Preface

The theory of random measures is a key area of modern probability theory, arguably as rich and important as martingale theory, ergodic theory, or probabilistic potential theory, to mention only three of my favorite areas. The purpose of this book is to give a systematic account of the basic theory, and to discuss some areas where the random measure point of view has been especially fruitful.

The subject has often been dismissed by the ignorant as an elementary and narrowly specialized area of probability theory, mainly of interest for some rather trite applications. Standard textbooks on graduate-level probability often contain massive chapters on Brownian motion and related subjects, but only a cursory mention of Poisson processes, along with a short discussion of their most basic properties. This is unfortunate, since random measures occur everywhere in our discipline and play a fundamental role in practically every area of stochastic processes.

Classical examples include the Lévy–Itô representation of stochastically continuous processes with independent increments, where the jump component may be described in terms of a Poisson random measure on a suitable product space. Poisson processes and their mixtures also arise naturally in such diverse areas as continuous-time Markov chains, Palm and Gibbs measures, the ergodic theory of particle systems, processes of lines and flats, branching and super-processes, just to name a few.

But there is so much more, and once you become aware of the random measure point of view, you will recognize such objects everywhere. On a very basic level, regular conditional distributions are clearly random measures with special properties. Furthermore, just as measures on the real line are determined by their distribution functions, every non-decreasing random process determines a random measure. In particular, it is often useful to regard the local time of a process at a fixed point as a random measure. Similarly, we may think of the Doob–Meyer decomposition as associating a predictable random measure with every sub-martingale.

The random measure point of view is not only useful in leading to new insights, quite often it is also the only practical one. For example, the jump structure of a general semi-martingale may be described most conveniently in terms of the generated jump point process, defined on a suitable product space. The associated compensator is a predictable random measure on the same space, and there is no natural connection to increasing processes. A similar situation arises in the context of super-processes, defined as diffusion limits of classical branching processes under a suitable scaling. Though there is indeed a deep and rather amazing description in terms of discrete particle systems, the process itself must still be understood as a randomly evolving family of diffuse random measures.

Though the discovery of Poisson processes goes back to the early 1900's, in connection with the modeling of various phenomena in physics, telecommunication, and finance, their fundamental importance in probability theory may not have become clear until the work of Lévy (1934–35). More general point processes were considered by Palm (1943), whose seminal thesis on queuing theory contains the germs of Palm distributions, renewal theory, and Poisson approximation. Palm's ideas were extended and made rigorous by Khinchin (1955), and a general theory of random measures and point processes emerged during the 1960's and 70's through the cumulative efforts of Rényi (1956/67), Grigelionis (1963), Matthes (1963), Kerstan (1964a/b), Mecke (1967), Harris (1968/71), Papangelou (1972/74a/b), Jacod (1975), and many others. A milestone was the German monograph by Kerstan, Matthes, & Mecke (1974), later appearing in thoroughly revised editions (1978/82) in other languages.

My own interest in random measures goes back to my student days in Gothenburg—more specifically to October 1971—when Peter Jagers returned from a sabbatical leave in the US, bringing his lecture notes on random measures, later published as Jagers (1974), which became the basis for our weakly seminar. Inspired by the author's writings and encouragement, I wrote my own dissertation on the subject, which was later published in extended form as my first random measure book K(1975/76), subsequently extended to double length in K(1983/86), through the addition of new material.

Since then so much has happened, so many exciting new discoveries have been made, and I have myself been working and publishing in the area, on and off, for the last four decades. Most of the previous surveys and monographs on random measures and point processes are today totally outdated, and it is time for a renewed effort to organize and review the basic results, and to bring to light material that would otherwise be lost or forgotten on dusty library shelves. In view of the vastness of current knowledge, I have been forced to be very selective, and my choice of topics has naturally been guided by personal interests, knowledge, and taste. Some omitted areas are covered by Daley & Vere-Jones (2003/08) or Last & Penrose (2017), which may serve as complements to the present text (with surprisingly little overlap).

Acknowledgments

This book is dedicated to *Peter Jagers*, without whose influence I would never have become a mathematician, or at best a very mediocre one. His lecture notes, and our ensuing 1971–72 seminar, had a profound catalytic influence on me, for which I am forever grateful. Since then he has supported me in so many ways. Thank you Peter!

Among the many great mathematicians that I have been privileged to know and learn from through the years, I want to mention especially the late *Klaus Matthes*—the principal founder and dynamic leader behind the modern developments of random measure theory. I was also fortunate for many years to count *Peter Franken* as a close friend, up to his sudden and tragic death. Both of them offered extraordinary hospitality and friendly encouragement, in connection with my many visits to East-Germany during the 1970's and early 80's. Especially the work of Matthes was a constant inspiration during my early career.

My understanding of topics covered by this book has also benefited from interactions with many other admirable colleagues and friends, including

> *David Aldous, Tim Brown, Daryl Daley, Alison Etheridge, Karl-Heinz Fichtner, Klaus Fleischmann, Daniel Gentner, Jan Grandell, Xin He, P.C.T. van der Hoeven, Martin Jacobsen, Jean Jacod, Klaus Krickeberg, Günter Last, Ross Leadbetter, Jean-François LeGall, Joseph Mecke, Gopalan Nair, Fredos Papangelou, Jurek Szulga, Hermann Thorisson, Anton Wakolbinger, Martina Zähle, and Ulrich Zähle*†.

I have also been lucky to enjoy the interest and encouragement of countless other excellent mathematicians and dear friends, including especially

> *Robert Adler, Istvan Berkes, Stamatis Cambanis*†, *Erhan Çinlar, Kai Lai Chung*†, *Donald Dawson, Persi Diaconis, Cindy Greenwood, Gail Ivanoff, Gopi Kallianpur*†, *Alan Karr, David Kendall*†, *Sir John Kingman, Ming Liao, Torgny Lindvall, Erkan Nane, Jim Pitman, Balram Rajput, Jan Rosinski, Frank Spitzer*†, *and Wim Vervaat*†.

I apologize for the unintentional omission of any names that ought to be on my lists.

During the final stages of preparation of my files, Günter Last kindly sent me a preliminary draft of his forthcoming book with Mathew Penrose, which led to some useful correspondence about history and terminology. Anders Martin-Löf helped me to clarify the early contributions of Lundberg and Cramér.

Though it may seem farfetched and odd to include some musical and artistic influences, the truth is that every theorem I ever proved has been inspired by music, and also to a lesser extent by the visual arts. The pivotal event was when, at age 16, I reluctantly agreed to join my best friend in high school to attend a recital in the Stockholm concert hall. This opened my eyes—or rather ears—to the wonders of classical music, making me an addicted concert goer and opera fan ever since. Now I am constantly listening to music, often leaving the math to mature in my mind during hours of piano practice. How can I ever thank the great composers, all dead, or the countless great performers who have so enriched my life and inspired my work?

Among the great musicians I have known personally, I would like to mention especially my longtime friend *Per Enflo*—outstanding pianist and also a famous mathematician—and the Van Cliburn gold medalist *Alexander Kobrin* with his fabulous students at the Schwob music school in Columbus, GA. Both belong to the exquisite group of supreme musicians who have performed at home recitals in our house. Somehow, experiences like those have inspired much of the work behind this book.

Whatever modest writing skills I have acquired through the years may come from my passionate reading, beginning 30 years ago with the marvelous *The Story of Civilization* by Will & Ariel Durant, eleven volumes of about a thousand pages each. Since then I have kept on buying countless books on especially cultural history and modern science, now piling up everywhere in our house, after the space in our bookcases has long been used up. I owe my debt to their numerous authors.

Let me conclude with two of my favorite quotations, beginning with one that I copied long ago from a Chinese fortune cookie:

> *Behind every successful man is a surprised mother-in-law.*

Though I truly appreciate the support of family and in-laws through the years, I admit that, in my case, the statement may have limited applicability. If I was ever lucky enough to stumble upon some interesting mathematical truths, I have utterly failed to convey any traces of those to my family or non-mathematical friends, who may still think that I am delving in a boring and totally incomprehensible world of meaningless formulas. They have no idea what treasures of sublime beauty they are missing!

My second quote, this time originating with the late American comedian *Groucho Marx*, may be a lot more relevant:

> *Man does not control his own fate—the women in his life do that for him.*

I am still struggling to navigate through the thorny thickets of life. Some wonderful people have demonstrated the meaning of true friendship by offering their encouragement and support when I needed them the most. Their generous remarks I will never forget.

Olav Kallenberg
January 2017

Introduction

This book is divided into thirteen chapters, each dealing with a different aspect of the theory and applications of random measures. Here we will give a general, informal introduction to some basic ideas of the different chapters, and indicate their significance for the subsequent development. A more detailed introduction will be given at the beginning of each chapter.

Informally, we may think of a *random measure*[1] as a randomly chosen measure ξ on a measurable space $(S, \mathcal{S})$. From this point of view, ξ is simply a measure depending on an extra parameter ω, belonging to some abstract probability space $(\Omega, \mathcal{A}, P)$. To ensure that the mass ξB assigned to a set B will be a random variable for every $B \in \mathcal{S}$, we need the function $\xi(\omega, B)$ on the product space $\Omega \times \mathcal{S}$ to be $\mathcal{A}$-measurable in ω for fixed B and a measure in B for fixed ω. In other words, ξ has to be a *kernel* from Ω to S. This condition is strong enough to ensure that even the integral $\xi f = \int f \, d\xi$ is a random variable, for every measurable function $f \geq 0$ on S.

The state space S is taken to be an abstract *Borel space*[2], defined by the existence of a bi-measurable $1-1$ mapping between S and a Borel set $B \subset \mathsf{R}$. This covers most cases of interest[3], since every measurable subset of a Polish space is known to be Borel. We also need to equip S with a *localizing structure*, consisting of a ring $\hat{\mathcal{S}} \subset \mathcal{S}$ of *bounded* measurable subsets. When S is a separable and complete metric space, we may choose $\hat{\mathcal{S}}$ as the class of bounded Borel sets, and if S is further assumed to be locally compact, we may take $\hat{\mathcal{S}}$ to consist of all relatively compact Borel sets.

A fixed or random measure ξ on a localized Borel space $(S, \hat{\mathcal{S}})$ is said to be *locally finite*, if $\xi B < \infty$ a.s. for all $B \in \hat{\mathcal{S}}$. This will henceforth be taken as part of our definition. Thus, we *define* a random measure on $(S, \hat{\mathcal{S}})$ as a locally finite kernel from Ω to S. Equivalently, it may be defined as a random element in the space $\mathcal{M}_S$ of all locally finite measures on S, endowed with the σ-field generated by all *evaluation* maps $\pi_B \colon \mu \mapsto \mu B$ with $B \in \mathcal{S}$. The space $\mathcal{M}_S$ is again known to be Borel.

The additional structure enables us to prove more. Thus, if ξ is a locally finite random measure on a localized Borel space S, then the integral $\xi Y = \int Y d\xi$ is a random variable for every product-measurable process $Y \geq 0$ on

[1]Often confused with L^0-valued vector measures on S, such as white noise. In K(05) those are called *continuous linear random functionals*, or simply *CLRFs*.

[2]also known as a *standard* space

[3]The theory has often been developed under various metric or topological assumptions, although such a structure plays no role, except in the context of weak convergence.

S, and if Y is further assumed to be bounded, then the measure $Y \cdot \xi$, given by $(Y \cdot \xi)f = \xi(fY)$, is again a random measure on S. We may also derive an essentially unique *atomic decomposition* $\xi = \alpha + \sum_k \beta_k \, \delta_{\sigma_k}$, in terms of a *diffuse* (non-atomic) random measure α, some distinct random elements $\sigma_1, \sigma_2, \ldots$ in S, and some random weights $\beta_1, \beta_2, \ldots \geq 0$. Here δ_s denotes a unit mass[4] at $s \in S$, so that $\delta_s B = 1_B(s)$, where 1_B is the *indicator function*[5] of the set B.

When the random measure ξ is integer-valued, its atomic decomposition reduces to $\xi = \sum_k \beta_k \, \delta_{\sigma_k}$, where the coefficients β_k are now integer-valued as well. Then ξ is called a *point process* on S, the elements $\sigma_1, \sigma_2, \ldots$ are the *points* of ξ, and $\beta_1, \beta_2, \ldots$ are the corresponding *multiplicities*. We may think of ξ as representing a random *particle system*, where several particles may occupy the same site σ_k. Since ξ is locally finite, there are only finitely many particles in every bounded set. A point process ξ is said to be *simple*, if all multiplicities equal 1. Then ξ represents a locally finite random set Ξ in S, and conversely, any such set Ξ may be represented by the associated *counting random measure* ξ, where ξB denotes the number of points of Ξ in the set B. The correspondence becomes an equivalence through a suitable choice of σ-fields.

The distribution of a random measure ξ on S is determined by the class of finite-dimensional distributions $\mathcal{L}(\xi B_1, \ldots, \xi B_n)$, and hence by the distributions of all integrals $\xi f = \int f d\xi$, for any measurable functions $f \geq 0$. When ξ is a simple point process, its distribution is determined by the *avoidance* probabilities $P\{\xi B = 0\}$ for arbitrary $B \in \hat{\mathcal{S}}$, and for diffuse random measures it is given by the set of all one-dimensional distributions $\mathcal{L}(\xi B)$.

Partial information about ξ is provided by the *intensity measure* $E\xi$ and the higher order *moment measures* $E\xi^n$, and for point processes we may even consider the *factorial moment measures* $E\xi^{(n)}$, defined for simple ξ as the restrictions of $E\xi^n$ to the non-diagonal parts of S^n. In particular, ξ is a.s. diffuse iff $E\xi^2 D = 0$, and a point process ξ is a.s. simple iff $E\xi^{(2)} D = 0$, where D denotes the diagonal in S^2.

— — —

So far we have summarized the main ideas of the first two chapters, omitting some of the more technical topics. In **Chapter 3** we focus on some basic processes of special importance. Most important are of course the *Poisson processes*, defined as point processes ξ, such that the random variables $\xi B_1, \ldots, \xi B_n$ are independent and Poisson distributed, for any disjoint sets $B_1, \ldots, B_n \in \hat{\mathcal{S}}$. In fact, when ξ is simple, independence of the increments[6] alone guarantees the Poisson property of all ξB, and likewise, the Poisson

[4]often called *Dirac measure*

[5]The term *characteristic function* should be avoided here, as it has a different meaning in probability theory.

[6]sometimes called *complete randomness* or *complete independence*

property alone guarantees the independence of the increments. The distribution of a Poisson process ξ is clearly determined by the intensity $E\xi$, which is automatically locally finite.

Closely related to the Poisson processes are the *binomial processes*[7], defined as point processes of the form $\xi = \delta_{\sigma_1} + \cdots + \delta_{\sigma_n}$, where $\sigma_1, \ldots, \sigma_n$ are i.i.d. random elements in S. In particular, a Poisson process on a bounded set is a mixture[8] of binomial processes based on a common distribution, and any Poisson process can be obtained by patching together mixed binomial processes of this kind. Mixtures of Poisson processes with different intensities, known as *Cox processes*[9], play an equally fundamental role. The classes of Poisson and Cox processes are preserved under measurable transformations and randomizations, and they arise in the limit under a variety of thinning and displacement operations.

Apart from the importance of Poisson processes to model a variety of random phenomena, such processes also form the basic building blocks for construction of more general processes, similar to the role of Brownian motion in the theory of continuous processes. Most striking is perhaps the representation of an *infinitely divisible* random measure or point process ξ as a *cluster process* $\int \mu\, \eta(d\mu)$ (in the former case apart from a trivial deterministic component), where the clusters μ are generated by a Poisson process η on $\mathcal{M}_S$ or $\mathcal{N}_S$, respectively. Thus, the distribution of ξ is essentially determined by the intensity $\lambda = E\eta$, known as the *Lévy measure*[10] of ξ. This leads to a simple interpretation of the celebrated *Lévy–Khinchin* representation of infinitely divisible distributions. Cluster processes of various kinds, in their turn, play a fundamental role within the theory of branching processes.

— — —

Chapter 4 deals with convergence in distribution of random measures, which is where the metric or topological structure of S comes in. The basic assumption is to take S to be a separable and complete metric space with Borel σ-field $\mathcal{S}$, and let $\hat{\mathcal{S}}$ be the subclass of bounded Borel sets. The metric topology on S induces the *vague topology* on $\mathcal{M}_S$, generated by the integration maps $\pi_f \colon \mu \mapsto \mu f$ for all bounded continuous functions $f \geq 0$ on S with bounded support, so that $\mu_n \xrightarrow{v} \mu$ iff $\mu_n f \to \mu f$ for any such f.

The vague topology makes even $\mathcal{M}_S$ a Polish space, which allows us to apply the standard theory of weak convergence to random measures ξ_n and ξ on S. The associated convergence in distribution[11], written as $\xi_n \xrightarrow{vd}$

[7] in early literature often referred to as *sample* or *Bernoulli processes*, or, when $S = \mathsf{R}$, as processes with the *order statistics property*

[8] Strictly speaking, we are mixing the *distributions*, not the processes.

[9] originally called *doubly stochastic Poisson processes*

[10] sometimes called the *canonical* or *KLM measure*, or in German *Schlangemaß*

[11] Often confused with *weak convergence*. Note that $\xi_n \xrightarrow{d} \xi$ iff $\mathcal{L}(\xi_n) \xrightarrow{w} \mathcal{L}(\xi)$. The distinction is crucial here, since the distribution of a random measure is a measure on a measure space.

ξ, means that $Eg(\xi_n) \to Eg(\xi)$ for any bounded and vaguely continuous function g on $\mathcal{M}_S$. In particular, $\xi_n \xrightarrow{vd} \xi$ implies $\xi_n f \xrightarrow{d} \xi f$ for any bounded continuous function $f \geq 0$ with bounded support. Quite surprisingly, the latter condition is also sufficient for the convergence $\xi_n \xrightarrow{vd} \xi$. Thus, no extra tightness condition is needed, which makes applications of the theory pleasingly straightforward and convenient.

We may now derive criteria for the convergence $\sum_j \xi_{nj} \xrightarrow{vd} \xi$, when the ξ_{nj} form a *null array*[12] of random measures on S, in the sense that the ξ_{nj} are independent in j for fixed n, and $\xi_{nj} \xrightarrow{vd} 0$ as $n \to \infty$, uniformly in j. When the ξ_{nj} are point processes and ξ is Poisson with $E\xi = \lambda$, we get in particular the classical criteria

$$\sum\nolimits_j P\{\xi_{nj}B = 1\} \to \lambda B, \qquad \sum\nolimits_j P\{\xi_{nj}B > 1\} \to 0,$$

for arbitrary $B \in \hat{\mathcal{S}}$ with $\lambda \partial B = 0$. More generally, we may characterize the convergence to any infinitely divisible random measure ξ, where the conditions simplify in various ways under assumptions of simplicity or diffuseness.

Beside the topological notion of distributional convergence, $\xi_n \xrightarrow{vd} \xi$, there is also a strong, non-topological notion of *locally uniform convergence in distribution*, written as $\xi_n \xrightarrow{uld} \xi$, and defined by the condition $\|\mathcal{L}(1_B\xi_n) - \mathcal{L}(1_B\xi)\| \to 0$ for arbitrary $B \in \hat{\mathcal{S}}$, where $\|\cdot\|$ denotes the total variation norm for signed measures on $\mathcal{M}_S$. Again we may derive some necessary and sufficient conditions for convergence, which reveal a striking analogy between the two cases. Both modes of convergence are of great importance in subsequent chapters.

— — —

In **Chapter 5** we specialize to stationary random measures on a Euclidean space $S = \mathsf{R}^d$, where *stationarity*[13] means that $\theta_r \xi \overset{d}{=} \xi$ for all $r \in S$. Here the *shift operators* θ_r on $\mathcal{M}_S$ are given by $(\theta_r \mu)f = \mu(f \circ \theta_r)$, where $\theta_r s = s + r$ for all $r, s \in S$. Note that stationarity of ξ implies *invariance* of the intensity measure $E\xi$, in the sense that $\theta_r E\xi = E\xi$ for all $r \in S$. Invariant measures on R^d are of course proportional to the d-dimensional Lebesgue measure λ^d.

Our first aim is to develop the theory of *Palm measures* Q_ξ, here defined by the formula

$$Q_\xi f = E \int_{I_1} f(\theta_{-r})\, \xi(dr), \quad f \geq 0,$$

where $I_1 = [0,1]^d$, and the function f is understood to be measurable. The measure Q_ξ is always σ-finite, and when $0 < E\xi I_1 < \infty$ it can be normalized into a *Palm distribution* $\hat{Q}_\xi$.

---

[12]Feller's term, here preferred to Loève's *uniformly asymptotically negligible* (u.a.n.), Chung's *holospoudic*, and the term *infinitesimal* used in MKM_2

[13]Often confused with *invariance*. Note that ξ is stationary iff $\mathcal{L}(\xi)$ is invariant.

The latter is especially important when ξ is a simple, stationary point process on R with finite and positive intensity. Writing η for a point process on R with distribution $\hat{Q}_\xi$, and letting $\cdots < \tau_{-1} < \tau_0 < \tau_1 < \cdots$ be the points of η with $\tau_0 = 0$, we show that η is *cycle-stationary*, in the sense that the sequence of *spacing variables* $\tau_k - \tau_{k-1}$ is again stationary. In fact, the Palm transformation essentially provides a $1-1$ correspondence between the distributions of stationary and cycle-stationary point processes on R.

This suggests that, for general random measures ξ on R, we may introduce an associated *spacing random measure* $\tilde{\xi}$, such that ξ and $\tilde{\xi}$ are simultaneously stationary, and the spacing transformation of $\tilde{\xi}$ essentially leads back to ξ. For simple point processes on R^d, we further derive some basic approximation properties, justifying the classical interpretation of the Palm distribution of ξ as the conditional distribution, given that ξ has a point at 0.

A second major theme of the chapter is the ergodic theory, for stationary random measures ξ on R^d. Using the multi-variate ergodic theorem, we show that the averages $\xi B_n/\lambda^d B_n$ converge a.s. to a random limit $\bar{\xi} \geq 0$, known as the *sample intensity* of ξ, for any increasing sequence of convex sets $B_n \subset \mathsf{R}^d$ with inner radii $r_n \to \infty$. This provides a point of departure for a variety of weak or strong limit theorems, involving stationary random measures, along with their Palm and spacing measures. The ergodic theorem also yields the most general version to date of the classical *ballot theorem.*

— — —

Palm measures remain important for general random measures, even without any assumption of stationarity or invariance. The general notion is treated in **Chapter 6**, where for any random measure ξ on S and random element η in T, we define the associated *Palm measures* $\mathcal{L}(\eta \,\|\, \xi)_s$ by disintegration of the *Campbell measure* $C_{\xi,\eta}$ on $S \times T$, as in[14]

$$\begin{aligned} C_{\xi,\eta} f &= E \int \xi(ds)\, f(s,\eta) \\ &= \int E\xi(ds)\, E\{f(s,\eta) \,\|\, \xi\}_s, \end{aligned}$$

for any measurable function $f \geq 0$ on $S \times T$. When ξ is a simple point process with σ-finite intensity $E\xi$, we may think of $\mathcal{L}(\eta \,\|\, \xi)_s$ as the conditional distribution of η, given that ξ has an atom at s, and when $\xi = \delta_\sigma$ it agrees with the conditional distribution $\mathcal{L}(\eta \,|\, \sigma)_s$.

Replacing ξ by the product measure ξ^n on S^n, we obtain the associated *n-th order Palm measures* $\mathcal{L}(\eta \,\|\, \xi^n)_s$, for $s \in S^n$. When ξ is a point process and $\eta = \xi$, the latter are a.e. confined to measures $\mu \in \mathcal{N}_S$ with atoms at $s_1, \ldots, s_n$, which suggests that we consider instead the *reduced Palm measures*, obtained by disintegration of the *reduced Campbell measures*

$$C_\xi^{(n)} f = E \int \xi^{(n)}(ds)\, f\Big(s,\, \xi - \textstyle\sum_{k\leq n} \delta_{s_k}\Big), \quad f \geq 0.$$

[14]A version of this formula has often been referred to as the *refined Campbell theorem.*

By symmetry, we may regard the multi-variate Palm measures as functions of the bounded point measures $\mu = \sum_{k\le n} \delta_{s_k}$, and by combination we see that all such measures $\mathcal{L}(\xi \,\|\, \xi)_\mu$ arise by disintegration of the *compound Campbell measure*

$$C_\xi f = E \sum_{\mu \le \xi} f(\mu, \xi - \mu) = \sum_{n=0}^{\infty} \frac{C_\xi^{(n)} f}{n!}, \quad f \ge 0,$$

which will also play an important role on Chapter 8.

Of the many remarkable properties of general Palm measures, we may comment specifically on some basic duality principles. Then consider the *dual disintegrations*

$$\begin{aligned} C_{\xi,\eta} &= \nu \otimes \mathcal{L}(\eta \,\|\, \xi) \\ &\cong \mathcal{L}(\eta) \otimes E(\xi \,|\, \eta), \end{aligned}$$

where ν is a supporting measure of ξ. Here $E(\xi|\,\eta) \ll \nu$ a.s. iff $\mathcal{L}(\eta \,\|\, \xi)_s \ll \mathcal{L}(\eta)$ a.e. ν, in which case we may choose the two density functions on $S \times T$ to agree. Taking η to be the identity map on Ω with filtration $(\mathcal{F}_t)$, we conclude from the stated equivalence that

$$E(\xi \,|\, \mathcal{F}_t) = M_t \cdot E\xi \ \text{ a.s.} \quad \Leftrightarrow \quad P(\mathcal{F}_t \,\|\, \xi)_s = M_t^s \cdot P \ \text{ a.e.},$$

for some product-measurable function M_t^s on $S \times \mathsf{R}_+$. Assuming S to be Polish, we prove that $P(\mathcal{F}_t \| \, \xi)_s$ is continuous in total variation in s for fixed t iff M_t^s is L^1-continuous in s, whereas $P(\mathcal{F}_t \| \, \xi)_s$ is consistent in t for fixed s iff M_t^s is a martingale in t. Such results will play a crucial role in some later chapters.

— — —

Chapter 7 deals with random measures ξ on S that are stationary under the action of some abstract measurable group G. Under suitable regularity conditions, we may then choose the Palm measures of ξ to form an invariant kernel from S to $\mathcal{M}_S$. More generally, assuming G to act measurably on S and T, and considering a jointly stationary pair of a random measure ξ on S and a random element η in T, we may look for an invariant kernel $\mu\colon S \to T$, representing the Palm measures of η with respect to ξ. Here the *invariance* is defined by $\mu_{rs} = \mu_s \circ \theta_r^{-1}$, or in explicit notation

$$\int \mu_{rs}(dt) f(t) = \int \mu_s(dt) f(rt), \quad r \in G, \ s \in S,$$

where $f \ge 0$ is an arbitrary measurable function on T.

When $S = G$, we may construct the Palm measure at the identity element $\iota \in G$ by a simple *skew transformation*. Similar methods apply when $S = G \times S'$ for some Borel space S', in which case the entire kernel is determined by the invariance relation. Various devices are helpful to deal with more general spaces, including the notion of *inversion kernel*, which maps every invariant measure on $S \times T$ into an invariant measure on a space $G \times A \times T$.

Many striking properties and identities are known for invariant disintegration kernels, translating into properties of invariant Palm measures. In particular, we may provide necessary and sufficient conditions for a given kernel to be the Palm kernel of some stationary random measure, in which case there is also an explicit inversion formula. Another classical result is the celebrated *exchange formula*, relating the Palm distributions $\mathcal{L}(\xi \,\|\, \eta)$ and $\mathcal{L}(\eta \,\|\, \xi)$, for any jointly stationary random measures ξ and η.

From invariant disintegrations, we may proceed to the more challenging problem of *stationary disintegrations*. Given some jointly stationary random measures ξ on S and η on $S \times T$ satisfying $\eta(\cdot \times T) \ll \xi$ a.s., we are then looking for an associated disintegration kernel ζ, such that the triple (ξ, η, ζ) is stationary. Using the representation of G in terms of projective limits of Lie groups, we show that such a kernel ζ exists when G is locally compact. In particular, we conclude that if ξ and η are jointly stationary random measures on S satisfying $\eta \ll \xi$ a.s., then $\eta = Y \cdot \xi$ a.s. for some product-measurable process $Y \geq 0$ on S such that (ξ, η, Y) is stationary.

— — —

The reduced Palm measures Q_μ of a point process ξ were defined in Chapter 6 through disintegration of the compound Campbell measure C, as in

$$\begin{aligned} Cf &= E \sum_{\mu \leq \xi} f(\mu, \xi - \mu) \\ &= \int \nu(d\mu)\, Q_\mu(d\mu')\, f(\mu, \mu'), \end{aligned}$$

for any measurable function $f \geq 0$ on $\hat{\mathcal{N}}_S \times \mathcal{N}_S$. When ξ is simple, we may interpret Q_μ as the conditional distribution of $\xi - \mu$ given $\mu \leq \xi$ a.s., which suggests the remarkable formula

$$\mathcal{L}(1_{B^c}\xi \mid 1_B \xi) = Q_{1_B \xi}(\,\cdot \mid \mu B = 0) \;\text{ a.s.}, \quad B \in \hat{\mathcal{S}}.$$

In this sense, the Palm measures Q_μ govern the laws of *interior conditioning* $\mathcal{L}(1_{B^c}\xi \mid 1_B\xi)$. In **Chapter 8** we study the corresponding *exterior laws* $\mathcal{L}(1_B\xi \mid 1_{B^c}\xi)$, of special importance in statistical mechanics.

Though for bounded S we may simply interchange the roles of B and B^c, the general construction relies on the notion of *Gibbs kernel* $\Gamma = G(\xi, \cdot)$, defined by the *maximal dual disintegration*

$$\begin{aligned} E\,\Gamma f(\cdot, \xi) &= E \int G(\xi, d\mu)\, f(\mu, \xi) \\ &\leq E \sum_{\mu \leq \xi} f(\mu, \xi - \mu). \end{aligned}$$

The exterior conditioning may now be expressed by the equally remarkable formula

$$\mathcal{L}(1_B\xi \mid 1_{B^c}\xi) = \Gamma(\,\cdot \mid \mu B^c = 0) \;\text{ a.s. on }\; \{\xi B = 0\}, \quad B \in \hat{\mathcal{S}},$$

showing how the conditional distributions on the left can be recovered, by elementary conditioning, from a single random measure Γ.

Restricting Γ to the set of unit masses $\mu = \delta_s$, we are essentially led to the notion of *Papangelou kernel* η, which is again a random measure on S. We may often use η to draw conclusions about some distributional properties of ξ. In particular, fixing a diffuse measure λ on S, we show that if η is a.s. invariant under any λ-preserving transformation of S, then ξ itself is λ-*symmetric*, in the sense that $\xi \circ f^{-1} \stackrel{d}{=} \xi$ for any such transformation f. For unbounded λ, it follows that ξ is a Cox process directed by η. Such results have proved to be especially useful in stochastic geometry.

From the Papangelou kernel η, we may proceed to the closely related notion of *external intensity*[15] ζ, which also arises in the limit from various sums of conditional probabilities and expectations. We may also characterize ζ as the *dual external projection* of ξ, in the sense that $E\xi Y = E\zeta Y$, for any *externally measurable*[16] process $Y \geq 0$ on S. To avoid technicalities, we postpone the precise definitions.

— — —

Chapter 9 deals with the dynamic or martingale aspects of random measures ξ on a product space $\mathsf{R}_+ \times S$. Assuming ξ to be adapted to a filtration $\mathcal{F} = (\mathcal{F}_t)$, we may introduce the associated *compensator* η, defined as a predictable random measure on the same space specifying the rate of random evolution of ξ. The compensator η of a simple point process ξ on R_+ plays a similar role as the quadratic variation $[M]$ of a continuous local martingale M. Thus, if ξ is further assumed to be *quasi-leftcontinuous*, in the sense that $\xi\{\tau\} = 0$ a.s. for every predictable time τ, it may be reduced to Poisson through a *random time change* determined by η. In particular, ξ itself is then Poisson iff η is a.s. non-random.

A related but deeper result is the *predictable mapping theorem*, asserting that, whenever ξ is a random measure on $[0,1]$ or R_+ satisfying $\xi \circ f^{-1} \stackrel{d}{=} \xi$ for every measure-preserving transformation f, the same relation holds with f replaced by any predictable map V with measure-preserving paths. Further invariance properties of this kind may be stated most conveniently in terms of the *discounted compensator* ζ, obtainable from η as the unique solution to *Doléans' differential equation* $Z = 1 - Z_- \cdot \eta$, where $Z_t = 1 - \zeta(0,t]$.

The simplest case is when ξ is a single point mass $\delta_{\tau,\chi}$, for some optional time[17] τ with associated mark χ in S. We may then establish a unique integral representation $\mathcal{L}(\tau,\chi,\eta) = \int P_\mu \, \nu(d\mu)$, where P_μ is the distribution when $\mathcal{L}(\tau,\chi) = \mu$, and η is the compensator of (τ,χ) with respect to the *induced* filtration. Equivalently, ζ can be extended to a random probability measure ρ on $\mathsf{R}_+ \times S$ satisfying $\mathcal{L}(\tau,\chi \,|\, \rho) = \rho$ a.s., in which case $\mathcal{L}(\rho) = \nu$.

[15]sometimes called the *stochastic intensity*, often confused with the Papangelou kernel
[16]also called *visible* or *exvisible*
[17]also called a *stopping time*

We finally consider some basic results for *tangential processes*[18], defined as pairs of processes with the same *local characteristics.* Here the main results are the *tangential existence* and *comparison* theorems, where the former guarantees the existence, for every semi-martingale X, of a tangential process $\tilde{X}$ with conditionally independent increments, whereas the latter shows how some basic asymptotic properties are related for tangential processes. Combining those results, we may often reduce the study of general random measures ξ to the elementary case where ξ has independent increments.

— — —

The purpose of **Chapter 10** is to study multiple integrals of the form

$$\xi_1 \cdots \xi_d f = \int \cdots \int \xi_1(ds_1) \cdots \xi_d(ds_d)\, f(s_1, \ldots, s_d),$$

where $\xi_1, \ldots, \xi_d$ are random measures on a Borel space S and f is a measurable function on S^d. When $\xi_k = \xi$ for all k, we may write the integral as $\xi^d f$. Starting with the case of Poisson or more general point processes $\xi_1, \ldots, \xi_d$ with independent increments, we proceed first to symmetric point processes, then to positive or symmetric Lévy processes, and finally to broad classes of more general processes.

Our main problems are to find necessary and sufficient conditions for the existence of the integral $\xi_1 \cdots \xi_d f$, and for the convergence to 0 of a sequence of such integrals. In the case of independent Poisson processes $\xi_1, \ldots, \xi_d$ and nonnegative integrands f or f_n, we can use elementary properties of Poisson processes to give exact criteria in both cases, expressed in terms of finitely many Lebesgue-type integrals. The same criteria apply to any point processes with independent increments. A simple decoupling argument yields an immediate extension to the integrals $\xi^d f$.

Using some basic properties of *random multi-linear forms*, we can next derive similar criteria for the integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$, where the $\tilde{\xi}_k$ are conditionally independent symmetrizations of $\xi_1, \ldots, \xi_d$. It now becomes straightforward to handle the case of positive or symmetric Lévy processes. The extension to more general processes requires the sophisticated methods of tangential processes, developed in the previous chapter. Since multiple series can be regarded as special multiple integrals, we can also derive some very general criteria for the former.

We also include a section dealing with *escape* conditions of the form $|\xi^d f_n| \xrightarrow{P} \infty$ or $|\xi_1 \cdots \xi_d f_n| \xrightarrow{P} \infty$. Here it is often difficult, even in simple cases, to find precise criteria, and we are content to provide some partial results and comparison theorems, using concentration inequalities and other subtle tools of elementary probability theory.

— — —

[18] often confused with *tangent processes*

Much of the theory developed so far was originally motivated by applications. The remainder of the book deals with specific applications of random measure theory to three broad areas, beginning in **Chapter 11** with some aspects of random *line* and *flat processes* in Euclidean spaces, a major subfield of *stochastic geometry*. Here the general idea is to regard any random collection of geometrical objects as a point process on a suitable parameter space. The technical machinery of the previous chapters then leads inevitably to some more general random measures on the same space.

A *line process* ξ in R^d is a random collection of straight lines in R^d. More generally, we may consider *flat processes* ξ in R^d, consisting of random, k-dimensional, affine subspaces, for arbitrary $1 \leq k < d$. We always assume ξ to be *locally finite*, in the sense that at most finitely many lines or flats pass through any bounded Borel set in R^d. We say that ξ is *stationary*, if its distribution is invariant under shifts on the underlying space. Identifying the lines or flats with points in the parameter space S, we may regard ξ as a point process on S. The choice of parametrization is largely irrelevant and may depend on our imminent needs.

Already for stationary line processes ξ in R^2 satisfying some mild regularity conditions, we can establish a remarkable moment identity with surprising consequences. In particular, it implies the existence of a Cox line process ζ with the same first and second order moment measures, which suggests that ξ might have been a Cox process to begin with. Though this may not be true in general, it does hold under additional regularity assumptions. The situation for more general flat processes is similar, which leads to the fundamental problem of finding minimal conditions on a stationary k-flat process ξ in R^d, ensuring ξ to be a Cox process directed by some invariant random measure.

Most general conditions known to date are expressed in terms of the Papangelou kernel η of ξ. From Chapter 8 we know that ξ is a Cox process of the required type, whenever η is a.s. invariant, where the latter property is again defined with respect to shifts on the underlying space R^d. This leads to the simpler—though still fiercely difficult—problem of finding conditions on a stationary random measure η on the set of k-flats in R^d that will ensure its a.s. invariance. Here a range of methods are available.

The easiest approach applies already under some simple *spanning conditions*, which can only be fulfilled when $k \geq d/2$. In the harder case of $k < d/2$, including the important case of line processes in R^d with $d \geq 3$, the desired a.s. invariance can still be established, when the projections of η onto the linear subspace of directions are a.s. absolutely continuous with respect to some sufficiently regular fixed measure. The general case leads to some subtle considerations, involving certain *inner* and *outer degeneracies*.

We may finally comment on the obvious connection with random particle system. Here we consider an infinite system of particles in R^d, each moving indefinitely with constant velocity. The particles will then trace out straight lines in a $(d+1)$-dimensional space-time diagram, thus forming a line process in R^{d+1}. Assuming the entire system ξ of positions and velocities to be stationary in R^d at time $t = 0$, we may look for conditions ensuring that ξ will approach a steady-state distribution as $t \to \infty$. The limiting configuration is then stationary in all $d+1$ directions, hence corresponding to a stationary line process in R^{d+1}. This provides a useful dynamical approach to the previous invariance problems for line processes.

— — —

Another case where applications of the previous theory have led to important developments is for regenerative processes, treated in **Chapter 12**. Here the general setup involves a suitably regular process X on R_+ that is *regenerative* at some fixed state a, in the sense that for any optional time $\tau < \infty$ with $X_\tau = a$ a.s., we have $\mathcal{L}(\theta_\tau X \,|\, \mathcal{F}_\tau) = \mathcal{L}(X)$ a.s. In other words, X is assumed to satisfy the strong Markov property at visits to a. This is of course true when the entire process X is strong Markov, but the theory applies equally to the general case. A familiar example is provided by a standard Brownian motion, which is clearly regenerative under visits to 0.

Most elementary is the case of *renewal processes*, where the *regenerative set* $\Xi = \{t \geq 0;\, X_t = a\}$ is discrete and unbounded. Here the central result is of course the classical *renewal theorem*, which can be extended to a statement about the *occupation measure* of any transient random walk in R. Assuming this case to be well known, we move on to the equally important and more challenging case, where the closure of Ξ is perfect, unbounded, and nowhere dense. For motivation, we may keep in mind our favorite example of Brownian motion.

In this case, there exists a *local time* random measure ξ with support $\bar{\Xi}$, enjoying a similar regenerative property. Furthermore, the excursion structure of X is described by a stationary Poisson process η on the product space $\mathsf{R}_+ \times D_0$ with intensity measure $\lambda \otimes \nu$, such that a point of η at (s, x) encodes an excursion path $x \in D_0$, spanning the time interval where $\xi[0, t] = s$. Here ν is the celebrated *Itô excursion law*, a σ-finite measure on the space D_0 of excursion paths, known to be unique up to a normalization.

Note that we are avoiding the traditional understanding of local time as a non-decreasing process L, favoring instead a description in terms of the associated random measure ξ. This somewhat unusual viewpoint opens the door to applications of the previous random measure theory, including the powerful machinery of Palm distributions. The rewards are great, since the latter measures of arbitrary order turn out to play a central role in the analysis of local hitting and conditioning, similar to their role for simple point processes in Chapter 6.

To indicate the nature of those results, fix any times $0 = t_0 < t_1 < \ldots < t_n$, such that $E\xi$ has a continuous density p around each point $t_k - t_{k-1}$. The hitting probabilities $P \bigcap_k \{\xi I_k > 0\}$ are then given, asymptotically as $I_k \downarrow \{t_k\}$ for each k, by some simple expressions involving p, and the corresponding conditional distributions of X agree asymptotically with the associated multivariate Palm distributions. In this context, the latter measures can also be factored into univariate components, which enjoy a range of useful continuity properties.

— — —

It is only fitting that the book ends, in **Chapter 13**, with a long discourse on branching and super-processes, an area where all aspects of the previous random measure theory come into play. Since any comprehensive account would require a book-length treatment, we are forced to state some basic existence and structural propositions without proof, focusing instead on areas where the central ideas of random measure theory play an especially prominent role.

Our starting point is a *branching Brownian motion* in R^d, where the life lengths are independent and exponentially distributed with rate 2, and each particle either dies or splits into two, with equal probability $\frac{1}{2}$. We also assume the spatial movements of the individual particles to be given by independent Brownian motions.

We may now perform a scaling, where the particle density and branching rate are both increased by a factor n, whereas the weight of each particle is reduced by a factor n^{-1}. As $n \to \infty$, we get in the limit a *Dawson–Watanabe super-process* (or *DW-process* for short), which may be thought of as a randomly evolving diffuse cloud. The original discrete tree structure is gone in the limit, and when $d \geq 2$, the mass distribution at time $t > 0$ is given by an a.s. diffuse, singular random measure ξ_t of Hausdorff dimension 2.

It is then quite remarkable that the discrete genealogical structure of the original discrete process persists in the limit, leading to a fundamental cluster structure of the entire process. Thus, for fixed $t > 0$, the *ancestors* of ξ_t at an earlier time $s = t - h$ form a Cox process ζ_s^t directed by $h^{-1}\xi_s$. Even more amazingly, the collection of ancestral processes ζ_s^t with $s < t$ forms an inhomogeneous *Yule branching Brownian motion*, approximating ξ_t as $s \to t$. The individual ancestors give rise to i.i.d. *clusters*, and the resulting cluster structure constitutes a powerful tool for analysing the process.

Among included results, we note in particular the basic *Lebesgue approximation*, which shows how ξ_t can be approximated, up to a normalizing factor, by the restriction of Lebesgue measure λ^d to an ε-neighborhood of the support Ξ_t. We can also establish some local hitting and conditioning properties of the DW-process, similar to those for simple point processes in Chapter 6, and for regenerative processes in Chapter 12. For $d \geq 3$, the approximating random measure $\tilde{\xi}$ is a space-time stationary version of the process. Though no such version exists when $d = 2$, the indicated approximations still hold with $\tilde{\xi}$ replaced by a stationary version $\tilde{\eta}$ of the canonical cluster.

Our proofs of the mentioned results rely on a careful analysis of the multivariate moment measures with associated densities. The local conditioning property also requires some sufficiently regular versions of the multi-variate

Palm distributions, derived from the conditional moment densities via the general duality theory of Chapter 6.

The moment measures exhibit some basic recursive properties, leading to some surprising and useful representations in terms of certain *uniform Brownian trees*, established by various combinatorial and martingale arguments. A deeper analysis, based on an extension of Le Gall's *Brownian snake*, reveals an underlying *Palm tree*[19] *representation*, which provides a striking connection between higher order historical Campbell measures and appropriate conditional distributions of the uniform Brownian tree.

— — — ◇ — — — ◇ — — —

This is where the book ends, but certainly not the subject. The quest goes on.

[19] named after Conny Palm—no relation to subtropical forestry

We conclude with a short list of some commonly used notation. A more comprehensive list will be found at the end of the volume.

$\mathsf{N} = \{1, 2, \ldots\}$, $\mathsf{Z}_+ = \{0, 1, 2, \ldots\}$, $\mathsf{R}_+ = [0, \infty)$, $\bar{\mathsf{R}} = [-\infty, \infty]$,

$(S, \mathcal{S}, \hat{\mathcal{S}})$: localized Borel space, classes of measurable or bounded sets,

$\mathcal{S}_+$: class of $\mathcal{S}$-measurable functions $f \geq 0$,

$S^{(n)}$: non-diagonal part of S^n,

$\mathcal{M}_S$, $\hat{\mathcal{M}}_S$: class of locally finite or bounded measures on S,

$\mathcal{N}_S$, $\hat{\mathcal{N}}_S$: class of integer-valued measures in $\mathcal{M}_S$ or $\hat{\mathcal{M}}_S$,

$\mathcal{M}_S^*$, $\mathcal{N}_S^*$: classes of diffuse measures in $\mathcal{M}_S$ and simple ones in $\mathcal{N}_S$,

G, λ: measurable group with Haar measure, Lebesgue measure on R,

$\delta_s B = 1_B(s) = 1\{s \in B\}$: unit mass at s and indicator function of B,

$\mu f = \int f \, d\mu$, $(f \cdot \mu)g = \mu(fg)$, $(\mu \circ f^{-1})g = \mu(g \circ f)$, $1_B \mu = 1_B \cdot \mu$,

$\mu^{(n)}$: when $\mu \in \mathcal{N}_S^*$, the restriction of μ^n to $S^{(n)}$,

$(\theta_r \mu)f = \mu(f \circ \theta_r) = \int \mu(ds) f(rs)$,

$(\nu \otimes \mu)f = \int \nu(ds) \int \mu_s(dt) f(s, t)$, $(\nu\mu)f = \int \nu(ds) \int \mu_s(dt) f(t)$,

$(\mu * \nu)f = \int \mu(dx) \int \nu(dy) f(x + y)$,

$(E\xi)f = E(\xi f)$, $E(\xi|\mathcal{F})g = E(\xi g|\mathcal{F})$,

$\mathcal{L}(\,)$, $\mathcal{L}(\,|\,)_s$, $\mathcal{L}(\,\|\,)_s$: distribution, conditional or Palm distribution,

$C_\xi f = E \sum_{\mu \leq \xi} f(\mu, \xi - \mu)$ with bounded μ,

$\perp\!\!\!\perp$, $\perp\!\!\!\perp_{\mathcal{F}}$: independence, conditional independence given $\mathcal{F}$,

$\overset{d}{=}$, $\xrightarrow{d}$: equality and convergence in distribution,

$\xrightarrow{w}$, $\xrightarrow{v}$, $\xrightarrow{u}$: weak, vague, and uniform convergence,

$\xrightarrow{wd}$, $\xrightarrow{vd}$: weak or vague convergence in distribution,

$\|f\|$, $\|\mu\|$: supremum of $|f|$ and total variation of μ.

http://www.springer.com/978-3-319-41596-3

Random Measures, Theory and Applications
Kallenberg, O.
2017, XXVIII, 680 p., Hardcover
ISBN: 978-3-319-41596-3