackrel{d}{=} \varepsilon$. If so, it will provide a simple proof to the 'only if' part of Lemma 3.6.3.

To restate the question, observe the following well-known result

$$\frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} \sim \text{Unif}(0, 1). \tag{3.125}$$

Thus, it is equivalent to ask

Open Problem 3.9.2 *Suppose J_1, J_2 are i.i.d. copies of a positive random variable J and*

$$\frac{J_1}{J_1 + J_2} \sim \textit{Unif}(0, 1). \tag{3.126}$$

Is it always true that J is exponentially distributed?

Or more generally,

Open Problem 3.9.3 *Suppose J_1, J_2 are i.i.d. copies of a positive random variable J. With what extra condition does the distribution of $\frac{J_1}{J_1+J_2}$ characterize the distribution of J?*

These questions are related to Lukacs' characterization of the gamma distribution [74], where the independence between the sum $J_1 + J_2$ and the ratio $\frac{J_1}{J_1+J_2}$ is assumed. See also Chou and Huang [13].

3.10 Historical remarks

Selfdecomposable laws and self-similar processes The class of selfdecomposable laws was first studied by Lévy [71] and Khintchine [65], as an extension to stable laws as the limit distributions for sums of identically distribute random variables. Self-similar processes were first studied Lamperti [69] as a generalization to stable processes, which he called *semi-stable processes.*

Theorem 3.10.1 (Lamperti [69]) *Let* $(T(s), s > 0)$ *be an additive process. Suppose* $(T(s), s > 0)$ *is also* semi-stable*, that is, for every* $c > 0$*, there exists a constant* $b(c)$ *such that*

$$(T(cs), s \geq 0) \stackrel{d}{=} (b(c)T(s), s \geq 0), \tag{3.127}$$

then for some $H \geq 0$

$$b(c) = c^H \qquad (c > 0). \tag{3.128}$$

Note that if $H = 0$, the process is trivial; otherwise, it is an H-SSA process.

The integral representation of selfdecomposable laws was studied by Wolfe [107] for $\mathbb{R}$-valued random variables and then by Jurek and Vervaat [61] for more general Banach space-valued random variables. Although they did not consider self-similar processes, the integral representation

$$T(1) = \int_0^\infty e^{-s} Y(ds), \tag{3.129}$$

for $Y(\cdot)$ the background driving Lévy process (BDLP) associated with $T(1)$, is essentially the same as our integral representation (3.27) by treating e^{-s} as our index s. See also Jurek [58] for a recent study of selfdecomposable laws and the associated BDLP.

Sato [96, 95] investigated in detail and built the connection between the class L of selfdecomposable distributions and SSA processes as stated in Theorem 3.3.1. There is also a detailed background and a comprehensive list of references to earlier results on selfdecomposable laws in [96].

Jeanblanc, Pitman and Yor [56] pointed out that either of the two representations by Wolfe [107] and Sato [96] follows easily from the other. That reference also provides further background theory of selfdecomposable laws and their representations. Bertoin [5] treats the entrance law of self-similar processes. Tudor [104] studied the variations of self-similar processes from a stochastic calculus approach.

Summation representation of $T(1)$ In the setup of Theorem 3.4.1, the identity (3.25) can be viewed as a decomposition of the selfdecomposable random variable $T(s)$. For simplicity, consider only $T(1)$. Now index S_z such that

$$0 < \cdots < S_{-2} < S_{-1} < S_0 \leq 1 < S_1 < S_2 < \cdots < \infty. \tag{3.130}$$

By Lemma 3.2.1

$$S_{1-n} = \prod_{k=1}^{n} \beta_k \qquad (n \geq 1), \tag{3.131}$$

where β_k are i.i.d. beta$(\theta, 1)$ random variables.

Corollary 3.10.2 *Suppose $T(1)$ is a non-negative, selfdecomposable random variable. Then there exists a sequence of i.i.d. non-negative random variables J_z such that*

$$T(1) \stackrel{d}{=} \sum_{n=1}^{\infty} \left(\prod_{k=1}^{n} \beta_k \right) J_{1-n}, \tag{3.132}$$

where β_k are i.i.d. beta$(\theta, 1)$ random variables independent of J_z.

The distribution of a random variable $T(1)$ admitting the representation (3.132) for i.i.d. sequences β_k (not necessarily beta) and J_k was studied first by Vervaat [106, example 3.8] as the solution of the stochastic difference equation

$$X \stackrel{d}{=} A(X + C), \tag{3.133}$$

with X, A, C independent and $|A| < 1$. By iteration,

$$X \stackrel{d}{=} \sum_{n=1}^{\infty} \left(\prod_{k=1}^{n} A_k \right) C_n, \tag{3.134}$$

where A_k, C_n are i.i.d. copies of A and C, independent of each other. This is identical to (3.132) by setting $A_n = \beta_n$ and $C_n = J_n$.

The following corollary of the fact that a SSA gamma process is associated with an exponential generic jump, giving a representation of a gamma distributed random variable, is also given in [106, example 3.8.2].

Corollary 3.10.3 (Vervaat [106]) *If $C \sim exp(\lambda)$ and $A \sim beta(\theta, 1)$ for some $\lambda, \theta > 0$, then the unique solution to the stochastic difference equation (3.133) is $X \sim gamma(\theta, \lambda)$.*

However, Vervaat did not make any connection with SSA processes in his work on stochastic difference equations, as he did not treat $(\prod_{k=1}^{n} A_k)$ as a time index, and he did not discuss selfdecomposability of X, only infinitely divisibility.

Extremal processes As mentioned in Section 3.1, the 1-SSA exponential process with rate $\theta = 1$ arises in the theory of extremal process introduced by Dwass [23].

Starting from an i.i.d. sequence $(X_n, n = 1, 2, \ldots)$ of continuous random variables, it is elementary that the *record sequence* $(M_n := \max_{1 \leq k \leq n} X_k, n = 1, 2, \ldots)$ is a Markov chain with state space $\mathbb{R}$ and transition probabilities specified by

$$\mathbb{P}(M_{n+m} \leq y \mid M_n = x) = F^m(y) 1(x \leq y), \tag{3.135}$$

where $F^t(y)$ is the t-th power of the common cumulative distribution function (c.d.f.) of the X_i. It was observed in 1964 by Dwass [23] and Lamperti [68] that the record process can be generalized to a time-homogeneous pure jump-type Markov process, known as the *extremal process* $M := (M(t), t > 0)$, described by its entrance law

$$\mathbb{P}(M(t) \leq y) = F^t(y), \tag{3.136}$$

and the hold-jump description: conditional on $M(s) = x$,

- 'Hold' - at level x for an exponential time H_x with rate $Q(x) := -\log F(x)$ so that

$$\mathbb{P}(H_x > t) = \mathbb{P}(M(t) \leq x) = F^t(x) = e^{-tQ(x)}; \tag{3.137}$$

- 'Jump' - at time H_x to a state $L_x := M(s + H_x)$ with distribution

$$\mathbb{P}(L_x > b) = \frac{Q(b)}{Q(x)} \qquad (b \geq x). \tag{3.138}$$

See Shorrock [98] or Kallenberg [63, Chapter 13] for more details.

Proposition 3.10.4 (Dwass [23] Resnick and Rubinovitch [91] Shorrock [98]) *Let $((T_z, M_z), z \in \mathbb{Z})$ be a listing indexed by integers $\mathbb{Z}$ of the times T_z of jumps of $(M(t), t \geq 0)$ and the corresponding record levels $M_z := M(T_z)$, with $T_z < T_{z+1}$. Then*

$$range((M_z, T_{z+1} - T_z), z \in \mathbb{Z}) \text{ is } PPP(Q(dm)e^{-Q(m)t}dt). \tag{3.139}$$

In particular,

(I) the random set of record times T_z is a scale invariant PPP on $\mathbb{R}_+$ with rate 1;

(II) the random set of record levels $range(M_z, z \in \mathbb{Z})$ is $PPP(Q(dm)/Q(m))$;

(III) the random set of holding times at these record levels $range(T_{z+1} - T_z, z \in \mathbb{Z})$ is a scale invariant PPP on $\mathbb{R}_+$ with rate 1.

As indicated by Dwass and Lamperti, the extremal Markov process associated with each continuous distribution F on the line is essentially the same as that associated with every other continuous distribution F', via the monotonic transformation

$$M' = \psi(M), \quad \text{for } \psi \text{ with } F'(\cdot) = F(\psi \in \cdot). \tag{3.140}$$

To illustrate the point and see its relation with our Theorems 3.1.3 and 3.1.2, consider the case of Gnedenko's extreme value distribution, that is the distribution of ε^{-1} for ε exponential with mean 1:

$$F(y) = \exp(-y^{-1}) = \mathbb{P}(\varepsilon^{-1} \leq y) \qquad (y > 0), \tag{3.141}$$

hence the rate in (3.137) is

$$Q(x) = -\log F(x) = \frac{1}{x}, \tag{3.142}$$

and the jump has distribution

$$\mathbb{P}(J_x > b) = \frac{Q(b)}{Q(x)} = \frac{x}{b} \qquad (b \geq x), \tag{3.143}$$

namely $x/J_x \sim \text{beta}(1,1)$. It is easy to check that the functional inverse (which means the 'hold' and 'jump' are swapped) coincides with the hold-jump description (3.56) (3.57) of a 1-SSA gamma process $(T(s), s > 0)$ with rate 1 and $T(s) \stackrel{d}{=} s\varepsilon \sim \exp(s^{-1})$.

Therefore, the path of an extremal process associated with (3.141) has the same distribution as the functional inverse of the path of a standard 1-SSA exponential process. This matches Proposition 3.10.4(I) with Theorem 3.1.3, and Proposition 3.10.4(III) with Theorem 3.1.2(III). Lastly, in this case the $PPP(Q(dm)/Q(m))$ in Proposition 3.10.4(II) is a scale invariant PPP on $\mathbb{R}_+$ with rate 1, as in Theorem 3.1.2(II).

It may also be of interest to replace the i.i.d. sequence above with an inhomogeneous Markov sequence $(X_n, n = 1, 2, \dots)$ defined as follows:

- The first term X_1 has c.d.f. F^θ;
- Conditionally on $M_n = x$, let F_{x+} be the c.d.f. of X given $X > x$

  $$F_{x+}(y) := (F(y) - F(x))/(1 - F(x)) \qquad (y > x), \tag{3.144}$$

  and F_{x-}be the one given $X \leq x$

  $$F_{x-}(y) := F(y)/F(x) \qquad (y \leq x). \tag{3.145}$$

  Then the distribution of X_{n+1} is the mixture with weights $1 - F(x)$ and $F(x)$ of the distribution of a random variable with c.d.f. F^θ_{x+}, and distribution of a random variable with c.d.f. F_{x-}, i.e.

  $$\mathbb{P}(X_{n+1} \leq y \mid M_n = x) = (1 - F(x))F^\theta_{x+}(y) + F(x)F_{x-}(y). \tag{3.146}$$

An example of this sequence is obtained by putting $X \sim \text{Uniform}(0,1)$. Then $X_1 \sim \text{beta}(1,\theta)$ and conditionally on $M_n = x$, X_{n+1} is the mixture with weights $1 - x$ and x of $\text{Uniform}(0, 1 - x)$ and $x + (1 - x)\beta$ for $\beta \sim \text{beta}(1,\theta)$.

Theorem 3.10.5 *Suppose X is a real-valued random variable without atoms, and the sequence $(X_n, n = 1, 2, \dots)$ follows the inductive construction above. Then the maximum indicators $(B_n, n = 1, 2, \dots)$ form a sequence of independent Bernoulli random variables with $\mathbb{E}B_n = \theta/(\theta + n - 1)$.*

This is a generalization to Najnudel and Pitman [79, Corollary 1.4] where they proved the same result for $X \sim \text{Uniform}(0, 1)$. This generalization works because the distribution of X has no atoms, hence has no influence on record times.

It is not hard to construct a time-inhomogeneous extremal process M associated with $(X_n, n = 1, 2, \ldots)$ and check that the path of such extremal process has the same distribution as the functional inverse of the path of a 1-SSA gamma process $(T(s), s > 0)$ with $T(1) \sim \text{gamma}(\theta, 1)$. In this case, Theorem 3.10.5 can be read as a discrete-time analogue of Theorem 3.1.3, and an analogue of Proposition 3.10.4 can also be easily given by replacing the rate in (I) and (III) by θ, and replacing the PPP in (II) by $PPP(\theta\ Q(dm)/Q(m))$.

Scale invariant point processes Scale invariant point processes, and scale invariant random sets, including the scale invariant PPP, were studied in [87, 38] by considering the random partitions of $(0, 1)$ related to the Poisson-Dirichlet distribution. It was observed in [87] that a random closed set is scale invariant if and only if the associated age process is 1-self-similar, while [38] remarked on the relationship between scale invariant PPP and records.

The scale invariant Poisson spacings lemma for general $\theta \neq 1$ The formulation of the Theorem 3.1.3 for general $\theta > 0$ was suggested by proofs of the scale invariant Poisson spacings Theorem 3.1.1, first indicated by Arratia [3, 2], and detailed by Arratia, Barbour and Tavaré [4], where they make use of a Poisson point process on $(0, \infty)^2$ which turns out to be our η in (3.27). Based on [4], Gnedin [37, Section 4] pointed out that the scale invariant Poisson spacings theorem for positive integer θ follows from a specialization of Ignatov's theorem in the form of [39, Corollary 5.1].

Feller's coupling and the Ewens sampling formula The parameter θ in the Poisson spacings theorem is related to the Ewens(θ) distribution [28] as a generalization to the uniform random permutation of $[n]$. Feller [30] provided a coupling between the counts of cycles of various sizes in a uniform random permutation of $[n]$ and the spacings between successes in a sequence of n independent Bernoulli(k^{-1}) trials at the kth trial. Informally, this sequence of independent Bernoulli trials is a discrete analogue of the scale invariant PPP with rate 1 relative to $x^{-1}dx$.

Ignatov [54] proved that in an infinite sequence of independent Bernoulli(n^{-1}) trials, as the indicators of record values in an i.i.d. sequence, the numbers of spacings of length k between successes/records are independent Poisson variables with means k^{-1}. This Poisson sequence provides another discrete analogue of the scale invariant PPP with rate 1. It is interesting that this discrete result was obtained many years after theory of extremal processes by Dwass [23]. Ignatov's result was generalized by Arratia, Barbour and Tavaré [4] in the study of cycles of (non-uniform) random permutations governed by the Ewens(θ) distribution, i.e. a permutation is weighted θ^k if there are k cycles. See Najnudel and Pitman [79] for details of the coupling between the random permutations governed by the Ewens(θ)

distribution and the sequence of inhomogeneous Bernoulli($\theta(\theta+n-1)^{-1}$) trials, as mentioned in Theorem 3.10.5.

Chapter 4

Bivariate generating functions

In this chapter, we give a formula for the bivariate generating function of a stationary 1-dependent counting process in terms of its run probability generating function, with a probabilistic proof. The formula reduces to the well known bivariate generating function of the Eulerian distribution in the case of descents of a sequence of independent and identically distributed random variables. The formula is then compared with alternative expressions from the theory of determinantal point processes and the combinatorial enumeration formula of sequences.

4.1 Definitions

Counting one-dependent processes

A discrete time stochastic process $(X_1, X_2, \cdots)$ is said to be 1-*dependent* if

$$(X_1, \ldots, X_{m-1}) \text{ is independent of } (X_{m+1}, \ldots, X_{m+n}) \tag{4.1}$$

for all positive integers m and n. In contrast to a Markov chain, this independence requires no knowledge of current position. This dependence structure has been widely investigated in probability theory [6, 48, 47], and as a tool in statistics [64] and queuing systems [46, 99]. Examples of 1-dependent processes are provided by 2-*block factors* [53] generated by a function of two successive terms in an independent sequence. But not all 1-dependent processes can be constructed this way: Aaronson et al. [1] explicitly gave a two-parameter family of 1-dependent processes which cannot be expresses as 2-block factors. Other examples of this kind arise in the theory of random colorings of integers developed by Holroyd and Liggett [48, 47].

Here, we restrict attention to 1-dependent processes $(X_1, X_2, \cdots)$ which are also *stationary*, i.e. for all positive integers n,

$$(X_1, X_2, \ldots, X_n) \stackrel{d}{=} (X_2, X_3, \ldots, X_{n+1}). \tag{4.2}$$

In the case of an independent and identically distributed (*i.i.d.*) sequence, the distribution of the count $S_n(A) = \sum_{k=1}^n 1(X_k \in A)$ is binomial with parameters n and $\mathbb{P}(X_n \in A)$, for any measurable subset A of the state space of the sequence. We describe here a bivariate generating function which determines the distribution of this counting variable $S_n(A)$ for any stationary 1-dependent process $(X_1, X_2, \cdots)$.

Bivariate generating functions

Following the work of de Moivre on the distribution of the number of spots in a number of die rolls, the encoding of a sequence by its *generating function* was exploited by Euler [27] and many subsequent authors for combinatorial enumeration [41, 33]. To describe the distribution of an integer-valued random variable, Laplace [70] introduced the *probability generating function* [21]. For a sequence of non-negative-integer-valued random variables S_n, let Q_{S_n} denote the probability generating function of S_n:

$$Q_{S_n}(z) := \mathbb{E}z^{S_n} = \sum_{k=0}^{\infty} \mathbb{P}(S_n = k) z^k, \tag{4.3}$$

and let $Q(z, v)$ be the *bivariate generating function* of distributions of S_n

$$Q(z, v) := \sum_{n \geq 0} \sum_{k \geq 0} \mathbb{P}(S_n = k) z^k v^n = \sum_{n \geq 0} Q_{S_n}(z) v^n, \tag{4.4}$$

for all z, v such that the summation converges, including $|z| \leq 1$ and $|v| < 1$. This bivariate generating function determines the distribution of S_n for every n by extraction of the coefficient of $z^k v^n$ from $Q(z, v)$:

$$\mathbb{P}(S_n = k) = [z^k v^n] Q(z, v), \quad n, k = 0, 1, 2, \cdots \tag{4.5}$$

In our set up for counting processes, $S_n := X_1 + \cdots + X_n$, where $(X_n, n \geq 1)$ is an indicator sequence, so each count S_n takes values in $\{0, 1, 2, \cdots, n\}$, and the series (4.4) is absolutely convergent for $|zv| < 1$. See e.g. [33, Chapter III] for further background on bivariate generating functions.

Run probability generating functions

For an indicator sequence $(X_n, n \geq 1)$, define its 0-*run probabilities*

$$q_0 := 1 \text{ and } q_n := \mathbb{P}(S_n = 0) = \mathbb{P}(X_1 = X_2 = \cdots = 0), \quad n = 1, 2, \ldots \tag{4.6}$$

and the associated 0-*run probability generating function*

$$Q(v) := \sum_{n=0}^{\infty} q_n v^n = \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0) v^n = Q(0, v). \tag{4.7}$$

The 1-*run probability* sequence can be similarly defined, treated as the 0-run probability sequence for the *dual* indicator sequence ($\hat{X}_n := 1 - X_n, n \geq 1$), with counts $\hat{S}_n = n - S_n$:

$$p_0 := 1 \text{ and } p_n := \mathbb{P}(S_n = n) = \mathbb{P}(X_1 = X_2 = \cdots = 1), \quad n = 1, 2, \ldots \tag{4.8}$$

The associated 1-*run probability generating function* is then for $0 \leq v < 1$

$$P(v) := \sum_{n=0}^{\infty} p_n v^n = \sum_{n=0}^{\infty} \mathbb{P}(S_n = n) v^n = \hat{Q}(0, v) = \lim_{z \to 0} Q(z^{-1}, zv), \tag{4.9}$$

where $\hat{Q}(z, v)$ is the bivariate generating function of $\hat{S}_n$, and the last equality is by dominated convergence as $z \to 0$, using the evaluation for $|zv| < 1$:

$$\hat{Q}(z, v) = \sum_{n \geq 0} \sum_{k=0}^{n} \mathbb{P}(S_n = n - k) z^k v^n \tag{4.10}$$

$$= \sum_{n \geq 0} \sum_{k=0}^{n} \mathbb{P}(S_n = k) z^{-k} (zv)^n = Q(z^{-1}, zv). \tag{4.11}$$

In our case of a stationary 1-dependent sequence of indicator variables ($X_n, n \geq 1$), it is known [90, Chapter 7.4][1, Theorem 1] that the distribution of $S_n = X_1 + \cdots + X_n$, is uniquely determined by its sequence of 1-run probabilities, or just as well by its sequence of 0-run probabilities, through a determinantal formula for the probability function of the random vector $(X_1, \ldots, X_n)$. Our main result, presented in Section 4.2, gives a formula for the bivariate generating function $Q(z, v)$ of distributions of S_n in this case, which is simpler than might be expected from this determinantal formula. The rest of this chapter is organized as follows.

- Section 4.3 shows how the Eulerian bivariate generating function is obtained from our result in the case of descents.
- Section 4.4 displays the bivariate generating function of some other stationary 1-dependent processes.
- Section 4.5 verifies our result from the perspective of determinantal point processes.
- Section 4.6 makes connection with a combinatorial result in Goulden and Jackson [41] and provides Corollary 4.6.1 which is suitable for counting a particular pattern in 2-block factors.
- Section 4.7 compares our formula for the bivariate generating function in the stationary 1-dependent case to similar formulae for exchangeable or renewal processes, which are either known or easily derived from known results.

4.2 Main result

Theorem 4.2.1 *For a stationary* 1*-dependent indicator sequence* $(X_n, n \geq 1)$*, the bivariate generating function* $Q(z, v)$ *of distributions of its partial sums* S_n *is determined either by the* 0*-run probability generating function* $Q(v)$*, or by the* 1*-run probability generating function* $P(v)$*, via the formulae*

$$Q(z,v) = \frac{Q((1-z)v)}{1 - zvQ((1-z)v)} = \frac{P(-(1-z)v)}{1 - vP(-(1-z)v)}. \tag{4.12}$$

The particular case $z = 0$ of (4.12) reduces to the following known result:

Corollary 4.2.2 (Involution [6, Proposition 7.4]) *In the setting of the previous theorem, for any stationary* 1*-dependent indicator sequence, the* 0*-run generating function* $Q(v)$ *and the* 1*-run generating function* $P(v)$ *determine each other via the involution of formal power series*

$$Q(v) = \frac{P(-v)}{1 - vP(-v)}; \qquad P(v) = \frac{Q(-v)}{1 - vQ(-v)}. \tag{4.13}$$

Proof (of Theorem 4.2.1 and Corollary 4.2.2)

We will first prove the left equality in (4.12), rearranged as

$$Q(z,v) = Q((1-z)v) + zv\,Q((1-z)v)Q(z,v), \tag{4.14}$$

by establishing the corresponding identity of coefficients of powers of v, that is,

$$Q_{S_n}(z) = [v^n]Q((1-z)v) + z\sum_{k=0}^{n-1}[v^k]Q((1-z)v)[v^{n-1-k}]Q(z,v). \tag{4.15}$$

Recall that $q_j := [v^j]Q(v)$, whence

$$[v^j]Q((1-z)v) = (1-z)^j[v^j]Q(v) = (1-z)^j q_j, \quad j = 0, 1, \ldots. \tag{4.16}$$

So (4.15) for each $n = 1, 2, \ldots$, with $j = k - 1$, reduces to

$$Q_{S_n}(z) = (1-z)^n q_n + \sum_{k=1}^{n}\left((1-z)^{k-1}z\right) q_{k-1} Q_{S_{n-k}}(z), \tag{4.17}$$

which has the following interpretation. For $0 \leq z \leq 1$, let $(Y_n, n \geq 1)$ be a sequence of i.i.d. Bernoulli(z) random variables, also independent of $(X_n, n \geq 1)$. Employing van Dantzig's method of marks [19], treat Y_n as a *mark* on X_n: say the n-th item X_n is *z-marked* if $Y_n = 1$, and *non-z-marked* if $Y_n = 0$. By construction, $Q_{S_n}(z)$ is the probability that every success among the first n trials is z-marked. In particular, if $z = 0$, every success in non-z-marked. Then the only way every success in the first n trials is z-marked is if there are no successes. Hence $Q_{S_n}(0) = q_n$ is the probability of no successes in the first n trials. The identity (4.17), decomposes the event that every success in the first n trials is z-marked according to the value of $T_z := \min\{n : Y_n = 1\}$, the index of the first z-mark. So

- T_z has the geometric(z) distribution $\mathbb{P}(T_z = k) = (1-z)^{k-1}z$ for $k = 1, 2, \ldots$;
- On the event of probability $(1-z)^n$, that the first z-mark occurs at $T_z > n$, no trial among the first n is allowed to be success, with probability q_n;
- On the event of probability $(1-z)^{k-1}z$, that the first z-mark occurs at $T_z = k$ for some $1 \le k \le n$, no trial among the first $k-1$ is allowed to be success, with probability q_{k-1}, and all success after the k-th (excluding the k-th) are z-marked, with probability $Q_{S_{n-k}}(z)$, with independence before and after the k-th trial by the assumption that $(X_n, n \ge 1)$ is 1-dependent.

This proves the left equality of (4.12). To prove the right, it is easiest to prove Corollary 4.2.2. Recall (4.9),

$$P(v) = \lim_{z\to 0} Q(z^{-1}, zv) = \lim_{z\to 0} \frac{Q((z-1)v)}{1 - vQ((z-1)v)} = \frac{Q(-v)}{1 - vQ(-v)}, \tag{4.18}$$

which yields the P identity in (4.13). To see the Q identity, simply replace v with $-v$ in the last equation.

Lastly, the right equation of (4.12) is obtained from the left one and the involution

$$Q(z,v) = \frac{P(-(1-z)v)/(1-(1-z)vP(-(1-z)v))}{1 - zvP(-(1-z)v)/(1-(1-z)vP(-(1-z)v))} \tag{4.19}$$

$$= \frac{P(-(1-z)v)}{1 - vP(-(1-z)v)}. \tag{4.20}$$

■

4.3 Application to descents

In this section, we present the example of *Eulerian numbers*. We were led to the general formula for the bivariate generating function of counts of a 1-dependent indicator sequence by the algebraically simple form of the bivariate generating function of Eulerian numbers, whose probabilistic meaning is not immediately obvious, but nicely explained by the above proof of Theorem 4.2.1.

It is well known that a large class of stationary 1-dependent indicator sequences (though not all, see [1, 7]) may be constructed from an independent and identically distributed *background sequence* $(Y_1, Y_2, \ldots)$, as *two-block factors*

$$X_n := 1((Y_n, Y_{n+1}) \in B), \tag{4.21}$$

for some product-measurable subset B of the space of pairs of Y-values, say $[0,1]^2$ for $Y_i \sim \text{Uniform}(0,1)$.

An important example is provided by the sequence of *descents* $X_n := 1(Y_n > Y_{n+1})$ for real-valued Y_i. In particular, for $Y_i \sim \text{Uniform}(0,1)$ (or any continuous distribution) and $S_n := D_{n+1}$ counting the number of descents $Y_i > Y_{i+1}$ with $1 \le i \le n$:

$$\mathbb{P}(S_n = 0) = \mathbb{P}(S_n = n) = \mathbb{P}(Y_1 > \cdots > Y_{n+1}) = \frac{1}{(n+1)!}. \tag{4.22}$$

So the run generating functions $Q(v)$ and $P(v)$ in this case are easily evaluated as

$$Q(v) = P(v) = \sum_{n\ge 0} \frac{v^n}{(n+1)!} = \frac{e^v - 1}{v}. \tag{4.23}$$

Eulerian numbers

The *Eulerian numbers* $\left\langle {n \atop k} \right\rangle$ are commonly defined by the numbers of permutations of $[n] := \{1, 2, \cdots, n\}$ with exactly k descents, i.e. k adjacent pairs with first larger than the second [42]. So the count $\hat{S}_{n-1}$ of descents in a uniform random permutation of $[n]$ has the *Eulerian distribution*

$$\mathbb{P}(\hat{S}_{n-1} = k) = \frac{1}{(n)!}\left\langle {n \atop k} \right\rangle. \tag{4.24}$$

Observe that this uniform permutation can be done by taking the *ranks* of the i.i.d. background sequence $(Y_1, Y_2, \cdots, Y_n)$. Here, we say the rank of Y_i is k if and only if Y_i is the k-th smallest among $Y_1, Y_2, \cdots Y_n$. Then, for $Y_i \sim \text{Uniform}(0,1)$, the ranks are almost surely a uniform permutation of $[n]$. Therefore, $\hat{S}_n$ has the same distribution as S_n in (4.22). Now, applying Theorem 4.2.1 to the descents $X_n := 1(Y_n > Y_{n+1})$ implies the following bivariate generating function

$$Q(z,v) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} \mathbb{P}(S_n = k) z^k v^n = \frac{e^{(1-z)v} - 1}{v(1 - ze^{(1-z)v})} = \frac{e^v - e^{zv}}{v(e^{zv} - ze^v)}, \tag{4.25}$$

which is the classical bivariate generating function of the Eulerian numbers [8, 16, 42, 76, 52].

4.4 More examples

To simplify displays, we work in this section with the *shifted run generating functions*

$$\tilde{Q}(v) = 1 + vQ(v), \qquad \tilde{P}(v) = 1 + vP(v), \tag{4.26}$$

and the *shifted bivariate generating function*

$$\tilde{Q}(z,v) = 1 + vQ(z,v). \tag{4.27}$$

For reasons which do not seem obvious, the algebraic form of the generating functions associated with a 1-dependent indicator sequence is typically simpler when they are shifted like this. The shifted generating functions of some selected models are shown in the table below, with detailed explanation later.

There are some benefits for using the shifted generating functions. Firstly, they are simpler, especially in the Eulerian case; secondly, for 2-block factors, the shifted generating functions are actually 'standard' in combinatorics, since n is set to be the length of background sequence; thirdly, the formulae in Theorem 4.2.1 and Corollary 4.2.2 are also slightly simplified: the involution formula (4.13) becomes (see, e.g. [75, 6] for earlier occurrences, and [35, 9] where this formula was first discovered in the study of 2-block factors)

$$\tilde{Q}(v) = \frac{1}{\tilde{P}(-v)}; \qquad \tilde{P}(v) = \frac{1}{\tilde{Q}(-v)}, \tag{4.28}$$

while our main theorem (4.12) is re-written as

$$\tilde{Q}(z,v) = \frac{(1-z)\tilde{Q}((1-z)v)}{1-z\tilde{Q}((1-z)v)} = \frac{1-z}{\tilde{P}(-(1-z)v)-z}. \tag{4.29}$$

Table 1

Model	Shifted 0-run pgf $\tilde{Q}(v)$	Shifted 1-run pgf $\tilde{P}(v)$	Shifted bgf $\tilde{Q}(z,v)$
Eulerian	e^v	e^v	$\frac{1-z}{e^{-(1-z)v}-z}$
I.i.d.	$\frac{1+pv}{1-(1-p)v}$	$\frac{1+v-pv}{1-pv}$	$\frac{1+pv-pzv}{1-v+pv-pzv}$
One-pair	$\frac{1+pv}{1-v+pv-pv^2+p^2v^2}$	$\frac{1+v-pv-pv^2+p^2v^2}{1-pv}$	$\frac{1+pv-pvz}{1-v+pv-pv^2+p^2v^2-pvz+pv^2z-p^2v^2z}$
Carries	$\left(1-\frac{v}{b}\right)^{-b}$	$\left(1+\frac{v}{b}\right)^{b}$	$\frac{1-z}{\left(1-\frac{(1-z)v}{b}\right)^b-z}$
Flipping	$\sqrt{\frac{q}{p}}\tan\left[v\sqrt{pq} - \arctan\left(\frac{q}{p}\right)\right]$	$\sqrt{\frac{p}{q}}\Big/\tan\left[-v\sqrt{pq} - \arctan\left(\frac{q}{p}\right)\right]$	$\frac{(1-z)\tan\left[(1-z)v\sqrt{pq}-\arctan(q/p)\right]}{\sqrt{p/q}-z\tan\left[(1-z)v\sqrt{pq}-\arctan(q/p)\right]}$
Non-2BF	$\frac{1}{1-v+\alpha v^2-\beta v^3}$	$1+v+\alpha v^2+\beta v^3$	$\frac{1}{1-v+\alpha(1-z)v^2-\beta(1-z)^2v^3}$

We already discussed the *Eulerian* model in Section 4.3. The rest of Table 1 will be briefed here row by row. Usually, we only say how one of the run probability generating functions is obtained, since the other one and the bivariate generating function can then be found easily through (4.26), (4.28) and (4.29).

Independent and identically distributed trials (I.i.d.) The classical example of i.i.d. Bernoulli(p) trials is for sure an example of 1-dependent sequence.

Indicator of two consecutive ones (One-pair) Consider the simplest 2-block factors $X_n := 1(Y_n = Y_{n+1} = 1)$, where $(Y_n, n \geq 1)$ is i.i.d. Bernoulli(p) trials. Its 1-run probability generating function is easy to compute.

Considering the coefficient of $z^k v^n$ in its bivariate generating function gives the recursion

$$q_{n,k} = (1-p)q_{n-1,k} + pq_{n-1,k-1} + p(1-p)(q_{n-2,k} - q_{n-2,k-1}), \qquad n \geq 2, k \geq 0, \tag{4.30}$$

where $q_{n,k} := [z^k v^n]Q(z,v) = [z^k v^{n+1}]\tilde{Q}(z,v) = \mathbb{P}(S_n = k)$ with initial values $q_{0,0} = 1, q_{1,0} = 1 - p^2, q_{1,1} = p^2$ and convention $q_{n,k} = 0$ for $k > n$ or $k < 0$.

Setting $p = 1/2$ recovers the *Fibonacci* sequence as its 0-run probabilities:

$$F_{n+2} =: 2^n q_{n,0} = 2^{n-1} q_{n-1,0} + 2^{n-2} q_{n-2,0} = F_{n+1} + F_n, \tag{4.31}$$

where $F_0 = F_1 = 1$ is the first two terms of the Fibonacci sequence we use here. This can also be interpreted as the chance of not getting any consecutive heads in a row of coin tosses, which has been recognized by many others, see [78, 26, 32, 49].

Carries when adding a list of digits (Carries) Adding a list of digits using carries is discussed in [6] as a stationary 1-dependent process. This example also falls into the category of 2-block factors with i.i.d. Uniform$(\{0, 1, \cdots, b-1\})$'s as its background sequence, and $B = \{(x, y) : b > x > y \geq 0\}$. Its 0-run probabilities and generating function are explicitly given in [6].

Edge flipping on the integers (Flipping) Chung and Graham [14] introduced the following discrete time model of a random pattern in $\{0,1\}^V$ indexed by the vertex set V of a finite simple graph (V, E): pick an edge uniformly at random from E, then the pattern is updated by replacing its values on the two vertices joined with the picked edge by 11 with probability p and by 00 otherwise, where the choices of edges and update of values on vertices are assumed to be all independent of the others, for some $0 < p < 1$.

They offered an analysis of discrete-time edge flipping on an n-cycle, which can be generalized in terms of a stationary continuous-time (1-dependent indicator) process indexed by the integers.

In short, we may sample its stationary distribution in the following manner:

- first, generate a sequence $(U_n, n \in Z)$ of i.i.d. Uniform $(0,1)$'s, which works as the time of last update on the edge $\{n, n+1\}$. That is to say, if $U_n > U_{n-1}$, then the last update on edge $\{n, n+1\}$ happened later than the one on $\{n-1, n\}$;
- secondly, generate a sequence $(W_n, n \in \mathbb{Z})$ of i.i.d. Bernoulli (p)'s, independent of $(U_n, n \in Z)$, which stands for whether the last update on the edge $\{n, n+1\}$ is 11 or 00.
- lastly,
$$X_n := 1(U_n > U_{n-1})W_n + 1(U_n < U_{n-1})W_{n-1}, \quad n \in \mathbb{Z}. \tag{4.32}$$

This is apparently a stationary 2-block factor. Its shifted 0-run probability generating function is given in [14, Theorem 6].

A non-2-block-factor example (Non-2BF) Aaronson, Gilat, Keane and de Valk [1] first discovered this family of non-2-block-factor stationary 1-dependent indicator processes. In short, they forbid the appearance of three consecutive ones, hence the 1-run probability generating functions are as simple as quadratic functions. Note that not all value pairs (α, β) make this process not a 2-block factor – only some work, while some others do not yield stationary 1-dependent processes at all. See [1, Fig. 2].

4.5 Determinantal representation

Any indicator process can be treated as a point process by regarding the indicated events as points. It was shown in [6, Theorem 7.1] that any 1-dependent point process on a segment of $\mathbb{Z}$ is a *determinantal* process, as discussed further in [17, 18, 6].

Given a finite set $\mathcal{X}$, a point process on $\mathcal{X}$ is a probability measure $\mathbb{P}$ on $2^{\mathcal{X}}$. Its *correlation function* is the function of subsets $A \subseteq \mathcal{X}$ defined by $\rho(A) := \mathbb{P}(S : S \supseteq A)$. A point process is said to be *determinantal* with *kernel* $K(x, y)$ if

$$\rho(A) = \det(K(x,y))_{x,y\in A}, \tag{4.33}$$

where the right hand side is the determinant of the $|A| \times |A|$ matrix with $K(x, y)$ on its (x, y)-th entry for $x, y \in A$. Now, we may state the following theorem from [6].

Theorem 4.5.1 (Theorem 7.1 [6]) *Every* 1*-dependent point process on a segment of* $\mathbb{Z}$ *is determinantal with kernel*

$$K(x,y) = \sum_{r=1}^{y-x+1} (-1)^{r-1} \sum_{x=l_0<l_1<\cdots<l_r=y+1} \prod_{k=1}^{r} \rho([l_{k-1}, l_k) \cap \mathbb{Z}), \quad x \le y, \tag{4.34}$$

$K(x, y) = -1$ *for* $x = y + 1$*, and* $K(x, y) = 0$ *for* $x \ge y + 2$.

In the stationary case,

$$\rho([a,b) \cap \mathbb{Z}) = \mathbb{P}(S_{b-a} = b-a) = p_{b-a}, \tag{4.35}$$

where S_n and p_n inherit the setup in Section 4.1. It is easy to see that then the kernel is also stationary, i.e.

$$K(x,y) := k(y-x) = K(x+c, y+c), \qquad \forall c \in \mathbb{Z}. \tag{4.36}$$

To better describe the kernel, consider the *kernel generating function*

$$G_k(v) := \sum_{n\in\mathbb{Z}} k(n) v^n = -v^{-1} + p_1 + (p_2 - p_1^2)v + \cdots \tag{4.37}$$

Borodin, Diaconis and Fulman[6, Corollary 7.3] give the following relationship between the kernel generating function G_k and the 1-run probability generating function P

$$G_k(v)P(v) = -\frac{1}{v}. \tag{4.38}$$

One last interesting (but not hard) result [6, Theorem 4.1] is that we may write the probability of any string pattern as a determinant. Consider $\mathcal{X} = [n] := \{1, 2, \cdots, n\}$, then the string with exactly k zeros at $0 < w_1 < w_2 < \cdots < w_k \leq n$, which corresponds to the subset $A := \{w_1, w_2, \cdots, w_k\}^c$, has probability

$$\mathbb{P}(A) = \det(p_{w_{j+1}-w_i-1})_{0\leq i,j\leq k}, \tag{4.39}$$

where $p_n = 0, \forall n < -1, p_0 = p_{-1} = 1$ and $w_0 = 0, w_{k+1} = n + 1$.

As shown in [6], this formula (4.39) is obtained by application of inclusion-exclusion formula on the correlation function. Hence, we may recover our bivariate generating function (4.12) as follows:

Corollary 4.5.2 *The multivariate probability generating function of the indicators I_i of location $i, i \in \mathcal{X} = [n]$ is*

$$G_n(\boldsymbol{z}) := G_{I_1,I_2,\cdots,I_n}(z_1, z_2, \cdots, z_n) = \det(g_{i,j})_{0\leq i,j\leq n}, \tag{4.40}$$

where $g_{i,j} = p_{j-i}, \forall i - j \neq 1$ and $g_{j+1,j} = p_{-1} - z_j = 1 - z_j$.

Corollary 4.5.3 *The ordinary probability generating function of the number S_n of ones in $\mathcal{X} = [n]$ is*

$$G_{S_n}(z) = \det(\hat{p}_{j-i})_{0\leq i,j\leq n}, \tag{4.41}$$

where $\hat{p}_k = p_k, \forall k \neq -1$ and $\hat{p}_{-1} = p_{-1} - z = 1 - z$.

The proof to Corollary 4.5.2 is quite straightforward: just multiply (4.39) by $\prod_{i\in A} z_i$ and then sum it up over all subsets $A \subseteq [n]$. Corollary 4.5.3 is then obvious by considering the ordinary generating function of $S_n = \sum_{i\in[n]} I_i$, i.e. treat all z_i in (4.40) as z. We are also aware of the following known determinantal generating function formula for any 1-dependent process (not necessarily stationary)

$$G_{S_n}(z) = \det(I + (z - 1)K), \tag{4.42}$$

where I is the identity matrix and K is the determinantal correlation kernel given by Theorem 4.5.1. This is not easy to simplify even for the stationary case. But from hindsight, one may check that this is essentially equivalent to (4.41). Lastly, one may also show Theorem 4.2.1 by the Laplace expansion of the last column of the determinant on the right of (4.41).

4.6 Enumeration of sequences

We have derived our main Theorem 4.2.1 from a probabilistic point of view, without assuming the sequences to be 2-block factors. But its enumerative Corollary 4.6.1 on integer-valued 2-block factors can be deduced from a combinatorial result in Goulden and Jackson [41, Section 4].

Recall (4.21), we defined 2-block factors on V based on the background sequence $(Y_1, Y_2, \ldots)$ as indicators of adjacent pairs (Y_i, Y_{i+1}) falling in some subset $B \subseteq V^2$, which obviously is a 1-dependent indicator process. For example, in section 4.3, we discussed the case where $V = [0, 1]$ and $B = \{(x, y) : 0 \le y \le x \le 1\}$.

Say a string on $\mathbb{Z}_+$ is a *B-string* if each adjacent pair belongs to B. Suppose further that the background sequence $(Y_1, Y_2, \ldots)$ has i.i.d. uniform distribution on $[m] = \{1, 2, \cdots, m\}$. Then, the 1-run probability p_k can be used to count the number m_k of B-strings of length k, i.e.

$$m_k = m^k p_{k-1}, \quad k \ge 1, \tag{4.43}$$

and $m_0 = 1$ by convention. Define the *B-string generating function* as

$$G(v) = \sum_{k=0}^{\infty} m_k v^k = 1 + \sum_{k=1}^{\infty} p_{k-1} m^k v^k = 1 + mvP(mv), \tag{4.44}$$

where $P(v)$ is the 1-run probability generating function. Now the following corollary follows from Theorem 4.2.1 by elementary algebra.

Corollary 4.6.1 *For $V = [m]$, the number of sequences of length $n \ge 1$ on V with exactly $k \in [0, n-1]$ occurrences of pairs of adjacent components in B is*

$$[z^k v^n] \frac{z - 1}{z - G((z-1)v)}. \tag{4.45}$$

To see the connection between Corollary 4.6.1 and [41], we need the following definition from [41]. Let the *B-string of length k enumerator* be

$$\gamma_k(\mathbf{v}) = \gamma_k(v_1, v_2, \cdots) := \sum_{|\sigma|=k} \prod_i v_i^{\tau_i(\sigma)}, \tag{4.46}$$

where the sum is over all B-string σ of length k and $\tau_i(\sigma)$ is the number of times that component $i \in \mathbb{Z}_+$ appears, with v_i being the counting variable associated with i.

Using the B-string of length k enumerator, the following result for enumerating sequences with a certain number of occurrences of B is shown in [41]:

Theorem 4.6.2 ([41, Corollary 4.2.12]) *The number of sequences of length $n = \sum_i \tau_i$ with exactly τ_i i's and $k (\le n-1)$ occurrences of pairs of adjacent components in B is*

$$[z^k \prod_i v_i^{\tau_i}] \left\{ 1 - \sum_{j=1}^{\infty} (z-1)^{j-1} \gamma_j(\boldsymbol{v}) \right\}^{-1}. \tag{4.47}$$

By setting $v_1 = v_2 = \cdots = v$, this becomes equivalent to (4.45) since $\gamma_j(v\mathbf{1}) = m_j v^j$. A more generalized version of Theorem 4.6.2, known as the Goulden-Jackson cluster theorem, allows enumeration of adjacent components of any length (not just pairs), see [40, 36].

In the case of descents, treated Section 4.3, the above discussion does not apply directly, as since the background sequence is continuously distributed. However, this can be circumvented by considering the ranks in the partial sequence instead of the absolute values, hence it can be treated combinatorially as in [41, Section 2.4.21] by considering only permutations of $\{1, 2, \cdots, n\}$ instead of all possible sequences.

An application of Corollary 4.6.1 can be found in [34].

Theorem 4.6.3 ($a = 4$: Florez [34, Theorem 9]) *The number of sequences on $V = \{0, 1, \cdots, a-1\}$ of length $n + m$ with exactly m occurrences of adjacent pair 01 is*

$$f(n, m) = [x^m y^n]F(x, y) = [x^m y^n]\frac{1}{1 - (a + x)y + y^2}. \tag{4.48}$$

To check this from Corollary 4.6.1, note first that the 1-run probability generating function is $P(v) = 1 + \frac{1}{a^2}v$, hence the B-string generating function is $G(v) = 1 + av + v^2$. (4.48) is then proved by substitutions $k \to m, n - k \to n, v \to y, zv \to x$.

4.7 Comparison with other dependence structures

Table 2		
Dependence structure	$Q(z, v)$ in terms of $Q(v)$	$Q(z, v)$ in terms of $P(v)$
Stationary 1-dependent	$\frac{Q((1-z)v)}{1-zvQ((1-z)v)}$	$\frac{P(-(1-z)v)}{1-vP(-(1-z)v)}$
Exchangeable	$\frac{Q\left(\frac{(1-z)v}{1-zv}\right)}{1-zv}$	$\frac{P\left(-\frac{(1-z)v}{1-v}\right)}{1-v}$
Renewal	$\frac{Q(v)}{1-z(1+(v-1)Q(v))}$	N/A
Stationary renewal	$\frac{-z+(z-v+(1-z)vQ'(0))Q(v)}{z(v-1)+(1-z)v(Q'(0)-1)+z(v-1)^2Q(v)}$	N/A

Apart from stationary 1-dependent processes we have discussed in this chapter, there are several other dependence structures whose bivariate generating functions can be written in terms of run probability generating functions. It is interesting to compare these formulas, but no current theory seems to unify them. Some examples may even belong to more than one dependence structure, see Table 2 above and explanations below.

Exchangeable A sequence $(X_n, n \geq 1)$ is *exchangeable* iff for each $n \geq 1$ and each permutations π of $[n]$

$$(X_{\pi(1)}, X_{\pi(2)}, \cdots, X_{\pi(n)}) \stackrel{d}{=} (X_1, X_2, \cdots, X_n). \tag{4.49}$$

Properties of exchangeable sequences were first studied by de Finetti [31]. The i.i.d. example in Section 4.4 is also exchangeable.

Renewal and stationary renewal Consider the partial sum $S_n = X_1 + X_2 + \cdots + X_n$ of the sequence $(X_n, n \geq 1)$, and its inter-arrival times $T_1 = \min\{n : S_n = 1\}, T_k = \min\{n : S_n = k\} - T_{k-1}$. We say $(X_n, n \geq 1)$ is *delayed renewal*, if the *renewals* $T_2, T_3, \cdots$ are i.i.d., and the *delay* T_1 is independent of $T_2, T_3, \cdots$. In particular, we say the indicator sequence is

1. *renewal conditional on* $X_0 = 1$ or *renewal*, if the delay T_1 has the same distribution as the renewals $T_2, T_3, \cdots$;
2. *stationary renewal*, if it is both stationary and delayed renewal.

For these sequences it is not possible to recover the bivariate generating function from the 1-run probability generating function $P(v)$, since in the delayed renewal case

$$P(v) = 1 + q_1(v + r_1 v^2 + r_1^2 v^3 + \cdots)$$

contains only the information about immediate renewals, which does not determine the distribution of T_1 or the 0-run probabilities.

The bivariate generating formula is an indirect corollary of [51, Theorem 3.5.4]. See [29] for introductions to renewal theory.

Three examples which appeared earlier in this chapter are both stationary 1-dependent and stationary renewal: the indicator of two consecutive ones in Section 4.4, the case $b = 2$ of carries when adding a list of digits in Section 4.4, and the generalized Florez' example in Section 4.6. All these three examples can be treated as indicators of time-homogeneous Markov chains visiting a particular state.

Chapter 5

A positivity problem

In this chapter, we discuss a positivity problem related to a renewal process in integer times and the associated bivariate probability generating function, whose coefficients forms a Riordan array introduced by Rogers [92]. The renewal sequence is generalized such that some of the coefficients are allowed to be negative, while the entries of the Riordan array remain non-negative. We provide some necessary conclusions bounding the negative coefficients. Additionally, we provide graphs as well as qualitative observational results about the coefficients of powers of polynomials. However, it is challenging to give a quantitative description or find known results to explain them.

5.1 Definitions

Consider a renewal process on $\{0, 1, 2, \cdots\}$ with its 0-th renewal at 0. Suppose the *holding times* $T_1, T_2, \cdots$ are independent and identically distributed $\mathbb{Z}^+$ -valued random variable with distribution

$$\mathbb{P}(T_1 = k) = f_k, \qquad (k = 1, 2, \cdots). \tag{5.1}$$

Let $X_n := \mathbb{1}$(there is a renewal at time n), i.e. the indicator if there is a renewal at time n. In particular $X_0 \equiv 1$ by our assumption that there is a renewal at time 0. Call X the *renewal indicator process*. Let $S_n := \sum_{i=1}^{n} X_n$ be the number of renewals up to time n. Call S the *renewal process*. (Sometimes X is also called the renewal process, but we will avoid this name for X.)

Other than f_k, the 0-*run probabilities*

$$q_k = \mathbb{P}(S_k = 0) = 1 - f_1 - f_2 - \cdots - f_k, \qquad (k = 1, 2, \cdots) \tag{5.2}$$

also characterize the distribution of this renewal process.

In terms of Riordan array, the ordinary probability generating functions of $S_1, S_2, \cdots$, as row vectors, forms a *Riordan array* $(g(x), f(x))$ with

$$f(x) = f_1 x + f_2 x^2 + f_3 x^3 + \cdots = \sum_{n=1}^{\infty} f_n x^n \tag{5.3}$$

and

$$g(x) = \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0)x^n = 1 + (1 - f_1)x + (1 - f_1 - f_2)x^2 + \cdots = \frac{1 - f(x)}{1 - x}. \tag{5.4}$$

In other words, the k-th column of this array is the coefficients of the polynomial $g(x)f^k(x)$

$$\mathbb{P}(S_n = k) = [x^n]g(x)f^k(x). \tag{5.5}$$

Riordan array was first studied by Rogers [92] as a generalization of the Pascal triangle, under the name *renewal array*. Shapiro et al. further generalized the same concept and used the name of Riordan array. See also, Sprugnoli [100] and Merlini et al. [77] for more on the theory of Riordan arrays.

5.2 Positivity

We say a polynomial is *positive*, if all of its coefficients are non-negative. We say a Riordan array is *positive*, if all of its entries are non-negative. By natural probabilistic explanation, we know that for each probability generating function $f(\cdot)$ of holding times, the associated Riordan array $\left(g(x) := \frac{1-f(x)}{1-x}, f(x)\right)$ is positive, namely the probability generating functions of all S_n's are non-negative if

- $f(\cdot)$ is positive;
- or equivalently, the associated sequence of 0-run probabilities $1 = q_0 \geq q_1 \geq q_2 \geq \cdots \geq 0$ is non-increasing.

However, the converse statement is not true: without assuming all f_k's to be positive, or the 0-run probability to be non-decreasing, the associated Riordan array can still be positive!

Example 5.2.1 *Set*

$$f_1 = \frac{1}{5},\ f_2 = \frac{2}{5},\ f_3 = -\frac{1}{5},\ f_4 = \frac{2}{5},\ f_5 = \frac{1}{5}, \qquad \textit{and} \qquad f_k = 0, k \geq 6, \tag{5.6}$$

or in terms of Riordan array,

$$f(x) = \frac{x}{5}(1 + 2x - x^2 + 2x^3 + x^4), \qquad g(x) = \frac{1}{5}(5 + 4x + 2x^2 + 3x^3 + x^4). \tag{5.7}$$

We check the positivity of all coefficients of $f^n(x)g(x)$, to examine the positivity of the associated Riordan array. Since the constant term of $f(\cdot)$ is zero, consider

$$h(x) = \frac{f(x)}{x} = f_1 + f_2 x + f_3 x^2 + \cdots, \tag{5.8}$$

for simplicity. Namely, it is sufficient to check the positivity of $h^n(x)g(x)$ instead.

Observe that

$$h^4(x) = \frac{1}{5^4}(1 + 8x + 20x^2 + 16x^3 + 26x^4 + 88x^5 + 48x^6 + 24x^7 + 163x^8 + 24x^9 + 48x^{10} + 88x^{11} + 26x^{12} + 16x^{13} + 20x^{14} + 8x^{15} + x^{16}), \quad (5.9)$$

and

$$g(x) = \frac{1}{5}(5 + 4x + 2x^2 + 3x^3 + x^4); \quad (5.10)$$

$$h(x)g(x) = \frac{1}{25}(5 + 14x + 6x^2 + 15x^3 + 17x^4 + 9x^5 + 8x^6 + 5x^7 + x^8); \quad (5.11)$$

$$h^2(x)g(x) = \frac{1}{625}(5 + 24x + 29x^2 + 23x^3 + 74x^4 + 54x^5 + 45x^6 + 61x^7 + 38x^8 + 22x^9 + 17x^{10} + 7x^{11} + x^{12}); \quad (5.12)$$

$$h^3(x)g(x) = \frac{1}{5^4}(5 + 34x + 72x^2 + 67x^3 + 144x^4 + 261x^5 + 154x^6 + 268x^7 + 297x^8 + 181x^9 + 190x^{10} + 156x^{11} + 80x^{12} + 51x^{13} + 30x^{14} + 9x^{15} + x^{16}). \quad (5.13)$$

The positivity of $h^n(x)g(x)$ then follows immediately, since every $n \geq 4$ can be written as $n = 4m + k$ for some $k \in \{0, 1, 2, 3\}$ and $m \in \mathbb{Z}^+$, hence

$$h^n(x)g(x) = \left[h^4(x)\right]^m \cdot h^k(x)g(x) \quad (5.14)$$

has all its coefficients positive.

Note that $[x^3]h^3(x) = [x^6]f^3(x) = -\frac{7}{625} < 0$. Thus it is necessary to check $h^4(x)$ and $h^k(x)g(x)$ for $k = 0, 1, 2, 3$.

Definition 5.2.2 *We say a polynomial $f(x) = f_1x + f_2x^2 + \cdots$ is a* semi-positive renewal probability generating function *if*

1. *f_1 is not zero;*
2. *$\sum_{n=1}^{\infty} f_n = 1$;*
3. *There exists at least one n such that $f_n < 0$;*
4. *The associated Riordan array $\left(g(x) = \frac{1-f(x)}{1-x}, f(x)\right)$ is positive, namely $g(x)f^k(x)$ is positive for $k = 0, 1, \cdots$.*

In fact, Example 5.2.1 has the lowest possible degree among all semi-positive renewal probability generating functions.

Theorem 5.2.3 *If a semi-positive renewal probability generating function $f(\cdot)$ has finite degree d, then $d \geq 5$.*

Proof We show by proving the first two and last two coefficients cannot be negative.

f_1 cannot be negative since $[x]g(x)f(x) = f_1 \geq 0$.

f_2 cannot be negative as

$$[x]g(x)h^n(x) = f_1^n(1-f_1) + nf_1^{n-1}f_2 = nf_1^{n-1}\left(\frac{(1-f_1)f_1}{n} + f_2\right) \geq 0, \tag{5.15}$$

hence

$$f_2 \geq -\lim_{n\to\infty}\frac{(1-f_1)f_1}{n} = 0. \tag{5.16}$$

The above arguments are true without assuming that $f(\cdot)$ has finite degree. Now take the assumption that $f(\cdot)$ has finite *degree* d into consideration, i.e. $f_d \neq 0$ while $f_k = 0, k \geq d+1$, whence

$$g(x) = \frac{1-f(x)}{1-x} = \cdots + (f_{d-1}+f_d)x^{d-2} + f_d x^{d-1}. \tag{5.17}$$

Hence, f_d cannot be negative since

$$[x^{3d-1}]g(x)f^2(x) = f_d^3 \geq 0. \tag{5.18}$$

Lastly, f_{d-1} cannot be negative as

$$[x^{(n+1)d-2}]g(x)f^n(x) = f_d^n(f_{d-1}+f_d)+nf_{d-1}f_d^{n-1}f_d = (n+1)f_d^n\left(f_{d-1}+\frac{f_d}{n+1}\right) \geq 0, \tag{5.19}$$

hence

$$f_{d-1} \geq -\lim_{n\to\infty}\frac{f_d}{n+1} = 0. \tag{5.20}$$

■

5.3 Numerical result on a simple case

Based on the discussion of Theorem 5.2.3, consider the following 'minimal' case:

$$h(x) = f_1 + f_2 x - (2f_1 + 2f_2 - 1)x^2 + f_2 x^3 + f_1 x^4, \tag{5.21}$$

and

$$g(x) = 1 + (1-f_1)x + (1-f_1-f_2)x^2 + (f_1+f_2)x^3 + f_1 x^4. \tag{5.22}$$

In spirit of Example 5.2.1, I write a computer program brutally searching for values of f_1, f_2 such that $g(x)h^k(x)$ is positive for every $k \leq 100$. See Figure 5.1 below. $f_3 = 1 - 2f_1 - 2f_2$ is minimized when $f_1 \approx 0.222, f_2 \approx 0.445$.

Conjecture 5.3.1 *For $h(x)$ and $g(x)$ so defined as in (5.21) and (5.22), the minimal value of f_3 such that $g(x)h^k(x)$ is positive for every k is reached when $f_1 = 2/9$ and $f_2 = 4/9$, i.e.*

$$h(x) = \frac{1}{9}(2 + 4x - 3x^2 + 4x^3 + 2x^4), \tag{5.23}$$

and

$$g(x) = \frac{1}{9}(9 + 7x + 3x^2 + 6x^3 + 2x^4). \tag{5.24}$$

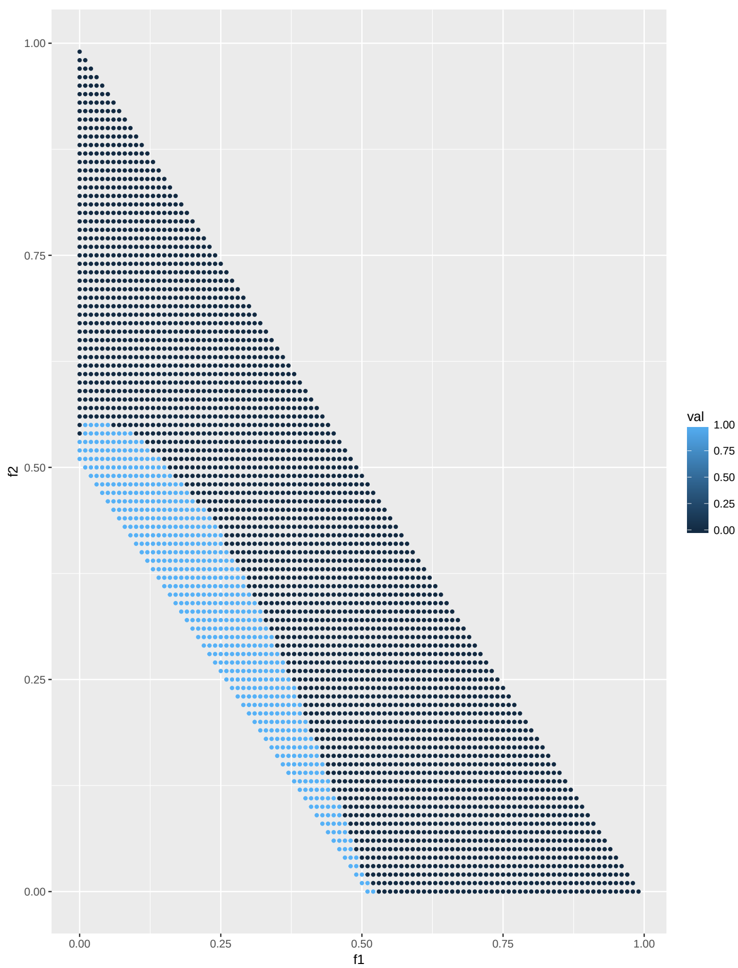


Figure 5.1: The dot is blue if $g(x)h^k(x)$ is positive for every $k \leq 100$, otherwise black.

It is still unclear with two parameters f_1 and f_2, so let us further simplify (5.21) and (5.22) with $f_1 = f_2/2 = a$, namely

$$h(x) = a + 2ax - (6a - 1)x^2 + 2ax^3 + ax^4, \tag{5.25}$$

and

$$g(x) = 1 + (1 - a)x + (1 - 3a)x^2 + 3ax^3 + ax^4, \tag{5.26}$$

which matches Example 5.2.1 when $a = \frac{1}{5}$ and Conjecture 5.3.1 when $a = \frac{2}{9}$.

In this case, it seems most interesting when $a \approx \frac{2}{9}$. Again here are some results from computer programming:

- When a approaches $\frac{2}{9}$ from above, the first k such that $h^n(x)g(x)$ is no longer positive increases. For example, when $a = 0.2223$, $h^n(x)g(x)$ is positive for all $n \leq 85$ but not $n = 86$. This pattern is true until $a = 0.2222223$ where $h^n(x)g(x)$ remains positive until n is so large that the accuracy is not enough to detect a negative coefficient.

- When a approaches $\frac{2}{9}$ from below, it seems always possible to give a proof by finding the minimal $N := N(a)$ such that $h^n(x)$ and $h^k(x)g(x), k = 0, 1, \cdots n-1$ are all positive, just like the proof of Example 5.2.1. The following numbers are accurate:

$$N(0.21) = 6, N(0.22) = 22, N(0.222) = 180, N(0.2222) = 1740. \tag{5.27}$$

  It is worth-noticing that positivity of $h^N(x)$ does not imply positivity of $h^{N+1}(x)$. In fact, when $a = \frac{1}{5}$ in Example 5.2.1, $h^k(x)$ for every $k > N$; but for each of the four a values above, $h^{N+1}(x)$ is not positive.

- When $a = \frac{2}{9}$, $h^n(x)g(x)$ is positive while $h^n(x)$ is not for every n as long as the accuracy allows.

In addition, when plotting the coefficients (normalized such that the maximum is 1) of $h^n(x)$ for different choices of h, different patterns show up, as displayed in Figures 5.2 through 5.5 below. For n large, the contours are all well approximated by one or two normal curves, except the negative one in Figure 5.4 which accounts for the case when $a = \frac{2}{9}$.

The most interesting case occurs when $f_1 = 0.232 > 2/9, f_2 = 0.434 < 4/9$, shown in Figure 5.5, where the pattern clearly consists of three branches of functions, each of which is approximately the same normal curve multiplied by a sin function with a phase shift $\frac{2\pi}{3}$ among one another. We will see this pattern again in Section 5.5

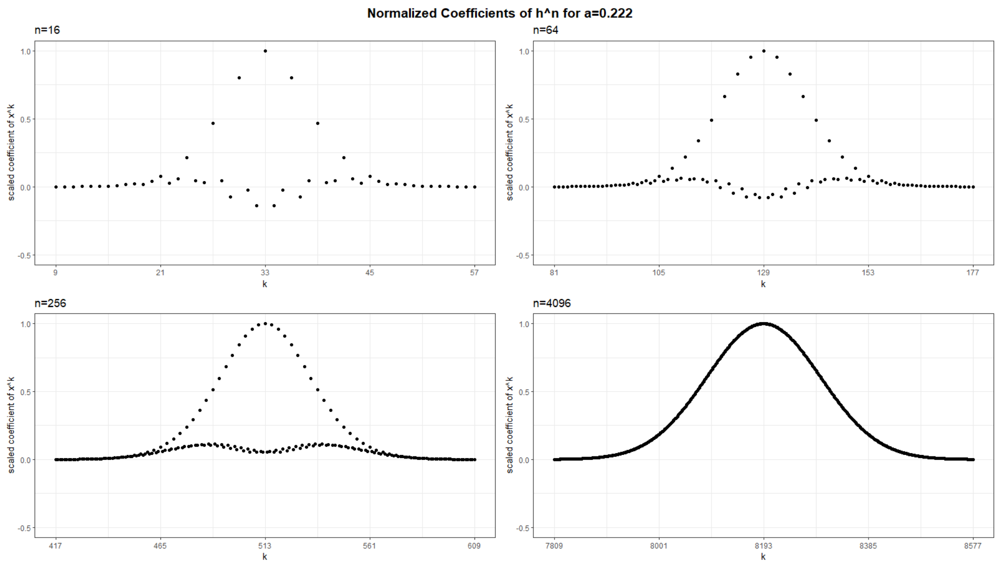


Figure 5.2: $a = 0.222 < 2/9$

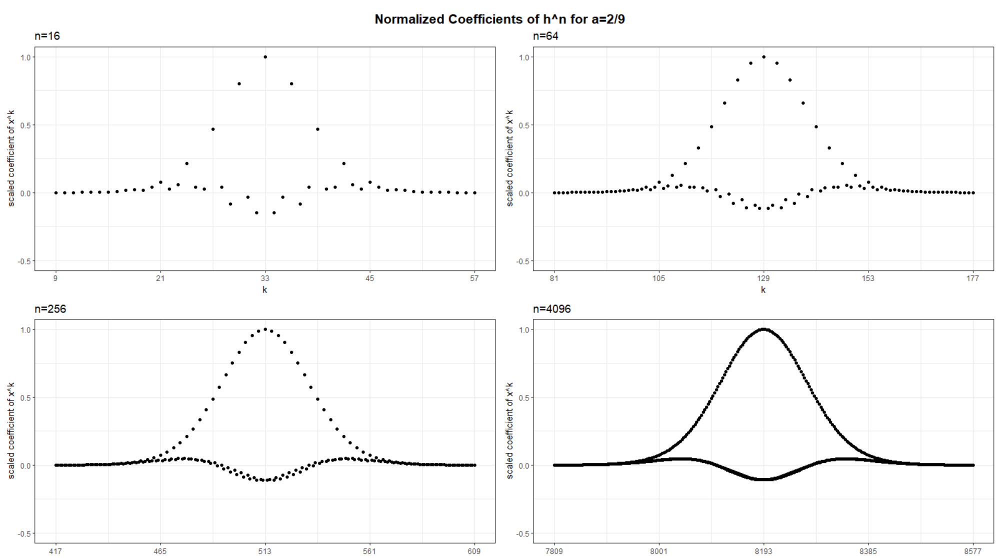


Figure 5.3: $a = 2/9$

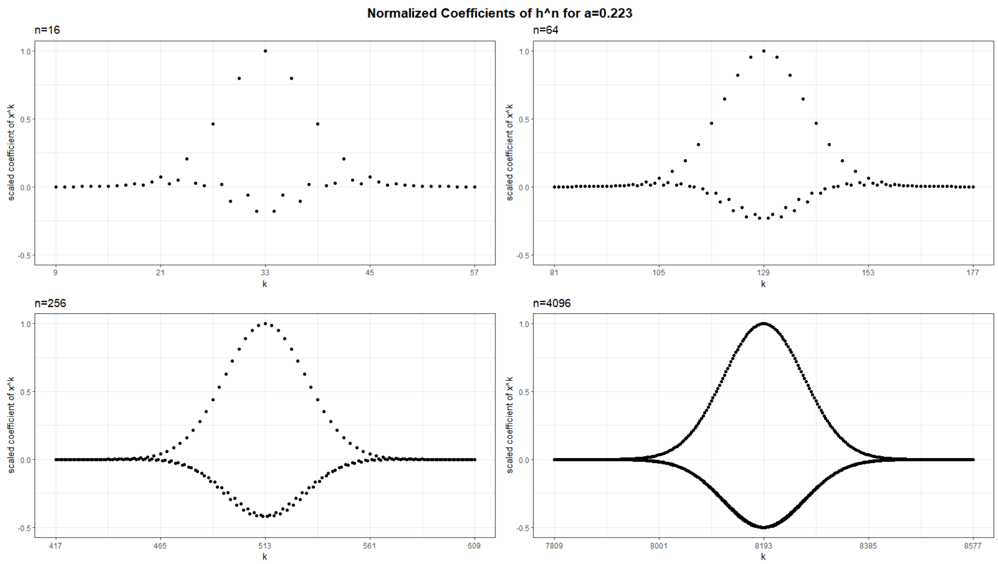


Figure 5.4: $a = 0.223 > 2/9$

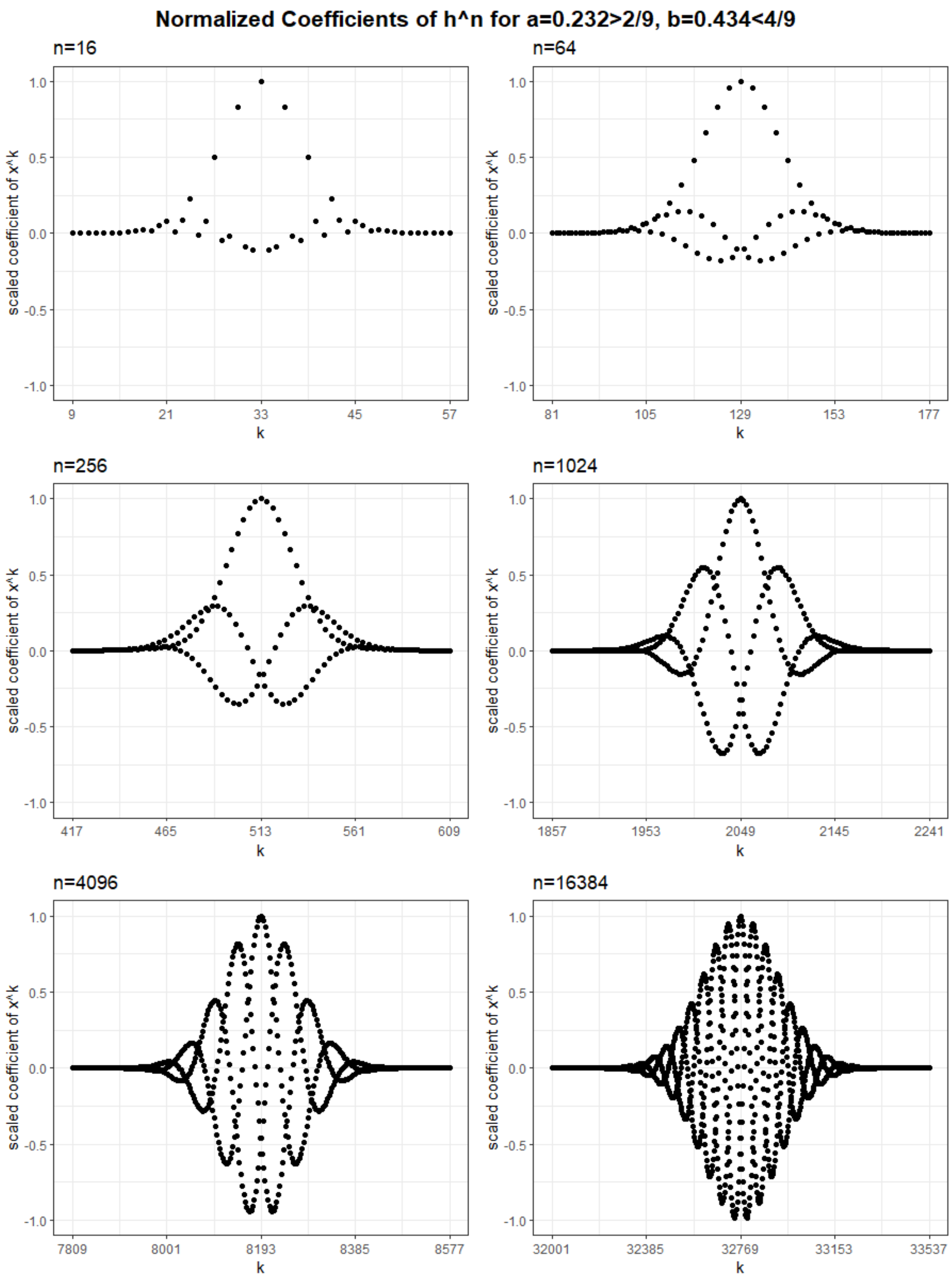


Figure 5.5: $f_1 = 0.232 > 2/9, f_2 = 0.434 < 4/9$

5.4 A lower bound on f_3

In this section, we focus on a semi-positive renewal probability generating function

$$f(x) = f_1 x + f_2 x^2 + f_3 x^3 + f_4 x^4 + f_5 x^5. \tag{5.28}$$

We know in the proof of Theorem 5.2.3 that f_3 is the only negative coefficient. We will show that

Theorem 5.4.1 *If a semi-positive renewal probability generating function $f(\cdot)$ has degree $d = 5$, then $-\frac{1}{3} \leq f_3 < 0$.*

Proof

The proof is done by considering the convolution on the *quotient polynomial ring*

$$Q_3[x] := R[x]/R[x^3 - 1] \tag{5.29}$$

instead of the regular polynomial ring $R[x]$. In other words, for any polynomial $a(x) = \sum_{n=1}^{\infty} a_n x^n$, we consider only the lowest order representative of its equivalence class

$$a_{(3)}(x) = a_{(0)} + a_{(1)}x + a_{(2)}x^2 := \sum_{n=0}^{\infty} a_{3n} + \sum_{n=0}^{\infty} a_{3n+1}x + \sum_{n=0}^{\infty} a_{3n+2}x^2. \tag{5.30}$$

Hence, we write

$$f(x) + R[x^3 - 1] = f_3 + (f_1 + f_4)x + (f_2 + f_5)x^2 + R[x^3 - 1], \tag{5.31}$$

or

$$f(x) \equiv f_3 + (f_1 + f_4)x + (f_2 + f_5)x^2 := c_0 + c_1 x + c_2 x^2 = c_{(3)} \tag{5.32}$$

for short.

Note that in the quotient ring $Q_3[x]$, multiplication with $f(x) \equiv c_{(3)}(x)$ can be treated as left multiplication by the following 3×3 *circulant matrix*

$$C_3 := \begin{pmatrix} c_0 & c_2 & c_1 \\ c_1 & c_0 & c_2 \\ c_2 & c_1 & c_0 \end{pmatrix} \qquad \text{with} \qquad \sum_{j=0}^{2} c_j = 1, \tag{5.33}$$

while the k-th power $f^k(x) \equiv c_{(3)}^k(x)$ can be written as

$$(1, x, x^2) C_3^k (1, 0, 0)^t, \qquad (k = 0, 1, 2, \cdots), \tag{5.34}$$

where $(\cdots)^t$ means transpose of a vector or a matrix. The above power representation naturally brings our interest to the eigenvalues and eigenvectors of the circulant matrix C_3.

It is not hard to check that for an $n \times n$ circulant matrix

$$C_n := \begin{pmatrix} c_0 & c_{n-1} & \cdots & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_{n-2} & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_1 & c_0 \end{pmatrix} \qquad \text{with} \qquad \sum_{j=0}^{n-1} c_j = 1, \tag{5.35}$$

its normalized eigenvectors are

$$v_j = \frac{1}{\sqrt{n}}(1, \omega_j, \omega_j^2, \cdots, \omega_j^{n-1})^t, \qquad j = 0, 1, 2, \cdots, n-1, \tag{5.36}$$

where i is the imaginary unit and $w_j = e^{\frac{2j\pi i}{n}}, j = 0, 1, \ldots, n-1$ are the n-th root of unity. The corresponding eigenvalues are

$$\lambda_j = c_0 + c_{n-1}\omega_j + c_{n-2}\omega_j^2 + \cdots + c_2\omega_j^{n-2} + c_1\omega_j^{n-1}, \qquad j = 0, 1, 2, \cdots, n-1. \tag{5.37}$$

In particular, $\lambda_0 = 1$.

Note that

$$(1, 0, 0, \cdots, 0)^t = \frac{1}{\sqrt{n}}(v_0 + v_1 + v_2 + \cdots + v_{n-1}). \tag{5.38}$$

Thus,

$$C_n^k(1, 0, 0, \cdots, 0)^t = \frac{1}{\sqrt{n}} C_n^k \sum_{j=0}^{n-1} v_j = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \lambda_j^k v_j, \tag{5.39}$$

which means when $\lambda_0 = 1$ is the only dominant eigenvalue, this limit converges to

$$\lim_{k\to\infty} C_n^k(1, 0, 0, \cdots, 0)^t = \frac{1}{\sqrt{n}} v_0 = \frac{1}{n} v_0, \tag{5.40}$$

and if there exists some $|\lambda_j| > 1$, this limit diverges, since the summation over all entries of $C_n^k(1, 0, 0, \cdots, 0)^t$ is always 1.

See Ingleton [55] for a comprehensive study on circulant matrices.

Now getting back to our case of interest when

$$f(x) \equiv c_{(3)}(x) = f_3 + (f_1 + f_4)x + (f_2 + f_5)x^2, \tag{5.41}$$

$$\begin{aligned} g(x) &= 1 + (1 - f_1)x + (1 - f_1 - f_2)x^2 + (f_4 + f_5)x^3 + f_5x^4 \\ &\equiv d(x) = (1 + f_4 + f_5) + (1 - f_1 + f_5)x + (1 - f_1 - f_2)x^2. \end{aligned} \tag{5.42}$$

and the three eigenvalues are $\lambda_0 = 1$ and

$$\lambda_1 = \overline{\lambda_2} = f_3 + (f_1 + f_4)\omega_1 + (f_2 + f_5)\omega_1^2 = \frac{3}{2}f_3 - \frac{1}{2} + \frac{\sqrt{3}}{2}(-f_1 + f_2 - f_4 + f_5)i, \tag{5.43}$$

where $\omega_1 := e^{\frac{2\pi i}{3}} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$.

Let $A_k := Re(\lambda_1^k), B_k := \sqrt{3}Im(\lambda_1^k)$ where $Re(\cdot)$ and $Im(\cdot)$ stand for the real and imaginary part, repectively, then

$$c_{(3)}^k(x) = (1,x,x^2)C_3^k(1,0,0)^t \equiv (1,x,x^2)\frac{1}{\sqrt{3}}(v_0 + \lambda_1^k v_1 + \lambda_2^k v_2) \tag{5.44}$$

$$= (1,x,x^2)\frac{1}{3}(1+2A_k, 1-A_k-B_k, 1-A_k+B_k)^t \tag{5.45}$$

$$= \frac{1}{3}\left[(1+x+x^2) + A_k(2-x-x^2) + B_k(-x+x^2)\ \right]. \tag{5.46}$$

Then the lowest order equivalent of $d(x)c_3^k(x)$ is

$$d(x)c_{(3)}^k(x) = [(1+f_4+f_5) + (1-f_1+f_5)x + (1-f_1-f_2)x^2]\times \tag{5.47}$$

$$\frac{1}{3}\left[(1+x+x^2) + A_k(2-x-x^2) + B_k(-x+x^2)\ \right] \tag{5.48}$$

$$\equiv \frac{1}{3}\{[u + (f_2+f_5)A_k + (2f_1+f_2+2f_4+f_5)B_k] + \tag{5.49}$$

$$[u + (-f_1-f_2-f_4-f_5)A_k + (-f_1+f_2-f_4+f_5)B_k]\ x + \tag{5.50}$$

$$[u + (f_1+f_4)A_k + (-f_1-2f_2-f_4-2f_5)B_k]\ x^2\Big\} := \frac{1}{3}L_k(x), \tag{5.51}$$

where $u = 3 - 2f_1 - f_2 + f_4 + 2f_5$ remains unchanged when k changes.

We show here by contradiction that $|\lambda_1| \le 1$. Suppose $|\lambda_1| > 1$. When k is large, $\max\{|A_k|, |B_k|\} \gg 1$ hence we may omit the constant u when considering non-negativity of $L_k(x)$.

Recall that we say a real polynomial $L_k(x)$ is positive when each of its coefficient is non-negative. Since $[1]L_k(x) + [x]L_k(x) + [x^2]L_k(x) = 3u$, each coefficient can only range within $[0, 3u]$.

It is easy to see the following fact since f_1, f_5 are positive and f_2, f_4 are non-negative: when $|A_k| > |B_k|$, $|[x]L_k(x)| \gg 3u$ for k large; otherwise, $|[x^2]L_k(x)| \gg 3u$ for k large. Hence, it is impossible for $L_k(x)$ to be positive for all $k \in \mathbb{Z}_+$.

Therefore, for f to be a semi-positive renewal probability generating function, it is necessary that $|\lambda_1| \le 1$, i.e.

$$\left[\frac{3}{2}f_3 - \frac{1}{2}\right]^2 + \left[\frac{\sqrt{3}}{2}(-f_1+f_2-f_4+f_5)\right]^2 \le 1. \tag{5.52}$$

Then $f_3 \ge -\frac{1}{3}$ follows. ∎

Remark 5.4.2 *From the last inequality* (5.52)*, we also obtain a necessary condition for the extreme case $f_3 = 1/3$ to happen:*

$$f_1 + f_4 = f_2 + f_5. \tag{5.53}$$

5.5 The twisting branches in powers of polynomials - more examples

In this section, we display several more examples of powers of a polynomial with similar pattern as shown in Figure 5.5.

There are actually a whole family of such functions, in the form

$$h(x) = h_3x^3 + h_2x^2 + h_1x + h_0 \tag{5.54}$$

where $h_0 + h_1 + h_2 + h_3 = 1$, $h_3 > 0, h_0 > 0$ and at least one of h_2 or h_1 is strictly negative. For simplicity, only the following cases are displayed:

$$h(x) = ax^3 - (2a-1)x + a \tag{5.55}$$

for $a = 0.53, 0.6, 0.8, 1, 3, 10$. But these facts are true for many other more general functions so defined in (5.54) as well.

Central limit theorem does not apply here since $h(x)$ is not positive. Nevertheless, the positive and negative contours of $h^n(x)$ are well approximated by the normal density of $N(c_1n, c_2n)$ multiplied by c_3^n for c_1, c_2 determined by $h(x)$ and

$$c_3 = \max_t |h(e^it)|. \tag{5.56}$$

For each $h(x)$, the three branches are approximately from the same normal curve multiplied by a sin function with a phase shift $\frac{2\pi}{3}$ among one another. The frequency of the sin function is different for different $h(x)$, but does not change when n is getting larger.

It is known by Greville [44, 43] that when $|h(e^{it})| < 1$ for $0 < t < 2\pi$, the convolution power will be convergent asymptotically, which is not our case of interest. See also Diaconis [22] for a recent study on convolution powers of polynomials.

Natural open questions include:

- Fix $a > 0$ and consider $h(x)$ given in (5.55). Are there preferably linear functions $C_1(n), C_2(n) > 0$, a preferably exponential function $C_3(n)$ and preferably constants $C_4(n) > 0$ and $0 < C_5(n) < 2\pi$ such that the distance between

$$([x^k]h^n(x), k \geq 0) \tag{5.57}$$

and

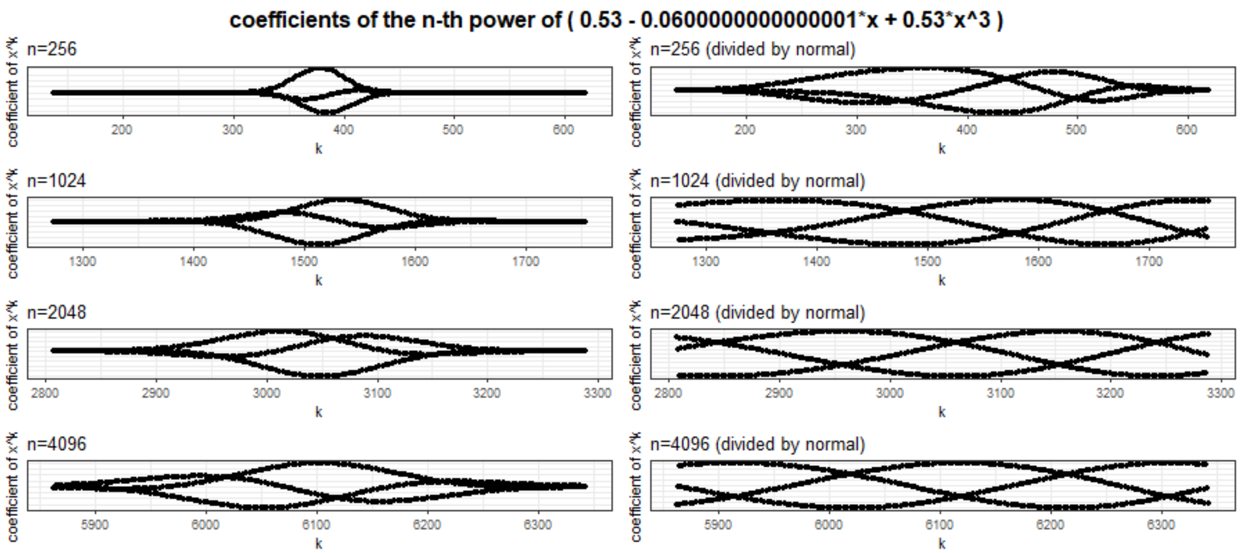
$$\frac{C_3(n)}{\sqrt{C_2(n)}}\phi\left(\frac{k - C_1(n)}{\sqrt{C_2(n)}}\right) sin\left(C_4(n)k + C_5(n) + \frac{2\pi}{3}(k\%3)\right) \tag{5.58}$$

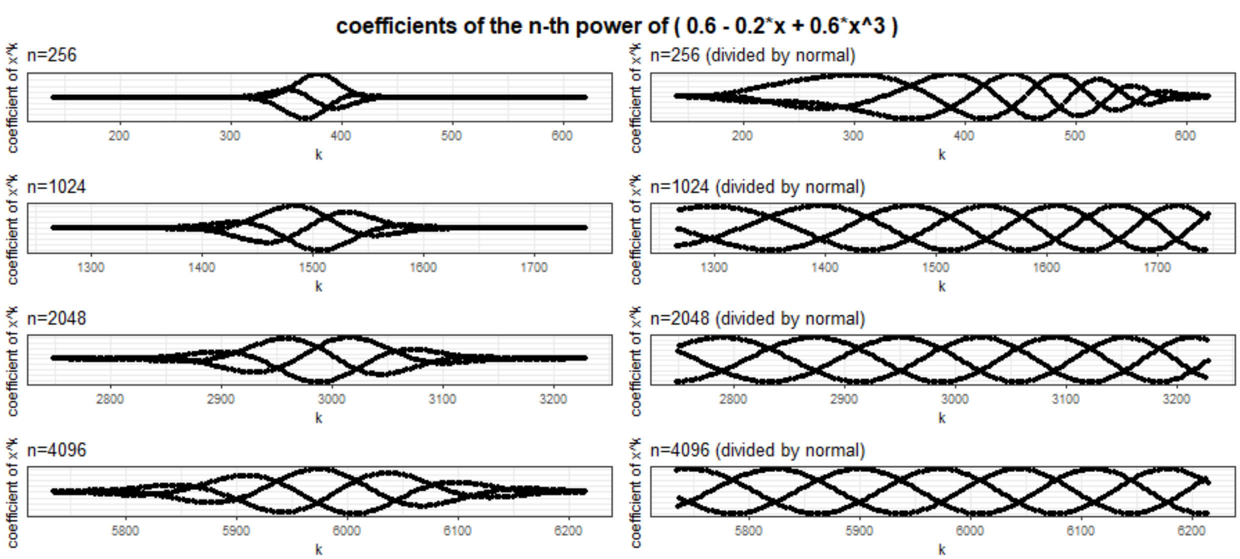
is vanishing when $n \to +\infty$, in some proper sense such as the L^1 or L^2 norm. Here $\phi(\cdot)$ is the usual standard normal density function, and $(k\%3)$ is the remainder of k divided by 3.

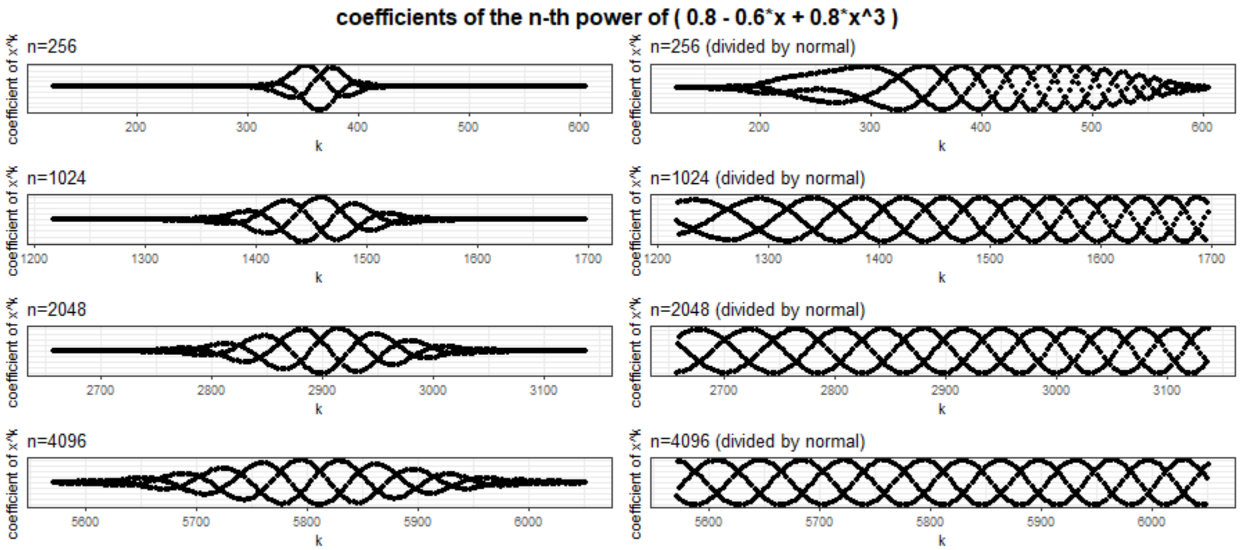
- Does the above result generalize to (5.54)? If so, what is the condition for the coefficients h_k such that this result holds?
- More generally, does this generalize to $h(x)$ of higher odd/prime order p such that there are p twisting branches appearing? A first example to check could be

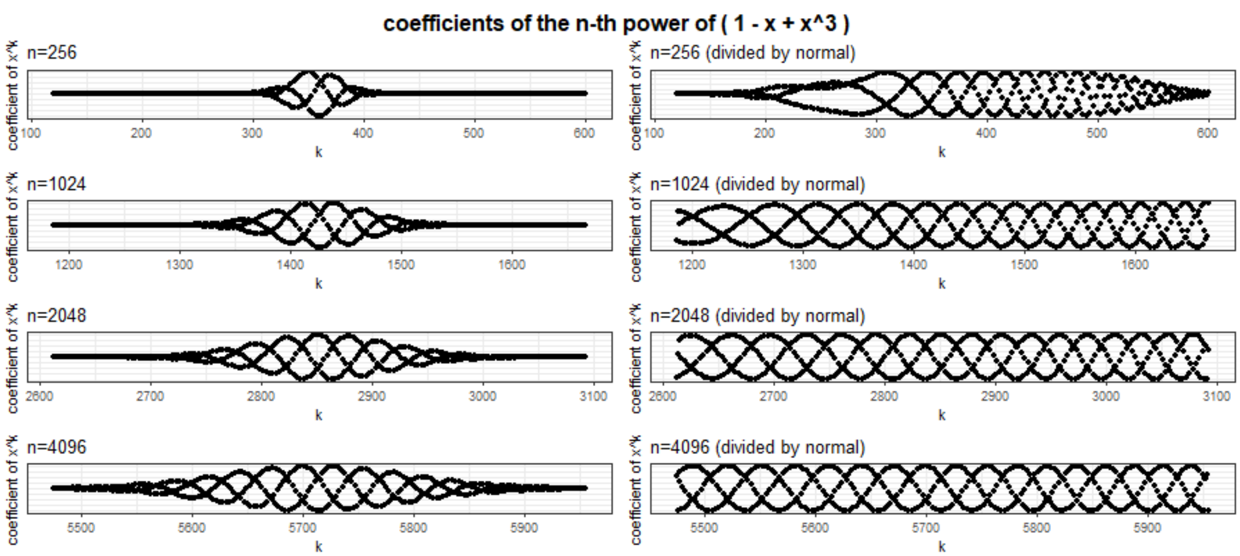
$$h(x) = ax^p - (2a-1)x + a \tag{5.59}$$

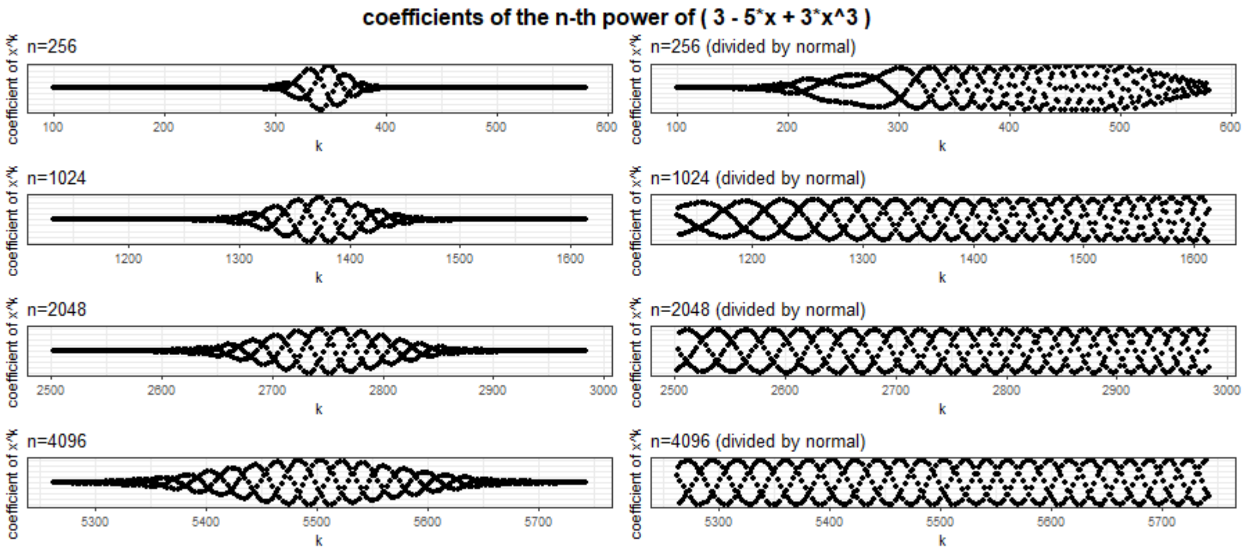
for prime number $p \geq 5$ and $a > 0$.

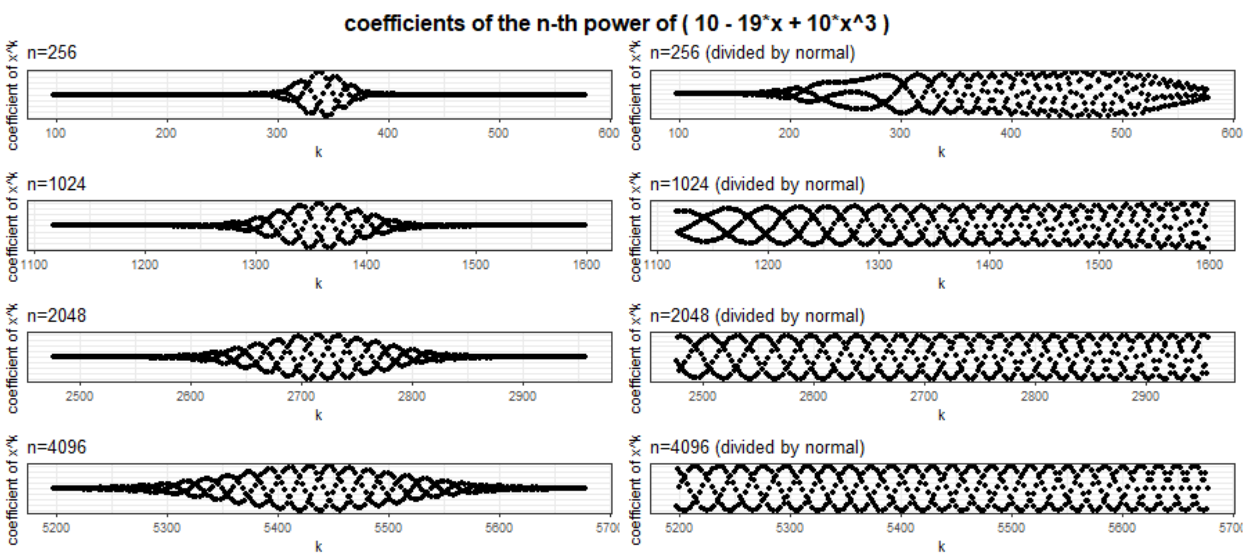





coefficients of the n-th power of (0.8 - 0.6*x + 0.8*x^3)
n=256
n=256 (divided by normal)
n=1024
n=1024 (divided by normal)
n=2048
n=2048 (divided by normal)
n=4096
n=4096 (divided by normal)
coefficient of x^k
k


coefficients of the n-th power of (1 - x + x^3)
n=256
n=256 (divided by normal)
n=1024
n=1024 (divided by normal)
n=2048
n=2048 (divided by normal)
n=4096
n=4096 (divided by normal)
coefficient of x^k
k


coefficients of the n-th power of (3 - 5*x + 3*x^3)
n=256
n=256 (divided by normal)
n=1024
n=1024 (divided by normal)
n=2048
n=2048 (divided by normal)
n=4096
n=4096 (divided by normal)
coefficient of x^k
k


coefficients of the n-th power of (10 - 19*x + 10*x^3)
n=256
n=256 (divided by normal)
n=1024
n=1024 (divided by normal)
n=2048
n=2048 (divided by normal)
n=4096
n=4096 (divided by normal)
coefficient of x^k
k

Bibliography

[1] Jon Aaronson et al. “An algebraic construction of a class of one-dependent processes”. In: *Ann. Probab.* 17.1 (1989), pp. 128–143. ISSN: 0091-1798. URL: `http://links.jstor.org/sici?sici=0091-1798(198901)17:1%3C128:AACOAC%3E2.0.CO;2-6&origin=MSN`.

[2] Richard Arratia. “On the amount of dependence in the prime factorization of a uniform random integer”. In: *Contemporary combinatorics*. Vol. 10. Bolyai Soc. Math. Stud. János Bolyai Math. Soc., Budapest, 2002, pp. 29–91.

[3] Richard Arratia. “On the central role of scale invariant Poisson processes on $(0, \infty)$”. In: *Microsurveys in discrete probability (Princeton, NJ, 1997)*. Vol. 41. DIMACS Ser. Discrete Math. Theoret. Comput. Sci. Amer. Math. Soc., Providence, RI, 1998, pp. 21–41.

[4] Richard Arratia, A. D. Barbour, and Simon Tavaré. “A tale of three couplings: Poisson-Dirichlet and GEM approximations for random permutations”. In: *Combin. Probab. Comput.* 15.1-2 (2006), pp. 31–62. ISSN: 0963-5483. DOI: `10.1017/S0963548305007054`. URL: `https://doi.org/10.1017/S0963548305007054`.

[5] Jean Bertoin and Marc Yor. “The entrance laws of self-similar Markov processes and exponential functionals of Lévy processes”. In: *Potential Anal.* 17.4 (2002), pp. 389–400. ISSN: 0926-2601. DOI: `10.1023/A:1016377720516`. URL: `https://doi.org/10.1023/A:1016377720516`.

[6] Alexei Borodin, Persi Diaconis, and Jason Fulman. “On adding a list of numbers (and other one-dependent determinantal processes)”. In: *Bull. Amer. Math. Soc. (N.S.)* 47.4 (2010), pp. 639–670. ISSN: 0273-0979. DOI: `10.1090/S0273-0979-2010-01306-9`. URL: `https://doi.org/10.1090/S0273-0979-2010-01306-9`.

[7] Robert M. Burton, Marc Goulet, and Ronald Meester. “On 1-dependent processes and k-block factors”. In: *Ann. Probab.* 21.4 (1993), pp. 2157–2168. ISSN: 0091-1798. URL: `http://links.jstor.org/sici?sici=0091-1798(199310)21:4%3C2157:O1PAF%3E2.0.CO;2-B&origin=MSN`.

[8] L. Carlitz and Richard Scoville. "Generalized Eulerian numbers: combinatorial applications". In: *J. Reine Angew. Math.* 265 (1974), pp. 110–137. ISSN: 0075-4102. DOI: 10.1515/crll.1976.288.218. URL: https://doi.org/10.1515/crll.1976.288.218.

[9] L. Carlitz, Richard Scoville, and Theresa Vaughan. "Enumeration of pairs of sequences by rises, falls and levels". In: *Manuscripta Math.* 19.3 (1976), pp. 211–243. ISSN: 0025-2611. DOI: 10.1007/BF01170773. URL: https://doi.org/10.1007/BF01170773.

[10] Peter Carr et al. "Self-decomposability and option pricing". In: *Math. Finance* 17.1 (2007), pp. 31–57. ISSN: 0960-1627. DOI: 10.1111/j.1467-9965.2007.00293.x. URL: https://doi.org/10.1111/j.1467-9965.2007.00293.x.

[11] Augustin-Louis Cauchy. *Cours d'analyse de l'École Royale Polytechnique*. Cambridge Library Collection. Reprint of the 1821 original, Previously published by Cooperativa Libraria Universitaria Editrice Bologna, Bologna, 1992 [MR1267567]. Cambridge University Press, Cambridge, 2009, pp. ii+xvi+576. ISBN: 978-1-108-00208-0. DOI: 10.1017/CBO9780511693328. URL: https://doi.org/10.1017/CBO9780511693328.

[12] Loïc Chaumont and Marc Yor. *Exercises in probability: a guided tour from measure theory to random processes, via conditioning*. Second. Vol. 35. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2012, pp. xx+279. ISBN: 978-1-107-60655-5. DOI: 10.1017/CBO9781139135351. URL: https://doi.org/10.1017/CBO9781139135351.

[13] Chao-Wei Chou and Wen-Jang Huang. "Characterizations of the gamma distribution via conditional moments". In: *Sankhyā* 65.2 (2003), pp. 271–283. ISSN: 0972-7671.

[14] Fan Chung and Ron Graham. "Edge flipping in graphs". In: *Adv. in Appl. Math.* 48.1 (2012), pp. 37–63. ISSN: 0196-8858. DOI: 10.1016/j.aam.2011.06.002. URL: https://doi.org/10.1016/j.aam.2011.06.002.

[15] Alain Comtet, Cécile Monthus, and Marc Yor. "Exponential functionals of Brownian motion and disordered systems". In: *Journal of applied probability* 35.2 (1998), pp. 255–271.

[16] Louis Comtet. *Advanced combinatorics*. enlarged. The art of finite and infinite expansions. D. Reidel Publishing Co., Dordrecht, 1974, pp. xi+343. ISBN: 90-277-0441-4.

[17] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I*. Second. Probability and its Applications (New York). Elementary theory and methods. Springer-Verlag, New York, 2003, pp. xxii+469. ISBN: 0-387-95541-0.

[18] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II*. Second. Probability and its Applications (New York). General theory and structure. Springer, New York, 2008, pp. xviii+573. ISBN: 978-0-387-21337-8. DOI: 10.1007/978-0-387-49835-5. URL: https://doi.org/10.1007/978-0-387-49835-5.

[19] D. van Dantzig. "Sur la méthode des fonctions génératrices". In: *Le Calcul des Probabilités et ses Applications. Colloques International du Centre National de la Recherche Scientifique, no. 13.* Centre National de la Recherche Scientifique, Paris, 1949, pp. 29–45.

[20] Gaston Darboux. "Mémoire sur les fonctions discontinues". In: vol. 4. 1875, pp. 57–112. URL: http://www.numdam.org/item?id=ASENS_1875_2_4__57_0.

[21] Michael A. B. Deakin. "The development of the Laplace transform, 1737–1937. I. Euler to Spitzer, 1737–1880". In: *Arch. Hist. Exact Sci.* 25.4 (1981), pp. 343–390. ISSN: 0003-9519. DOI: 10.1007/BF01395660. URL: https://doi.org/10.1007/BF01395660.

[22] Persi Diaconis and Laurent Saloff-Coste. "Convolution powers of complex functions on $\mathbb{Z}$". In: *Math. Nachr.* 287.10 (2014), pp. 1106–1130. ISSN: 0025-584X. DOI: 10.1002/mana.201200163. URL: https://doi.org/10.1002/mana.201200163.

[23] Meyer Dwass. "Extremal processes". In: *Ann. Math. Statist.* 35 (1964), pp. 1718–1725. ISSN: 0003-4851. DOI: 10.1214/aoms/1177700394. URL: https://doi.org/10.1214/aoms/1177700394.

[24] Meyer Dwass. "Extremal processes. II". In: *Illinois J. Math.* 10 (1966), pp. 381–391. ISSN: 0019-2082. URL: http://projecteuclid.org/euclid.ijm/1256054990.

[25] Meyer Dwass. "Extremal processes. III". In: *Bull. Inst. Math. Acad. Sinica* 2 (1974). Collection of articles in celebration of the sixtieth birthday of Ky Fan, pp. 255–265. ISSN: 0304-9825.

[26] Richard A. Epstein. *The theory of gambling and statistical logic.* Revised. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1977, pp. xv+450.

[27] Leonhard Euler and Christian Goldbach. "Leonhard Euler und Christian Goldbach. Briefwechsel 1729–1764". In: *Abh. Deutsch. Akad. Wiss. Berlin Kl. Philos., Gesch., Staats-, Rechts-, Wirtschaftswiss* 1965.1 (1965), pp. ix+420.

[28] W. J. Ewens. "The sampling theory of selectively neutral alleles". In: *Theoret. Population Biol.* 3 (1972). ISSN: 0040-5809. DOI: 10.1016/0040-5809(72)90035-4. URL: https://doi.org/10.1016/0040-5809(72)90035-4.

[29] William Feller. *An introduction to probability theory and its applications. Vol. I.* 2nd ed. John Wiley, Sons, Inc., New York; Chapman, and Hall, Ltd., London, 1957, pp. xv+461.

[30] William Feller. "The fundamental limit theorems in probability". In: *Bull. Amer. Math. Soc.* 51 (1945), pp. 800–832. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1945-08448-1. URL: https://doi.org/10.1090/S0002-9904-1945-08448-1.

[31] Bruno de Finetti. *Theory of probability*. Wiley Series in Probability and Statistics. A critical introductory treatment, Translated from the 1970 Italian original by Antonio Machıé and Adrian Smith and with a preface by Smith, Combined edition of Vol. 1, 1974 and Vol. 2, 1975. John Wiley & Sons, Ltd., Chichester, 2017, pp. xiii+582. ISBN: 978-1-119-28637-0. DOI: 10.1002/9781119286387. URL: https://doi.org/10.1002/9781119286387.

[32] Mark Finkelstein and Robert Whitley. "Fibonacci Numbers in Coin Tossing Sequences". In: *Fibonacci Quart.* 16.6 (1978), pp. 539–541.

[33] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009, pp. xiv+810. ISBN: 978-0-521-89806-5. DOI: 10.1017/CBO9780511801655. URL: https://doi.org/10.1017/CBO9780511801655.

[34] Rigoberto Flórez, Leandro Junes, and José L. Ramıérez. "Further results on paths in an n-dimensional cubic lattice". In: *J. Integer Seq.* 21.1 (2018), Art. 18.1.2, 27.

[35] Ralph Fröberg. "Determination of a class of Poincaré series". In: *Math. Scand.* 37.1 (1975), pp. 29–39. ISSN: 0025-5521. DOI: 10.7146/math.scand.a-11585. URL: https://doi.org/10.7146/math.scand.a-11585.

[36] Ira M. Gessel. *An application of the Goulden-Jackson cluster theorem*. 2020. DOI: 10.48550/ARXIV.2011.03171. URL: https://arxiv.org/abs/2011.03171.

[37] Alexander V. Gnedin. "Corners and records of the Poisson process in quadrant". In: *Electron. Commun. Probab.* 13 (2008), pp. 187–193. ISSN: 1083-589X. DOI: 10.1214/ECP.v13-1351. URL: https://doi.org/10.1214/ECP.v13-1351.

[38] Alexander V. Gnedin and Jim Pitman. "Self-similar and Markov composition structures". In: *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 326.Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 13 (2005), pp. 59–84, 280–281. ISSN: 0373-2703. DOI: 10.1007/s10958-007-0447-0. URL: https://doi-org.libproxy.berkeley.edu/10.1007/s10958-007-0447-0.

[39] Charles M. Goldie and L. C. G. Rogers. "The k-record processes are i.i.d". In: *Z. Wahrsch. Verw. Gebiete* 67.2 (1984), pp. 197–211. ISSN: 0044-3719. DOI: 10.1007/BF00535268. URL: https://doi.org/10.1007/BF00535268.

[40] Ian P. Goulden and David M. Jackson. "An inversion theorem for cluster decompositions of sequences with distinguished subsequences". In: *J. London Math. Soc. (2)* 20.3 (1979), pp. 567–576. ISSN: 0024-6107. DOI: 10.1112/jlms/s2-20.3.567. URL: https://doi.org/10.1112/jlms/s2-20.3.567.

[41] Ian P. Goulden and David M. Jackson. *Combinatorial enumeration*. With a foreword by Gian-Carlo Rota, Reprint of the 1983 original. Dover Publications, Inc., Mineola, NY, 2004, pp. xxvi+569. ISBN: 0-486-43597-0.

[42] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics*. A foundation for computer science. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1989, pp. xiv+625. ISBN: 0-201-14236-8.

[43] T. N. E. Greville. "On a problem of E. L. De Forest in iterated smoothing". In: *SIAM J. Math. Anal.* 5 (1974), pp. 376–398. ISSN: 0036-1410. DOI: 10.1137/0505041. URL: https://doi.org/10.1137/0505041.

[44] T. N. E. Greville. "On stability of linear smoothing formulas". In: *SIAM J. Numer. Anal.* 3.1 (1966), pp. 157–170. ISSN: 0036-1429. DOI: 10.1137/0703011. URL: https://doi.org/10.1137/0703011.

[45] Piet Groeneboom. "The concave majorant of Brownian motion". In: *Ann. Probab.* 11.4 (1983), pp. 1016–1027. ISSN: 0091-1798. URL: http://links.jstor.org/sici?sici=0091-1798(198311)11:4%3C1016:TCMOBM%3E2.0.CO;2-7&origin=MSN.

[46] Shane G Henderson and Peter W Glynn. "Regenerative steady-state simulation of discrete-event systems". In: *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 11.4 (2001), pp. 313–345.

[47] Alexander E. Holroyd. "One-dependent coloring by finitary factors". In: vol. 53. 2. 2017, pp. 753–765. DOI: 10.1214/15-AIHP735. URL: https://doi.org/10.1214/15-AIHP735.

[48] Alexander E. Holroyd and Thomas M. Liggett. "Finitely dependent coloring". In: *Forum Math. Pi* 4 (2016), e9, 43. DOI: 10.1017/fmp.2016.7. URL: https://doi.org/10.1017/fmp.2016.7.

[49] Ross Honsberger. "A second look at the Fibonacci and Lucas numbers". In: *Mathematical gems III. Washington, DC: Math Assoc Amer* (1985).

[50] Bamdad Hosseini. "Two Metropolis-Hastings algorithms for posterior measures with non-Gaussian priors in infinite dimensions". In: *SIAM/ASA J. Uncertain. Quantif.* 7.4 (2019), pp. 1185–1223. DOI: 10.1137/18M1183017. URL: https://doi.org/10.1137/18M1183017.

[51] Jeffrey J. Hunter. *Mathematical techniques of applied probability. Vol. 2.* Operations Research and Industrial Engineering. Discrete time models: techniques and applications. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1983, pp. xii+286. ISBN: 0-12-361802-9.

[52] Hsien-Kuei Hwang, Hua-Huai Chern, and Guan-Huei Duh. "An asymptotic distribution theory for Eulerian recurrences with applications". In: *Adv. in Appl. Math.* 112 (2020), pp. 101960, 125. ISSN: 0196-8858. DOI: 10.1016/j.aam.2019.101960. URL: https://doi.org/10.1016/j.aam.2019.101960.

[53] I. A. Ibragimov and Yu. V. Linnik. *Independent and stationary sequences of random variables.* With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman. Wolters-Noordhoff Publishing, Groningen, 1971, p. 443.

[54] Zvetan G. Ignatov. "Point processes generated by order statistics and their applications". In: *Point processes and queuing problems (Colloq., Keszthely, 1978)*. Vol. 24. Colloq. Math. Soc. János Bolyai. North-Holland, Amsterdam-New York, 1981, pp. 109–116.

[55] Aubrey William Ingleton. "The rank of circulant matrices". In: *J. London Math. Soc.* 31 (1956), pp. 632–635. ISSN: 0024-6107. DOI: 10.1112/jlms/s1-31.4.445. URL: https://doi.org/10.1112/jlms/s1-31.4.445.

[56] Monique Jeanblanc, Jim Pitman, and Marc Yor. "Self-similar processes with independent increments associated with Lévy and Bessel". In: *Stochastic Process. Appl.* 100 (2002), pp. 223–231. ISSN: 0304-4149. DOI: 10.1016/S0304-4149(02)00098-4. URL: https://doi.org/10.1016/S0304-4149(02)00098-4.

[57] Zbigniew J Jurek. "The classes L m (Q) of probability measures on Banach spaces". In: *Bull. Polish Acad. Sci. Math* 31 (1983), pp. 51–62.

[58] Zbigniew J. Jurek. "Background driving distribution functions and series representations for log-gamma selfdecomposable random variables". In: *Teor. Veroyatn. Primen.* 67.1 (2022), pp. 134–149. ISSN: 0040-361X. DOI: 10.4213/tvp5422. URL: https://doi.org/10.4213/tvp5422.

[59] Zbigniew J. Jurek. "Limit distributions and one-parameter groups of linear operators on Banach spaces". In: *J. Multivariate Anal.* 13.4 (1983), pp. 578–604. ISSN: 0047-259X. DOI: 10.1016/0047-259X(83)90042-8. URL: https://doi.org/10.1016/0047-259X(83)90042-8.

[60] Zbigniew J. Jurek. "Selfdecomposability: an exception or a rule?" In: *Ann. Univ. Mariae Curie-Skłodowska Sect. A* 51.1 (1997), pp. 93–107. ISSN: 0365-1029.

[61] Zbigniew J. Jurek and Wim Vervaat. "An integral representation for self-decomposable Banach space valued random variables". In: *Z. Wahrsch. Verw. Gebiete* 62.2 (1983), pp. 247–262. ISSN: 0044-3719. DOI: 10.1007/BF00538800. URL: https://doi.org/10.1007/BF00538800.

[62] Zbigniew J. Jurek and Marc Yor. "Selfdecomposable laws associated with hyperbolic functions". In: *Probab. Math. Statist.* 24.1, Acta Univ. Wratislav. No. 2646 (2004), pp. 181–190. ISSN: 0208-4147.

[63] Olav Kallenberg. *Foundations of modern probability.* Vol. 99. Probability Theory and Stochastic Modelling. Third edition. Springer, Cham, 2021, pp. xii+946. ISBN: 978-3-030-61871-1. DOI: 10.1007/978-3-030-61871-1. URL: https://doi.org/10.1007/978-3-030-61871-1.

[64] Udo Kamps. "Chebyshev polynomials and least squares estimation based on one-dependent random variables". In: *Linear Algebra Appl.* 112 (1989), pp. 217–230. ISSN: 0024-3795. DOI: 10.1016/0024-3795(89)90597-1. URL: https://doi.org/10.1016/0024-3795(89)90597-1.

[65] Aleksandr Yakolevich Khintchine. “Limit laws of sums of independent random variables”. In: *ONTI, Moscow,(Russian)* (1938).

[66] Davar Khoshnevisan. *Multiparameter processes*. Springer Monographs in Mathematics. An introduction to random fields. Springer-Verlag, New York, 2002, pp. xx+584. ISBN: 0-387-95459-7. DOI: 10.1007/b97363. URL: https://doi.org/10.1007/b97363.

[67] Arunod Kumar and Bertram M Schreiber. “Representation of certain infinitely divisible probability measures on Banach spaces”. In: *Journal of Multivariate Analysis* 9.2 (1979), pp. 288–303.

[68] John Lamperti. “On extreme order statistics”. In: *Ann. Math. Statist.* 35 (1964), pp. 1726–1737. ISSN: 0003-4851. DOI: 10.1214/aoms/1177700395. URL: https://doi.org/10.1214/aoms/1177700395.

[69] John Lamperti. “Semi-stable stochastic processes”. In: *Trans. Amer. Math. Soc.* 104 (1962), pp. 62–78. ISSN: 0002-9947. DOI: 10.2307/1993933. URL: https://doi.org/10.2307/1993933.

[70] Pierre-Simon Laplace. *Théorie analytique des probabilités. Vol. I.* Introduction: Essai philosophique sur les probabilités. [Introduction: Philosophical essay on probabilities], Livre I: Du calcul des fonctions génératrices. [Book I: On the calculus of generating functions], Reprint of the 1819 fourth edition (Introduction) and the 1820 third edition (Book I). Éditions Jacques Gabay, Paris, 1995, pp. clxx+194. ISBN: 2-87647-161-2.

[71] Paul Lévy. *Théorie de l’addition des variables aléatoires*. Paris: Gauthier-Villars, 1937.

[72] Paul Lévy. “Wiener’s random function, and other Laplacian random functions”. In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*. University of California Press, Berkeley and Los Angeles, Calif., 1951, pp. 171–187.

[73] Michel Loeve. “Elementary probability theory”. In: *Probability Theory I*. Springer, 1977, pp. 1–52.

[74] Eugene Lukacs. “A characterization of the gamma distribution”. In: *Ann. Math. Statist.* 26 (1955), pp. 319–324. ISSN: 0003-4851. DOI: 10.1214/aoms/1177728549. URL: https://doi.org/10.1214/aoms/1177728549.

[75] Ian G. Macdonald. *Symmetric functions and Hall polynomials*. Second. Oxford Mathematical Monographs. With contributions by A. Zelevinsky, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1995, pp. x+475. ISBN: 0-19-853489-2.

[76] Percy Alexander MacMahon. *Combinatory analysis. Vol. I, II (bound in one volume)*. Dover Phoenix Editions. Reprint of ıt An introduction to combinatory analysis (1920) and ıt Combinatory analysis. Vol. I, II (1915, 1916). Dover Publications, Inc., Mineola, NY, 2004, pp. ii+761. ISBN: 0-486-49586-8.

[77] Donatella Merlini et al. "On some alternative characterizations of Riordan arrays". In: *Canad. J. Math.* 49.2 (1997), pp. 301–320. ISSN: 0008-414X. DOI: 10.4153/CJM-1997-015-x. URL: https://doi.org/10.4153/CJM-1997-015-x.

[78] Sri Gopal Mohanty. "An r-coin tossing game and the associated partition of generalized Fibonacci numbers". In: *Sankhyā Ser. A* 29 (1967), pp. 207–214. ISSN: 0581-572X.

[79] Joseph Najnudel and Jim Pitman. "Feller coupling of cycles of permutations and Poisson spacings in inhomogeneous Bernoulli trials". In: *Electron. Commun. Probab.* 25 (2020), Paper No. 73, 11. DOI: 10.1214/20-ecp352. URL: https://doi-org.libproxy.berkeley.edu/10.1214/20-ecp352.

[80] Thu Nguyen. "Multiply self-decomposable probability measures on Banach spaces". In: *Studia Mathematica* 66 (1979), pp. 161–175.

[81] Mehdi Ouaki and Jim Pitman. *Markovian structure in the concave majorant of Brownian motion*. 2021. DOI: 10.48550/ARXIV.2105.11042. URL: https://arxiv.org/abs/2105.11042.

[82] Mihael Perman, Jim Pitman, and Marc Yor. "Size-biased sampling of Poisson point processes and excursions". In: *Probab. Theory Related Fields* 92.1 (1992), pp. 21–39. ISSN: 0178-8051. DOI: 10.1007/BF01205234. URL: https://doi.org/10.1007/BF01205234.

[83] Jim Pitman. "Random weighted averages, partition structures and generalized arcsine laws". In: *arXiv preprint arXiv:1804.07896* (2018).

[84] Jim Pitman and Nathan Ross. "The greatest convex minorant of Brownian motion, meander, and bridge". In: *Probab. Theory Related Fields* 153.3-4 (2012), pp. 771–807. ISSN: 0178-8051. DOI: 10.1007/s00440-011-0385-0. URL: https://doi.org/10.1007/s00440-011-0385-0.

[85] Jim Pitman and Marc Yor. "Arcsine laws and interval partitions derived from a stable subordinator". In: *Proc. London Math. Soc. (3)* 65.2 (1992), pp. 326–356. ISSN: 0024-6115. DOI: 10.1112/plms/s3-65.2.326. URL: https://doi.org/10.1112/plms/s3-65.2.326.

[86] Jim Pitman and Marc Yor. "Infinitely divisible laws associated with hyperbolic functions". In: *Canadian Journal of Mathematics* 55.2 (2003), pp. 292–330.

[87] Jim Pitman and Marc Yor. "Random discrete distributions derived from self-similar random sets". In: *Electron. J. Probab.* 1 (1996), no. 4, approx. 28 pp. ISSN: 1083-6489. DOI: 10.1214/EJP.v1-4. URL: https://doi-org.libproxy.berkeley.edu/10.1214/EJP.v1-4.

[88] Jim Pitman and Zhiyi You. *Stationary 1-dependent Counting Processes: from Runs to Bivariate Generating Functions*. 2021. DOI: 10.48550/ARXIV.2105.08255. URL: https://arxiv.org/abs/2105.08255.

[89] Jim Pitman and Zhiyi You. *The range of a self-similar additive gamma process is a scale invariant Poisson point process*. 2021. DOI: 10.48550/ARXIV.2111.09409. URL: https://arxiv.org/abs/2111.09409.

[90] Alexander Polishchuk and Leonid Positselski. *Quadratic algebras*. Vol. 37. University Lecture Series. American Mathematical Society, Providence, RI, 2005, pp. xii+159. ISBN: 0-8218-3834-2. DOI: 10.1090/ulect/037. URL: https://doi.org/10.1090/ulect/037.

[91] Sidney I. Resnick and Michael Rubinovitch. "The structure of extremal processes". In: *Advances in Appl. Probability* 5 (1973), pp. 287–307. ISSN: 0001-8678. DOI: 10.2307/1426037. URL: https://doi.org/10.2307/1426037.

[92] Douglas G. Rogers. "Pascal triangles, Catalan numbers and renewal arrays". In: *Discrete Math.* 22.3 (1978), pp. 301–310. ISSN: 0012-365X. DOI: 10.1016/0012-365X(78)90063-8. URL: https://doi.org/10.1016/0012-365X(78)90063-8.

[93] Bernard Roynette and Marc Yor. "Couples de Wald indéfiniment divisibles. Exemples liés à la fonction gamma d'Euler et à la fonction zeta de Riemann". In: *Annales de l'institut Fourier*. Vol. 55. 4. Chartres: L'Institut, 1950-. 2005, pp. 1219–1284.

[94] Ken-iti Sato. "Class L of multivariate distributions and its subclasses". In: *J. Multivariate Anal.* 10.2 (1980), pp. 207–232. ISSN: 0047-259X. DOI: 10.1016/0047-259X(80)90014-7. URL: https://doi.org/10.1016/0047-259X(80)90014-7.

[95] Ken-iti Sato. *Lévy processes and infinitely divisible distributions*. Vol. 68. Cambridge Studies in Advanced Mathematics. Translated from the 1990 Japanese original, Revised by the author. Cambridge University Press, Cambridge, 1999, pp. xii+486. ISBN: 0-521-55302-4.

[96] Ken-iti Sato. "Self-similar processes with independent increments". In: *Probab. Theory Related Fields* 89.3 (1991), pp. 285–300. ISSN: 0178-8051. DOI: 10.1007/BF01198788. URL: https://doi.org/10.1007/BF01198788.

[97] Ken-iti Sato and Makoto Yamazato. "On distribution functions of class L". In: *Z. Wahrsch. Verw. Gebiete* 43.4 (1978), pp. 273–308. ISSN: 0178-8051. DOI: 10.1007/BF00534763. URL: https://doi.org/10.1007/BF00534763.

[98] Richard W. Shorrock. "On discrete time extremal processes". In: *Advances in Appl. Probability* 6 (1974), pp. 580–592. ISSN: 0001-8678. DOI: 10.2307/1426235. URL: https://doi.org/10.2307/1426235.

[99] Karl Sigman. "One-dependent regenerative processes and queues in continuous time". In: *Math. Oper. Res.* 15.1 (1990), pp. 175–189. ISSN: 0364-765X. DOI: 10.1287/moor.15.1.175. URL: https://doi.org/10.1287/moor.15.1.175.

[100] Renzo Sprugnoli. "Riordan arrays and combinatorial sums". In: *Discrete Mathematics* 132.1-3 (1994), pp. 267–290.

[101] Frederik Willem Steutel and Klaas van Harn. "Discrete analogues of self-decomposability and stability". In: *Ann. Probab.* 7.5 (1979), pp. 893–899. ISSN: 0091-1798. URL: `http://links.jstor.org/sici?sici=0091-1798(197910)7:5%3C893:DAOSAS%3E2.0.CO;2-F&origin=MSN`.

[102] Joseph Talacko. "Perks' distributions and their role in the theory of Wiener's stochastic variables". In: *Trabajos de estadıéstica* 7.2 (1956), pp. 159–174.

[103] Mathias Trabs. "Calibration of self-decomposable Lévy models". In: *Bernoulli* 20.1 (2014), pp. 109–140. ISSN: 1350-7265. DOI: `10.3150/12-BEJ478`. URL: `https://doi.org/10.3150/12-BEJ478`.

[104] Ciprian A. Tudor. *Analysis of variations for self-similar processes: A stochastic calculus approach*. Probability and its Applications (New York). Springer, Cham, 2013, pp. xii+268. ISBN: 978-3-319-00935-3. DOI: `10.1007/978-3-319-00936-0`. URL: `https://doi.org/10.1007/978-3-319-00936-0`.

[105] Kazimierz Urbanik. "Limit laws for sequences of normed sums satisfying some stability conditions". In: *Multivariate Analysis–III*. Elsevier, 1973, pp. 225–237.

[106] Wim Vervaat. "On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables". In: *Adv. in Appl. Probab.* 11.4 (1979), pp. 750–783. ISSN: 0001-8678. DOI: `10.2307/1426858`. URL: `https://doi.org/10.2307/1426858`.

[107] Stephen James Wolfe. "On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$". In: *Stochastic Process. Appl.* 12.3 (1982), pp. 301–312. ISSN: 0304-4149. DOI: `10.1016/0304-4149(82)90050-3`. URL: `https://doi.org/10.1016/0304-4149(82)90050-3`.

[108] Vladimir Mikhailovich Zolotarev. "Mellin-Stieltjes transforms in probability theory". In: *Theory of Probability & Its Applications* 2.4 (1957), pp. 433–460.