Probability and Its Applications

A series published in association with the Applied Probability Trust

Editors: J. Gani, C.C. Heyde, T.G. Kurtz

Olav Kallenberg

Probabilistic Symmetries and Invariance Principles

Springer

Olav Kallenberg
Department of Mathematics and Statistics
Auburn University
Auburn, AL 36849-5310
USA
kalleoh@auburn.edu

Mathematics Subject Classification (2000): 60G09, 60-02

Library of Congress Control Number: 2005923579

ISBN-10: 0-387-25115-4 Printed on acid-free paper.
ISBN-13: 978-0387-25115-8

Printed in the United States of America. (MP)

9 8 7 6 5 4 3 2 1

springeronline.com

Preface

This book is about random objects—sequences, processes, arrays, measures, functionals—with interesting symmetry properties. Here symmetry should be understood in the broad sense of invariance under a family (not necessarily a group) of measurable transformations. To be precise, it is not the random objects themselves but rather their distributions that are assumed to be symmetric.

Though many probabilistic symmetries are conceivable and have been considered in various contexts, four of them—*stationarity, contractability, exchangeability,* and *rotatability*—stand out as especially interesting and important in several ways: Their study leads to some deep structural theorems of great beauty and significance, they are intimately related to some basic areas of modern probability theory, and they are mutually connected through a variety of basic relationships. The mentioned symmetries may be defined as invariance in distribution under shifts, contractions, permutations, and rotations. Stationarity being a familiar classical topic, treated extensively in many standard textbooks and monographs, most of our attention will be focused on the remaining three basic symmetries.

The study of general probabilistic symmetries essentially originated with the work of de Finetti (1929–30), who proved by elementary means (no advanced tools being yet available) the celebrated theorem named after him—the fact that every infinite sequence of exchangeable events is mixed i.i.d. Though the statistical and philosophical implications were thoroughly explored by de Finetti himself and his followers, the result was long regarded by probabilists as an isolated curiosity. (The attitude still prevails among the ignorant!) The mathematical study of exchangeability and related topics was not revived until the work of Hewitt and Savage (1955), Ryll-Nardzewski (1957), Bühlmann (1960), and Freedman (1962–63). The area eventually became fashionable, owing to some remarkable discoveries of Aldous (1977, 1981), Hoover (1979), Kingman (1978), and others, which led to a vigorous further development. By the time of the 1981 Rome conference on exchangeability, honoring de Finetti on his 75th birthday, the accumulated knowledge was already so vast and the literature so scattered that a quick overview was getting difficult. Though the situation was partially remedied by a brilliant review paper of Aldous (1985), a serious student of the subject would still need to consult the original research papers.

To me personally, probabilistic symmetries have been a major interest—even an obsession—throughout my career, and at least half of my publications during the past thirty years have been in this area. More than twenty years ago, I began (foolishly) to announce my intention of writing a monograph on the subject. What held me back was the constant discovery of new results and basic connections that needed to be explored before the subject could be said to have reached a critical level of maturity. In the meantime, the relevant literature has kept growing exponentially, and a single volume would now seem totally inadequate to cover the entire area. Regrettably, this situation has forced me to be very selective, and I hope that my readers will forgive me for giving preference to the areas that I know best. Entire theories, often of great significance, are either ignored altogether or deferred to appendices for brief summaries, with their major proofs omitted.

de Finetti's theorem—once regarded as a deep result, requiring 50 pages of proof—can now be established in just a few lines. What remains, then, to fill the pages of a whole book? The answer may be surprising to the novice: The area of distributional symmetries is one of the richest of modern probability theory, exhibiting an abundance of deep, beautiful, and often astounding results. Though some proofs remain long and intricate, even with the best tools currently available, the patient reader will be greatly rewarded. As in any area of mathematics, the interest lies as much in the overall, logical structure of the edifice as in the individual statements of theorems. I have therefore spent countless hours (indeed weeks, months, or years) reworking the proofs to make them as accessible and transparent as possible, in the hope that young researchers will be inspired to continue the work where previous generations have left off. As should be clear from a glance at the bibliography, the current theory has emerged through the collective efforts of countless authors of many nationalities.

The present exposition is divided into nine chapters, each devoted to a major aspect of the theory. The first chapter introduces the basic symmetries of contractability, exchangeability, and rotatability and gives the representations that are accessible by elementary methods. Important martingale connections are introduced in Chapter 2, and martingale methods are used to study the paths of exchangeable processes. The general representation of such processes is derived by weak convergence methods in Chapter 3, which also contains a variety of limit theorems and approximation results. In Chapter 4 we present the predictable sampling theorem and its continuous-time counterpart, and in Chapter 5 we consider the closely related decoupling identities for exchangeable sums and integrals. Exchangeable random sets and the associated excursion theory are considered in Chapter 6. The remaining Chapters 7–9 are devoted to multi-variate symmetries of different kinds. Thus, we deal in Chapter 7 with exchangeable or contractable arrays of arbitrary dimension, in Chapter 8 with separately or jointly rotatable arrays or functionals, and in Chapter 9 with separately or jointly exchangeable random measures on a finite or infinite rectangle.

Special, though related, topics include the L^p-symmetries of Chapter 1, some invariance results for Palm measures in Chapter 2, the sub-sequence principles in Chapter 3, the time-change reductions in Chapter 4, some decompositions of the strong Markov property in Chapter 6, the paint-box representation of symmetric partitions in Chapter 7, and the theory of exchangeable random sheets in Chapter 8. Though most results are previously known, there are many new results (or new versions of old ones) scattered throughout the book.

For motivation and general guidance, we begin with a short introduction, highlighting some main results from the various chapters. Some auxiliary results of a more technical nature are treated in appendices, and the book concludes with some detailed historical notes, along with references to the original papers. A long, though far from exhaustive, bibliography lists publications closely related to the included material. Though most proofs should be accessible to readers with a good knowledge of graduate level probability and real analysis, detailed references are often provided, for the reader's convenience, to my earlier book *Foundations of Modern Probability* (2nd ed.), where FMP $a.b$ refers to Proposition b in Chapter a.

Acknowledgments: My interest in symmetries goes back to my student days in Gothenburg, Sweden, where I had the good fortune of being nurtured by an unusually stimulating research environment. I am especially indebted to Peter Jagers for some crucial influences during the formative year of 1971–72, when I wrote my dissertation.

In later years I have met most of the key players in the development of the subject, many of whom have had a profound influence on my own work. For helpful or stimulating discussions through the years, I would like to mention especially David Aldous, Istvan Berkes, Persi Diaconis, Douglas Hoover, Gail Ivanoff, John Kingman, Jim Pitman, and Paul Ressel. I am also grateful for the interest and encouragement of countless other friends and colleagues, including the late Stamatis Cambanis, Kai Lai Chung, Erhan Çinlar, the late Peter Franken, Cindy Greenwood, Gopi Kallianpur, Klaus Krickeberg, Ross Leadbetter, Ming Liao, the late Klaus Matthes, Balram Rajput, and Hermann Thorisson. (I apologize for any unintentional omission of names that ought to be on my list.)

Special thanks go to Gail Ivanoff and Neville Weber, who allowed me to include the part of Theorem 1.13 conjectured by them. Generous help was offered by Ulrich Albrecht and my son Matti, when my computer broke down at a crucial stage. I am also grateful for some excellent remarks of an anonymous reviewer, and for the truly professional work of John Kimmel and the production team at Springer-Verlag.

My greatest thanks, as always, extend to my eternally youthful and wondrously clever wife Jinsoo, for her constant love and support, without which this book could never have been finished. As a devoted wife, she has gracefully accepted that the most useful mathematical insights are often coming

at times that were meant to be spent with the family. On a personal note, I adore her for being such a wonderful mother to our children, and I owe her special thanks for tolerating my daily piano practice and for never complaining, as I am filling our house with piles and piles of books on every aspect of cultural history.

Olav Kallenberg
January 2005

Contents

Introduction

The hierarchy of distributional symmetries considered in this book, along with the associated classes of transformations, may be summarized as follows:

invariant objects	*transformations*
stationary	shifts
contractable	sub-sequences
exchangeable	permutations
rotatable	isometries

where each invariance property is clearly stronger than the preceding one. All four symmetries may be considered in discrete or continuous time (or space) and in one or several dimensions. There is also the distinction between bounded and unbounded index sets.

The most obvious problem in this area is to characterize the class of objects of a given type with a specified symmetry property. Thus, for example, de Finetti's theorem describes the infinite, exchangeable sequences of random variables as mixed i.i.d. This classical result is rather easy to prove by modern martingale or weak convergence arguments. Other characterizations may be a lot harder. Thus, it takes about 30 pages of tight mathematical reasoning to derive the characterization of contractable arrays of arbitrary dimension, and for the multi-variate rotatable case another 40 pages may be required. In other cases again, no simple representation seems to exist. Thus, for example, stationary sequences are unique mixtures of ergodic ones, but there is no (known) representation of a (strictly) stationary and ergodic sequence in terms of simpler building blocks. The situation for finite, contractable sequences is even worse, since here the integral representation in terms of extreme points is not even unique.

The next step might be to explore the relationship between the various symmetries. For example, Ryll-Nardzewski's theorem shows that every infinite, contractable sequence of random variables is even exchangeable, so that, for infinite sequences, the two symmetries are in fact equivalent. (The equivalence fails for finite sequences.) A higher-dimensional counterpart is the much deeper fact that every contractable array on a tetrahedral index set can be extended (non-uniquely) to an exchangeable array on the corresponding product set. For a connection with stationarity, it is easy to show that an infinite sequence is contractable iff (if and only if) it is strongly stationary, in the sense of invariance in distribution under optional shifts. Let us finally mention the fundamental and nontrivial fact that every continuous and contractable (in the sense of the increments) process on $\mathbb{R}_+ = [0,\infty)$ with zero drift is also rotatable.

Connections can also take the form of limit theorems or approximation properties. Thus, for example, we may approximate a finite exchangeable sequence by infinite ones. This, of course, is nothing else than the familiar asymptotic equivalence between sampling with or without replacement from a finite population, explored in scores of statistical papers. Similar approximation theorems can be established in continuous time, where modern weak convergence and coupling methods play a prominent role.

Our investigation may continue with a more detailed study of the various symmetric random objects. For example, though on $\mathbb{R}_+$ the exchangeable (increment) processes are just mixtures of the familiar Lévy processes—this is the continuous-time counterpart of de Finetti's theorem, first noted by Bühlmann—on $[0,1]$ one gets a much broader class of exchangeable processes, and it becomes interesting to explore the path and other properties of the latter. It may then be natural to relate the various symmetries to a filtration of σ-fields and to employ the powerful machinery of modern martingale theory and stochastic calculus.

This new dynamical approach has led to some startling new discoveries, opening up entirely new domains of study. We have already mentioned the elementary connection between exchangeability and strong stationarity. Further instances are given by the wide range of characterizations involving direct or reverse martingales in discrete or continuous time. Less obvious are the predictable sampling and mapping theorems, where the defining properties of contractable or exchangeable sequences and processes are extended to suitably predictable random transformations. Apart from their intrinsic interest, those results also serve as valuable general tools, providing short and streamlined proofs of the arcsine and related theorems for random walks and Lévy processes. A connected, indeed even stronger, class of theorems are the decoupling identities for sums and integrals with respect to exchangeable processes, discussed in further detail below.

All the mentioned problems continue to make sense in higher dimensions, and the last third of the book deals with multi-variate symmetries of various kind. As already noted, already the basic characterization problems here become surprisingly hard, and the picture is still incomplete.

— — —

We turn to a more detailed summary of the contents of the book, introducing at the same time some crucial definitions and notation. Let us first define the basic symmetries, as they appear already in Chapter 1. Given an infinite sequence of random elements $\xi = (\xi_1, \xi_2, \ldots)$, we say that ξ is *contractable* (sometimes even called *spreading invariant* or *spreadable*) if every sub-sequence has the same distribution, so that

$$(\xi_{k_1}, \xi_{k_2}, \ldots) \stackrel{d}{=} (\xi_1, \xi_2, \ldots) \qquad (1)$$

for any positive integers $k_1 < k_2 < \cdots$, where $\stackrel{d}{=}$ denotes equality in distribution. It is clearly enough to consider sub-sequences obtained by omitting a single element. (Our term comes from the fact that we may form a sub-sequence by omitting some elements and then *contracting* the resulting sequence to fill in the resulting gaps. This clearly corresponds to a *spreading* of the associated index set.) A finite sequence $\xi_1, \ldots, \xi_n$ is said to be contractable if all sub-sequences of equal length have the same distribution.

For the stronger property of *exchangeability*, we require (1) to hold for any distinct (not necessarily increasing) elements $k_1, k_2, \ldots$ of the index set. Thus, for infinite sequences, the k_n are required to form an injective (but not necessarily surjective) transformation on $\mathbb{N} = \{1, 2, \ldots\}$. However, it is clearly enough to require that (1) be fulfilled for any *finite permutation* of $\mathbb{N}$, defined as a bijective transformation $n \mapsto k_n$ such that $k_n = n$ for all but finitely many n. (Indeed, it suffices to consider transpositions of pairs of adjacent elements.) For technical purposes, it is useful to note that the finite permutations form a countable group of transformations of the index set. (By contrast, the classes of shifts or sub-sequences of $\mathbb{N}$ are only semigroups, which leads occasionally to some technical problems.)

Finally, when the ξ_k are real-valued (or take values in a suitable linear space), we may define *rotatability* of ξ by requiring that every finite sub-sequence be invariant in distribution under arbitrary orthogonal transformations. Thus, for every n, the distribution μ_n of $(\xi_1, \ldots, \xi_n)$ is assumed to be spherically symmetric on $\mathbb{R}^n$. It is then clear, at least intuitively, that each μ_n is a mixture of uniform distributions over concentric spherical shells around the origin. The situation for infinite sequences may be less obvious. Since permutations are special cases of rotations, we note that every rotatable sequence is exchangeable. Similarly, by the injectivity of contractions, every exchangeable sequence is clearly contractable. Finally, shifts are special contractions, so every contractable sequence is stationary.

In continuous time, we may define the corresponding symmetries in terms of the increments. The initial value playing no role, we may then restrict our attention to processes starting at 0, and by suitable scaling and shifting we may assume that the index set I is either $[0, 1]$ or $\mathbb{R}_+$. Considering any disjoint intervals $I_1, I_2, \ldots \subset I$ of equal length, listed in the order from left to right, we may say that a process X on I is contractable, exchangeable, or rotatable if the increments of X over $I_1, I_2, \ldots$ have the corresponding property (for all choices of intervals). A problem with this definition is that the underlying transformations are only applied to the increments of X (rather than to X itself).

A preferable approach would be to base the definition on suitable pathwise transformations. Thus, for any points $a < b$ in I, we may form the *contraction* $C_{a,b}X$ by deleting the path on (a, b) and attaching the continued path starting at (b, X_b) to the loose end at (a, X_a), employing a suitable parallel displacement in space and time. By a similar construction, for any three points $a < b < c$ in I, we may form the *transposition* $T_{a,b,c}X$ by swapping the paths over the intervals $[a, b]$ and $[b, c]$. Then we say that X is *contractable* if $C_{a,b}X \overset{d}{=} X$ for all $a < b$ and *exchangeable* if $T_{a,b,c}X \overset{d}{=} X$ for all $a < b < c$. For rotations, a Hilbert space approach often seems more appropriate, as explained below.

— — —

Chapter 1 begins with a proof of de Finetti's theorem. In its original form, the theorem says that every infinite, exchangeable sequence of random variables $\xi = (\xi_1, \xi_2, \ldots)$ is *mixed i.i.d.* In other words, there exists a probability measure ν on the set of distributions m on $\mathbb{R}$ such that

$$\mathcal{L}(\xi) \equiv P\{\xi \in \cdot\} = \int m^\infty \nu(dm). \tag{2}$$

(Here m^∞ denotes the distribution of an i.i.d. sequence based on the measure m.) The condition is clearly even sufficient, so it characterizes the class of exchangeable sequences.

A more sophisticated version of the theorem is in terms of conditioning. Thus, an infinite sequence ξ as above is exchangeable iff it is *conditionally i.i.d.* In other words, there exists a random probability measure ν on $\mathbb{R}$ such that, conditionally on ν, the ξ_k are i.i.d. with the common distribution ν. This we can write conveniently as

$$P[\xi \in \cdot|\nu] = \nu^\infty \text{ a.s.} \tag{3}$$

Even this characterization is superseded by the stronger statement of Ryll-Nardzewski, the fact that every infinite, contractable sequence of random variables is conditionally i.i.d. Thus, for infinite sequences of random variables, the four stated properties—contractable, exchangeable, mixed i.i.d., and conditionally i.i.d.—are all equivalent. This is the modern statement of de Finetti's theorem, proved in Section 1.1.

de Finetti's theorem suggests a corresponding result in continuous time, characterizing exchangeable processes on $\mathbb{R}_+$ as mixtures of Lévy processes. The proposed statement, first noted by Bühlmann, requires some qualifications, owing to some technical difficulties associated with the uncountable nature of the index set. One way to make the claim precise is to consider only exchangeable processes defined on the rationals, and assert that any such process X can be extended a.s. to a mixture of Lévy processes. Another way is to require X to be continuous in probability and claim that X has then a version with the stated property. A third option is to assume from the outset that X is right-continuous with left-hand limits (rcll), which ensures the validity of the original claim. Whatever the approach, the conclusion may be stated in either a mixing form, in the format of (2), or a conditional form, akin to (3). We finally note that, by Ryll-Nardzewski's theorem, even Bühlmann's result remains true in a stronger contractable version.

The results of de Finetti and Bühlmann are no longer true for bounded index sets. The obvious counterpart in discrete time is the fact that finite exchangeable sequences are mixtures of so-called *urn sequences*, obtained by sampling without replacement from a finite collection (that may be represented by tickets in an urn). The general result in continuous time is much harder and will be discussed later. However, the special case of random measures is accessible by elementary means, and we may prove that a random measure ξ on $[0,1]$ is exchangeable iff it can be represented in the form

$$\xi = \alpha\lambda + \sum\nolimits_k \beta_k \delta_{\tau_k} \text{ a.s.,} \tag{4}$$

where α and $\beta_1, \beta_2, \ldots$ are non-negative random variables and $\tau_1, \tau_2, \ldots$ is an independent sequence of i.i.d. $U(0,1)$ random variables, $U(0,1)$ being the uniform distribution on $[0,1]$. Here λ denotes Lebesgue measure on $[0,1]$, and δ_t is the measure assigning a unit mass to the point t.

In particular, we see that a simple point process ξ on $[0,1]$ is exchangeable iff it is a *mixed binomial process* of the form $\xi = \sum_{k \le \kappa} \delta_{\tau_k}$, where $\tau_1, \tau_2, \ldots$ are i.i.d. $U(0,1)$ and κ is an independent random variable. This may be contrasted with the infinite-interval case, where ξ is mixed Poisson by Bühlmann's theorem. Those rather elementary results, of considerable importance in their own right, are interesting also because of their connections with classical analysis, as they turn out to be essentially equivalent to Bernstein's characterizations of completely monotone functions. Likewise, de Finetti's original theorem for exchangeable events is equivalent to Hausdorff's celebrated moment representation of completely monotone sequences.

We turn to the rotatable case. Here Freedman's theorem states that an infinite sequence of random variables $\xi = (\xi_1, \xi_2, \ldots)$ is rotatable iff it is mixed i.i.d. centered Gaussian. In its conditional form the condition says that, given a suitable random variable $\sigma \ge 0$, the ξ_k are conditionally i.i.d. $N(0, \sigma^2)$. The latter description corresponds to the a.s. representation $\xi_k = \sigma\zeta_k$, $k \in \mathbb{N}$, where $\zeta_1, \zeta_2, \ldots$ are i.i.d. $N(0,1)$ and independent of σ. In either form, the result is equivalent to Schoenberg's theorem in classical analysis—the remarkable fact that a continuous function φ on $\mathbb{R}_+$ with $\varphi(0) = 1$ is completely monotone, hence a Laplace transform, iff for every $n \in \mathbb{N}$ the function $f_n(x) = \varphi(|x|^2) = \varphi(x_1^2 + \cdots + x_n^2)$ on $\mathbb{R}^n$ is non-negative definite, hence a characteristic function.

The rotational invariance may be expressed in the form $\sum_k c_k \xi_k \overset{d}{=} \xi_1$, where $c_1, c_2, \ldots$ are arbitrary constants satisfying $\sum_k c_k^2 = 1$. It is natural to consider the more general case of l^p-*invariance*, where we require instead that

$$\sum\nolimits_k c_k \xi_k \overset{d}{=} \|c\|_p \, \xi_1, \quad c = (c_1, c_2, \ldots) \in l^p.$$

This makes sense for arbitrary $p \in (0,2]$, and the property is equivalent to the a.s. representation $\xi_k = \sigma\zeta_k$, where the ζ_k are now i.i.d. symmetric p-

stable and σ is an independent random variable. This is another classical result, due to Bretagnolle, Dacunha-Castelle, and Krivine.

— — —

The main purpose of Chapter 2 is to introduce and explore the basic martingale connections in discrete and continuous time. Given a finite or infinite random sequence $\xi = (\xi_1, \xi_2, \ldots)$ adapted to a filtration $\mathcal{F} = (\mathcal{F}_n)$, we say that ξ is *$\mathcal{F}$-contractable* or *$\mathcal{F}$-exchangeable* if for every $n \in \mathbb{N}$ the shifted sequence $\theta_n \xi = (\xi_{n+1}, \xi_{n+2}, \ldots)$ is conditionally contractable or exchangeable, respectively, given $\mathcal{F}_n$. When $\mathcal{F}$ is the filtration induced by ξ, we note that the two properties reduce to the corresponding elementary versions. However, the added generality is often useful for applications. For infinite sequences, either property holds iff ξ is *strongly stationary*, in the sense that $\theta_\tau \xi \stackrel{d}{=} \xi$ for every finite optional (or stopping) time τ. It is also equivalent that the *prediction sequence*

$$\mu_n = P[\theta_n \xi \in \cdot | \mathcal{F}_n], \quad n \geq 0,$$

be a measure-valued martingale on $\mathbb{Z}_+ = \{0, 1, \ldots\}$. Similar results hold in continuous time.

Even more striking are perhaps the reverse martingale connections. To state the basic discrete-time version, let $\xi = (\xi_1, \xi_2, \ldots)$ be a finite or infinite random sequence in an arbitrary measurable space, and consider the associated sequence of *empirical distributions*

$$\eta_n = n^{-1} \sum\nolimits_{k \leq n} \delta_{\xi_k}, \quad n \geq 1.$$

Then ξ turns out to be exchangeable iff the η_n form a reverse, measure-valued martingale. In continuous time, we consider any integrable semi-martingale X on $I = [0, 1]$ or $\mathbb{R}_+$ with jump point process ξ and let $[X^c]$ denote the quadratic variation of the continuous component X^c. Then X is exchangeable iff the process

$$Y_t = t^{-1}(X_t,\, \xi[0, t],\, [X^c]_t), \quad t \in I \setminus \{0\},$$

is a reverse martingale.

Returning to the associated forward martingales, we show that any integrable and contractable process X on $\mathbb{Q}_{[0,1]} = \mathbb{Q} \cap [0, 1]$ can be extended to a special semi-martingale on $[0, 1)$ with associated jump point process ξ, such that $[X^c]$ is linear and X and ξ have compensators of the form

$$\hat{X} = M \cdot \lambda, \qquad \hat{\xi} = \lambda \otimes \eta, \tag{5}$$

where M and η denote the martingales

$$M_t = \frac{E[X_1 - X_t | \mathcal{F}_t]}{1 - t}, \qquad \eta_t = \frac{E[\xi_1 - \xi_t | \mathcal{F}_t]}{1 - t}.$$

In particular, this allows us to extend X to a process with rcll paths. We may henceforth add the latter property to the defining characteristics of a contractable or exchangeable process.

By a much deeper analysis, it can be shown that every integrable and exchangeable process X on $[0,1]$ has an a.s. representation

$$X_t = \alpha t + \sigma B_t + \sum\nolimits_k \beta_k (1\{\tau_k \le t\} - t), \quad t \in [0,1], \tag{6}$$

for some i.i.d. $U(0,1)$ random variables τ_k, an independent Brownian bridge B, and an independent collection of random coefficients α, σ, and $\beta_1, \beta_2, \ldots$, where $1\{\cdot\} = 1_{\{\cdot\}}$. This clearly generalizes the representation (4) of exchangeable random measures on $[0,1]$.

The martingale description involving (5) has a surprising partial converse. Here we consider any uniformly integrable, special semi-martingale X on $[0,1]$ with jump point process ξ such that the end values X_1, $[X^c]_1$, and ξ_1 are a.s. non-random. Then X is exchangeable iff the compensating processes $\hat{X}$, $[X^c]$, and $\hat{\xi}$ are absolutely continuous and admit densities that form martingales on $(0,1)$. This result is related to Grigelionis' characterization of mixed Lévy processes—hence of exchangeable processes on $\mathbb{R}_+$—as semi-martingales whose local characteristics are a.s. linear.

The various martingale descriptions enable us to prove some powerful norm relations for contractable and related processes. For example, for any L^p-valued, exchangeable processes X on $[0,1]$, we have the relations

$$\|X_t\|_p \asymp \|X_1^*\|_p \asymp \left\|([X]_1 + X_1^2)^{1/2}\right\|_p,$$

uniformly in X and for $t \in (0,1)$ and $p > 0$ in compacts. (Here $a \asymp b$ means that the ratio a/b is bounded above and below by positive constants, and $X_t^* = \sup_{s \le t} |X_s|$.) We can also use martingale methods to estimate the local growth rates of arbitrary exchangeable processes.

The chapter closes with a discussion of Palm measures. For simple point processes ξ on S, the *Palm distribution* Q_s at a point $s \in S$ can be thought of as the conditional distribution of ξ, given that ξ has a point at s. (For the existence, we need to assume that the intensity measure $E\xi$ is σ-finite.) Each measure Q_s is again the distribution of a simple point process ξ_s on S, which is clearly such that one of the points lies at s. The *reduced Palm distribution* Q_s' can be defined as the law of the point process $\xi_s' = \xi_s - \delta_s$, obtained from ξ_s by omitting the trivial point at s. Now a central result says that Q_s' is (or, rather, can be chosen to be) independent of s iff ξ is a mixed binomial or Poisson process based on $E\xi$. Recall that the latter are precisely the exchangeable point processes on S, except that exchangeability is now defined in the obvious way with respect to the measure $E\xi$. As a special case, one recovers a celebrated result of Slivnyak—the fact that ξ is Poisson iff the measures Q_s' all agree with the original distribution $\mathcal{L}(\xi)$.

— — —

Chapter 3 deals with weak convergence and related approximation properties. To introduce the subject, recall that by de Finetti's theorem every infinite exchangeable sequence $\xi = (\xi^1, \xi^2, \ldots)$ can be described in terms of a *directing random measure* ν, such that conditionally on ν, the ξ^k are i.i.d. with the common distribution ν. It is easy to see that the distributions of ξ and ν determine each other uniquely. Now consider a whole sequence of such infinite exchangeable sequences $\xi_n = (\xi_n^1, \xi_n^2, \ldots)$, $n \in \mathbb{N}$, with associated directing random measures $\nu_1, \nu_2, \ldots$. We may then ask for conditions ensuring the distributional convergence $\xi_n \xrightarrow{d} \xi$, where ξ is again exchangeable and directed by ν. Here one naturally expects that $\xi_n \xrightarrow{d} \xi$ iff $\nu_n \xrightarrow{d} \nu$, which is indeed true for a suitable choice of topology in the space of measures.

In a similar way, we can go on and consider exchangeable sequences, measures, or processes on any bounded or unbounded index set. The first step is then to identify the corresponding *directing random elements*, which may be summarized as follows:

	bounded	*unbounded*
sequences	ν	ν
measures	(α, β)	(α, ν)
processes	(α, σ, β)	(α, σ, ν)

For unbounded sequences, ν is the directing measure in de Finetti's theorem, and for bounded sequences $\xi = (\xi^1, \ldots, \xi^m)$ we may choose ν to be the empirical distribution $m^{-1} \sum_k \delta_{\xi^k}$. For exchangeable random measures ξ on $[0,1]$, we have a representation (4) in terms of some non-negative random variables α and $\beta_1, \beta_2, \ldots$, and we may choose the directing random elements to be α and β, where β denotes the point process $\sum_k \delta_{\beta_k}$ on $(0, \infty)$. Similarly, exchangeable processes X on $[0,1]$ have a representation as in (6), which suggests that we choose the directing elements α, σ, and $\beta = \sum_k \delta_{\beta_k}$. For exchangeable processes on $\mathbb{R}_+$, we may derive the directing triple (α, σ, ν) from the characteristics of the underlying Lévy processes, and for exchangeable random measures on $\mathbb{R}_+$ the choice of directing pair (α, ν) is similar.

In each of the mentioned cases, there is a corresponding limit theorem, similar to the one for exchangeable sequences. For example, if X and $X_1, X_2, \ldots$ are exchangeable processes on $[0,1]$ with directing triples (α, σ, β) and $(\alpha_n, \sigma_n, \beta_n)$, respectively, then

$$X_n \xrightarrow{d} X \qquad \Longleftrightarrow \qquad (\alpha_n, \sigma_n, \beta_n) \xrightarrow{d} (\alpha, \sigma, \beta), \tag{7}$$

for carefully chosen topologies on the appropriate function and measure spaces. But much more is true. To indicate the possibilities, consider some finite, exchangeable sequences $\xi_n = (\xi_n^k;\, k \le m_n)$ with associated summation processes

$$X_n(t) = \sum\nolimits_{k \le m_n t} \xi_n^k, \quad t \in [0,1],$$

and introduce the random triples $(\alpha_n, \sigma_n, \beta_n)$, where

$$\alpha_n = \sum\nolimits_k \xi_n^k, \qquad \sigma_n = 0, \qquad \beta_n = \sum\nolimits_k \delta_{\xi_n^k}.$$

Then (7) remains true for any exchangeable process X on $[0, 1]$ directed by (α, σ, β), provided only that $m_n \to \infty$. The corresponding result for summation processes on $\mathbb{R}_+$ generalizes Skorohod's functional limit theorem for i.i.d. random variables. We can also use a similar approach to establish the representation (6) of exchangeable processes on $[0, 1]$, now in full generality, without imposing the previous integrability condition.

The processes in (6) are similar to, but more general than Lévy processes, which leads to the obvious challenge of extending the wide range of path properties known for the latter to the broader class of general exchangeable processes. The problems in the general case are often much harder, owing to the lack of simple independence and Markov properties. One powerful method for obtaining such results is by *coupling*. Given an exchangeable process X on $[0, 1]$ with constant directing triple (α, σ, β), we may then try to construct a Lévy process Y approximating X in a suitable path-wise sense, to ensure that at least some of the path properties of Y will carry over to X. In particular, such an approach allows us to extend some delicate growth results for Lévy processes, due to Khinchin, Fristedt, and Millar, to a general exchangeable setting.

The chapter concludes with a discussion of sub-sequence principles. Here we note that, given any tight sequence of random elements $\xi = (\xi_1, \xi_2, \ldots)$ in an arbitrary Polish space S, we can extract an asymptotically exchangeable sub-sequence $\xi \circ p = (\xi_{p_1}, \xi_{p_2}, \ldots)$, in the sense that the shifted sequences $\theta_n(\xi \circ p)$ tend in distribution, as $n \to \infty$, toward a fixed exchangeable sequence ζ. A stronger result, established in various forms by several people, including Dacunha-Castelle and Aldous, is the *weak sub-sequence principle*, where the asserted limit holds in the sense of the weak convergence

$$E[\eta;\, \theta_n(\xi \circ p) \in \cdot\,] \xrightarrow{w} E\eta\nu^\infty, \quad \eta \in L^1,$$

ν being the directing random measure of ζ. In other words, the previously noted convergence $\theta_n(\xi \circ p) \xrightarrow{d} \zeta$ is *stable*, in the sense of Rényi. A related and even more powerful result is the *strong sub-sequence principle*, due to Berkes and Péter, ensuring that for any $\varepsilon > 0$ we can choose the sub-sequence $\xi \circ p$ and an approximating exchangeable sequence ζ such that

$$E[\rho(\xi_{p_n}, \zeta_n) \wedge 1] \le \varepsilon, \quad n \in \mathbb{N},$$

where ρ denotes a fixed, complete metrization of S.

— — —

Chapter 4 deals with various properties of invariance under predictable transformations in discrete or continuous time. To explain the discrete-time results, recall that a finite or infinite random sequence $\xi = (\xi_1, \xi_2, \ldots)$ is exchangeable if its distribution is invariant under non-random permutations of the elements, as in (1). The *predictable sampling* theorem extends the distributional invariance to certain random permutations. More precisely, letting $\tau_1, \tau_2, \ldots$ be a.s. distinct predictable times, taking values in the (finite or infinite) index set of ξ, we have

$$(\xi_{\tau_1}, \xi_{\tau_2}, \ldots) \stackrel{d}{=} (\xi_1, \xi_2, \ldots).$$

(Recall that a random time τ is said to be *predictable* if $\tau - 1$ is optional, hence an ordinary stopping time.) If ξ is only assumed to be contractable, then the same property holds for any strictly increasing sequence of predictable times $\tau_1, \tau_2, \ldots$. The latter result is a version of the *optional skipping* theorem, first established for i.i.d. sequences by Doob. The more general result for exchangeable sequences yields simple proofs of some classical arcsine laws and fluctuation identities for random walks and Lévy processes.

The corresponding continuous-time results are much harder. In the exchangeable case, we may define the random transformations in terms of predictable processes V on the interval $I = [0,1]$ or $\mathbb{R}_+$, taking values in the same set I and such that $\lambda \circ V^{-1} = \lambda$ a.s., where λ denotes Lebesgue measure on I. In other words, we assume the paths of V to be measure-preserving transformations on I, corresponding to the permutations considered in discrete time. Given a suitable process X on I, we now define the transformed process $X \circ V^{-1}$ by

$$(X \circ V^{-1})_t = \int_I 1\{V_s \le t\}\, dX_s, \quad t \in I, \tag{8}$$

in the sense of stochastic integration of predictable processes with respect to general semi-martingales. To motivate the notation, we note that if $X_t = \xi[0,t]$ for some random measure ξ, then $X \circ V^{-1}$ is the distribution function of the transformed measure $\xi \circ V^{-1}$. The *predictable mapping* theorem states that if X is an exchangeable process on I, then $X \circ V^{-1} \stackrel{d}{=} X$.

In the contractable case, we need to consider predictable subsets A of the index set I with $\lambda A \ge h$. The corresponding time-change process τ is given by

$$\tau_t = \inf\{s \in I;\ \lambda(A \cap [0,s]) > t\}, \quad t \le h,$$

and we may define the associated *contraction* $C_A X$ of X by

$$(C_A X)_t = X(A \cap [0, \tau_t]) = \int_0^{\tau_t} 1_A(s)\, dX_s, \quad t \le h,$$

again in the sense of general stochastic integration. (A technical difficulty is to establish the existence of the integral on the right, since in general contractable processes are not known to be semi-martingales.) We can then show that if X is contractable, then $C_A X \stackrel{d}{=} X$ on $[0, h)$.

Even stronger invariance properties can be established when X is a stable Lévy process. In the simplest case, we assume that X is strictly p-stable for some $p \in (0,2]$ and consider a predictable process $U \geq 0$ such that U^p is locally integrable. Let V be another predictable process such that

$$(U^p \cdot \lambda) \circ V^{-1} = \lambda \ \text{a.s.}, \tag{9}$$

where the left-hand side is defined as in (8) in terms of the integral process $(U^p \cdot \lambda)_t = \int_0^t U_s^p ds$. Then we can show that

$$(U \cdot X) \circ V^{-1} \stackrel{d}{=} X,$$

where $U \cdot X$ denotes the stochastic integral $\int_0^t U dX$ and the mapping by V is again defined as in (8). In particular, the result leads to time-change representations for stable integrals. If X is symmetric p-stable, we can drop the condition $U \geq 0$, provided we replace U by $|U|$ in (9). Even stronger results are obtainable when X is a Brownian motion or bridge.

The quoted results may be regarded as far-reaching extensions of the classical time-change reductions of continuous local martingales to Brownian motion, due to Dambis, Dubins and Schwarz, and Knight, and the corresponding reductions of quasi-left-continuous, simple point processes to Poisson, due to Papangelou and Meyer. In fact, an abstract version that combines all those classical results into a single theorem plays the role of a universal tool in this chapter.

Time-change reductions of optional times are closely related to the theory of exchangeable processes. To see the connection, let ξ be an exchangeable random measure on $[0,1]$, given by (4) for some i.i.d. $U(0,1)$ random variables $\tau_1, \tau_2, \ldots$, and assume for simplicity that the coefficients are a.s. non-random with $\beta_1 > \beta_2 > \cdots$. Then for any random mapping V of $[0,1]$ into itself, we have

$$\xi \circ V^{-1} = \alpha\lambda \circ V^{-1} + \sum\nolimits_k \beta_k \, \delta_{V(\tau_k)},$$

and we see that $\xi \circ V^{-1} \stackrel{d}{=} \xi$ iff V is a.s. λ-preserving and satisfies

$$(V_{\tau_1}, V_{\tau_2}, \ldots) \stackrel{d}{=} (\tau_1, \tau_2, \ldots). \tag{10}$$

In fact, the predictable mapping theorem shows that (10) holds automatically as soon as V is predictable with $\lambda \circ V^{-1} = \lambda$ a.s.

Now consider an arbitrary sequence of optional times $\tau_1, \tau_2, \ldots$ with associated marks $\kappa_1, \kappa_2, \ldots$ in some measurable space K, and introduce the compensators $\eta_1, \eta_2, \ldots$ of the random pairs (τ_j, κ_j). The corresponding *discounted compensators* $\zeta_1, \zeta_2, \ldots$ are random sub-probability measures on $\mathbb{R}_+ \times K$, defined as the unique solutions to the Doléans differential equations

$$d\zeta_j = -Z_-^j \, d\eta_j, \quad Z_0^j = 1, \qquad j \in \mathbb{N},$$

where $Z_t^j = 1 - \zeta_j([0,t] \times K)$. Consider any predictable mappings $T_1, T_2, \ldots$ on $\mathbb{R}_+ \times K$ taking values in a space S, and fix some probability measures $\mu_1, \mu_2, \ldots$ on S. Then the conditions

$$\zeta_j \circ T_j^{-1} \le \mu_j \text{ a.s.}, \quad j \in \mathbb{N}, \tag{11}$$

ensure that the images $\gamma_j = T_j(\tau_j, \kappa_j)$ will be independent with distributions μ_j. This may be surprising, since the inequalities in (11) are typically strict and should allow considerable latitude for tinkering with any initial choice of mappings T_j.

— — —

Chapter 5 deals with the closely related family of *decoupling identities.* To explain the background, consider a sequence of i.i.d. random variables $\xi_1, \xi_2, \ldots$ with finite mean $E\xi_1$ and a predictable sequence of random variables $\eta_1, \eta_2, \ldots$. Under suitable integrability conditions on the η_k, we note that

$$E \sum\nolimits_k \xi_k \eta_k = E(\xi_1)\, E \sum\nolimits_k \eta_k.$$

Similarly, assuming $E\xi_1 = 0$ and $E\xi_1^2 < \infty$, we have

$$E\Big(\sum\nolimits_k \xi_k \eta_k\Big)^2 = E(\xi_1^2)\, E \sum\nolimits_k \eta_k^2,$$

under appropriate integrability conditions. The remarkable thing about these formulas is that the right-hand side is the same regardless of the dependence between the sequences (ξ_k) and (η_k). In other words, if we choose $(\tilde\eta_k) \overset{d}{=} (\eta_k)$ to be independent of (ξ_k), then (under the previous conditions)

$$\begin{aligned} E \sum\nolimits_k \xi_k \eta_k &= E \sum\nolimits_k \xi_k \tilde\eta_k, \\ E\Big(\sum\nolimits_k \xi_k \eta_k\Big)^2 &= E\Big(\sum\nolimits_k \xi_k \tilde\eta_k\Big)^2. \end{aligned}$$

This is the idea of *decoupling:* Whatever the dependence may be between the variables ξ_k and η_k, we can evaluate the expressions on the left as if the two sequences were independent.

The situation in continuous time is similar. Thus, assuming X to be a Lévy process with associated filtration $\mathcal{F} = (\mathcal{F}_t)$, letting V be an $\mathcal{F}$-predictable process on $\mathbb{R}_+$, and choosing $\tilde V \overset{d}{=} V$ to be independent of X, we have (again under suitable integrability conditions)

$$\begin{aligned} E \textstyle\int_0^\infty V dX = E(X_1)\, E \int_0^\infty V_s\, ds = E \int_0^\infty \tilde V dX, \\ E(\textstyle\int_0^\infty V dX)^2 = E(X_1^2)\, E \int_0^\infty V_s^2\, ds = E\Big(\int_0^\infty \tilde V dX\Big)^2, \end{aligned}$$

where, for the latter equations, we need to assume that $EX_1 = 0$. The formulas follow from standard facts in stochastic calculus. Specializing to the case where $\eta_k = 1\{\tau \ge k\}$ or $V_t = 1\{\tau \ge t\}$ for some optional time $\tau < \infty$ and writing $X_k = \sum_{j \le k} \xi_j$, we get the classical *Wald identities*

$$E(X_\tau) = E(X_1)\, E(\tau), \qquad E(X_\tau^2) = E(X_1^2)\, E(\tau),$$

where the latter formula requires $EX_1 = 0$ and both equations are valid, in discrete and continuous time, under suitable integrability conditions.

The relations mentioned so far are quite elementary and rather obvious. With some further effort, one can derive higher-dimensional formulas of the same kind, such as

$$E\Big(\sum\nolimits_k \xi_k \eta_k\Big)^d = E\Big(\sum\nolimits_k \xi_k \tilde{\eta}_k\Big)^d,$$
$$E\Big(\int_0^\infty V dX\Big)^d = E\Big(\int_0^\infty \tilde{V} dX\Big)^d,$$

which, in addition to appropriate integrability conditions, require that the sums or integrals

$$S_m = \sum\nolimits_k \eta_k^m, \qquad I_m = \int_0^\infty V_s^m \, ds,$$

be non-random (or at least $\mathcal{F}_0$-measurable) for $1 \le m < d$. From here on, we may easily proceed to general product moments, which leads to decoupling identities of the form

$$E \prod\nolimits_{j \le d} \sum\nolimits_k \xi_{jk} \, \eta_{jk} = E \prod\nolimits_{j \le d} \sum\nolimits_k \xi_{jk} \, \tilde{\eta}_{jk}, \tag{12}$$
$$E \prod\nolimits_{j \le d} \int_0^\infty V_j \, dX_j = E \prod\nolimits_{j \le d} \int_0^\infty \tilde{V}_j \, dX_j, \tag{13}$$

requiring non-randomness of the sums or integrals

$$S_J = \sum\nolimits_k \prod\nolimits_{j \in J} \eta_{jk}, \qquad I_J = \int_0^\infty \prod\nolimits_{j \in J} V_j(s) \, ds, \tag{14}$$

for all nonempty, proper subsets $J \subset \{1, \ldots, d\}$.

Up to this point, we have assumed the sequence $\xi = (\xi_{jk})$ to be i.i.d. (in the second index k) or the process $X = (X_j)$ to be Lévy. The truly remarkable fact is that moment formulas of similar type remain valid for finite exchangeable sequences (ξ_{jk}) and for exchangeable processes (X_j) on $[0, 1]$. Thus, in the general exchangeable case we can still prove decoupling identities such as (12) or (13), the only difference being that the sums or integrals in (14) are now required to be non-random even for $J = \{1, \ldots, d\}$.

These innocent-looking identities are in fact quite amazing, already when $d = 1$. For a simple gambling illustration, suppose that the cards of a well-shuffled deck are drawn one by one. You are invited to bet an amount η_k on the kth card, based on your knowledge of previous outcomes, and the bank will return the double amount if the card is red, otherwise nothing. Also assume that, before entering the game, you must fix your total bet $\sum_k \eta_k$. Then (12) shows that your expected total gain is 0. This is surprising, since you might hope to improve your chances by betting most of your money when the proportion of red cards in the remaining deck is high. If $\sum_k \eta_k^2$ is also fixed in advance, then even the variance of your total gain is independent of your strategy, and so on for higher moments.

Though the methods of proof in Chapter 5 are very different from those in the previous chapter, the results are actually closely related. Indeed, it is not hard to derive the predictable sampling theorem in Chapter 4 from the corresponding decoupling identities, and similarly in continuous time. We can even prove contractable versions of the decoupling identities that are strong enough to imply the optional skipping theorem for contractable sequences, and similarly for contractable processes on $[0,1]$.

— — —

In Chapter 6 we consider exchangeable random sets and related processes on $[0,1]$ or $\mathbb{R}_+$. To motivate the basic ideas, consider an rcll process X on $\mathbb{R}_+$ with $X_0 = 0$ such that the random set $\Xi = \{t \geq 0;\ X_t = 0\}$ is a.s. unbounded. Then X is said to be *regenerative* at 0 if for any optional time τ taking values in Ξ, the shifted process $\theta_\tau X$ is independent of $\mathcal{F}_\tau$ with the same distribution as X, where $\mathcal{F} = (\mathcal{F}_t)$ denotes the right-continuous filtration induced by X. *Local homogeneity* is the same property without the independence condition. Thus, X is said to be locally homogeneous at 0 if $\theta_\tau X \overset{d}{=} X$ for any optional time τ in Ξ.

In view of de Finetti's theorem and the characterization of infinite exchangeable sequences by strong stationarity, it is easy to believe (but not quite so easy to prove) that X is locally homogeneous iff it is a mixture of regenerative processes. Since the latter can be described in terms of a local time random measure ξ supported by Ξ, along with a homogeneous Poisson point process η of excursions, as specified by Itô's excursion law, in the general locally homogeneous case we obtain a conditional representation of the same type. In particular, the cumulative local time process $L_t = \xi[0,t]$ is a mixture of inverse subordinators, hence the inverse of a non-decreasing, exchangeable process on $\mathbb{R}_+$.

For processes on $[0,1]$, we need to replace the local homogeneity by a suitable reflection property. To make this precise, let us first consider the case of a random, closed subset $\Xi \subset [0,1]$ containing 0. For any optional time τ in Ξ, we may construct a reflected set $R_\tau \Xi$ by reversing the restriction of Ξ to $[\tau,1]$. The *strong reflection property* is defined by the condition $R_\tau \Xi \overset{d}{=} \Xi$ for every optional time τ in Ξ. The definition is similar for processes X on $[0,1]$, except that the initial reflection needs to be combined with a reversal of each excursion.

With reflection invariance defined in this way, the theory becomes analogous to the one for processes on $\mathbb{R}_+$. Thus, under the stated condition, we have again a local time random measure ξ supported by Ξ, along with an exchangeable point process η of excursions, such that X admits a conditional Itô-type representation in terms of ξ and η. In particular, the cumulative local time L, normalized such that $L_1 = 1$, now becomes the inverse of a non-decreasing exchangeable process on $[0,1]$.

We proceed to describe an interesting sampling property of the local time process L. First suppose that Ξ is regenerative. Combining the regenerative property with the loss-of-memory property characterizing the exponential distribution, we note that if the time τ is exponentially distributed and independent of Ξ, then $\sigma = L_\tau$ has again an exponential distribution. Iterating this result, we see that if $\tau_1 < \tau_2 < \cdots$ form a homogeneous Poisson process on $\mathbb{R}_+$ independent of Ξ, then the variables $L_{\tau_1}, L_{\tau_2}, \ldots$ form a compound Poisson process. (Multiplicities occur when several times τ_k hit the same excursion interval.) Ignoring repetitions in the sequence (L_{τ_k}), we get another homogeneous Poisson process $\sigma_1, \sigma_2, \ldots$, and by suitably normalizing L we can arrange that

$$(\sigma_1, \sigma_2, \ldots) \stackrel{d}{=} (\tau_1, \tau_2, \ldots). \tag{15}$$

As stated in this form, the result clearly carries over to the general locally homogeneous case.

So far things are quite simple and elementary. Now consider instead a random subset Ξ of $[0, 1]$, satisfying the strong reflection property and admitting a normalized local time process L, as described above. Here we take the variables $\tau_1, \tau_2, \ldots$ to be i.i.d. $U(0, 1)$, independently of Ξ, and let $\sigma_1, \sigma_2, \ldots$ be the distinct elements of the sequence $L_{\tau_1}, L_{\tau_2}, \ldots$. Then, surprisingly, the σ_k are again i.i.d. $U(0, 1)$, and so (15) remains fulfilled.

Still considering an exchangeable random set Ξ in $[0, 1]$, as described above, we now assume that $\lambda\Xi = 0$ a.s. Continuing Ξ periodically and shifting by an independent $U(0, 1)$ random variable, we obtain a stationary random set $\tilde{\Xi}$ in $\mathbb{R}$. Under suitable regularity conditions, the distributions of Ξ and $\tilde{\Xi}$ can be shown to resemble each other locally, apart from a normalizing factor. Here only some simple aspects of this similarity are discussed. For any fixed interval $I \subset [0, 1]$, we may consider the probabilities that I intersects Ξ or $\tilde{\Xi}$, which will typically tend to 0 as I shrinks to a single point $t \in (0, 1)$. Under mild restrictions on the underlying parameters, we can show that the local time intensity $E\xi$ is absolutely continuous with a nice density p, and that for almost every $t \in (0, 1)$

$$P\{\Xi \cap I \neq \emptyset\} \sim p_t \, P\{\tilde{\Xi} \cap I \neq \emptyset\}, \quad I \downarrow \{t\}.$$

Similar relations hold for regenerative sets in $\mathbb{R}_+$. To appreciate these formulas, we note that the probabilities on the left are very difficult to compute. For those on the right the computation is easy.

If X is locally homogeneous at several states, it is clearly conditionally regenerative at each of them. It may be less obvious that these properties can be combined, under suitable conditions, into a conditional strong Markov property on the corresponding part of the state space. Strengthening the local homogeneity into a property of *global homogeneity*, we may even deduce the Markov property in its usual, unconditional form. This leads us naturally to regard the strong Markov property of a process X at an optional time $\tau < \infty$ as the combination of two properties,

$$P[\theta_\tau X \in \cdot | \mathcal{F}_\tau] = P[\theta_\tau X \in \cdot | X_\tau] = \mu(X_\tau, \cdot) \text{ a.s.},$$

a conditional independence and a global homogeneity condition. We have already indicated how the latter property at every τ implies the strong Markov property, hence also the conditional independence. Under certain regularity conditions, we can even obtain an implication in the opposite direction. Thus, again quite surprisingly, the homogeneity and independence components of the strong Markov property are then equivalent.

— — —

The last three chapters deal with certain multi-variate symmetries. Crucial for our subsequent developments is the discussion of exchangeable and contractable arrays in Chapter 7. Here our main aim is to derive representations of separately or jointly contractable or exchangeable arrays of arbitrary dimension. For motivation we note that, by de Finetti's theorem, any infinite sequence of exchangeable random variables $X_1, X_2, \ldots$ has a representation

$$X_n = f(\alpha, \xi_n) \text{ a.s.}, \quad n \in \mathbb{N},$$

in terms of a measurable function f on $[0,1]^2$ and some i.i.d. $U(0,1)$ random variables α and $\xi_1, \xi_2, \ldots$.

A two-dimensional array $X = (X_{ij};\, i,j \in \mathbb{N})$ is said to be *separately exchangeable* if

$$X \circ (p,q) \equiv (X_{p_i,q_j};\, i,j \in \mathbb{N}) \stackrel{d}{=} (X_{ij};\, i,j \in \mathbb{N}) \equiv X,$$

for any (finite) permutations $p = (p_i)$ and $q = (q_j)$ of $\mathbb{N}$, and *jointly exchangeable* if the same property holds when $p = q$, so that for any permutation p of $\mathbb{N}$,

$$X \circ p \equiv (X_{p_i,p_j};\, i,j \in \mathbb{N}) \stackrel{d}{=} (X_{ij};\, i,j \in \mathbb{N}) \equiv X.$$

Restricting p and q to sub-sequences $p_1 < p_2 < \cdots$ and $q_1 < q_2 < \cdots$ yields the corresponding properties of *separate* or *joint contractability*. However, since any separately contractable array is also separately exchangeable, by Ryll-Nardzewski's theorem, it is enough to consider the jointly contractable case.

For jointly exchangeable arrays it is often more natural to consider the index set N_2, consisting of all pairs $i,j \in \mathbb{N}$ with $i \neq j$. This is because an array $X = (X_{ij})$ on $\mathbb{N}^2$ is jointly exchangeable iff the same property holds for the array of pairs (X_{ii}, X_{ij}) indexed by N_2. Similarly, for (jointly) contractable arrays, we may prefer the index set T_2 of pairs $i,j \in \mathbb{N}$ with $i < j$, since an array $X = (X_{ij})$ on $\mathbb{N}^2$ is contractable iff the same property holds for the array of triples (X_{ii}, X_{ij}, X_{ji}) on T_2. It is also convenient to think of T_2 as the class of sets $\{i,j\} \subset \mathbb{N}$ of cardinality 2.

The first higher-dimensional representation theorems were obtained, independently, by Aldous and Hoover, who proved that an array $X = (X_{ij})$ on $\mathbb{N}^2$ is separately exchangeable iff

$$X_{ij} = f(\alpha, \xi_i, \eta_j, \zeta_{ij}) \text{ a.s.}, \quad i,j \in \mathbb{N},$$

for some measurable function f on $[0,1]^4$ and some i.i.d. $U(0,1)$ random variables α, ξ_i, η_j, ζ_{ij}, $i,j \in \mathbb{N}$. Hoover also settled the more general case of jointly exchangeable arrays of arbitrary dimension. In particular, he showed that a two-dimensional array $X = (X_{ij})$ is jointly exchangeable iff

$$X_{ij} = f(\alpha, \xi_i, \xi_j, \zeta_{ij}) \text{ a.s.}, \quad i,j \in \mathbb{N}, \tag{16}$$

for some measurable function f as above and some i.i.d. $U(0,1)$ random variables α, ξ_i, and $\zeta_{ij} = \zeta_{ji}$. These results and their higher-dimensional counterparts are quite deep, and their proofs occupy much of Chapter 7.

In the same chapter, we also derive the corresponding representations of (jointly) contractable arrays. In two dimensions we have the same representation as in (16), except that i and j should now be restricted to the triangular index set T_2. In higher dimensions, it is natural to choose as our index set the class $\widetilde{\mathbb{N}}$ of all finite subsets $J \subset \mathbb{N}$. The representation can then be stated compactly in the form

$$X_J = f(\xi_I;\, I \subset J), \quad J \in \widetilde{\mathbb{N}}, \tag{17}$$

where $\xi = (\xi_J)$ is an array of i.i.d. $U(0,1)$ random variables, also indexed by $\widetilde{\mathbb{N}}$, and f is a measurable function on a suitable space. This formula can be used to extend X in a natural way to a jointly exchangeable array on the index set $\overline{\mathbb{N}}$, consisting of all finite sequences $(k_1, \ldots, k_d)$ in $\mathbb{N}$ with distinct entries. The surprising conclusion is that *an array on* $\widetilde{\mathbb{N}}$ *is contractable iff it admits an exchangeable extension to* $\overline{\mathbb{N}}$. Note that this extension is far from unique, owing to the non-uniqueness of the representing function in (17).

After a detailed study of some matters of uniqueness and conditioning, too technical to describe here, we conclude the chapter with a discussion of symmetric partitions. Informally, a random partition of $\mathbb{N}$ into disjoint subsets $A_1, A_2, \ldots$ is said to be exchangeable if an arbitrary permutation of $\mathbb{N}$ yields a partition with the same distribution. To formalize this, we may introduce a random array X on $\mathbb{N}$ with entries 0 and 1, such that $X_{ij} = 1$ iff i and j belong to the same set A_k. The partition $\{A_k\}$ is then defined to be exchangeable if X is jointly exchangeable, in the sense of the previous discussion. The classical result, due to Kingman, states that $\{A_k\}$ is exchangeable iff it admits a *paint-box* representation

$$X_{ij} = 1\{\xi_i = \xi_j\} \text{ a.s.}, \quad i,j \in \mathbb{N}, \tag{18}$$

in terms of a sequence of exchangeable random variables $\xi_1, \xi_2, \ldots$. The term comes from the interpretation of the variables ξ_j as colors, chosen at random from a possibly infinite paint box, which determine a partition of $\mathbb{N}$ into subsets of different colors.

There is nothing special about exchangeable partitions. Letting $\mathcal{T}$ be an arbitrary family of injective maps $p\colon \mathbb{N} \to \mathbb{N}$, we can show that a random partition $\{A_k\}$ is $\mathcal{T}$-invariant (in distribution) iff it admits a representation as in (18) in terms of a $\mathcal{T}$-invariant sequence of random variables. To indicate the possibilities, we can apply the stated result, along with the previous representations of arrays, to obtain representations of exchangeable or contractable partitions of $\overline{\mathbb{N}}$ or $\widetilde{\mathbb{N}}$, respectively.

— — —

In Chapter 8 we turn our attention to multi-variate rotatability. For arrays, the definitions are analogous to those in the contractable and exchangeable cases. Thus, for suitable two-dimensional arrays X, U, and V indexed by $\mathbb{N}^2$, we define the transformed array $Y = (U \otimes V)X$ by

$$Y_{ij} = \sum\nolimits_{h,k} U_{ih}\, V_{jk}\, X_{hk}, \quad i, j \in \mathbb{N}.$$

To rotate X in the first index, we choose U to be an orthogonal matrix on a finite index set $I^2 \subset \mathbb{N}^2$, and extend U to $\mathbb{N}^2$ by putting $U_{ij} = \delta_{ij} \equiv 1_{\{i=j\}}$ when either i or j lies in I^c. Similarly, we get a rotation in the second index by choosing V to be an orthogonal matrix of the same type. Then X is said to be *separately rotatable* if

$$(U \otimes V)X \overset{d}{=} X, \quad U, V \in \mathcal{O}, \tag{19}$$

where $\mathcal{O}$ denotes the class of infinite, orthogonal matrices as above, rotating only finitely many coordinates. It is *jointly rotatable* if (19) holds whenever $U = V \in \mathcal{O}$. The definitions carry over immediately to arrays of arbitrary dimension.

We have already seen that, by Freedman's theorem, an infinite sequence $X = (X_1, X_2, \ldots)$ is rotatable iff it can be represented in the form $X_j = \sigma\zeta_j$, where $\zeta_1, \zeta_2, \ldots$ are i.i.d. $N(0,1)$ and σ is an independent random variable. For separately rotatable arrays $X = (X_{ij})$ on $\mathbb{N}^2$, we get the a.s. representation

$$X_{ij} = \sigma\zeta_{ij} + \sum\nolimits_k \alpha_k\, \xi_{ki}\, \eta_{kj}, \quad i, j \in \mathbb{N}, \tag{20}$$

in terms of some i.i.d. $N(0,1)$ random variables ξ_{ki}, η_{kj}, and ζ_{ij} and an independent collection of random coefficients σ and $\alpha_1, \alpha_2, \ldots$, where the latter need to be such that $\sum_k \alpha_k^2 < \infty$ a.s. This is a quite deep result, first conjectured by Dawid, and then proved (under a moment condition) by Aldous. For jointly rotatable arrays on $\mathbb{N}^2$, we get instead an a.s. representation of the form

$$X_{ij} = \rho\delta_{ij} + \sigma\zeta_{ij} + \sigma'\zeta_{ji} + \sum\nolimits_{h,k} \alpha_{hk}\, (\xi_{ki}\, \xi_{kj} - \delta_{ij}\, \delta_{hk}), \quad i, j \in \mathbb{N}, \tag{21}$$

for some i.i.d. $N(0,1)$ random variables ξ_{ki} and ζ_{ij} and an independent collection of random coefficients ρ, σ, σ', and α_{hk}, where the latter must satisfy $\sum_{h,k} \alpha_{hk}^2 < \infty$ a.s., to ensure convergence of the double series in (21).

The representations (20) and (21) of separately or jointly rotatable arrays are easily extended to the continuous parameter case, as follows. Here we consider processes X on $\mathbb{R}_+^2$ with $X(s,t) = 0$ when $s = 0$ or $t = 0$, and we may define rotatability in the obvious way in terms of the two-dimensional increments

$$\Delta_{h,k}X(s,t) = X(s+h,t+k) - X(s+h,t) - X(s,t+k) + X(s,t),$$

where $s,t,h,k \geq 0$ are arbitrary. Assuming X to be continuous in probability, we get in the separately rotatable case an a.s. representation

$$X(s,t) = \sigma Z(s,t) + \sum\nolimits_k \alpha_k \, B_k(s) \, C_k(t), \quad s,t \geq 0,$$

for a Brownian sheet Z, some independent Brownian motions $B_1, B_2, \ldots$ and $C_1, C_2, \ldots$, and an independent collection of random coefficients σ and $\alpha_1, \alpha_2, \ldots$ with $\sum_k \alpha_k^2 < \infty$ a.s. In the jointly rotatable case, we get instead an a.s. representation of the form

$$\begin{aligned} X(s,t) \;&=\; \rho(s \wedge t) + \sigma Z(s,t) + \sigma' Z(t,s) \\ &\quad + \sum\nolimits_{h,k} \alpha_{hk} \, (B_h(s) \, B_k(t) - \delta_{hk} \, (s \wedge t)). \end{aligned} \tag{22}$$

The last two representation formulas are best understood in their measure-valued or functional versions

$$X \;=\; \sigma Z + \sum\nolimits_k \alpha_k \, (B_k \otimes C_k), \tag{23}$$

$$X \;=\; \rho \lambda_D + \sigma Z + \sigma' \tilde{Z} + \sum\nolimits_{h,k} \alpha_{hk} \, (B_h \otimes B_k), \tag{24}$$

where λ_D denotes Lebesgue measure along the main diagonal D in $\mathbb{R}_+^2$, $\tilde{Z}$ is the reflection of Z given by $\tilde{Z}_{s,t} = Z_{t,s}$, and $B_h \otimes B_k$ denotes the double stochastic integral formed by the processes B_h and B_k. In particular, we note that

$$(B_k \otimes B_k)([0,s] \times [0,t]) = B_k(s)B_k(t) - s \wedge t, \quad s,t \geq 0,$$

by the expansion of multiple Wiener–Itô integrals in terms of Hermite polynomials, which explains the form of the last term in (22).

The representations in the discrete and continuous parameter cases can be unified by a Hilbert-space approach, which also clears the way for extensions to higher dimensions. Here we consider any *continuous linear random functional (CLRF)* X on an infinite-dimensional, separable Hilbert space H, where linearity means that

$$X(ah + bk) = aXh + bXk \;\text{ a.s.}, \quad h,k \in H, \;\; a,b \in \mathbb{R},$$

and continuity is defined by $Xh \xrightarrow{P} 0$ (or $E[|Xh| \wedge 1] \to 0$) as $h \to 0$ in H. Rotatability of X means that $X \circ U \stackrel{d}{=} X$ for any unitary operator U on H, where $(X \circ U)h = X(Uh)$, and Freedman's theorem shows that X is rotatable iff $X = \sigma\eta$ for some isonormal Gaussian process (G-process) η on H and an independent random variable σ.

For CLRFs X on $H^{\otimes 2} = H \otimes H$ we may define separate and joint rotatability in the obvious way in terms of tensor products of unitary operators, and we get the a.s. representations

$$Xh = \sigma\zeta h + (\xi \otimes \eta)(\alpha \otimes h), \tag{25}$$
$$Xh = \sigma\zeta h + \sigma'\tilde{\zeta}h + \eta^{\otimes 2}(\alpha \otimes h), \tag{26}$$

in terms of some independent G-processes ξ, η, and ζ on $H^{\otimes 2}$ and an independent pair (σ, α) or triple $(\sigma, \sigma', \alpha)$, where σ and σ' are real random variables and α is a random element of $H^{\otimes 2}$. Here (25) is just an abstract version of (20) and (23). However, (26) is less general than (21) or (24), since there is no term in (26) corresponding to the diagonal terms $\rho\delta_{ij}$ in (21) or $\rho(s \wedge t)$ in (24). This is because these terms have no continuous extension to $H^{\otimes 2}$.

Those considerations determine our strategy: First we derive representations of separately or jointly rotatable CLRFs on the tensor products $H^{\otimes d}$, which leads to simple and transparent formulas, independent of the choice of ortho-normal basis in H. Applying the latter representations to the various diagonal terms, we can then deduce representation formulas for rotatable arrays of arbitrary dimension.

It may come as a surprise that the representations of rotatable random functionals on the tensor products $H^{\otimes d}$ can also be used to derive representation formulas for separately or jointly exchangeable or contractable random sheets of arbitrary dimension. (By a *random sheet* on $\mathbb{R}_+^d$ or $[0,1]^d$ we mean a continuous random process X such that $X = 0$ on each of the d coordinate hyper-planes.) To see the connection, recall that a continuous process X on $\mathbb{R}_+$ or $[0,1]$ with $X_0 = 0$ is exchangeable iff $X_t = \alpha t + \sigma B_t$ a.s., where B is a Brownian motion or bridge, respectively, independent of the random coefficients α and σ. Omitting the drift term αt gives $X_t = \sigma B_t$, which we recognize, for processes on $\mathbb{R}_+$, as the general form of a rotatable process. When X is defined on $[0,1]$, we may apply the scaling transformation

$$Y(t) = (1+t)\, X\left(\frac{t}{1+t}\right), \quad t \geq 0, \tag{27}$$

to convert X into a rotatable process on $\mathbb{R}_+$.

In higher dimensions, we can decompose X into drift terms associated with the different coordinate subspaces of $\mathbb{R}_+^d$, and then apply transformations of type (27), if necessary, to get a description of X in terms of rotatable processes of different dimension. The previously established representations in the rotatable case can then be used to yield the desired representation formulas for exchangeable random sheets. For contractable random sheets on $\mathbb{R}_+^d$, we may easily reduce to the exchangeable case by means of the general extension property from Chapter 7. The indicated approach and resulting formulas provide a striking confirmation of the close relationship between the various symmetries in our general hierarchy.

— — —