Probability Theory and Stochastic Modelling 77

Olav Kallenberg

Random Measures, Theory and Applications

Springer

Probability Theory and Stochastic Modelling

Volume 77

The **Probability Theory and Stochastic Modelling** series is a merger and continuation of Springer's two well established series Stochastic Modelling and Applied Probability and Probability and Its Applications series. It publishes research monographs that make a significant contribution to probability theory or an applications domain in which advanced probability methods are fundamental. Books in this series are expected to follow rigorous mathematical standards, while also displaying the expository quality necessary to make them useful and accessible to advanced students as well as researchers. The series covers all aspects of modern probability theory including

- Gaussian processes
- Markov processes
- Random fields, point processes and random sets
- Random matrices
- Statistical mechanics and random media
- Stochastic analysis

as well as applications that include (but are not restricted to):

- Branching processes and other models of population growth
- Communications and processing networks
- Computational methods in probability and stochastic processes, including simulation
- Genetics and other stochastic models in biology and the life sciences
- Information theory, signal processing, and image synthesis
- Mathematical economics and finance
- Statistical methods (e.g. empirical processes, MCMC)
- Statistics for stochastic processes
- Stochastic control
- Stochastic models in operations research and stochastic optimization
- Stochastic models in the physical sciences

More information about this series at http://www.springer.com/series/13205

Olav Kallenberg

Random Measures, Theory and Applications

Springer

Olav Kallenberg
Department of Mathematics and Statistics
Auburn University
Auburn, Alabama
USA

ISSN 2199-3130 ISSN 2199-3149 (electronic)
Probability Theory and Stochastic Modelling
ISBN 978-3-319-41596-3 ISBN 978-3-319-41598-7 (eBook)
DOI 10.1007/978-3-319-41598-7

Library of Congress Control Number: 2017936349

Mathematics Subject Classification (2010): 60G55, 60G57

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

I am dedicating this book to my early mentor, colleague, and friend Peter Jagers, whose influence inspired me long ago to get started in mathematics, and whose faithful support has sustained me through more than four decades.

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

Contents

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 $\pi_B : \mu \mapsto \mu B$ 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 $\pi_f : \mu \mapsto \mu f$ 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 $\theta_r \xi \stackrel{d}{=} \xi$ 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 $\bar{\xi} \ge 0$, 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 $f \ge 0$ 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 $s_1, \dots, s_n$, which suggests that we consider instead the *reduced Palm measures*, obtained by disintegration of the *reduced Campbell measures*

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

[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 $1 \le k < d$. 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 $k \ge d/2$. In the harder case of $k < d/2$, including the important case of line processes in R^d with $d \ge 3$, 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 $\mathcal{L}(\theta_\tau X \mid \mathcal{F}_\tau) = \mathcal{L}(X)$ 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.

Chapter 1

Spaces, Kernels, and Disintegration

The purpose of this chapter is to introduce some underlying framework and machinery, to form a basis for our subsequent development of random measure theory. The chapter also contains some more technical results about differentiation and disintegration, needed only for special purposes. The impatient reader may acquire some general familiarity with the basic notions and terminology from the following introduction, and then return for further details and technical proofs when need arises.

To ensure both sufficient generality and technical flexibility, we take the underlying space S to be an abstract *Borel space*, defined as a measurable space that is *Borel isomorphic* to a Borel set B on the real line R. In other words, we assume the existence of a $1-1$, bi-measurable mapping between S and B. In Theorem 1.1 we prove that every Polish space S is Borel, which implies the same property for every Borel set in S. Recall that a topological space is said to be *Polish* if it admits a separable and complete metrization. The associated σ-field $\mathcal{S}$ is understood to be the one generated by the topology, known as the *Borel σ-field.*

For technical reasons, we restrict our attention to *locally finite* measures on S. In the absence of any metric or topology on S, we then need to introduce a *localizing structure*, consisting of a ring $\hat{\mathcal{S}} \subset \mathcal{S}$ of *bounded* sets with suitable properties. In fact, it is enough to specify a sequence $S_n \uparrow S$ in $\mathcal{S}$, such that a set is bounded iff it is contained in one of the sets S_n. For metric spaces S, we may choose the S_n to be concentric balls of radii n, and when S is *lcscH* (locally compact, second countable, and Hausdorff), we may choose $\hat{\mathcal{S}}$ to consist of all relatively compact subsets of S.

The space $\mathcal{M}_S$ of all locally finite measures μ on S may be equipped with the σ-field generated by all *evaluation* maps $\pi_B : \mu \mapsto \mu B$ with $B \in \hat{\mathcal{S}}$, or equivalently by all integration maps $\pi_f : \mu \mapsto \mu f = \int f d\mu$, with $f \geq 0$ a measurable function on S. The space $\mathcal{M}_S$ is again Borel by Theorem 1.5, as is the subspace $\mathcal{N}_S$ of all integer-valued measures on S.

Every measure $\mu \in \mathcal{M}_S$ has an *atomic decomposition*

$$\mu = \alpha + \sum_{k \leq \kappa} \beta_k \, \delta_{\sigma_k}, \tag{1}$$

in terms of a *diffuse* (non-atomic) measure α, some distinct *atoms* $\sigma_k \in S$, and some real weights $\beta_k > 0$. Here δ_s denotes a unit mass at s, defined by

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_1

$\delta_s B = 1_B(s)$, where 1_B is the *indicator function* of the set B. The sum is unique up to the order of terms, and we can choose α and all the σ_k and β_k to be measurable functions of μ. When $\mu \in \mathcal{N}_S$, we have $\alpha = 0$, and $\beta_k \in \mathsf{N} = \{1, 2, \ldots\}$ for all k.

To prove such results, we may consider partitions of S into smaller and smaller subsets. To formalize the idea, we may introduce a *dissection system*, consisting of some nested partitions of S into subsets $I_{nj} \in \hat{\mathcal{S}}$, such that every bounded set is covered by finitely many sets I_{nj}, and the whole family generates the σ-field $\mathcal{S}$. The collection $\{I_{nj}\}$ clearly forms a *semi-ring*, defined as a class $\mathcal{I}$ of subsets closed under finite intersections and such that every proper difference in $\mathcal{I}$ is a finite union of disjoint $\mathcal{I}$-sets. Generating rings and semi-rings will play important roles in the sequel.

Let $\mathcal{N}_S^*$ denote the class of *simple* point measures $\mu \in \mathcal{N}_S$, where all *multiplicities* β_k in (1) equal 1. Such a measure μ may be identified with its support $\operatorname{supp}(\mu)$, which yields a $1-1$ correspondence between $\mathcal{N}_S^*$ and the class $\hat{\mathcal{F}}_S$ of all locally finite subsets of S. The correspondence becomes a Borel isomorphism, if we endow $\hat{\mathcal{F}}_S$ with the σ-field generated by all *hitting* maps $h_B \colon F \mapsto 1\{F \cap B \neq \emptyset\}$ with $B \in \hat{\mathcal{S}}$.

Given two measurable spaces $(S, \mathcal{S})$ and $(T, \mathcal{T})$, we define a *kernel* from S to T as a function $\mu \colon S \times \mathcal{T} \to [0, \infty]$, such that $\mu(s, B)$ is measurable in $s \in S$ for fixed B and a measure in $B \in \mathcal{T}$ for fixed s. It is said to be *locally finite* if $\mu(s, B) < \infty$ for all $B \in \hat{\mathcal{T}}$ and a *probability kernel* if $\mu(s, T) \equiv 1$. Though often neglected in real analysis, kernels play a fundamental role in all areas of probability theory. In the context of our present exposition, they are needed already for the definition of random measures and their transforms, and they figure prominently especially in connection with Palm measures and Gibbs kernels.

For any two kernels $\mu \colon S \to T$ and $\nu \colon T \to U$, we define their *composition* $\mu \otimes \nu$ and *product* $\mu\nu$ as kernels from S to $T \times U$ or U, respectively, given by

$$\begin{aligned}(\mu \otimes \nu)_s f &= \int \mu_s(dt) \int \nu_t(du)\, f(t, u),\\ (\mu\nu)_s f &= \int \mu_s(dt) \int \nu_t(du)\, f(u),\end{aligned}$$

so that $\mu\nu = (\mu \otimes \nu)(T \times \cdot)$. Measures on T may be regarded as kernels from a singleton space S.

Conversely, kernels often arise through the *disintegration* of measures ρ on a product space $S \times T$. Thus, when ρ is σ-finite and S and T are Borel, we may form the *dual* disintegrations $\rho = \nu \otimes \mu \cong \nu' \otimes \mu'$, in terms of some σ-finite measures ν on S and ν' on T and some kernels $\mu : S \to T$ and $\mu' \colon T \to S$, or more explicitly

$$\begin{aligned}\rho f &= \int \nu(ds) \int \mu_s(dt)\, f(s, t)\\ &= \int \nu'(dt) \int \mu'_t(ds)\, f(s, t).\end{aligned}$$

When the *projections* $\rho(\cdot \times T)$ and $\rho(S \times \cdot)$ are σ-finite, they may be chosen as our *supporting measures* ν and ν', in which case we may choose μ and μ' to be probability kernels. In general, we can choose any σ-finite measures $\nu \sim \rho(\cdot \times T)$ and $\nu' \sim \rho(S \times \cdot)$. Note that if $\rho = \mathcal{L}(\xi, \eta)$ for some random elements ξ in S and η in T with marginal distributions ν and ν', respectively, then μ and μ' are versions of the regular conditional distributions $\mathcal{L}(\eta \,|\, \xi)_s$ and $\mathcal{L}(\xi \,|\, \eta)_t$. Despite its fundamental importance, the subject of disintegration has often been neglected, or even dismissed as a technical nuisance[1].

Disintegration of kernels is more subtle. In particular, regarding the measures ρ and ν above as measurable functions of the pair (ρ, ν), hence as kernels from $\mathcal{M}_{S\times T} \times \mathcal{M}_S$ to $S \times T$ and S, respectively, we may want to choose even the disintegration kernel μ to depend measurably on ρ and ν, hence as a kernel from $S \times \mathcal{M}_{S\times T} \times \mathcal{M}_S$ to T. The existence of such a measurable disintegration is ensured by Corollary 1.26.

Even more subtle is the subject of *iterated disintegration*, clarified by Theorem 1.27, which plays an important role in the context of general Palm measures. To explain the idea in a simple case, consider some random elements ξ, η, and ζ in arbitrary Borel spaces, and introduce their marginal and conditional distributions, denoted as in

$$\mu_{12} = \mathcal{L}(\xi, \eta), \qquad \mu_{23|1} = \mathcal{L}(\eta, \zeta \,|\, \xi) \qquad \mu_{3|12} = \mathcal{L}(\zeta \,|\, \xi, \eta).$$

We wish to form versions of the kernel $\mu_{3|12}$, via disintegration of $\mu_{23|1}$ or $\mu_{13|2}$. Denoting the resulting disintegration kernels by $\mu_{3|2|1}$ and $\mu_{3|1|2}$, we prove that, under suitable hypotheses,

$$\mu_{3|12} = \mu_{3|2|1} \overset{\sim}{=} \mu_{3|1|2} \ \text{ a.e. } \mu_{12}.$$

This should not be confused with the much more elementary chain rule for conditional expectations.

Differentiation may be regarded as a special case of disintegration. Here we explain the classical dissection approach to the Lebesgue decomposition and Radon-Nikodym theorem, which leads directly to some product-measurable versions, required for the previous disintegration theory. We also review the classical theory of general *differentiation bases*, and prove some general approximation properties, useful in subsequent chapters.

1.1 Borel and Measure Spaces

Two measurable spaces $(A, \mathcal{A})$ and $(B, \mathcal{B})$ are said to be *Borel isomorphic*, if they are related by a bijective map $f : A \to B$ such that f and f^{-1} are both measurable. A *Borel space* is a measurable space $(S, \mathcal{S})$ that is Borel

[1] "The theorem about disintegration of measures has a bad reputation, and probabilists rarely employ it without saying 'we are obliged to' ..." (quoted from Dellacherie & Meyer (1975), p 125).

isomorphic to a Borel set in R. Most measures, random or not, considered in this book are defined on Borel spaces. The following result ensures this setting to be general enough for most applications of interest. Yet, it is simple and flexible enough to avoid some distracting technical subtleties. In particular, the family of Borel spaces is closed under countable unions and intersections, and any measurable subset of a Borel space is again Borel.

If nothing else is said, any topological space S is endowed with its Borel σ-field, often denoted by $\mathcal{S}$ or $\mathcal{B}_S$. A topological space S is said to *Polish* if it is separable and allows a complete metrization. This holds in particular when S is *lcscH* (locally compact, second countable, and Hausdorff), in which case we can choose the metric to be such that the relatively compact sets are precisely the metrically bounded ones.

Theorem 1.1 *(Polish and Borel spaces) Any Polish space S is*
(i) *homeomorphic to a Borel set in* $[0,1]^\infty$,
(ii) *Borel isomorphic to a Borel set in* $[0,1]$.

Proof: (i) Fix a complete metric ρ in S. We may assume that $\rho \leq 1$, since we can otherwise replace ρ by the equivalent metric $\rho \wedge 1$, which is again complete. Since S is separable, we may choose a dense sequence $x_1, x_2, \ldots \in S$. Then the mapping

$$x \mapsto \{\rho(x,x_1), \rho(x,x_2), \ldots\}, \quad x \in S,$$

defines a homeomorphic embedding of S into the compact space $K = [0,1]^\infty$, and we may regard S as a subset of K. In K we introduce the metric

$$d(x,y) = \sum\nolimits_n 2^{-n}|x_n - y_n|, \quad x, y \in K,$$

and we define $\bar{S}$ as the closure of S in K.

Writing $|B_x^\varepsilon|_\rho$ for the ρ-diameter of the d-ball $B_x^\varepsilon = \{y \in S;\ d(x,y) < \varepsilon\}$ in S, we define

$$U_n(\varepsilon) = \{x \in \bar{S};\ |B_x^\varepsilon|_\rho < n^{-1}\}, \quad \varepsilon > 0,\ n \in N,$$

and put $G_n = \bigcup_\varepsilon U_n(\varepsilon)$. The G_n are open in $\bar{S}$, since $x \in U_n(\varepsilon)$ and $y \in \bar{S}$ with $d(x,y) < \varepsilon/2$ implies $y \in U_n(\varepsilon/2)$, and $S \subset G_n$ for each n by the equivalence of the metrics ρ and d. This gives $S \subset \tilde{S} \subset \bar{S}$, where $\tilde{S} = \bigcap_n G_n$.

For any $x \in \tilde{S}$, we may choose some $x_1, x_2, \ldots \in S$ with $d(x, x_n) \to 0$. By definitions of $U_n(\varepsilon)$ and G_n, the sequence (x_k) is Cauchy even for ρ, and so by completeness $\rho(x_k, y) \to 0$ for some $y \in S$. Since ρ and d are equivalent on S, we have even $d(x_k, y) \to 0$, and therefore $x = y$. This gives $\tilde{S} \subset S$, and so $\tilde{S} = S$. Finally, $\tilde{S}$ is a Borel set in K, since the G_n are open subsets of the compact set $\bar{S}$.

(ii) Write 2^∞ for the countable product $\{0,1\}^\infty$, and let B denote the subset of binary sequences with infinitely many zeros. Then $x = \sum_n x_n 2^{-n}$ defines a $1-1$ correspondence between $I = [0,1)$ and B. Since $x \mapsto (x_1, x_2, \ldots)$

is clearly bi-measurable, I and B are Borel isomorphic, written as $I \sim B$. Furthermore, $B^c = 2^\infty \setminus B$ is countable, so that $B^c \sim \mathsf{N}$, which implies

$$\begin{aligned} B \cup \mathsf{N} &\sim B \cup \mathsf{N} \cup (-\mathsf{N}) \\ &\sim B \cup B^c \cup \mathsf{N} = 2^\infty \cup \mathsf{N}, \end{aligned}$$

and hence $I \sim B \sim 2^\infty$. This gives $I^\infty \sim (2^\infty)^\infty \sim 2^\infty \sim I$, and the assertion follows by (i). □

For any space S, a non-empty class $\mathcal{U}$ of subsets of S is called a *ring* if it is closed under finite unions and intersections, as well as under proper differences. A *semi-ring* is a non-empty class $\mathcal{I}$ of subsets closed under finite intersections, and such that any proper difference between sets in $\mathcal{I}$ is a finite, disjoint union of $\mathcal{I}$-sets. Every ring is clearly a semi-ring, and for any semi-ring $\mathcal{I}$, the finite unions of sets in $\mathcal{I}$ form a ring. Typical semi-rings in R^d consist of rectangular boxes $I_1 \times \cdots \times I_d$, where $I_1, \ldots, I_d$ are real intervals.

Given a measurable space $(S, \mathcal{S})$, we say that a class $\hat{\mathcal{S}} \subset \mathcal{S}$ is a *localizing ring* in S if

- $\hat{\mathcal{S}}$ is a subring of $\mathcal{S}$,
- $B \in \hat{\mathcal{S}}$ and $C \in \mathcal{S}$ imply $B \cap C \in \hat{\mathcal{S}}$,
- $\hat{\mathcal{S}} = \bigcup_n (\mathcal{S} \cap S_n)$ for some sets $S_n \uparrow S$ in $\hat{\mathcal{S}}$.

In this case, $S_1, S_2, \ldots$ is called a *localizing sequence* in S, and we say that a set $B \in \mathcal{S}$ is *bounded* if it belongs to $\hat{\mathcal{S}}$. Note that $\hat{\mathcal{S}}$ is closed under countable intersections, and that if $B_1, B_2, \ldots \in \hat{\mathcal{S}}$ with $B_n \subset C \in \hat{\mathcal{S}}$, then even $\bigcup_n B_n \in \hat{\mathcal{S}}$.

The class $\mathcal{S}$ itself is clearly a localizing ring. More generally, given any sequence $S_n \uparrow S$ in $\mathcal{S}$, we can *define* a localizing ring by $\hat{\mathcal{S}} = \bigcup_n (\mathcal{S} \cap S_n)$. If S is a metric space, we can take $\hat{\mathcal{S}}$ to be the class of all metrically bounded Borel sets in S, and if S is lcscH, we may choose $\hat{\mathcal{S}}$ to consist of all relatively compact sets in $\mathcal{S}$. A *localized Borel space* is defined as a triple $(S, \mathcal{S}, \hat{\mathcal{S}})$, where $(S, \mathcal{S})$ is a Borel space and $\hat{\mathcal{S}}$ is a localizing ring in $\mathcal{S}$. To simplify our statements, every Borel space below is understood to be localized.

Given any space S, we define a *π-system* in S as a non-empty class of subsets closed under finite intersections, whereas a *λ-system* is defined as a class of subsets that contains S and is closed under proper differences and increasing limits. For any class $\mathcal{C}$, the intersection of all σ-fields containing $\mathcal{C}$ is again a σ-field, denoted by $\sigma(\mathcal{C})$ and called the σ-field *generated* by $\mathcal{C}$.

Next let $(S, \mathcal{S})$ be a measurable space with a localizing ring $\hat{\mathcal{S}}$. By a *local monotone class* we mean a non-empty subclass $\mathcal{M} \subset \hat{\mathcal{S}}$ that is closed under bounded monotone limits, in the sense that $B_1, B_2, \ldots \mathcal{M}$ with $B_n \downarrow B$ or $B_n \uparrow B \in \hat{\mathcal{S}}$ implies $B \in \mathcal{M}$. For any class $\mathcal{C} \subset \hat{\mathcal{S}}$, the intersection of all local monotone rings in $\hat{\mathcal{S}}$ containing $\mathcal{C}$ is again a class of the same type, here denoted by $\hat{\sigma}(\mathcal{C})$ and called the local monotone ring *generated* by $\mathcal{C}$.

Lemma 1.2 *(monotone-class theorems, Sierpiński) Fix a measurable space $(S, \mathcal{S})$ with a localizing ring $\hat{\mathcal{S}}$. Then*

(i) *for any λ-system $\mathcal{D} \subset \mathcal{S}$ and π-system $\mathcal{I} \subset \mathcal{D}$, we have $\mathcal{D} \supset \sigma(\mathcal{I})$,*

(ii) *for any local monotone class $\mathcal{M} \subset \hat{\mathcal{S}}$ and ring $\mathcal{U} \subset \mathcal{M}$, we have $\mathcal{M} \supset \hat{\sigma}(\mathcal{U})$.*

Proof: Since the intersection of any family of λ-systems or local monotone classes is a class of the same type, we may assume in (i) that $\mathcal{D}$ is the smallest λ-system containing $\mathcal{I}$, and in (ii) that $\mathcal{M}$ is the smallest local monotone class containing $\mathcal{U}$, here written as $\mathcal{D} = \lambda(\mathcal{I})$ and $\mathcal{M} = M(\mathcal{U})$.

If a class $\mathcal{D}$ is both a π-system and a λ-system, it is clearly a σ-field. The corresponding property in (ii) is obvious. It is then enough in (i) to prove that $\mathcal{D}$ is a π-system and in (ii) that $\mathcal{M}$ is a ring. Thus, in (i) we need to show that if $A, B \in \mathcal{D}$, then also $A \cap B \in \mathcal{D}$. Similarly, in (ii) we need to show that if $A, B \in \mathcal{M}$, then even $A \cap B$, $A \cup B$, and $A \setminus B$ lie in $\mathcal{M}$. Here we consider only (i), the proof for (ii) being similar.

The relation $A \cap B \in \mathcal{D}$ holds for all $A, B \in \mathcal{I}$, since $\mathcal{I}$ is a π-system contained in $\mathcal{D}$. For fixed $A \in \mathcal{I}$, let $\mathcal{B}_A$ denote the class of sets $B \subset S$ with $A \cap B \in \mathcal{D}$. Since $\mathcal{D}$ is a λ-system, so is clearly $\mathcal{B}_A$. Hence, $\mathcal{B}_A \supset \lambda(\mathcal{I}) = \mathcal{D}$, which shows that $A \cap B \in \mathcal{D}$ for all $A \in \mathcal{I}$ and $B \in \mathcal{D}$. Interchanging the roles of A and B, we see that $A \cap B \in \mathcal{D}$ remains true for all $A, B \in \mathcal{D}$. □

Returning to the case of a general Borel space $(S, \mathcal{S})$ with a localizing ring $\hat{\mathcal{S}}$ of bounded Borel sets, we define a *dissection system* in S as an array $\mathcal{I}$ of subsets $I_{nj} \in \hat{\mathcal{S}}$, $n, j \in \mathsf{N}$, such that

- for every $n \in \mathsf{N}$, the sets I_{nj} form a countable partition of S,
- for any $m < n$ in N, the partition (I_{nj}) is a refinement of (I_{mj}),
- for every $n \in \mathsf{N}$, each set $B \in \hat{\mathcal{S}}$ is covered by finitely many I_{nj},
- the σ-field generated by all I_{nj} equals $\mathcal{S}$.

Since singletons in S belong to $\mathcal{S}$, the last property shows that $\mathcal{I}$ *separates points*, in the sense that for any $s \neq t$ in S, there exists an $n \in \mathsf{N}$ such that s and t lie in different sets I_{nj}. When $S = \mathsf{R}$, we may choose $\mathcal{I}$ as the class of all dyadic intervals $I_{nj} = 2^{-n}(j-1, j]$, $n \in \mathsf{Z}_+$, $j \in \mathsf{Z}$, and similarly in higher dimensions. We show that a dissection system always exists.

Lemma 1.3 *(dissection system) For any localized Borel space S, the ring $\hat{\mathcal{S}}$ contains a dissection system $\mathcal{I} = (I_{nj})$.*

Proof: Fix a localizing sequence $S_n \uparrow S$ in $\hat{\mathcal{S}}$, and put $S_0 = \emptyset$. The differences $B_n = S_n \setminus S_{n-1}$ form a partition of S into $\hat{\mathcal{S}}$-sets, such that every $B \in \hat{\mathcal{S}}$ is covered by finitely many sets B_n. Since the latter are Borel, they may be identified with Borel sets in $[n-1, n)$, $n \in \mathsf{N}$, so that S becomes a

Borel set in R_+. By the localizing property of the sets S_n, we see that $B \in \hat{\mathcal{S}}$ iff its image set in R_+ is a bounded Borel set. The sets

$$I_{nj} = 2^{-n}[j-1, j) \cap S, \quad n, j \in \mathsf{N},$$

clearly form a dissection system in S. □

A measure μ on S is said to be *s-finite* if it is a countable sum of bounded measures, and *σ-finite* if $\mu f < \infty$ for some measurable function $f > 0$ on S. For any measures μ and ν, the equivalence $\mu \sim \nu$ means that $\mu \ll \nu$ and $\nu \ll \mu$, in the sense of absolute continuity. Any σ-finite measure $\nu \sim \mu$ is called a *supporting measure* of μ. We note some elementary facts.

Lemma 1.4 *(boundedness, supporting measure) Fix any measurable space S.*

(i) *A measure μ on S is σ-finite iff there exists a measurable partition $B_1, B_2, \ldots$ of S, such that $\mu B_n < \infty$ for all n. In particular, every σ-finite measure is s-finite.*

(ii) *For any s-finite measure $\mu \neq 0$ on S, there exists a probability measure $\nu \sim \mu$. In particular, every σ-finite measure $\rho \neq 0$ on $S \times T$ has a supporting probability measure $\nu \sim \rho(\cdot \times T)$.*

Proof: (i) First let μ be σ-finite, so that $\mu f < \infty$ for some measurable function $f > 0$ on S. Then the sets $B_n = \{s \in S;\ f(s) \in (n-1, n]\}$, $n \in \mathsf{N}$, form a measurable partition of S, and $\mu B_n < \infty$ for all n. Conversely, if $B_1, B_2, \ldots \in \mathcal{S}$ form a partition of S with $\mu B_n < \infty$ for all n, then the function $f = \sum_n 2^{-n}(\mu B_n \vee 1)^{-1} 1_{B_n}$ is strictly positive with $\mu f \leq 1$, which shows that μ is σ-finite.

(ii) Assuming $\mu = \sum_n \mu_n$ with $0 < \|\mu_n\| < \infty$ for all $n \in \mathsf{N}$, we may choose $\nu = \sum_n 2^{-n} \mu_n / \|\mu_n\|$. If the measure ρ on $S \times T$ is s-finite, then so is the measure $\mu = \rho(\cdot \times T)$ on S, and the second assertion follows. □

A measure μ on S is said to be *locally finite* if $\mu B < \infty$ for every $B \in \hat{\mathcal{S}}$, and we write $\mathcal{M}_S$ for the class of locally finite measures on S. In $\mathcal{M}_S$ we introduce the σ-field $\mathcal{B}_{\mathcal{M}_S}$ generated by all *evaluation* maps $\pi_B\colon \mu \mapsto \mu B$ with $B \in \mathcal{S}$. Equivalently, $\mathcal{B}_{\mathcal{M}_S}$ is generated by the integration maps $\pi_f\colon \mu \mapsto \mu f$ for all $f \in \mathcal{S}_+$, the class of measurable functions $f \geq 0$ on S. The following basic result justifies our notation:

Theorem 1.5 *(measure space, Prohorov) For any Borel space $(S, \mathcal{S})$ with localizing ring $\hat{\mathcal{S}}$, the associated measure space $(\mathcal{M}_S, \mathcal{B}_{\mathcal{M}_S})$ is again Borel.*

Proof: Fix a localizing sequence $S_n \uparrow S$ in $\hat{\mathcal{S}}$, and put $B_n = S_n \setminus S_{n-1}$ with $S_0 = \emptyset$. Then $\mu \in \mathcal{M}_S$ iff $\mu_{B_n} \in \hat{\mathcal{M}}_{B_n}$ for all $n \in \mathsf{N}$, where $\mu_B = 1_B \mu$ denotes the restriction of the measure μ to the set B. Furthermore, the mapping $\mu \mapsto (\mu_{B_1}, \mu_{B_2}, \ldots)$ defines a Borel isomorphism between $\mathcal{M}_S$ and the product of all $\hat{\mathcal{M}}_{B_n}$. Since the Borel property is preserved by countable products, it it enough to consider one of those factors, which reduces the discussion to the measure space $\hat{\mathcal{M}}_S$. Since S is Borel, we may assume that $S \in \mathcal{B}_{[0,1]}$, and since $\{\mu S^c = 0\} \in \mathcal{B}(\hat{\mathcal{M}}_{[0,1]})$, we may even take $S = [0,1]$.

Now let $f_1, f_2, \ldots$ be dense in $C^+_{[0,1]}$, and introduce in $\hat{\mathcal{M}}_S$ the metric

$$\rho(\mu, \nu) = \sum\nolimits_k 2^{-k}(|\mu f_k - \nu f_k| \wedge 1), \quad \mu, \nu \in \hat{\mathcal{M}}_S.$$

This defines a homeomorphism between $\hat{\mathcal{M}}_S$ with the weak topology and a subset of R_+^∞ with the product topology, which makes $\hat{\mathcal{M}}_S$ a separable metric space. To see that the metric is also complete, let $\mu_1, \mu_2, \ldots$ be Cauchy in $\hat{\mathcal{M}}_S$, so that $\mu_m f_k - \mu_n f_k \to 0$ as $m, n \to \infty$ for each k. By Helly's selection theorem (FMP 5.19, or Lemma 4.4 below), $\mu_n \xrightarrow{w} \mu \in \hat{\mathcal{M}}_S$ along a sub-sequence, which gives $\mu_n f_k \to \mu f_k$ along the same sequence. By the Cauchy property, the convergence extends to the original sequence, which means that $\mu_n \xrightarrow{w} \mu$. Thus, $\hat{\mathcal{M}}_S$ is Polish in the weak topology. Since $\mathcal{B}_{\hat{\mathcal{M}}_S}$ agrees with the Borel σ-field in $\hat{\mathcal{M}}_S$, the latter is Borel by Theorem 1.1. $\square$

The *Dirac measure* δ_s at $s \in S$ is defined by $\delta_s B = 1_B(s)$ for all $B \in \mathcal{S}$, where 1_B denotes the *indicator function* of B, equal to 1 on B and to 0 on the complement B^c. For any measure $\mu \in \mathcal{M}_S$, we say that μ has an *atom* at s of size $b > 0$ if $\mu\{s\} = b$. Let $\mathcal{M}_S^*$ denote that class of *diffuse* (non-atomic) measures in $\mathcal{M}_S$.

Lemma 1.6 *(atomic decomposition) For any localized Borel space S, there exists a decomposition*

$$\mu = \alpha + \sum_{k \le \kappa} \beta_k \, \delta_{\sigma_k}, \quad \mu \in \mathcal{M}_S, \tag{2}$$

with $\alpha \in \mathcal{M}_S^*$, $\kappa \in \bar{\mathsf{Z}}_+$, $\beta_1, \beta_2, \ldots > 0$, and distinct $\sigma_1, \sigma_2, \ldots \in S$. The representation is unique up to the order of terms, and we may choose α, κ, and all (β_k, σ_k) to be measurable functions of μ. Conversely, any locally finite sum of this form, involving measurable functions α, κ, and (β_k, σ_k) as above, defined on some measurable space $(\Omega, \mathcal{A})$, determines μ as a measurable function on Ω.

Proof: Fixing a localizing sequence $S_n \uparrow S$ in $\hat{\mathcal{S}}$ and using the Borel property of the differences $S_n \setminus S_{n-1}$, we may reduce to the case where μ is a bounded measure on some set $S \in \mathcal{B}_{[0,1)}$. Extending μ to a measure on $[0,1)$, we may further reduce to the case where $S = [0,1)$. The desired decomposition now follows from the jump structure of the function $F_t = \mu[0,t]$, $t \in [0,1]$.

To construct a measurable representation, write $I_{nj} = 2^{-n}[j-1, j)$ for $j \le 2^n$, $n \in \mathsf{N}$. For any $n \in \mathsf{N}$ and $\varepsilon > 0$, we may enumerate the pairs $(\mu I_{nj}, j2^{-n})$ with $\mu I_{nj} \ge \varepsilon$, in the order of increasing j, as $(\beta_{nk}, \sigma_{nk})$, $k \le \kappa_n$.

Here κ_n and all β_{nk} and σ_{nk} converge as $n \to \infty$ to some limits κ^ε, β_k^ε, and σ_k^ε, where the measures $\beta_k^\varepsilon\,\delta_{\sigma_k^\varepsilon}$ with $k \le \kappa_k^\varepsilon$ are precisely the atoms in μ of size $\ge \varepsilon$. Subtracting the latter from μ and continuing recursively in countably many steps, we obtain a measurable representation of all atoms, along with the diffuse remainder α. The converse and uniqueness assertions are obvious. $\square$

In particular, the class $\mathcal{M}_S^*$ belongs to $\mathcal{B}_{\mathcal{M}_S}$ and may be regarded as a measure space in its own right, endowed with the trace σ-field $\mathcal{B}_{\mathcal{M}_S^*}$. By a *point measure* on S we mean a measure $\mu \in \mathcal{M}_S$ satisfying (2) with $\alpha = 0$ and $\beta_k \in \mathsf{N}$ for all $k \le \kappa$, and we say that μ is *simple* if $\beta_k = 1$ for all k. By Theorem 1.6, the classes of point measures and simple point measures again belong to $\mathcal{B}_{\mathcal{M}_S}$ and may be regarded as measurable spaces $\mathcal{N}_S$ and $\mathcal{N}_S^*$ in their own right, endowed with the trace σ-fields $\mathcal{B}_{\mathcal{N}_S}$ and $\mathcal{B}_{\mathcal{N}_S^*}$, respectively. The same thing is true for the class $\mathcal{P}_S$ of probability measures on S.

Now write $\hat{\mathcal{F}}_S$ for the class of locally finite subsets $M \subset S$. Thus, $M \in \hat{\mathcal{F}}_S$ iff $|M \cap B| < \infty$ for all $B \in \hat{\mathcal{S}}$, where $|M|$ denotes the cardinality of the set M. In $\hat{\mathcal{F}}_S$ we may introduce the σ-field $\mathcal{B}_{\hat{\mathcal{F}}_S}$ generated by all hitting maps $\chi_B\colon M \mapsto 1\{M \cap B \ne \emptyset\}$ with $B \in \hat{\mathcal{S}}$.

Lemma 1.7 *(sets and measures) The spaces $(\hat{\mathcal{F}}_S, \mathcal{B}_{\hat{\mathcal{F}}_S})$ and $(\mathcal{N}_S^*, \mathcal{B}_{\mathcal{N}_S^*})$ are Borel and related by the Borel isomorphism*

$$\begin{aligned} \mu_M B &= |M \cap B|, \qquad & B \in \hat{\mathcal{S}},\ M \in \hat{\mathcal{F}}_S, \\ M_\mu &= \{\sigma_k;\, k \le \kappa\}, \qquad & \mu = \textstyle\sum_{k \le \kappa} \delta_{\sigma_k} \in \mathcal{N}_S^*. \end{aligned}$$

Proof: If $M = \{\sigma_k;\, k \le \kappa\}$ in $\hat{\mathcal{F}}_S$, then clearly $\mu_M = \sum_k \delta_{\sigma_k} \in \mathcal{N}_S^*$. Conversely, for any $\mu \in \mathcal{N}_S^*$, the representation $\sum_k \delta_{\sigma_k}$ is unique up to the order of terms, and hence defines a unique set $M_\mu = \{\sigma_k\}$. Thus, the displayed mappings provide a $1-1$ correspondence between $\hat{\mathcal{F}}_S$ and $\mathcal{N}_S^*$. The measurability of the map $\mu \mapsto M_\mu$ is clear from FMP 1.4, since $M_\mu \cap B \ne \emptyset$ iff $\mu B > 0$. To see that even the map $M \mapsto \mu_M$ is measurable, we introduce as in Lemma 1.3 a dissection system (I_{nj}) in S, and note that

$$\mu_M B = \lim_{n \to \infty} \textstyle\sum_j 1\{M \cap B \cap I_{nj} \ne \emptyset\}, \quad B \in \hat{\mathcal{S}},\ M \in \hat{\mathcal{F}}_S,$$

which is $\mathcal{B}_{\hat{\mathcal{F}}_S}$-measurable in M for fixed B. Since $\mathcal{B}_{\mathcal{N}_S^*}$ is generated by the projection maps π_B with $B \in \hat{\mathcal{S}}$, the result follows again by FMP 1.4. $\square$

Every dissection system $\mathcal{I} = (I_{nj})$ induces a linear, lexicographical order on S, as follows. Letting $s, t \in S$ be arbitrary and writing $I_n(s)$ for the set I_{nj} containing s, we say that s *precedes* t and write $s \prec t$, if there exists an $n \in \mathsf{N}$ such that $I_m(s) = I_m(t)$ for all $m < n$, whereas $s \in I_{ni}$ and $t \in I_{nj}$ for some $i < j$. Since $\mathcal{I}$ separates points, any two elements $s \ne t$ satisfy either $s \prec t$ or $s \succ t$. We also write $s \preceq t$ if $s \prec t$ or $s = t$. An *interval* in S is a set

$I \subset S$ such that $s \prec t \prec u$ with $s, u \in I$ implies $t \in I$. For any points $s \preceq t$ in S, we may introduce the associated intervals $[s,t]$, $[s,t)$, $(s,t]$, and (s,t). Any finite set $B \neq \emptyset$ in S has a well-defined maximum and minimum.

Lemma 1.8 *(induced order) Every dissection system (I_{nj}) on S induces a linear order, such that*

(i) *the relations $\preceq$ and $\prec$ are measurable on S^2,*

(ii) *every interval in S is measurable,*

(iii) *for any $\mu \in \hat{\mathcal{N}}_S \setminus \{0\}$, the maximum and minimum of $\operatorname{supp} \mu$ are measurable functions of μ into S.*

Proof: (i) The sets $\Delta = \{(s,t);\, s \prec t\}$ and $\tilde{\Delta} = \{(s,t);\, s \succ t\}$ are countable unions of product sets in $\mathcal{I}^2$, and are therefore measurable. This implies measurability of the diagonal $D_S = \{(s,t);\, s = t\} = S^2 \setminus (\Delta \cup \tilde{\Delta})$ as well.

(ii) For any interval $J \subset S$, the set $J' = \bigcup\{I \in \mathcal{I};\, I \subset J\}$ is clearly measurable. The set $J \setminus J'$ contains at most two points, since if $s \prec u \prec t$ with $s, t \in J$, then u lies in a set $I \in \mathcal{I}$ with $s \prec r \prec t$ for all $r \in I$, which implies $u \in I \subset J'$. Since singletons are measurable, the assertion follows.

(iii) Let $\tau_\pm(\mu)$ denote the maximum and minimum of $\operatorname{supp} \mu$. Then for any $B \in \mathcal{I}$, the events $\{\tau_\pm(\mu) \in B\}$ are determined by countable set operations involving events $\{\mu I > 0\}$ with $I \in \mathcal{I}$. Hence, $\{\tau_\pm(\mu) \in B\} \in \mathcal{B}_{\mathcal{N}_S}$ for all $B \in \mathcal{I}$, which extends to arbitrary $B \in \hat{\mathcal{S}}$ by a monotone-class argument. □

For topological Borel spaces, we consider yet another type of classes of subsets. Given a separable and complete metric space S, localized by the metric, we say that a class $\mathcal{C} \subset \hat{\mathcal{S}}$ is *dissecting*, if

- every open set $G \subset S$ is a countable union of sets in $\mathcal{C}$,
- every set $B \in \hat{\mathcal{S}}$ is covered by finitely many sets in $\mathcal{C}$.

Similarly, a dissection system (I_{nj}) is said to be *topological*, if every open set $G \subset S$ is a countable union of sets I_{nj}. Note that every dissecting class generates the Borel σ-field $\mathcal{S}$.

For any measure μ on S, we introduce the class $\hat{\mathcal{S}}_\mu = \{B \in \hat{\mathcal{S}};\, \mu\partial B = 0\}$. Here ∂B denotes the boundary of B given by $\bar{B} \cap \bar{B^c} = \bar{B} \setminus B^o$, where $\bar{B}$ and B^o denote the closure and interior of B.

Lemma 1.9 *(dissecting rings and semi-rings) Let S be a separable and complete metric space, regarded as a localized Borel space. Then*

(i) *every dissecting ring or semi-ring in S contains a countable class of the same kind,*

(ii) *every dissecting semi-ring $\mathcal{I}$ in S contains a topological dissecting system (I_{nj}),*

(iii) *for any topological dissection system* (I_{nj}), *the entire family* $\{I_{nj}\}$ *is a dissecting semi-ring,*

(iv) *if* $\mathcal{U}$ *is a dissecting ring in* S, *then every bounded, closed set is a countable intersection of* $\mathcal{U}$*-sets,*

(v) *for every s-finite measure* μ *on* S, *the class* $\hat{\mathcal{S}}_\mu$ *is a dissecting ring,*

(vi) *if* $\mathcal{I}'$ *and* $\mathcal{I}''$ *are dissecting semi-rings in* S' *and* S'', *then* $\mathcal{I} = \mathcal{I}' \times \mathcal{I}''$ *is a dissecting semi-ring in* $S = S' \times S''$.

Proof: (i) Let $\mathcal{U}$ be a dissecting ring. Since S is separable, it has a countable topological base $G_1, G_2, \ldots$, and each G_n is a countable union of sets $U_{nj} \in \mathcal{U}$. Furthermore, the open n-ball B_n around the origin is covered by finitely many sets $V_{nj} \in \mathcal{U}$. The countable collections $\{U_{nj}\}$ and $\{V_{nj}\}$ form a dissecting class $\mathcal{C} \subset \mathcal{U}$, and the generated ring $\mathcal{V}$ is again countable, since its sets are obtained from those of $\mathcal{C}$ by finitely many set operations. A similar argument applies to semi-rings.

(ii) By (i) we may assume that $\mathcal{I}$ is countable, say $\mathcal{I} = \{I_1, I_2, \ldots\}$. Let $\mathcal{U}$ be the ring generated by $\mathcal{I}$, and choose some sets $U_n \supset B_n$ in $\mathcal{U}$. Since every bounded set is contained in some U_n, it is enough to choose a dissecting system in each U_n. Equivalently, we may assume $S \in \mathcal{U}$, so that $S = \bigcup_{k \le m} I_k$ for some m. For every $n \ge m$, let $\pi_n = (I_{nj})$ be the finite partition of S generated by $I_1, \ldots, I_n$, and note that the sets I_{nj} belong to $\mathcal{I}$, since they are finite intersections of $\mathcal{I}$-sets. We claim that the I_{nj} form a dissecting system. The nestedness is clear, since the collections $\{I_1, \ldots, I_n\}$ are increasing. The dissecting property holds since the π_n are finite partitions of S. The generating property holds since $\mathcal{I}$ is dissecting, so that every open set is a countable union of sets I_n.

(iii) This is clear from the topological dissection property, together with the obvious semi-ring property of the family $\mathcal{I} = \{I_{nj}\}$.

(iv) Let $F \in \hat{\mathcal{S}}$ be closed. Since $\mathcal{U}$ is a dissecting ring, we may choose some $\mathcal{U}$-sets $A_n \uparrow F^c$ and $B \supset F$. The $\mathcal{U}$-sets $U_n = B \setminus A_n$ are then non-increasing with

$$\begin{aligned} \textstyle\bigcap_n U_n &= \textstyle\bigcap_n (B \setminus A_n) = B \cap \textstyle\bigcap_n A_n^c \\ &= B \cap \left(\textstyle\bigcup_n A_n\right)^c \\ &= B \cap (F^c)^c = B \cap F = F. \end{aligned}$$

(v) To prove the ring property of $\hat{\mathcal{S}}_\mu$, we claim that

$$\begin{aligned} \partial(B \cup C) &\subset \partial B \cup \partial C, \\ \partial(B \cap C) &\subset \partial B \cup \partial C. \end{aligned}$$

Indeed, noting that $\overline{B \cup C} = \overline{B} \cup \overline{C}$ and $\overline{B \cap C} \subset \overline{B} \cap \overline{C}$, we get

$$\partial(B \cup C) = \overline{(B \cup C)} \cap \overline{(B \cup C)^c}$$

$$\begin{aligned}
&= (\overline{B} \cup \overline{C}) \cap \overline{(B^c \cap C^c)} \\
&\subset (\overline{B} \cup \overline{C}) \cap \overline{B^c} \cap \overline{C^c} \\
&= (\overline{B} \cap \overline{B^c} \cap \overline{C^c}) \cup (\overline{C} \cap \overline{B^c} \cap \overline{C^c}) \\
&\subset (\overline{B} \cap \overline{B^c}) \cup (\overline{C} \cap \overline{C^c}) = \partial B \cup \partial C.
\end{aligned}$$

Applying this relation to the complements gives

$$\begin{aligned}
\partial(B \cap C) &= \partial(B \cap C)^c = \partial(B^c \cup C^c) \\
&\subset \partial B^c \cup \partial C^c = \partial B \cup \partial C.
\end{aligned}$$

For every $s \in S$, the open r-ball B_s^r around s belongs to $\hat{\mathcal{S}}_\mu$ for all but countably many $r > 0$. The dissecting property then follows by the separability of S.

(vi) The semi-ring property of $\mathcal{I}' \times \mathcal{I}''$ follows easily from that of $\mathcal{I}'$ and $\mathcal{I}''$. The covering property is equally obvious, since a set $B \subset S' \times S''$ is bounded iff it is contained in a rectangle $B' \times B''$, such that B' and B'' are bounded in S' and S'', respectively. If $\mathcal{B}'$ and $\mathcal{B}''$ are topological bases in S' and S'', then $\mathcal{B}' \times \mathcal{B}''$ is a base for the product topology in $S' \times S''$. Since S' are S'' are both separable, so is $S' \times S''$, and every open set $G \subset S' \times S''$ is a countable union of sets $B'_k \times B''_k \in \mathcal{B}' \times \mathcal{B}''$. Since $\mathcal{I}'$ and $\mathcal{I}''$ are dissecting, B'_k and B''_k are themselves countable unions of sets $B'_{ki} \in \mathcal{I}'$ and $B''_{kj} \in \mathcal{I}''$. Thus, G is the countable union of all sets $B'_{ki} \times B''_{kj} \in \mathcal{I}' \times \mathcal{I}''$ with $i, j, k \in \mathsf{N}$. $\Box$

We conclude with a couple of technical results needed in later sections.

Lemma 1.10 *(completeness) For every measurable space S, the space $\hat{\mathcal{M}}_S$ is complete in total variation.*

Proof: Let $\mu_1, \mu_2, \ldots \in \hat{\mathcal{M}}_S$ with $\|\mu_m - \mu_n\| \to 0$ as $m, n \to \infty$. Assuming $\mu_n \neq 0$, we may define $\nu = \sum_n 2^{-n} \mu_n / \|\mu_n\|$, and choose some measurable functions $f_1, f_2, \ldots \in L_1(\nu)$ with $\mu_n = f_n \cdot \nu$. Then $\nu|f_m - f_n| = \|\mu_m - \mu_n\| \to 0$, which means that (f_n) is Cauchy in $L^1(\nu)$. Since L^1 is complete (FMP 1.31), we have convergence $f_n \to f$ in L^1, and so the measure $\mu = f \cdot \nu$ satisfies $\|\mu - \mu_n\| = \nu|f - f_n| \to 0$. $\Box$

Lemma 1.11 *(restriction) For any Borel spaces S and T, let μ_t denote the restriction of a measure $\mu \in \hat{\mathcal{M}}_{S \times T}$ to $S \times \{t_1, \ldots, t_d\}^c$, where $t = (t_1, \ldots, t_d) \in T^d$. Then the mapping $(\mu, t) \mapsto \mu_t$ is product measurable.*

Proof: We may take $T = \mathsf{R}$. Putting $I_{nj} = 2^{-n}(j-1, j]$ for $n, j \in \mathsf{Z}$, we define

$$U_n(t) = \bigcup\nolimits_j \{I_{nj};\, t_1, \ldots, t_d \notin I_{nj}\}, \quad n \in \mathsf{N},\ t \in \mathsf{R}^d.$$

The restriction μ_t^n of μ to $S \times U_n(t)$ is product measurable, and $\mu_t^n \uparrow \mu_t$ by monotone convergence. $\Box$

1.2 Product and Factorial Measures

For any Borel spaces S and T, the product space $S \times T$ is again Borel. Any localizing rings $\hat{\mathcal{S}}$ and $\hat{\mathcal{T}}$ in S and T induce a localizing ring in $S \times T$, consisting of all sets $B \in \mathcal{S} \otimes \mathcal{T}$ with $B \subset B' \times B''$ for some $B' \in \hat{\mathcal{S}}$ and $B'' \in \hat{\mathcal{T}}$. Thus, B is bounded in $S \times T$ iff it has bounded projections on S and T, and a measure μ on $S \times T$ is locally finite iff $\mu(B' \times B'') < \infty$ for any $B' \in \hat{\mathcal{S}}$ and $B'' \in \hat{\mathcal{T}}$. This extends immediately to any finite product spaces, and in particular S^n is again a localized Borel space.

Let $S^{(n)}$ denote the non-diagonal part of S^n, consisting of all n-tuples $(s_1, \ldots, s_n)$ with distinct $s_1, \ldots, s_n \in S$. For any point measure $\mu \in \mathcal{N}_S$, we proceed to introduce the associated *factorial measures* $\mu^{(n)}$ on S^n, $n \in \mathsf{Z}_+$. When μ is simple, we may define $\mu^{(n)} = 1_{S^{(n)}}\mu^{\otimes n}$ for all $n \in \mathsf{N}$. For general μ, any of the following descriptions may serve as a definition. For consistency we put $\mu^{(0)} = 1$.

Lemma 1.12 *(factorial measures) For any Borel spaces S and T, let $\mu = \sum_{i\in I} \delta_{\sigma_i}$ in $\mathcal{N}_S$, and fix some distinct $\tau_1, \tau_2, \ldots \in T$. Then the following descriptions of the measures $\mu^{(n)}$ are equivalent:*

(i) $\quad \mu^{(n)} = \sum_{i\in I^{(n)}} \delta_{\sigma_{i_1},\ldots,\sigma_{i_n}}$,

(ii) $\quad \mu^{(n)} = \tilde{\mu}^{(n)}(\cdot \times T^n)$ *with* $\tilde{\mu} = \sum_{i\in I} \delta_{\sigma_i,\tau_i}$,

(iii) $\quad \mu^{(m+n)} f = \int \mu^{(m)}(ds) \int \big(\mu - \sum_{i\le m} \delta_{s_i}\big)^{(n)}(dt)\, f(s,t)$.

Proof: When μ is simple, we have for any $n \in \mathsf{N}$

$$\mu^{(n)} = 1_{S^{(n)}}\mu^{\otimes n} = 1_{S^{(n)}} \sum_{i\in I^n} \delta_{\sigma_{i_1},\ldots,\sigma_{i_n}} = \sum_{i\in I^{(n)}} \delta_{\sigma_{i_1},\ldots,\sigma_{i_n}},$$

since the σ_i are distinct, so that $\sigma_i = \sigma_j$ iff $i = j$. Hence, our elementary definition then agrees with the description in (i). Note that (i) is also consistent with $\mu^{(1)} = \mu$. In general, we may then take (i) as our definition.

(i) $\Leftrightarrow$ (ii): Since the t_i are distinct, the measure $\tilde{\mu}$ is simple, and so

$$\tilde{\mu}^{(n)}(\cdot \times T^n) = \sum_{i\in I^{(n)}} \bigotimes_{k\le n} \delta_{\sigma_{i_k},\tau_{i_k}}(\cdot \times T) = \sum_{i\in I^{(n)}} \bigotimes_{k\le n} \delta_{\sigma_{i_k}} = \mu^{(n)}.$$

(i) $\Leftrightarrow$ (iii): Since the measures $\mu^{(n)}$ are uniquely determined by any recursive formula, together with the initial condition $\mu^{(1)} = \mu$, it suffices to show that the measures in (i) satisfy (iii). Writing $I_i = I \setminus \{i_1, \ldots, i_m\}$ for any $i = (i_1, \ldots, i_m) \in I^{(m)}$, we get from (i)

$$\begin{aligned}
&\int \mu^{(m)}(ds) \int \big(\mu - \textstyle\sum_{k\le m} \delta_{s_k}\big)^{(n)}(dt)\, f(s,t) \\
&\qquad = \sum_{i\in I^{(m)}} \int \big(\mu - \textstyle\sum_{k\le m} \delta_{\sigma_{i_k}}\big)^{(n)}(dt)\, f(\sigma_{i_1}, \ldots, \sigma_{i_m}, t)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i \in I^{(m)}} \int \Big(\textstyle\sum_{j \in I_i} \delta_{\sigma_j}\Big)^{(n)}(dt)\, f(\sigma_{i_1}, \ldots, \sigma_{i_m}, t) \\
&= \sum_{i \in I^{(m)}} \sum_{j \in I_i^{(n)}} f(\sigma_{i_1}, \ldots, \sigma_{i_m}, \sigma_{j_1}, \ldots, \sigma_{j_n}) \\
&= \sum_{i \in I^{(m+n)}} f(\sigma_{i_1}, \ldots, \sigma_{i_{m+n}}) = \mu^{(m+n)} f. \qquad \square
\end{aligned}
$$

Let $\mathcal{P}_n$ denote the class of partitions π of $\{1, \ldots, n\}$ into disjoint subsets J. Define $\mu^{(|\pi|)}$ as the factorial measure of μ on the space $S^\pi = \{s_J;\ J \in \pi\}$. Let D_π be the set of all $s \in S^n$ such that $s_i = s_j$ iff $i \overset{\pi}{\sim} j$, where $\overset{\pi}{\sim}$ denotes the equivalence relation induced by π, and put $\hat{D}_\pi = \bigcup_{\pi' \prec \pi} D_{\pi'}$, where $\pi' \prec \pi$ means that π' is courser than π. The natural projection $p_\pi : S^\pi \to \hat{D}_\pi$ is given by $p_\pi(s_J;\ J \in \pi) = (s_{J_1}, \ldots, s_{J_n})$, where J_i denotes the set $J \in \pi$ containing i. We may now define $\mu^{(\pi)} = \mu^{(|\pi|)} \circ p_\pi^{-1}$.

Theorem 1.13 *(factorial decompositions) Let $\mu \in \mathcal{N}_S$ be arbitrary.*

(i) *For any measurable function $f \geq 0$ on $\hat{\mathcal{N}}_S$, we have*

$$\sum_{n \geq 0} \frac{1}{n!} \int \mu^{(n)}(ds)\, f\Big(\textstyle\sum_{i \leq n} \delta_{s_i}\Big) = \sum_{\nu \leq \mu} f(\nu),$$

where the integral for $n = 0$ is interpreted as $f(0)$, and the summation on the right extends over all $\nu \leq \mu$ in $\hat{\mathcal{N}}_S$.

(ii) *For any partition $\pi \in \mathcal{P}_n$, let $\mu^{(\pi)}$ be the natural projection of $\mu^{(|\pi|)}$ onto the diagonal space $\hat{D}_\pi$. Then*

$$\mu^n = \sum_{\pi \in \mathcal{P}_n} \mu^{(\pi)}, \quad n \in \mathsf{N}.$$

For a precise interpretation of the right-hand side in (i), we may write the atomic decomposition of Lemma 1.6 in the form $\mu = \sum_{j \in J} \delta_{s_j}$, where the index set J is finite or countable. Then

$$\sum_{\nu \leq \mu} f(\nu) = \sum_{I \subset J} f \circ \sum_{i \in I} \delta_{s_i},$$

where the outer summation on the right extends over all finite subsets $I \subset J$.

Proof: (i) Choose an enumeration $\mu = \sum_{i \in I} \delta_{\sigma_i}$ with $I \subset \mathsf{N}$, and put $\hat{I}^{(n)} = \{i \in I^n;\ i_1 < \cdots < i_n\}$ for $n > 0$. Write $\pi \sim I_n$ to mean that π is a permutation of $I_n = \{1, \ldots, n\}$. Using the atomic representation of μ and the tetrahedral decomposition of $I^{(n)}$, we get

$$\int \mu^{(n)}(ds)\, f\Big(\textstyle\sum_{k \leq n} \delta_{s_k}\Big) = \sum_{i \in I^{(n)}} f\Big(\textstyle\sum_{k \leq n} \delta_{\sigma_{i_k}}\Big)$$

$$
\begin{aligned}
&= \sum_{i \in \hat{I}^{(n)}} \sum_{\pi \sim I_n} f\Big(\textstyle\sum_{k \le n} \delta_{\sigma_{\pi \circ i_k}}\Big) \\
&= n! \sum_{i \in \hat{I}^{(n)}} f\Big(\textstyle\sum_{k \le n} \delta_{\sigma_{i_k}}\Big) \\
&= n! \sum_{\nu \le \mu} \{f(\nu);\ \|\nu\| = n\}.
\end{aligned}
$$

Now divide by $n!$ and sum over n.

(ii) First let $\mu = \sum_k \delta_{\sigma_k}$ be simple, and fix any measurable rectangle $B = B_1 \times \cdots \times B_n$. Using the atomic representation and simplicity of μ, the identity $\delta_s B = 1_B(s)$, and the definitions of $\mu^{(|\pi|)}$ and $\mu^{(\pi)}$, we get

$$
\begin{aligned}
\mu^n(B \cap D_\pi) &= \sum_{k \in \mathsf{N}^n} \bigotimes_{i \le n} \delta_{\sigma_{k_i}}(B \cap D_\pi) \\
&= \sum_{k \in \mathsf{N}^{(\pi)}} \prod_{J \in \pi} 1\Big\{\sigma_{k_J} \in \bigcap_{i \in J} B_i\Big\} \\
&= \sum_{k \in \mathsf{N}^{(\pi)}} \bigotimes_{J \in \pi} \delta_{\sigma_{k_J}} \bigcap_{i \in J} B_i \\
&= \mu^{(|\pi|)} \prod_{J \in \pi} \bigcap_{i \in J} B_i \\
&= \mu^{(\pi)} \bigotimes_{i \le n} B_i = \mu^{(\pi)} B,
\end{aligned}
$$

which extends by a monotone-class argument to

$$
1_{D_\pi} \cdot \mu^n = \mu^{(\pi)}.
$$

The asserted decomposition now follows, since S^n is the disjoint union of the diagonal spaces D_π. To extend the formula to general point measures $\mu = \sum_k \beta_k \delta_{\sigma_k}$, we may apply the previous version to the simple point measure $\tilde{\mu} = \sum_k \sum_{j \le \beta_k} \delta_{\sigma_k, j}$ on $S \times \mathsf{N}$, and use Lemma 1.12 (ii). □

1.3 Kernels and Operators

Given some measurable spaces $(S, \mathcal{S})$ and $(T, \mathcal{T})$, we define a *kernel* from S to T as a measurable function μ from S to the space of measures on T, often written as $\mu\colon S \to T$. Thus, μ is a function of two variables $s \in S$ and $B \in \mathcal{T}$, such that $\mu(s, B)$ is $\mathcal{S}$-measurable in s for fixed B and a measure in B for fixed s. For a simple example, we note that the Dirac measures δ_s, $s \in S$, form a kernel *on* S (from S to itself), henceforth denoted by δ.

Just as any measure μ on S may be identified with a linear functional $f \mapsto \mu f$ on $\mathcal{S}_+$, we may identify any kernel $\mu\colon S \to T$ with a linear operator $A\colon \mathcal{T}_+ \to \mathcal{S}_+$, such that

$$
Af(s) = \mu_s f, \quad \mu_s B = A 1_B(s), \qquad s \in S, \tag{3}
$$

for any $f \in \mathcal{T}_+$ and $B \in \mathcal{T}$. When $\mu = \delta$, the associated operator is simply the identity map I on $\mathcal{S}_+$. In general, we often denote a kernel and its associated operator by the same letter.

When T is a localized Borel space, we say that the kernel μ is *locally finite* if μ_s is locally finite for each s. The following result characterizes the locally finite kernels from S to T.

Lemma 1.14 *(kernel criteria) Fix two Borel spaces S and T and a generating semi-ring $\mathcal{I} \subset \hat{\mathcal{T}}$. Consider a function $\mu: S \to \mathcal{M}_T$, and define the associated operator A by* (3). *Then these conditions are equivalent:*

(i) *μ is a kernel from S to T,*
(ii) *μ is a measurable function from S to $\mathcal{M}_T$,*
(iii) *$\mu_s I$ is $\mathcal{S}$-measurable for every $I \in \mathcal{I}$,*
(iv) *A maps $\mathcal{T}_+$ into $\mathcal{S}_+$.*

We also have:

(v) *for any function $f: S \to T$, the mapping $s \mapsto \delta_{f(s)}$ is a kernel from S to T iff f is measurable,*
(vi) *the identity map on $\mathcal{M}_S$ is a kernel from $\mathcal{M}_S$ to S.*

Proof: (i) ⇔ (ii): Since $\mathcal{M}_T$ is generated by all projection maps $\pi_B: \mu \mapsto \mu B$ with $B \in \mathcal{T}$, we see that (ii) implies (i). The converse is also true by FMP 1.4.

(i) ⇔ (iii): Since trivially (i) implies (iii), it remains to show that (iii) implies (i). Then note that the measurability in (iii) extends to the generated ring $\mathcal{U}$ by the finite additivity of μ and the semi-ring property of $\mathcal{I}$. Furthermore, the class $\mathcal{M}$ of sets $B \in \hat{\mathcal{S}}$ where $\mu_s B$ is measurable is a monotone class, by the countable additivity of μ. Hence, (iii) yields $\mathcal{M} \supset \hat{\sigma}(\mathcal{U}) = \hat{\mathcal{S}}$ by Lemma 1.2 (ii), which proves (i).

(i) ⇔ (iv): For $f = 1_B$, we note that $Af(s) = \mu_s B$. Thus, (i) is simply the property (iv) restricted to indicator functions. The extension to general $f \in \mathcal{T}_+$ follows immediately by linearity and monotone convergence.

(v) This holds since $\{s;\, \delta_{f(s)}B = 1\} = f^{-1}B$ for all $B \in \mathcal{T}$.

(vi) The identity map on $\mathcal{M}_S$ trivially satisfies (ii). □

A kernel $\mu: S \to T$ is said to be *finite* if $\|\mu_s\| < \infty$ for all $s \in S$, *s-finite* if it is a countable sum of finite kernels, and *σ-finite* if it is s-finite and satisfies $\mu_s f_s < \infty$ for some measurable function $f > 0$ on $S \times T$, where $f_s = f(s, \cdot)$. It is called a *probability kernel* if $\|\mu_s\| \equiv 1$ and a *sub-probability kernel* when $\|\mu_s\| \leq 1$. The kernels μ^i, $i \in I$, are said to be *(mutually) singular* if there exist some disjoint sets $B^i \in \mathcal{S} \otimes \mathcal{T}$ with $\mu_s^i (B_s^i)^c = 0$ for all i, where B_s^i denotes the s-section of B^i. We further say that μ is *uniformly σ-finite*, if there exists a measurable partition $B_1, B_2, \ldots$ of T, such that $\mu_s B_n < \infty$ for all s and n.

We list some characterizations and mapping properties of s- and σ-finite kernels.

Lemma 1.15 *(kernel properties) Consider any s-finite kernel $\mu\colon S \to T$.*

(i) *For measurable $f \geq 0$ on $S \otimes T$, the function $\mu_s f_s$ is $\mathcal{S}$-measurable, and $\nu_s = f_s \cdot \mu_s$ is again an s-finite kernel from S to T.*

(ii) *For measurable $f\colon S \times T \to U$, the function $\nu_s = \mu_s \circ f_s^{-1}$ is an s-finite kernel from S to U.*

(iii) *$\mu = \sum_n \mu_n$ for some sub-probability kernels $\mu_1, \mu_2, \ldots\colon S \to T$.*

(iv) *μ is σ-finite iff $\mu = \sum_n \mu_n$ for some finite, mutually singular kernels $\mu_1, \mu_2, \ldots\colon S \to T$.*

(v) *μ is σ-finite iff there exists a measurable function $f > 0$ on $S \times T$ with $\mu_s f_s = 1\{\mu_s \neq 0\}$ for all $s \in S$.*

(vi) *If μ is uniformly σ-finite, it is also σ-finite, and* (v) *holds with $f(s,t) = h(\mu_s, t)$ for some measurable function $h > 0$ on $\mathcal{M}_S \times T$.*

(vii) *If μ is a probability kernel and T is Borel, there exists a measurable function $f\colon S \times [0,1] \to T$ with $\mu_s = \lambda \circ f_s^{-1}$ for all $s \in S$.*

A function f as in (v) is called a *normalizing function* of μ.

Proof: (i) Since μ is s-finite and all measurability and measure properties are preserved by countable summations, we may assume that μ is finite. The measurability of $\mu_s f_s$ follows from the kernel property when $f = 1_B$ for some measurable rectangle $B \subset S \times T$. It extends by Lemma 1.2 to arbitrary $B \in \mathcal{S} \times \mathcal{T}$, and then by linearity and monotone convergence to any $f \in (\mathcal{S} \otimes \mathcal{T})_+$. Fixing f and applying the stated measurability to the functions $1_B f$ for arbitrary $B \in \mathcal{T}$, we see that $f_s \cdot \mu_s$ is again measurable, hence a kernel from S to T. The s-finiteness is clear if we write $f = \sum_n f_n$ for some bounded $f_1, f_2, \ldots \in (\mathcal{S} \otimes \mathcal{T})_+$.

(ii) Since μ is s-finite and the measurability, measure property, and inverse mapping property are all preserved by countable summations, we may again take μ to be finite. The measurability of f yields $f^{-1}B \in \mathcal{S} \otimes \mathcal{T}$ for every $B \in \mathcal{U}$. Since $f_s^{-1}B = \{t \in T;\ (s,t) \in f^{-1}B\} = (f^{-1}B)_s$, we see from (i) that $(\mu_s \circ f_s^{-1})B = \mu_s(f^{-1}B)_s$ is measurable, which means that $\nu_s = \mu_s \circ f_s^{-1}$ is a kernel. Since $\|\nu_s\| = \|\mu_s\| < \infty$ for all $s \in S$, the kernel ν is again finite.

(iii) Since μ is s-finite and N^2 is countable, we may assume that μ is finite. Putting $k_s = [\|\mu_s\|] + 1$, we note that $\mu_s = \sum_n k_s^{-1} 1\{n \leq k_s\}\, \mu_s$, where each term is clearly a sub-probability kernel.

(iv) First let μ be σ-finite, so that $\mu_s f_s < \infty$ for some measurable function $f > 0$ on $S \times T$. Then (i) shows that $\nu_s = f_s \cdot \mu_s$ is a finite kernel from S to T, and so $\mu_s = g_s \cdot \nu_s$ with $g = 1/f$. Putting $B_n = 1\{n-1 < g \leq n\}$ for

$n \in \mathsf{N}$, we get $\mu_s = \sum_n (1_{B_n} g)_s \cdot \nu_s$, where each term is a finite kernel from S to T. The terms are mutually singular since $B_1, B_2, \ldots$ are disjoint.

Conversely, let $\mu = \sum_n \mu_n$ for some mutually singular, finite kernels $\mu_1, \mu_2, \ldots$. By singularity we may choose a measurable partition $B_1, B_2, \ldots$ of $S \times T$ such that μ_n is supported by B_n for each n. Since the μ_n are finite, we may further choose some measurable functions $f_n > 0$ on $S \times T$, such that $\mu_n f_n \le 2^{-n}$ for all n. Then $f = \sum_n 1_{B_n} f_n > 0$, and

$$\begin{aligned} \mu f &= \sum_{m,n} \mu_m 1_{B_n} f_n \\ &= \sum_n \mu_n f_n \le \sum_n 2^{-n} = 1, \end{aligned}$$

which shows that μ is σ-finite.

(v) If μ is σ-finite, then $\mu_s g_s < \infty$ for some measurable function $g > 0$ on $S \otimes T$. Putting $f(s,t) = g(s,t)/\mu_s g_s$ with $0/0 = 1$, we note that $\mu_s f_s = 1\{\mu_s \ne 0\}$. The converse assertion is clear from the definitions.

(vi) The first assertion holds by (iv). Now choose a measurable partition $B_1, B_2, \ldots$ of T with $\mu_s B_n < \infty$ for all s and n, and define

$$g(s,t) = \sum\nolimits_n 2^{-n} (\mu_s B_n \vee 1)^{-1} 1_{B_n}(t), \quad s \in S,\ t \in T.$$

The function $f(s,t) = g(s,t)/\mu_s g_s$ has the desired properties.

(vii) Since T is Borel, we may assume that $T \in \mathcal{B}_{[0,1]}$. The function

$$f(s,r) = \inf\{x \in [0,1];\ \mu_s[0,x] \ge r\}, \quad s \in S,\ r \in [0,1],$$

is product measurable on $S \times [0,1]$, since the set $\{(s,r);\ \mu_s[0,x] \ge r\}$ is measurable for each x by FMP 1.12, and the infimum can be restricted to rational x. For $s \in S$ and $x \in [0,1]$, we have

$$\lambda \circ f_s^{-1}[0,x] = \lambda\{r \in [0,1];\ f(s,r) \le x\} = \mu_s[0,x],$$

and so $\lambda \circ f_s^{-1} = \mu_s$ on $[0,1]$ for all s by FMP 3.3. In particular, $f_s(r) \in T$ a.e. λ for every $s \in S$. On the exceptional set $\{f_s(r) \notin T\}$, we may redefine $f_s(r) = t_0$ for any fixed $t_0 \in T$. $\square$

The following consequences of part (vii) are of constant use.

Lemma 1.16 *(randomization and transfer) Fix any measurable space S and Borel space T.*

(i) *For any random element ξ in S and probability kernel $\mu\colon S \to T$, there exists a random element η in T with $\mathcal{L}(\eta \mid \xi) = \mu_\xi$ a.s.*

(ii) *For any random elements $\xi \overset{d}{=} \tilde{\xi}$ in S and η in T, there exists a random element $\tilde{\eta}$ in T with $(\xi, \eta) \overset{d}{=} (\tilde{\xi}, \tilde{\eta})$.*

Proof: (i) Lemma 1.15 (vii) yields a measurable function $\varphi\colon S\times[0,1]\to T$ with $\lambda\circ\varphi_s^{-1}=\mu_s$ for all $s\in S$. Letting $\vartheta\perp\!\!\!\perp\xi$ be $U(0,1)$, we define $\eta=\varphi(\xi,\vartheta)$. For any $f\in\mathcal{T}_+$, Fubini's theorem yields a.s.

$$\begin{aligned} E\{f(\eta)\,|\,\xi\} &= E\{f(\varphi(\xi,\vartheta))|\,\xi\} \\ &= \int\lambda(dr)\,f(\varphi(\xi,r))=\mu_\xi f, \end{aligned}$$

and so $\mathcal{L}(\eta\,|\,\xi)=\mu_\xi$ a.s.

(ii) Since T is Borel, there exists a probability kernel $\mu: S\to T$ with $\mu_\xi=\mathcal{L}(\eta\,|\,\xi)$ a.s. Then (i) yields a random element $\tilde\eta$ with $\mathcal{L}(\tilde\eta\,|\,\tilde\xi)=\mu_{\tilde\xi}$. Since $\tilde\xi\overset{d}{=}\xi$, we obtain $(\tilde\xi,\tilde\eta)\overset{d}{=}(\xi,\eta)$. □

For any s-finite kernels $\mu\colon S\to T$ and $\nu\colon S\times T\to U$, the *composition* $\mu\otimes\nu$ and *product* $\mu\nu$ are defined as the kernels from S to $T\times U$ and U, respectively, given by

$$(\mu\otimes\nu)_s f = \int\mu_s(dt)\int\nu_{s,t}(du)\,f(t,u), \tag{4}$$

$$\begin{aligned} (\mu\nu)_s f &= \int\mu_s(dt)\int\nu_{s,t}(du)\,f(u) \\ &= (\mu\otimes\nu)_s(1_T\otimes f). \end{aligned}$$

Note that $\mu\nu$ equals the projection of $\mu\otimes\nu$ onto U. To state the following result, we write $A_\mu f(s)=\mu_s f$ and $\hat\mu=\delta\otimes\mu$.

Lemma 1.17 *(compositions and products) For any s-finite kernels $\mu\colon S\to T$ and $\nu\colon S\times T\to U$, we have*

- (i) *$\mu\otimes\nu$ and $\mu\nu$ are s-finite kernels from S to $T\times U$ and U, respectively, and $\mu\otimes\nu$ is σ-finite whenever this holds for μ and ν,*
- (ii) *the kernel operations $\mu\otimes\nu$ and $\mu\nu$ are associative,*
- (iii) *$\mu\hat\nu=\mu\otimes\nu$ and $\hat\mu\hat\nu=(\mu\otimes\nu)\hat{}$,*
- (iv) *for kernels $\mu\colon S\to T$ and $\nu\colon T\to U$, we have $A_{\mu\nu}=A_\mu A_\nu$.*

Proof: (i) By Lemma 1.15 (i) applied twice, the inner integral in (4) is $\mathcal{S}\otimes\mathcal{T}$-measurable and the outer integral is $\mathcal{S}$-measurable. The countable additivity holds by repeated monotone convergence. This proves the kernel property of $\mu\otimes\nu$, and the result for $\mu\nu$ is an immediate consequence.

(ii) Consider any kernels $\mu\colon S\to T$, $\nu\colon S\times T\to U$, and $\rho\colon S\times T\times U\to V$. Letting $f\in(\mathcal{T}\otimes\mathcal{U}\otimes\mathcal{V})_+$ be arbitrary and using (4) repeatedly, we get

$$\begin{aligned} \{\mu\otimes(\nu\otimes\rho)\}_s f &= \int\mu_s(dt)\iint(\nu\otimes\rho)_{s,t}(du\,dv)\,f(t,u,v) \\ &= \int\mu_s(dt)\int\nu_{s,t}(du)\int\rho_{s,t,u}(dv)\,f(t,u,v) \\ &= \iint(\mu\otimes\nu)_s(dt\,du)\int\rho_{s,t,u}(dv)\,f(t,u,v) \\ &= \{(\mu\otimes\nu)\otimes\rho\}_s f, \end{aligned}$$

which shows that $\mu \otimes (\nu \otimes \rho) = (\mu \otimes \nu) \otimes \rho$. Projecting onto V yields $\mu(\nu\rho) = (\mu\nu)\rho$.

(iii) Consider any kernels $\mu\colon S \to T$ and $\nu\colon S \times T \to U$, and note that $\mu\hat{\nu}$ and $\mu \otimes \nu$ are both kernels from S to $T \times U$. For any $f \in (\mathcal{T} \otimes \mathcal{U})_+$, we have

$$\begin{aligned}
(\mu\hat{\nu})_s f &= \int \mu_s(dt) \int \delta_{s,t}(ds'dt') \int \nu_{s',t'}(du)\, f(t',u) \\
&= \int \mu_s(dt) \int \nu_{s,t}(du)\, f(t,u) \\
&= (\mu \otimes \nu)_s f,
\end{aligned}$$

which shows that $\mu\hat{\nu} = \mu \otimes \nu$. Hence, by (ii)

$$\begin{aligned}
\hat{\mu}\hat{\nu} &= (\delta \otimes \mu) \otimes \nu \\
&= \delta \otimes (\mu \otimes \nu) = (\mu \otimes \nu)\hat{}\,.
\end{aligned}$$

(iv) For any $f \in \mathcal{U}_+$, we have

$$\begin{aligned}
A_{\mu\nu} f(s) = (\mu\nu)_s f &= \int (\mu\nu)_s(du)\, f(u) \\
&= \int \mu_s(dt) \int \nu_t(du)\, f(u) \\
&= \int \mu_s(dt)\, A_\nu f(t) \\
&= A_\mu(A_\nu f)(s) = (A_\mu A_\nu) f(s),
\end{aligned}$$

which implies $A_{\mu\nu} f = (A_\mu A_\nu) f$, and hence $A_{\mu\nu} = A_\mu A_\nu$. □

The measurability of kernels can often be verified by means of Laplace transforms. Given a kernel $\nu\colon T \to \mathcal{M}_S$, we introduce the associated Laplace operator

$$L\nu_t(f) = \int e^{-\mu f} \nu_t(d\mu), \quad f \in \mathcal{S}_+.$$

Lemma 1.18 *(kernels into a measure space) For any measurable space T, consider a function $\nu\colon T \to \mathcal{M}_{\mathcal{M}_S}$ with $L\nu_t(f_0) < \infty$ for some fixed $f_0 \in \mathcal{S}_+$. Then ν is a kernel from T to $\mathcal{M}_S$ iff $L\nu_t(f)$ is $\mathcal{T}$-measurable in $t \in T$ for all $f \in \mathcal{S}_+$.*

Proof: Let $f_0 \in \mathcal{S}_+$ be such that $L\nu_t(f_0) < \infty$ for all $t \in T$. Then the measures $\nu'_t(d\mu) = e^{-\mu f_0}\, \nu_t(d\mu)$ are bounded and satisfy the same condition, and so we may assume that $\|\nu_t\| < \infty$ for all t. By the Stone-Weierstrass theorem, any continuous function $g : [0,\infty]^n \to \mathsf{R}_+$ admits a uniform approximation by linear combinations of functions e^{-rx}, $r \in \mathsf{R}_+^n$. Thus, for any $f_1, \dots, f_n \in \mathcal{S}_+$, the integral $\int g(\mu f_1, \dots, \mu f_n)\, \nu_t(d\mu)$ can be approximated by linear combinations of functions $L\nu_t(f)$, and is therefore measurable. By dominated convergence, $\nu_t \bigcap_k \{\mu;\, \mu f_k \le x_k\}$ is then measurable for all

$x_1, \dots, x_n \in \mathbb{R}_+$. Since the sets $\bigcap_k \{\mu;\ \mu f_k \le x_k\}$ form a π-system $\mathcal{C}$, whereas the sets $A \in \mathcal{B}_{\mathcal{M}_S}$ where $\nu_t A$ is measurable form a λ-system $\mathcal{D}$, Lemma 1.2 yields $\mathcal{D} \supset \sigma(\mathcal{C}) = \mathcal{B}_{\mathcal{M}_S}$, as required. □

We conclude with some simple norm relations. Here $\|\mu\|$ denotes the total variation of the signed measure μ.

Lemma 1.19 *(product comparison) For any measurable spaces $S_1, \dots, S_n$, let μ_j and ν_j be probability kernels from $S_1 \times \cdots \times S_{j-1}$ to S_j, $j = 1, \dots, n$. Then*

$$\Big\| \bigotimes\nolimits_j \mu_j - \bigotimes\nolimits_j \nu_j \Big\| \le \sum_{k \le n} (\mu_1 \otimes \cdots \otimes \mu_{k-1})\, \|\mu_k - \nu_k\|.$$

Proof: For any probability measures μ and μ' on S, probability kernels $\nu, \nu' \colon S \to T$, and measurable function f on $S \times T$ with $|f| \le 1$, we have

$$\begin{aligned} |((\mu - \mu') \otimes \nu) f| &= |(\mu - \mu')(\nu f)| \le \|\mu - \mu'\|, \\ |(\mu \otimes (\nu - \nu')) f| &= |\mu((\nu - \nu') f)| \\ &\le \mu|(\nu - \nu') f| \le \mu\|\nu - \nu'\|, \end{aligned}$$

which shows that

$$\begin{aligned} \|(\mu - \mu') \otimes \nu\| &\le \|\mu - \mu'\|, \\ \|\mu \otimes (\nu - \nu')\| &\le \mu\, \|\nu - \nu'\|. \end{aligned}$$

Hence,

$$\begin{aligned} \Big\| \bigotimes\nolimits_j \mu_j - \bigotimes\nolimits_j \nu_j \Big\| &\le \sum_{k \le n} \Big\| \Big(\bigotimes\nolimits_{i<k} \mu_i \Big) \otimes (\mu_k - \nu_k) \otimes \Big(\bigotimes\nolimits_{j>k} \nu_j \Big) \Big\| \\ &\le \sum_{k \le n} \Big\| \Big(\bigotimes\nolimits_{i<k} \mu_i \Big) \otimes (\mu_k - \nu_k) \Big\| \\ &\le \sum_{k \le n} \Big(\bigotimes\nolimits_{i<k} \mu_i \Big) \|\mu_k - \nu_k\|. \end{aligned}$$

□

We turn to a special case where the previous norm relation reduces to an equality. For organizational clarity, we may occasionally anticipate results from subsequent sections.

Lemma 1.20 *(norm identity) For any measurable space S and Borel space T, let $\nu \in \hat{\mathcal{M}}_S$, and consider some bounded kernels $\mu_1, \mu_2 : S \to T$. Then $\|\mu_1 - \mu_2\|$ is measurable, and*

$$\|\nu \otimes (\mu_1 - \mu_2)\| = \nu \|\mu_1 - \mu_2\|.$$

Proof: Write $\mu = \mu_1 - \mu_2$ and $\hat{\mu} = \mu_1 + \mu_2$. By Theorem 1.28 below there exist some measurable functions $f_1, f_2 \colon S \times T \to [-1, 1]$ such that $\mu_i = f_i \cdot \hat{\mu}$

for $i = 1, 2$. Writing $f = f_1 - f_2$, we get $\|\mu\| = \hat\mu|f|$, which is measurable by Lemma 1.15. Furthermore,

$$\begin{aligned}\|\nu \otimes \mu\| &= \|f \cdot (\nu \otimes \hat\mu)\| \\ &= (\nu \otimes \hat\mu)|f| \\ &= \nu(\hat\mu|f|) = \nu\|\mu\|. \end{aligned}$$ □

We conclude with a couple of technical results needed in Chapter 13.

Lemma 1.21 *(regularization) Let ν be a kernel from R to a Polish space S with* $\operatorname{supp}\nu = S$*, and let μ and $\mu_1, \mu_2, \ldots$ be bounded kernels from $S \times \mathsf{R}$ to a Borel space U, where each μ_n is continuous in total variation on $S \times G_n$ for some open sets $G_n \subset \mathsf{R}$. Assume $\|\mu - \mu_n\| \le h_n$ a.e. ν for some measurable functions $h_n \ge 0$ on $S \times \mathsf{R}$, where $h_n \to 0$ uniformly on bounded sets. Then $\mu = \mu'$ a.e. ν, where μ' is continuous in total variation on $S \times \limsup_n G_n$.*

Proof: First let the kernels $\nu, \mu, \mu_1, \mu_2, \ldots$ and functions $h_1, h_2, \ldots$ be independent of the real parameter, hence kernels or functions on S. Let $S' \subset S$ be the set where $\|\mu - \mu_n\| \le h_n$, so that $\nu(S')^c = 0$. For any $t, t' \in S'$, we have

$$\|\mu_t - \mu_{t'}\| \le \|\mu_t - \mu_t^n\| + \|\mu_t^n - \mu_{t'}^n\| + \|\mu_{t'}^n - \mu_{t'}\|,$$

where $\mu^n = \mu_n$. Fixing any $s \in S$ and letting $t, t' \to s$ and then $n \to \infty$, we get $\|\mu_t - \mu_{t'}\| \to 0$, and so Lemma 1.10 yields a bounded measure μ_s' on U with $\|\mu_t - \mu_s'\| \to 0$. Here μ_s' is well defined for every $s \in S$, and $\mu_s' = \mu_s$ when $s \in S'$.

To prove the stated continuity of μ', let $s_n \to s$ in S. Given a metrization d of S, we may choose some $t_1, t_2, \ldots \in S'$ with

$$d(s_k, t_k) + \|\mu'_{s_k} - \mu_{t_k}\| < 2^{-k}, \quad k \in \mathsf{N}.$$

In particular, we have $t_k \to s$, and so

$$\|\mu'_{s_k} - \mu'_s\| \le \|\mu'_{s_k} - \mu_{t_k}\| + \|\mu_{t_k} - \mu'_s\| \to 0,$$

as desired. The continuity of μ' implies its measurability, which means that μ' is again a locally bounded kernel from S to U. Further note that $\|\mu_n - \mu'\| \le h_n$ on S', which extends by continuity to $\|\mu_n - \mu'\| \le h_n'$ on S, where the functions h_n' are upper semi-continuous versions of h_n satisfying the same convergence condition.

Now allow $\nu, \mu, \mu_1, \mu_2, \ldots$ to depend on a parameter $x \in \mathsf{R}$. Constructing μ' as before for each x, we get $\|\mu_n - \mu'\| \to 0$, uniformly on bounded sets in $S \times \mathsf{R}$. Since each μ_n is continuous in total variation on $S \times G_n$, the same continuity holds for μ' on $S \times \limsup_n G_n$. □

Lemma 1.22 *(projection) Given three Borel spaces S, T, U, consider some σ-finite measures ν on $S \times U$ and $\hat\nu$ on S, along with some signed kernels $\mu\colon S \times U \to T$ and $\hat\mu\colon S \to T$, such that ν and $\nu \otimes \mu$ have projections $\hat\nu$ and $\hat\nu \otimes \hat\mu$ onto S and $S \times T$, respectively. Then $\|\hat\mu_s\| \le \sup_u \|\mu_{s,u}\|$ a.e. $\hat\nu$.*

Proof: Since $\hat{\nu}$ is σ-finite and U is Borel, Theorem 1.23 below yields $\nu = \hat{\nu} \otimes \rho$ for some probability kernel $\rho\colon S \to T$. Writing $\pi_{S\times T}$ for projection onto $S \times T$, we obtain

$$\begin{aligned}\hat{\nu} \otimes \hat{\mu} &= (\nu \otimes \mu) \circ \pi_{S\times T}^{-1} \\ &= (\hat{\nu} \otimes \rho \otimes \mu) \circ \pi_{S\times T}^{-1} \\ &= \hat{\nu} \otimes \rho\mu,\end{aligned}$$

and so $\hat{\mu} = \rho\mu$ a.e. $\hat{\nu}$. Hence, for any measurable function f on T with $|f| \le 1$, we get a.e.

$$\begin{aligned}|\hat{\mu} f| &= |(\rho\mu) f| = |\rho(\mu f)| \\ &\le \rho|\mu f| \le \rho\|\mu\|,\end{aligned}$$

which implies

$$\|\hat{\mu}_s\| \le \rho_s\|\mu_s\| \le \sup\nolimits_s \|\mu_{s,u}\| \ \text{ a.e. } \hat{\nu}. \qquad \square$$

1.4 Disintegration

In the previous section, we considered the composition $\rho = \nu \otimes \mu$ of two kernels $\nu\colon S \to T$ and $\mu\colon S \times T \to U$, resulting in a kernel $\rho\colon S \to T \times U$. For singleton S, the kernels ν and $\rho = \nu \otimes \mu$ reduce to measures on T and $T \times U$, respectively, and μ becomes a kernel from T to U.

Here we turn to the reverse problem of representing a given measure ρ on $S \times T$ in the form $\nu \otimes \mu$, for some measure ν on S and kernel $\mu\colon S \to T$. This so-called *disintegration* may be regarded as an infinitesimal decomposition of ρ into measures $\delta_s \otimes \mu_s$. Any σ-finite measure $\nu \sim \rho(\cdot \times T)$ is called a *supporting measure* of ρ, and we refer to μ as the associated *disintegration kernel.*

We begin with some basic properties of existence and uniqueness.

Theorem 1.23 *(disintegration) Let ρ be a σ-finite measure on $S \times T$, where T is Borel. Then there exist a σ-finite measure ν on S and a σ-finite kernel $\mu\colon S \to T$ such that $\rho = \nu \otimes \mu$. Here we may choose $\nu \sim \hat{\rho} \equiv \rho(\cdot \times T)$, in which case the μ_s are unique for $s \in S$ a.e. ν. They may further be chosen to be probability measures iff $\hat{\rho}$ is σ-finite and equal to ν.*

Proof: When $\|\rho\| = 1$, we may regard ρ as the distribution of a random pair (σ, τ) in $S \times T$. By FMP 6.3 (cf. Theorem 2.15 below), the set of conditional probabilities $P(\tau \in \cdot \mid \sigma)$ admits a regular version, in terms of a probability kernel $\mu\colon S \to T$ satisfying $\mathcal{L}(\tau \mid \sigma) = \mu_\sigma$ a.s. Writing $\nu = \mathcal{L}(\sigma)$, we have for any $f \in (\mathcal{S} \otimes \mathcal{T})_+$

$$\begin{aligned}\rho f = E f(\sigma,\tau) &= E\int \mu(\sigma, dt)\, f(\sigma, t) \\ &= \int \nu(ds) \int \mu(s, dt)\, f(s,t) \\ &= (\nu \otimes \mu) f.\end{aligned}$$

Indeed, the relation is obvious for indicator functions $1_{B\times A}$ with $B \in \mathcal{S}$ and $A \in \mathcal{T}$, and it extends to general f by a monotone-class argument and monotone convergence. This shows that $\rho = \nu \otimes \mu$.

For a general σ-finite ρ, we have $\rho = g \cdot \tilde{\rho}$ for some probability measure $\tilde{\rho}$ on $S \times T$ and function $g \in (\mathcal{S} \otimes \mathcal{T})_+$, and so $\tilde{\rho} = \nu \otimes \tilde{\mu}$ as before for some probability measure ν on S and a probability kernel $\mu\colon S \to T$. By Lemma 1.15 (i), we see that $\mu = g \cdot \tilde{\mu}$ is again a σ-finite kernel from S to T, and so for any $f \in (\mathcal{S} \otimes \mathcal{T})_+$,

$$\begin{aligned}(\nu \otimes \mu)f &= (\mu \otimes (g \cdot \tilde{\mu}))f \\ &= (\mu \otimes \tilde{\mu})(fg) \\ &= \tilde{\rho}(fg) = \rho f,\end{aligned}$$

which shows that $\rho = \nu \otimes \mu$.

If $\nu = \rho(\cdot \times T)$ is σ-finite and non-zero, we may choose a measurable function $g > 0$ on S, such that $\tilde{\nu} = g \cdot \nu$ becomes a probability measure on S. Then $\tilde{\rho} = \tilde{\nu} \otimes \mu$ is a probability measure on $S \times T$, and $\tilde{\rho} = \tilde{\nu} \otimes \mu$ holds as before for some probability kernel $\mu\colon S \to T$, which implies $\rho = \nu \otimes \mu$. Conversely, if $\rho = \nu \otimes \mu$ for some probability kernel μ, then $\rho(B \times T) = \int_B \nu(ds)\, \mu_s T = \nu B$, and so $\rho = \nu$, which is then σ-finite.

To prove the stated uniqueness, suppose that $\rho = \nu \otimes \mu = \nu' \otimes \mu'$ with $\nu \sim \nu' \sim \rho(\cdot \times T)$. Then

$$\begin{aligned}g \cdot \rho &= \nu \otimes (g \cdot \mu) \\ &= \nu' \otimes (g \cdot \mu'), \quad g \in (\mathcal{S} \otimes \mathcal{T})_+,\end{aligned}$$

which allows us to consider only bounded measures ρ. Since also $(h \cdot \nu) \otimes \mu = \nu' \otimes (h \cdot \mu')$ for any $h \in \mathcal{S}_+$, we may further assume that $\nu = \nu'$. Then

$$\int_A \nu(ds)\, (\mu_s B - \mu'_s B) = 0, \quad A \in \mathcal{S},\ B \in \mathcal{T},$$

which implies $\mu_s B = \mu'_s B$ a.e. ν. Since T is Borel, a monotone-class argument yields $\mu_s = \mu'_s$ a.e. ν. □

The following partial disintegration will be needed in Chapter 8, for the purpose of constructing Gibbs and Papangelou kernels.

Corollary 1.24 *(partial disintegration) Consider any σ-finite measures ν on S and ρ on $S \times T$, where T is Borel. Then there exists an a.e. unique maximal kernel $\mu\colon S \to T$ with $\nu \otimes \mu \le \rho$.*

Here the *maximality* of μ means that, whenever μ' is a kernel satisfying $\nu \otimes \mu' \le \rho$, we have $\mu'_s \le \mu_s$ for $s \in S$ a.e. ν.

Proof: Since ρ is σ-finite, we may choose a σ-finite measure γ on S with $\gamma \sim \rho(\cdot \times T)$. Consider its Lebesgue decomposition $\gamma = \gamma_a + \gamma_s$ with respect to ν (FMP 2.10), so that $\gamma_a \ll \nu$ and $\gamma_s \perp \nu$. Choose $A \in \mathcal{S}$ with $\nu A^c = \gamma_s A = 0$, and let ρ' denote the restriction of ρ to $A \times T$. Then $\rho'(\cdot \times T) \ll \nu$, and so Theorem 1.23 yields a σ-finite kernel $\mu \colon S \to T$ with $\nu \otimes \mu = \rho' \leq \rho$.

Now consider any kernel $\mu' : S \to T$ with $\nu \otimes \mu' \leq \rho$. Since $\nu A^c = 0$, we have $\nu \otimes \mu' \leq \rho' = \nu \otimes \mu$. Since μ and μ' are σ-finite, there exists a measurable function $g > 0$ on $S \times T$, such that the kernels $\tilde{\mu} = g \cdot \mu$ and $\tilde{\mu}' = g \cdot \mu'$ satisfy $\|\tilde{\mu}_s\| \vee \|\tilde{\mu}'_s\| \leq 1$. Then $\nu \otimes \tilde{\mu}' \leq \nu \otimes \tilde{\mu}$, and so $\tilde{\mu}'_s B \leq \tilde{\mu}_s B$, $s \in S$ a.e. ν, for every $B \in \mathcal{S}$. Hence, Lemma 2.1 (i) yields $\mu'_s \leq \mu_s$ a.e. ν, which shows that μ is maximal. In particular, μ is a.e. unique. □

Anticipating the basic differentiation Theorem 1.28 and its corollary, we proceed to the more difficult disintegration of kernels.

Theorem 1.25 *(disintegration of kernels) Consider a σ-finite kernel $\rho : S \to T \times U$, where U is Borel.*

(i) *If S is countable or T is Borel, there exist some σ-finite kernels $\nu \colon S \to T$ and $\mu \colon S \times T \to U$ such that $\rho = \nu \otimes \mu$.*

(ii) *For any probability measure P on S, there exist some kernels ν and μ as above, such that $P \otimes \rho = P \otimes \nu \otimes \mu$.*

The assertions remain true for any fixed, σ-finite kernel $\nu : S \to T$ with $\nu_s \sim \rho_s(\cdot \times U)$ for all $s \in S$.

Proof: (i) For any probability kernel $\rho : S \to T \times U$, the projections $\nu_s = \rho_s(\cdot \times U)$ form a probability kernel from S to T. If T is Borel, Theorem 1.28 yields for any $B \in \mathcal{U}$ a product-measurable function $h_B \geq 0$ on $S \times T$ with $\rho_s(\cdot \times B) = h_B(s, \cdot) \cdot \nu_s$ for all s. For fixed $s \in S$, we may proceed as in FMP 6.3 (cf. Theorem 2.15 below) to construct a probability kernel $\mu_s \colon T \to U$, such that $\mu_s(\cdot, B) = h_B(s, \cdot)$ a.s. ν_s for all $B \in \mathcal{B}$. This involves modifications of $h = (h_B)$ on countably many ν_s-null sets, and since the latter are all product-measurable, μ is indeed a kernel from $S \times T$ to U.

If S is countable, then for any T and fixed $s \in S$, we may apply the elementary disintegration in Theorem 1.23 to construct a kernel $\mu_{s,t} : T \to U$. Since any T-measurable function indexed by S is automatically product-measurable, μ is again a kernel from $S \times T$ to U. In either case, the relationship between ρ, ν, and h extends to $\rho = \nu \otimes \mu$.

For any σ-finite kernel ρ, we may choose a measurable function $g > 0$ on $S \times T \times U$, such that $\rho_s g(s, \cdot) = 1$ on the set $S' = \{s \in S;\, \rho_s \neq 0\}$. Then $g \cdot \rho$ is a probability kernel from S' to $T \times U$, and so $g \cdot \rho = \nu \otimes \mu$ holds as before for some probability kernels $\nu \colon S' \to T$ and $\mu \colon S' \times T \to U$. Writing $h = 1/g$, we get $\rho = \nu \otimes (h \cdot \mu)$ on S', where $h \cdot \mu$ is a σ-finite kernel from $S' \times T$ to U. Finally, we may choose $\nu_s = \mu_{s,t} = 0$ for $s \in S \setminus S'$.

Now fix any σ-finite kernels $\nu : S \to T$ and $\rho : S \to T \times U$ with $\nu_s \sim \rho_s(\cdot \times U)$ for all $s \in S$. As before, we may assume that $\|\nu_s\| = \|\rho_s\| = 1$ on S'. Putting $\nu'_s = \rho_s(\cdot \times U)$, we get a disintegration $\rho = (\nu + \nu') \otimes \mu$ in terms of a kernel $\mu\colon S\times T\to U$. In particular, $\nu'=h\cdot(\nu+\nu')$ with $h(s,t)=\mu_{s,t}U$, and so $\nu=(1-h)\cdot(\nu+\nu')$. Since $\nu\sim\nu'$, we obtain $\nu+\nu'=h'\cdot\nu$ with $h'=(1-h)^{-1}1\{h<1\}$, which implies $\rho=\nu\otimes\mu'$ with $\mu'=h'\mu$.

(ii) Again, we may take $\|\rho_s\|\equiv 1$. For any P, we may form a probability measure $\gamma=P\otimes\rho$ on $S\times T\times U$. Projecting onto $S\times T$ gives $\tilde\gamma=P\otimes\nu$ with $\tilde\gamma=\gamma(\cdot\times U)$ and $\nu=\rho(\cdot\times U)$, and so Theorem 1.23 yields $\gamma=\tilde\gamma\otimes\mu$ for some probability kernel $\mu\colon S\times T\to U$. Then by Lemma 1.17 (ii),

$$\begin{aligned}P\otimes\rho=\gamma&=\tilde\gamma\otimes\mu\\&=(P\otimes\nu)\otimes\mu\\&=P\otimes(\nu\otimes\mu).\end{aligned}$$

If even ν is fixed with $\nu_s\sim\rho_s(\cdot\times U)$, then $P\otimes\nu\sim(P\otimes\rho)(\cdot\times U)$, and we may choose μ as a kernel from $S\times T$ to U satisfying $P\otimes\rho=(P\otimes\nu)\otimes\mu$. The assertion now follows as before. □

Regarding the measures ρ and ν as measurable functions of the pair (ρ,ν), hence as kernels on $\mathcal{M}_{S\times T}\times\mathcal{M}_S$, we obtain the following measurable version of Theorem 1.23, needed repeatedly below.

Corollary 1.26 *(universal disintegration) For any Borel spaces S and T, there exists a kernel $\mu\colon S\times\mathcal{M}_{S\times T}\times\mathcal{M}_S\to T$, such that*

$$\rho=\nu\otimes\mu(\cdot,\rho,\nu),\quad \rho\in\mathcal{M}_{S\times T},\ \ \nu\in\mathcal{M}_S\ \ \textit{with}\ \ \rho(\cdot\times T)\ll\nu.$$

Here $\mu(\cdot,\rho,\nu)$ is unique a.e. ν for fixed ρ and ν.

For any $f\in(\mathcal{S}\otimes\mathcal{T})_+$, we may write the latter disintegration more explicitly as

$$\begin{aligned}\rho f&\equiv\iint\rho(ds\,dt)\,f(s,t)\\&=\int\nu(ds)\int\mu(s,\rho,\nu)(dt)\,f(s,t).\end{aligned}$$

We turn to the subject of iterated disintegration, needed in Chapter 6 for certain conditional and recursive constructions of Palm kernels. To explain the notation, consider some σ-finite measures μ_1, μ_2, μ_3, μ_{12}, μ_{13}, μ_{23}, μ_{123} on products of the Borel spaces S, T, U. They are said to form a *supportive* family, if all relations of the form $\mu_1\sim\mu_{12}(\cdot\times T)$ or $\mu_{12}\sim\mu_{123}(\cdot\times U)$ are satisfied. (For functions or constants $a,b\ge 0$, the relation $a\sim b$ means that $a=0$ iff $b=0$.) Then by Theorem 1.23 we have the disintegrations

$$\begin{aligned}\mu_{12}&=\mu_1\otimes\mu_{2|1}\cong\mu_2\otimes\mu_{1|2},\\\mu_{13}&=\mu_1\otimes\mu_{3|1},\\\mu_{123}&=\mu_{12}\otimes\mu_{3|12}=\mu_1\otimes\mu_{23|1}\\&\cong\mu_2\otimes\mu_{13|2},\end{aligned}$$

where $\mu_{2|1}$, $\mu_{3|12}$, $\mu_{12|3}$, ... are σ-finite kernels between appropriate spaces, and $\cong$ denotes equality apart from the order of component spaces. If $\mu_{2|1} \sim \mu_{23|1}(\cdot \times U)$ a.e. μ_1, we may proceed to form some iterated disintegration kernels such as $\mu_{3|2|1}$, where $\mu_{23|1} = \mu_{2|1} \otimes \mu_{3|2|1}$ a.e. μ_1. We show that the required support properties hold automatically, and that the iterated disintegration kernels $\mu_{3|2|1}$ and $\mu_{3|1|2}$ agree a.e. and are equal to the single disintegration kernel $\mu_{3|12}$.

Theorem 1.27 *(iterated disintegration) Consider a supportive family of σ-finite measures μ_1, μ_2, μ_3, μ_{12}, μ_{13}, μ_{23}, μ_{123} on suitable products of the Borel spaces S, T, U, and introduce the associated disintegration kernels $\mu_{1|2}$, $\mu_{2|1}$, $\mu_{3|1}$, $\mu_{3|12}$. Then*

(i) $\mu_{2|1} \sim \mu_{23|1}(\cdot \times U)$ *a.e.* μ_1,
$\mu_{1|2} \sim \mu_{13|2}(\cdot \times U)$ *a.e.* μ_2,

(ii) $\mu_{3|12} = \mu_{3|2|1} \cong \mu_{3|1|2}$ *a.e.* μ_{12},

(iii) *if* $\mu_{13} = \mu_{123}(\cdot \times T \times \cdot)$, *then also*
$\mu_{3|1} = \mu_{2|1}\, \mu_{3|1|2}$ *a.e.* μ_1.

Proof: (i) We may clearly assume that $\mu_{123} \neq 0$. Then choose a probability measure $\tilde{\mu}_{123} \sim \mu_{123}$, and let $\tilde{\mu}_1$, $\tilde{\mu}_2$, $\tilde{\mu}_3$, $\tilde{\mu}_{12}$, $\tilde{\mu}_{13}$, $\tilde{\mu}_{23}$ be projections of $\tilde{\mu}_{123}$ onto S, T, U, $S \times T$, $S \times U$, $T \times U$, respectively. Then for any $A \in \mathcal{S} \otimes \mathcal{T}$,

$$\begin{aligned}\tilde{\mu}_{12}A &= \tilde{\mu}_{123}(A \times U)\\ &\sim \mu_{123}(A \times U) \sim \mu_{12}A,\end{aligned}$$

which shows that $\tilde{\mu}_{12} \sim \mu_{12}$. A similar argument gives $\tilde{\mu}_1 \sim \mu_1$ and $\tilde{\mu}_2 \sim \mu_2$. Hence, the Radon–Nikodym theorem yields some measurable functions p_1, p_2, p_{12}, $p_{123} > 0$ on S, T, $S \times T$, and $S \times T \times U$, satisfying

$$\begin{aligned}\tilde{\mu}_1 &= p_1 \cdot \mu_1, & \tilde{\mu}_2 &= p_2 \cdot \mu_2,\\ \tilde{\mu}_{12} &= p_{12} \cdot \mu_{12}, & \tilde{\mu}_{123} &= p_{123} \cdot \mu_{123}.\end{aligned}$$

Inserting those densities into the disintegrations

$$\begin{aligned}\tilde{\mu}_{12} &= \tilde{\mu}_1 \otimes \tilde{\mu}_{2|1} \cong \tilde{\mu}_2 \otimes \tilde{\mu}_{1|2},\\ \tilde{\mu}_{123} &= \tilde{\mu}_1 \otimes \tilde{\mu}_{23|1} \cong \tilde{\mu}_2 \otimes \tilde{\mu}_{13|2},\end{aligned} \tag{5}$$

we obtain

$$\begin{aligned}\mu_{12} &= \mu_1 \otimes (p_{2|1} \cdot \tilde{\mu}_{2|1})\\ &\cong \mu_2 \otimes (p_{1|2} \cdot \tilde{\mu}_{1|2}),\\ \mu_{123} &= \mu_1 \otimes (p_{23|1} \cdot \tilde{\mu}_{23|1})\\ &\cong \mu_2 \otimes (p_{13|2} \cdot \tilde{\mu}_{13|2}),\end{aligned}$$

where

$$p_{2|1} = \frac{p_1}{p_{12}}, \quad p_{1|2} = \frac{p_2}{p_{12}}, \quad p_{23|1} = \frac{p_1}{p_{123}}, \quad p_{13|2} = \frac{p_2}{p_{123}}.$$

Comparing with the disintegrations of μ_{12} and μ_{123} and invoking the uniqueness in Theorem 1.23, we obtain a.e.

$$\begin{aligned}\tilde\mu_{2|1} &\sim \mu_{2|1}, & \tilde\mu_{1|2} &\sim \mu_{1|2},\\ \tilde\mu_{23|1} &\sim \mu_{23|1}, & \tilde\mu_{13|2} &\sim \mu_{13|2}.\end{aligned} \tag{6}$$

Furthermore, we get from (5)

$$\begin{aligned}\tilde\mu_1 \otimes \tilde\mu_{2|1} &= \tilde\mu_{12} = \tilde\mu_{123}(\cdot \times U)\\ &= \tilde\mu_1 \otimes \tilde\mu_{23|1}(\cdot \times U),\end{aligned}$$

and similarly with 1 and 2 interchanged, and so the mentioned uniqueness yields a.e.

$$\tilde\mu_{2|1} = \tilde\mu_{23|1}(\cdot \times U), \qquad \tilde\mu_{1|2} = \tilde\mu_{13|2}(\cdot \times U). \tag{7}$$

Combining (6) and (7), we get a.e.

$$\begin{aligned}\mu_{2|1} &\sim \tilde\mu_{2|1} = \tilde\mu_{23|1}(\cdot \times U)\\ &\sim \mu_{23|1}(\cdot \times U),\end{aligned}$$

and similarly with 1 and 2 interchanged.

(ii) By (i) and Theorem 1.25, we have a.e.

$$\mu_{23|1} = \mu_{2|1} \otimes \mu_{3|2|1}, \qquad \mu_{13|2} = \mu_{1|2} \otimes \mu_{3|1|2},$$

for some product-measurable kernels $\mu_{3|2|1}$ and $\mu_{3|1|2}$. Combining the various disintegrations and using the commutativity in Lemma 1.17 (ii), we get

$$\begin{aligned}\mu_{12} \otimes \mu_{3|12} &= \mu_{123} = \mu_1 \otimes \mu_{23|1}\\ &= \mu_1 \otimes \mu_{2|1} \otimes \mu_{3|2|1}\\ &= \mu_{12} \otimes \mu_{3|2|1},\end{aligned}$$

and similarly with 1 and 2 interchanged. It remains to use the a.e. uniqueness in Theorem 1.23.

(iii) Under the stated hypothesis,

$$\begin{aligned}\mu_1 \otimes \mu_{3|1} &= \mu_{13} = \mu_{123}(\cdot \times T \times \cdot)\\ &= \mu \otimes \mu_{23|1}(T \times \cdot),\end{aligned}$$

and so by (ii) and the uniqueness in Theorem 1.23, we have a.e.

$$\begin{aligned}\mu_{3|1} &= \mu_{23|1}(T \times \cdot)\\ &= \mu_{3|1} \otimes \mu_{3|2|1}(T \times \cdot)\\ &= \mu_{3|1}\, \mu_{3|2|1} = \mu_{3|1}\, \mu_{3|1|2},\end{aligned}$$

as required. $\square$

1.5 Differentiation

We begin with some kernel versions of the Lebesgue decomposition and Radon–Nikodym theorem. The result was used in the previous section to prove the more general Theorem 1.25.

Theorem 1.28 *(kernel density) Let μ and ν be σ-finite kernels from S to T, where T is Borel. Then there exists a measurable function $f: S\times T \to [0,\infty]$, such that $\mu = f\cdot\nu$ on $\{f<\infty\}$ and $\mu \perp \nu$ on $\{f=\infty\}$. Furthermore, the sets $\{\mu_s \ll \nu_s\}$ and $\{\mu_s \perp \nu_s\}$ are measurable, and f is unique a.e. $\mu+\nu$.*

Proof (de Possel, Doob): Define $\rho = \mu+\nu$. Discarding the set where $\rho_s = 0$, and noting that both hypothesis and assertion are invariant under changes to equivalent measures, we may assume that $\|\rho_s\| = 1$ for all s. Fixing a dissection system (I_{nj}) on T and writing $I_n(t)$ for the set I_{nj} containing t, we define

$$g_n(s,t) = \sum_{j\geq 1} \frac{\mu_s I_n(t)}{\rho_s I_n(t)}, \quad s\in S,\ t\in T,\ n\in \mathsf{N},$$

where we may take $0/0 = 0$. For fixed $s\in S$, the functions $g_n(s,t)$ form a bounded ρ_s-martingale for the filtration $\mathcal{T}_n = \sigma(I_{n1}, I_{n2}, \ldots)$. Since $\mathcal{T} = \sigma(\mathcal{T}_1, \mathcal{T}_2, \ldots)$, the g_n converge a.s. and in L^1 with respect to ρ_s toward a $\mathcal{T}$-measurable function g, admitting the product-measurable version

$$g(s,t) = \limsup_{n\to\infty} g_n(s,t), \quad s\in S,\ t\in T.$$

Since $\mu_s = g_n(s,\cdot)\cdot\rho_s$ on $\mathcal{T}_n$, the L^1-convergence yields $\mu = g\cdot\rho$ on $\mathcal{T}_n$ for every n, which extends to $\mathcal{T}$ by a monotone-class argument. Hence, $(1-g)\cdot\mu = g\cdot\nu$, and we may take $f = g/(1-g)$ with $1/0 = \infty$.

The absolutely continuous component $\mu^a = f\cdot\nu$ is measurable by Lemma 1.15 (i), and the same thing is true for the singular part $\mu^s = \mu - \mu^a$. In particular, this yields the measurability of the sets $\{\mu_r \ll \nu_r\} = \{\mu_r^s = 0\}$ and $\{\mu_r \perp \nu_r\} = \{\mu_r^a = 0\}$. The uniqueness of f is clear from the corresponding result in the classical case. □

In particular, we obtain the following measurable version of the classical result for measures.

Corollary 1.29 *(universal decomposition and density) For any localized Borel space S, there exists a measurable function $h\geq 0$ on $S\times \mathcal{M}_S^2$, such that*

$$\mu = h(\cdot,\mu,\nu)\cdot\nu + \mu_s \equiv \mu_a + \mu_s, \quad \mu,\nu\in\mathcal{M}_S,$$

where $\mu_a \ll \nu$ and $\mu_s \perp \nu$. Here the measures μ_a and μ_s and relations $\mu\ll\nu$ and $\mu\perp\nu$ depend measurably on μ and ν, and $h(\cdot,\mu,\nu)$ is unique a.e. ν, for any fixed μ and ν.

Proof: By Lemma 1.14 (vi), we may consider the identity map $\mu \mapsto \mu$ as a kernel from $\mathcal{M}_S$ to S. More generally, we may regard the maps $(\mu, \nu) \mapsto \mu$ and $(\mu, \nu) \mapsto \nu$ as kernels from $\mathcal{M}_S^2$ to S. It remains to apply Theorem 1.28 to the latter. □

Our next concern is to outline a general theory of differentiation of measures, needed for the construction of invariant disintegrations and densities, as well as for the approximation of Palm measures by ordinary conditional distributions. For suitable measures $\mu \ll \nu$ with $\mu = f \cdot \nu$ on a space S, we need conditions ensuring the convergence $\mu B/\nu B \to f(s)$, $s \in S$ a.e. ν, as $B \to \{s\}$ along a suitable family of sets $B \in \mathcal{S}$. Then for every $s \in S$ we consider a non-empty class $\mathcal{I}_s$ of sets $B \in \mathcal{S}$ containing s, assumed to be directed under downward inclusion, in the sense that for any $B_1, B_2 \in \mathcal{I}_s$, there exists a set $B \in \mathcal{I}_s$ with $B \subset B_1 \cap B_2$. The $\mathcal{I}_s$ are often generated by a universal class $\mathcal{I}$, such that $\mathcal{I}_s = \{B \in \mathcal{I};\, s \in B\}$ for all $s \in S$.

Given a σ-finite measure ν on S, we say that the classes $\mathcal{I}_s$ (or the universal class $\mathcal{I}$) form a *differentiation basis* for ν, if $\nu B < \infty$ for all $B \in \mathcal{I}_s$, and the convergence $\mu B/\nu B \to f(s)$ holds a.e. ν as $B \downarrow \{s\}$ along $\mathcal{I}_s$, whenever $\mu = f \cdot \nu$ with $f \in \mathcal{S}_+$. Here the limit is defined in the sense of nets indexed by $\mathcal{I}_s$. Thus, a set function $\varphi(B)$ is said to converge to $\varphi(s)$, written as $\varphi(B) \to \varphi(s)$, if for every $\varepsilon > 0$ there exists an $I \in \mathcal{I}_s$, such that $|\varphi(B) - \varphi(s)| < \varepsilon$ for all $B \in \mathcal{I}_s$ with $B \subset I$. Equivalently, we may define the upper and lower limits at s by

$$\overline{\varphi}(s) = \limsup_{B \downarrow \{s\}} \varphi(B) = \inf_{I \in \mathcal{I}_s} \sup_{B \in \mathcal{I}_s \cap I} \varphi(B),$$
$$\underline{\varphi}(s) = \liminf_{B \downarrow \{s\}} \varphi(B) = \sup_{I \in \mathcal{I}_s} \inf_{B \in \mathcal{I}_s \cap I} \varphi(B).$$

Then convergence holds at s whenever $\overline{\varphi}(s) = \underline{\varphi}(s)$, in which case the limit is given by

$$\varphi(s) = \lim_{B \downarrow \{s\}} \varphi(B) = \overline{\varphi}(s) = \underline{\varphi}(s).$$

We may distinguish three classical types of differentiation bases. The simplest case is that of a dissection system $\mathcal{I} = (I_{nj})$ on a Borel space S, which forms a differentiation basis for any locally finite measure on S, by the proof for Theorem 1.28 above. In this context, the entire dissection system $\mathcal{I}$, or the individual partitions (I_{nj}) for fixed n, are often referred to as *nets*. Unfortunately, this terminology conflicts with the topological notion of nets, employed in the definition of limits $\varphi(s)$ above.

When S is a Polish space, we say that the classes $\mathcal{I}_s$ form a *topological* differentiation basis for ν, if for any open neighborhood G of a point $s \in \operatorname{supp} \nu$, there exists a subset $B \subset G$ in $\mathcal{I}_s$ with $\nu B > 0$. Classical examples include the families of cubes or balls in R^d, which form differentiation bases for Lebesgue measure λ^d, as in Corollary A4.5 below.

A more sophisticated case is that of a *lifting* basis. Given a complete, σ-finite measure space $(S, \mathcal{S}, \nu)$, we say that two sets $B, B' \in \mathcal{S}$ are *ν-equivalent* and write $B \sim B'$ if $\nu(B\Delta B') = 0$. By a famous theorem of von Neumann, there exists a *lifting* map $B \mapsto \tilde{B} \sim B$ on $\mathcal{S}$, preserving the basic set operations and such that $B \sim B'$ implies $\tilde{B} = \tilde{B}'$. In other words, the lifting selects one element from each equivalence class of sets $B \in \mathcal{S}$, such that $B = \bigcup_n B_n$ implies $\tilde{B} = \bigcup_n \tilde{B}_n$ and $C = A \setminus B$ implies $\tilde{C} = \tilde{A} \setminus \tilde{B}$. Here the class $\mathcal{I} = \{\tilde{B};\, B \in \mathcal{S}\}$ of lifted sets is known to form a differentiation basis for ν. (The differentiation property typically fails for the original σ-field $\mathcal{S}$.)

In the Polish setting, we may also consider certain functional differentiation bases, consisting of indexed families $\mathcal{P}_s$ of functions $p_n^s \in \hat{\mathcal{S}}_+$ with locally bounded supports satisfying $\operatorname{supp} p_n^s \to \{s\}$, in the sense that $\operatorname{supp} p_n^s$ is contained, for large enough n, in any neighborhood G of s. The entire collection $\mathcal{P} = \bigcup_s \mathcal{P}_s$ is called a *functional differentiation basis* for $\nu \in \mathcal{M}_S$, if $\nu p_n^s < \infty$ for all n and s, and for any measure $\mu = f \cdot \nu$ in $\mathcal{M}_S$, we have $\mu p_n^s / \nu p_n^s \to f(s)$, $s \in S$ a.e. ν. For directed index sets, the convergence is again understood in the sense of nets. It is suggestive to write the stated approximation as $p_n^s \to \delta_s$, where δ_s denotes the Dirac function at s. The previous differentiation bases correspond to the special case where each p_n^s is proportional to the indicator function of a set in $\mathcal{I}_s$.

We refer to a differentiation basis of any of the mentioned types as a *standard differentiation basis* for ν. We need the following simple domination principle.

Lemma 1.30 *(domination) For any σ-finite measure μ on a Borel space S and a standard differentiation basis $\mathcal{I}$ or $\mathcal{P}$ for μ, there exists a dissection system (I_{nj}) in S, such that for $s \in I_{nj}$ a.e. μ, we may choose a $B \in \mathcal{I}_s$ or $p \in \mathcal{P}_s$ with $\mu(B \setminus I_{nj}) = 0$ or $\mu(p; I_{nj}^c) = 0$.*

A collection (I_{nj}) with the stated property will be referred to as a *dominating dissection system.*

Proof: The claim is trivial when $\mathcal{I}$ itself is a dissection basis. When $\mathcal{I}$ is a lifting basis, it holds for any dissection system (I_{nj}). Indeed, if $s \in I_{nj}$, we may choose B to be the associated lifted set $\tilde{I}_{nj}$. Since $\tilde{I}_{nj} \sim I_{nj}$, we have $\mu(\tilde{I}_{nj} \setminus I_{nj}) = 0$, and $s \in \tilde{I}_{nj}$ for s outside a fixed μ-null set.

Next let $\mathcal{I}$ be a topological differentiation basis on a Polish space S. Given a complete metrization of the topology, we may choose a topological base $B_1, B_2, \ldots$ consisting of balls with $\mu \partial B_n = 0$. For every n, let $I_{n1}, I_{n2}, \ldots$ be the partition of S generated by $B_1, \ldots B_n$, so that each I_{nj} is a finite intersection of sets B_k or B_k^c. Then $\mu \partial I_{nj} = 0$, and so for $s \in I_{nj}$ a.e. μ we have $s \in I_{nj}^\circ$, and we may choose $B \in \mathcal{I}_s$ with $\mu(B \setminus I_{nj}) = 0$. The case of functional differentiation bases is similar. □

Given a random element ξ with $\mathcal{L}(\xi) = \mu$ in a Borel space S, we may introduce, for any $I \in \hat{\mathcal{S}}$ or $p \in \hat{\mathcal{S}}_+$, the measures

$$\hat{\mu}_I = \frac{1_I \cdot \mu}{\mu I} = \mathcal{L}(\xi \mid \xi \in I),$$

$$\hat{\mu}_p = \frac{p \cdot \mu}{\mu p} = \frac{E\{p(\xi);\, \xi \in \cdot\}}{Ep(\xi)},$$

as long as the denominators are finite and positive. When the denominator vanishes, we may set $0/0 = 0$. For general measures $\mu \in \mathcal{M}_S$, we can still define $\hat{\mu}_I$ and $\hat{\mu}_p$ by the equalities on the left. In terms of this notation, the differentiation property of μ becomes $\hat{\mu}_I f \to f(s)$ or $\hat{\mu}_p f \to f(s)$ as $I \to \{s\}$ or $p \to \delta_s$, respectively.

We show that any two measures $\mu, \nu \in \mathcal{M}_S$ are asymptotically proportional near almost every point in the support of $\mu \wedge \nu$. More precisely, we have the following approximation property.

Lemma 1.31 *(local comparison) Let μ and ν be locally finite measures on a Borel space S, and fix a standard differentiation basis $\mathcal{I}$ for $\mu \vee \nu$. Then as $I \to \{s\}$ along $\mathcal{I}$, we have*

$$\|\hat{\mu}_I - \hat{\nu}_I\| \to 0, \quad s \in S \ \text{ a.e. } \mu \wedge \nu.$$

A similar statement holds for any functional differentiation basis $\mathcal{P}$.

Proof: Writing $p = 1_I$ for $I \in \mathcal{I}$, we may express the proofs for both $\mathcal{I}$ and $\mathcal{P}$ in a functional notation. We may further assume that $\mu \le \nu$, since the general result will then follow by application to the pairs $(\mu, \mu \vee \nu)$ and $(\nu, \mu \vee \nu)$. Finally, we may clearly take μ and ν to be bounded.

Now choose a measurable density $f\colon S \to [0,1]$ with $\mu = f \cdot \nu$, and fix a dominating dissection system (I_{nj}) in S. By a monotone-class argument, we may choose some simple functions $f_n\colon S \to [0,1]$ over (I_{nj}), $n \in \mathsf{N}$, such that $\nu|f - f_n| \to 0$. The assertion is trivial when μ is replaced by the measure $\mu_n = f_n \cdot \nu$ for fixed n. To estimate the difference, we may write

$$\begin{aligned}
\|\hat{\mu}_p - (\widehat{\mu_n})_p\| &\le \frac{\|p \cdot \mu - p \cdot \mu_n\|}{\mu p} + \|p \cdot \mu_n\| \left| \frac{1}{\mu p} - \frac{1}{\mu_n p} \right| \\
&\le 2\, \frac{\|p \cdot \mu - p \cdot \mu_n\|}{\mu p} = 2\, \frac{\hat{\nu}_p |f - f_n|}{\hat{\nu}_p f} \\
&\to 2\, \frac{|f(s) - f_n(s)|}{f(s)} = 2 \left| 1 - \frac{f_n(s)}{f(s)} \right|,
\end{aligned}$$

where the convergence holds for $s \in S$ a.e. μ as $p \to \delta_s$ along $\mathcal{P}_s$, by the differentiation property of $\mathcal{P}$. Since $f_n \to f > 0$ a.e. μ along a sub-sequence $N' \subset \mathsf{N}$, we get for $s \in S$ a.e. μ

$$\|\hat{\mu}_p - \hat{\nu}_p\| \le \|\hat{\mu}_p - (\widehat{\mu_n})_p\| + \|(\widehat{\mu_n})_p - \hat{\nu}_p\| \to 0,$$

as we let $p \to \delta_s$ along $\mathcal{P}$ and then $n \to \infty$ along N'. $\square$

To motivate the next result, let ξ and η be arbitrary random elements in S and T, respectively, and write their joint distribution as a composition

$\nu \otimes \mu$, where $\nu = \mathcal{L}(\xi)$ and $\mu_s = \mathcal{L}(\eta \,|\, \xi)_s$. For every $s \in S$, choose a random element $\eta_s \perp\!\!\!\perp \xi$ in T with $\mathcal{L}(\eta_s) = \mu_s$. We show that $(\xi, \eta) \overset{d}{\approx} (\xi, \eta_s)$ in total variation, near almost every point $s \in S$ in the support of ν. More precisely, we have the following asymptotic decoupling property:

Lemma 1.32 *(local decoupling) Fix any locally finite measures ν and ρ on some Borel spaces S and T, a probability kernel $\mu\colon S \to T$ with $\nu \otimes \mu \ll \nu \otimes \rho$, and a standard differentiation basis $\mathcal{I}$ for ν. Then as $I \downarrow \{s\}$ along $\mathcal{I}$,*

$$\| \hat{\nu}_I \otimes \mu - \hat{\nu}_I \otimes \mu_s \| \to 0, \quad s \in S \;\; a.e. \;\; \nu.$$

Though we are using the same symbol $\otimes$, the first term is a general probability measure on $I \times T$, written as the composition of a probability measure $\hat{\nu}_B$ on S with a probability kernel $\mu\colon S \to T$, whereas the second term is a product measure on $I \times T$, formed by the measures $\hat{\nu}_I$ on S and μ_s on T.

Proof: We may take ν and ρ to be bounded. Choose $f \in (\mathcal{S} \otimes \mathcal{T})_+$ with $\nu \otimes \mu = f \cdot (\nu \otimes \rho)$, and note that $\mu_s = f(s, \cdot) \cdot \rho$ a.e. ν, by the uniqueness in Theorem 1.23. We may then write the assertion as

$$\int \hat{\nu}_I(dr) \int \rho(dt)\, |f(r,t) - f(s,t)| \to 0, \quad s \in S \;\; \text{a.e. } \nu. \tag{8}$$

Now fix a dominating dissection system (I_{nj}) in S, as in Lemma 1.30. By a monotone-class argument, we may choose some functions $f_n \in (\mathcal{S} \times \mathcal{T})_+$ with $(\nu \otimes \rho)|f - f_n| \to 0$, such that each $f_n(s,t)$ is simple over $I_{n1}, I_{n2}, \ldots$, as a function of s for fixed t and n. Here (8) holds trivially with f_n in place of f. Writing

$$g_n(s) = \int \rho(dt)\, |f(s,t) - f_n(s,t)|, \quad s \in S, \; n \in \mathsf{N},$$

we see that the left-hand side of (8) is ultimately bounded by

$$\int \hat{\nu}_B(dr)\, |g_n(r) + g_n(s)| \to 2g_n(s), \quad s \in S \;\; \text{a.e. } \nu,$$

where the convergence holds as $I \downarrow \{s\}$ for fixed n, by the differentiation property of $\mathcal{I}$. Since also $\nu g_n \to 0$ by the choice of f_n, we have $g_n \to 0$ a.e. ν as $n \to \infty$ along a sub-sequence $N' \subset \mathsf{N}$, and (8) follows as we let $I \to \{s\}$ along $\mathcal{I}$, and then $n \to \infty$ along N'. □

The following result is useful to extend the differentiation property to a broader class of differentiation bases.

Theorem 1.33 *(extension principle) Fix a locally finite measure μ on a Borel space S, a standard differentiation basis $\mathcal{P}$ for μ, and a constant $c > 0$. For any $p \in \mathcal{P}$, let $\mathcal{C}_p$ be the class of functions $q \in \mathcal{S}_+$ with $q \le p$ and $\mu q \ge c \mu p$. Then the class $\mathcal{C} = (\mathcal{C}_p)$ is a differentiation basis for any locally finite measure $\nu \ll \mu$.*

Proof: Suppose that $\nu \ll \mu$, and consider some functions $h, f \in \mathcal{S}_+$ with $\nu = h \cdot \mu$ and $f \cdot \nu \in \mathcal{M}_S$. The differentiation property for μ yields

$$\begin{aligned} \hat{\nu}_p f &= \frac{\nu p f}{\nu p} = \frac{\mu h p f}{\mu h p} = \frac{\hat{\mu}_p h f}{\hat{\mu}_p h} \\ &\to \frac{h(s) f(s)}{h(s)} = f(s), \end{aligned}$$

valid for $s \in S$ a.e. ν, as $p \to \delta_s$ along $\mathcal{P}$. It is then enough to take $\nu = \mu$.

Now let $f \in \mathcal{S}_+$ with $f \cdot \mu \in \mathcal{M}_S$. Applying Lemma 1.31 to the measures μ and $\nu = f \cdot \mu$, we get a.e. μ on $\{f > 0\}$, as $q \le p \to \delta_s$ along $\mathcal{P}$,

$$\left| \frac{\mu q}{\mu p} - \frac{\mu q f}{\mu p f} \right| \le \|\hat{\mu}_p - \hat{\nu}_p\| \to 0.$$

Since $\mu q / \mu p \ge c > 0$, we obtain

$$\frac{\mu q}{\mu p} \sim \frac{\mu q f}{\mu p f} \quad \text{a.e. } \nu.$$

Hence, the differentiation property of $\mathcal{P}$ yields for $s \in S$ a.e. ν

$$\hat{\mu}_q f = \frac{\mu q f}{\mu q} \sim \frac{\mu p f}{\mu p} = \hat{\mu}_p f \to f(s).$$

Replacing f by $f + 1$, we get for $s \in S$ a.e. μ

$$\begin{aligned} \hat{\mu}_q f &= \hat{\mu}_q (f+1) - 1 \\ &\to f(s) + 1 - 1 = f(s), \end{aligned}$$

as required. □

We conclude with a technical comparison of densities, needed in Chapter 12. Given a measurable function f on a metric space S, we define

$$\hat{f}(x) = \inf_{G \ni x} \sup_{y \in G} f(y), \quad x \in S,$$

where the infimum extends over all neighborhoods G of x.

Lemma 1.34 *(density comparison) Let λ and $\mu \le \nu$ be σ-finite measures on a metric space S such that $\mu = p \cdot \lambda$ and $\nu = q \cdot \lambda$, where p is lower semi-continuous. Then $p \le \hat{q}$ on* $\operatorname{supp} \lambda$.

Proof: Since $\mu \le \nu$, we have $(q - p) \cdot \lambda \ge 0$, and so $p \le q$ a.e. λ. Fixing any $x \in \operatorname{supp} \lambda$, we may choose some $x_n \to x$ such that $p(x_n) \le q(x_n)$ for all n. Using the semi-continuity of p, we get for any neighborhood G of x

$$\begin{aligned} p(x) &\le \liminf_{n \to \infty} p(x_n) \\ &\le \liminf_{n \to \infty} q(x_n) \le \sup_{y \in G} q(y), \end{aligned}$$

and it remains to take the infimum over all G. □

Further results on differentiation of measures appear in Sections A4 and A5, with applications in Sections 7.5 and 7.6.

Chapter 2

Distributions and Local Structure

Given a localized Borel space $(S, \mathcal{S}, \hat{\mathcal{S}})$, we define a *random measure* on S as a locally finite kernel ξ from Ω to S, where $(\Omega, \mathcal{A}, P)$ denotes the underlying probability space. Thus, ξ is a function $\Omega \times \mathcal{S} \to [0, \infty]$, such that $\xi(\omega, B)$ is $\mathcal{A}$-measurable in $\omega \in \Omega$ for fixed B and a locally finite measure in $B \in \mathcal{S}$ for fixed ω. Equivalently, we may regard ξ as a random element in the space $\mathcal{M}_S$ of all locally finite measures μ on S, endowed with the σ-field generated by all projection maps $\pi_B : \mu \mapsto \mu B$ with $B \in \mathcal{S}$, or, equivalently, by all integration maps $\pi_f \colon \mu \mapsto \mu f = \int f d\mu$ with measurable $f \geq 0$.

By Theorem 1.6, any random measure ξ has an *atomic decomposition*

$$\xi = \alpha + \sum_{k \leq \kappa} \beta_k \, \delta_{\sigma_k} \ \text{ a.s.},$$

in terms of a diffuse random measure α, a $\overline{\mathsf{Z}}_+$-valued random variable κ, and some a.s. distinct random elements σ_k in S and random variables $\beta_k > 0$. By a *point process* we mean a random measure of the special form $\xi = \sum_k \beta_k \, \delta_{\sigma_k}$, where the β_k are integer-valued. Equivalently, it may be defined as an integer-valued random measure, or as a random element in the subspace $\mathcal{N}_S$ of integer-valued measures in $\mathcal{M}_S$.

A point process ξ is said to be *simple*, if all its weights β_k equal 1. In that case, ξ is the counting measure on its support $\Xi = \{\sigma_k\}$, which is a locally finite random set in S. For any point process ξ, the simple point process on its support is denoted by ξ^*. By a *marked point process* on S we mean a simple point process ξ on a product space $S \times T$, such that $\xi(\{s\} \times T) \leq 1$ for all $s \in S$. Note that the projection $\bar{\xi} = \xi(\cdot \times T)$ is not required to be locally finite.

Regarding a random measure ξ as a random process on $\hat{\mathcal{S}}$, we see that its distribution $\mathcal{L}(\xi)$ is uniquely determined by all finite-dimensional distributions $\mathcal{L}(\xi B_1, \ldots, \xi B_n)$ with $B_1, \ldots, B_n \in \hat{\mathcal{S}}$, $n \in \mathsf{N}$, hence also by the distributions of all integrals $\xi f = \int f d\xi$ with measurable integrands $f \geq 0$. Less obviously, we prove in Theorem 2.2 (ii) that the distribution of a simple point process ξ on S is determined by the set of *avoidance* probabilities $P\{\xi B = 0\}$ with $B \in \hat{\mathcal{S}}$.

The *intensity (measure)* $E\xi$ of a random measure ξ is given by $(E\xi)B = E(\xi B)$ or $(E\xi)f = E(\xi f)$. We may further introduce the higher order moment measures $E\xi^n$, where $\xi^n = \xi^{\otimes n}$ denotes the n-fold product measure on

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_2

S^n. For point processes ξ, we also consider the *factorial moment measures* $E\xi^{(n)}$, where the underlying factorial product measures $\xi^{(n)}$ are defined as in Lemma 1.12. When ξ is simple, we note that $E\xi^{(n)}$ equals the restriction of $E\xi^n$ to the non-diagonal part $S^{(n)}$ of S^n, where all coordinates are distinct. In general, Theorem 1.13 (ii) shows that the ordinary and factorial moment measures are related by $E\xi^n = \sum_\pi E\xi^{(\pi)}$, where π ranges over all partitions of $\{1, \dots, n\}$, and $\xi^{(\pi)}$ denotes the projection of $\xi^{(|\pi|)}$ onto the diagonal space $\hat{D}_\pi$.

With every random measure ξ on S, we may associate a bounded *supporting measure* ν, such that $\xi f = 0$ a.s. iff $\nu f = 0$. It follows easily that ξ has at most countably many fixed atoms. Writing D for the diagonal in S^2, we further show in Lemma 2.7 that a point process ξ on S is a.s. simple iff $E\xi^{(2)} D = 0$, whereas a random measure ξ on S is a.s. diffuse iff $E\xi^2 D = 0$.

Similar information is obtainable through the use of dissections. Assuming for simplicity that S is bounded with a dissection system (I_{nj}), we show in Corollary 2.9 that a point process ξ on S is a.s. simple whenever

$$\sum\nolimits_j P\{\xi I_{nj} > 1\} \to 0,$$

whereas a random measure ξ on S is a.s. diffuse when

$$\sum\nolimits_j E(\xi I_{nj} \wedge 1) \to 0.$$

In Theorem 2.13 we also show that, if a random measure ξ on S satisfies

$$\lim_{n\to\infty} \sum\nolimits_j \|\xi I_{nj}\|_p < \infty$$

for some constant $p > 1$, then $\xi \ll E\xi$ a.s., with a product-measurable density X satisfying $E\xi\|X\|_p < \infty$.

More generally, given a random measure ξ and a bounded, product-measurable process X on S, we may define a random measure $X \cdot \xi$ by $(X \cdot \xi)f = \xi(fX)$. In Theorem 2.11, we show that if $\xi \ll E\xi$ a.s. (but not in general!), the distribution of $X \cdot \xi$ is uniquely determined by the joint distribution $\mathcal{L}(\xi, X)$. In Theorem 2.12, we examine the implications of the more general condition $\xi \ll \eta$ a.s., for any random measures ξ and η.

As already noted, the distribution of a random measure is uniquely determined by the set of finite-dimensional distributions. This suggests that we look for conditions on the latter ensuring the existence of an associated random measure. In Theorem 2.15, we prove that any random set function ξ on $\hat{\mathcal{S}}$ satisfying the finite-additivity and continuity conditions

$$\begin{aligned} &\xi(A \cup B) = \xi A + \xi B \ \text{ a.s.}, \quad A, B \in \hat{\mathcal{S}} \ \text{ disjoint}, \\ &\xi A_n \xrightarrow{P} 0 \ \text{ a.s.}, \quad A_n \downarrow \emptyset \ \text{ along } \hat{\mathcal{S}}, \end{aligned}$$

admits a random measure version on S. This is remarkable, since the number of conditions of the stated form is typically uncountable. As a special case, we obtain the familiar existence theorem for regular conditional distributions.

We also noted earlier that the distribution of a simple point process ξ on S is determined by the set of avoidance probabilities $P\{\xi B = 0\}$ with $B \in \hat{\mathcal{S}}$. Here the corresponding finite-dimensional requirements are the *finite maxitivity* and continuity conditions

$$\eta(A \cup B) = \eta A \vee \eta B \ \text{ a.s.}, \quad A, B \in \hat{\mathcal{S}},$$
$$\eta A_n \xrightarrow{P} 0 \ \text{ a.s.}, \quad A_n \downarrow \emptyset \ \text{ along } \hat{\mathcal{S}}.$$

However, those conditions alone are not sufficient to guarantee the existence of an associated point process, and an additional tightness condition is needed to ensure local finiteness, as shown in Theorem 2.18. Without the demand of local finiteness, the theory becomes much more sophisticated, as evidenced by the characterization of random sets given in Theorem 2.22.

2.1 Uniqueness, Intensities, and Regularity

We begin with some simple extension principles.

Lemma 2.1 *(extension) Let ξ and η be random measures on S, and fix a generating semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}$. Then*

(i) *$\xi \le \eta$ a.s. iff $\xi I \le \eta I$ a.s. for all $I \in \mathcal{I}$,*

(ii) *ξ is a.s. non-random iff ξI is a.s. non-random for all $I \in \mathcal{I}$,*

(iii) *ξ is a point process iff $\xi I \in \mathsf{Z}_+$ a.s. for all $I \in \mathcal{I}$.*

Proof: Let $\mathcal{U}$ denote the ring of finite unions of $\mathcal{I}$-sets. By the semi-ring property, the latter sets can be taken to be disjoint, and so by finite additivity the stated conditions remain valid with $\mathcal{I}$ replaced by $\mathcal{U}$. Now let $\mathcal{D}$ denote the class of sets $B \in \hat{\mathcal{S}}$, such that the stated condition holds with I replaced by B. By dominated convergence we see that $\mathcal{D}$ is closed under bounded, monotone limits. Since $\mathcal{U}$ is again generating, Lemma 1.2 (ii) yields $\mathcal{D} \supset \hat{\sigma}(\mathcal{U}) = \hat{\mathcal{S}}$, which means that the stated conditions hold with $\mathcal{I}$ replaced by $\hat{\mathcal{S}}$. In particular, the condition in (ii) implies $E\xi \in \mathcal{M}_S$, and so (ii) follows from (i), applied to the random measures ξ and $E\xi$.

Now Lemma 1.3 ensures the existence of a dissection system $\mathcal{I}' \subset \hat{\mathcal{S}}$, and we note that each condition remains true for all $I \in \mathcal{I}'$, outside a fixed null set. In (i) and (iii) we may then assume that ξ and η are non-random. Since $\mathcal{I}'$ is again a generating semi-ring, the two conditions extend as before to arbitrary $B \in \hat{\mathcal{S}}$, which proves (i) and (iii). □

We turn to some basic uniqueness criteria for random measures. Further results will be obtained under special conditions in Theorem 3.8. Given a subclass $\mathcal{I} \subset \mathcal{S}$, we write $\hat{\mathcal{I}}_+$ for the class of simple, $\mathcal{I}$-measurable functions $f \ge 0$ on S.

Theorem 2.2 *(uniqueness) Fix any generating ring* $\mathcal{U} \subset \hat{\mathcal{S}}$ *and semi-ring* $\mathcal{I} \subset \hat{\mathcal{S}}$. *Then*[1]

(i) *for any random measures* ξ *and* η *on* S, *we have* $\xi \stackrel{d}{=} \eta$ *iff*

$$\xi f \stackrel{d}{=} \eta f, \quad f \in \hat{\mathcal{I}}_+,$$

(ii) *for any simple point processes* ξ *and* η *on* S, *we have* $\xi \stackrel{d}{=} \eta$ *iff*

$$P\{\xi U = 0\} = P\{\eta U = 0\}, \quad U \in \mathcal{U}.$$

When S *is Polish, it is equivalent in* (i) *to take* $f \in \hat{C}_S$.

Proof: (i) By the Cramér–Wold theorem, the stated condition implies

$$(\xi I_1, \dots, \xi I_n) \stackrel{d}{=} (\eta I_1, \dots, \eta I_n), \quad I_1, \dots, I_n \in \mathcal{I}, \; n \in \mathsf{N}.$$

By a monotone-class argument in $\mathcal{M}_S$, we then have $\xi \stackrel{d}{=} \eta$ on the σ-field $\Sigma_{\mathcal{I}}$ in $\mathcal{M}_S$ generated by all projections $\pi_I : \mu \to \mu I$, $I \in \mathcal{I}$. By a monotone-class argument in S, π_B remains $\Sigma_{\mathcal{I}}$-measurable for every $B \in \hat{\mathcal{S}}$. Hence, $\Sigma_{\mathcal{I}}$ contains the σ-field generated by the latter projections, and so $\Sigma_{\mathcal{I}}$ agrees with the total σ-field on $\mathcal{M}_S$, which shows that $\xi \stackrel{d}{=} \eta$. The statement for Polish spaces follows by a simple approximation.

(ii) Let $\mathcal{C}$ be the class of sets $\{\mu;\, \mu U = 0\}$ in $\mathcal{N}_S$ with $U \in \mathcal{U}$, and note that $\mathcal{C}$ is a π-system, since for any $B, C \in \mathcal{U}$,

$$\{\mu B = 0\} \cap \{\mu C = 0\} = \{\mu(B \cup C) = 0\} \in \mathcal{C},$$

by the ring property of $\mathcal{U}$. Furthermore, the sets $M \in \mathcal{B}_{\mathcal{N}_S}$ with $P\{\xi \in M\} = P\{\eta \in M\}$ form a λ-system $\mathcal{D}$, and $\mathcal{D} \supset \mathcal{C}$ by hypothesis. Hence, Lemma 1.2 (i) yields $\mathcal{D} \supset \sigma(\mathcal{C})$, which shows that $\xi \stackrel{d}{=} \eta$ on $\sigma(\mathcal{C})$.

Next let $\mathcal{S}'$ denote the class of sets $B \in \hat{\mathcal{S}}$ with $\{\mu;\, \mu B = 0\} \in \sigma(\mathcal{C})$, and note that $\mathcal{S}'$ contains the ring $\mathcal{U}$. Furthermore, $\mathcal{S}'$ is closed under bounded, monotone limits, since $B_n \uparrow B$ in $\hat{\mathcal{S}}$ implies $\{\mu B_n = 0\} \downarrow \{\mu B = 0\}$, whereas $B_n \downarrow B$ in $\hat{\mathcal{S}}$ implies $\{\mu B_n = 0\} \uparrow \{\mu B = 0\}$ in $\mathcal{N}_S$ (but not in $\mathcal{M}_S$). Since $\mathcal{U}$ is generating, Lemma 1.2 (ii) yields $\mathcal{S}' \supset \hat{\sigma}(\mathcal{U}) = \hat{\mathcal{S}}$, which gives $\xi \stackrel{d}{=} \eta$ on $\sigma(\mathcal{C}')$, where $\mathcal{C}'$ consists of all sets $\{\mu;\, \mu B = 0\}$ in $\mathcal{N}_S$ with $B \in \hat{\mathcal{S}}$. Since S is Borel, the ring $\hat{\mathcal{S}}$ contains a dissection system (I_{nj}) by Lemma 1.3, and we get

$$\lim_{n \to \infty} \sum\nolimits_j \{\mu(B \cap I_{nj}) \wedge 1\} = \mu^* B, \quad B \in \hat{\mathcal{S}}, \; \mu \in \mathcal{N}_S.$$

Since $\mathcal{N}_S$ is generated by the projection maps π_B with $B \in \hat{\mathcal{S}}$, we conclude that the mapping $\mu \mapsto \mu^*$ is $\sigma(\mathcal{C}')$-measurable, which implies $\xi^* \stackrel{d}{=} \eta^*$. $\Box$

This leads easily to a uniqueness criterion in terms of Laplace transforms[2].

[1] Here and below, we are using the compact integral notation $\mu f = \int f \, d\mu$.

[2] We omit the customary square brackets for expected values, as in $E[X]$, when there is no risk for confusion.

Corollary 2.3 *(Laplace criterion) For any generating semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}$ and random measures ξ and η on S, we have $\xi \overset{d}{=} \eta$ iff $Ee^{-\xi f} = Ee^{-\eta f}$ for all $f \in \hat{\mathcal{I}}_+$ with $f \le 1$. When S is Polish, it is equivalent to take $f \in \hat{C}_S$.*

Proof: For fixed $f \in \hat{\mathcal{I}}_+$ with $f \le 1$, the stated condition yields $Ee^{-t\,\xi f} = Ee^{-t\,\eta f}$ for all $t \in [0,1]$. Since both sides are analytic on $(0,\infty)$, the relation extends to all $t \ge 0$. The condition then implies (i), by the uniqueness theorem for multivariate Laplace transforms (FMP 5.3).

To prove the last assertion, we see as before that the condition $Ee^{-\xi f} = Ee^{-\eta f}$ for all $f \in \hat{C}_S$ implies

$$(\xi f_1, \dots, \xi f_n) \overset{d}{=} (\eta f_1, \dots, \eta f_n), \quad f_1, \dots, f_n \in \hat{C}_S, \ n \in \mathbb{N}.$$

Arguing as in the previous proof, we conclude that $\xi \overset{d}{=} \eta$. □

Given a random measure ξ on S, we define the associated *intensity (measure)* $E\xi$ by $(E\xi)B = E(\xi B)$ for all $B \in \mathcal{S}$. The definition is justified by the following statement. Say that a measure μ is *s-finite* if it is a countable sum of bounded measures. Note that every σ-finite measure is also s-finite, but not the other way around. The basic properties of σ-finite measures, including Fubini's theorem, remain valid in the s-finite case[3].

Lemma 2.4 *(intensity measure) For any random measure ξ on S, there exists an s-finite measure $E\xi$ on S, such that for every measurable function $f \ge 0$ on S,*

(i) $E(\xi f) = (E\xi)f$,
(ii) $E(f \cdot \xi) = f \cdot E\xi$.

Proof: (i) Define the set function $E\xi$ on $\mathcal{S}$ by $(E\xi)B = E(\xi B)$, and note that $E\xi$ is a measure by linearity and monotone convergence. For any simple measurable function $f = \sum_k c_k 1_{B_k}$, we get by the linearity of ξ and P

$$\begin{aligned}(E\xi)f &= (E\xi)\sum\nolimits_k c_k 1_{B_k} \\ &= \sum\nolimits_k c_k (E\xi)B_k = \sum\nolimits_k c_k E(\xi B_k) \\ &= E\Big(\xi \sum\nolimits_k c_k 1_{B_k}\Big) = E(\xi f).\end{aligned}$$

Finally, extend to general f by monotone convergence.

To see that $E\xi$ is s-finite, choose a partition $B_1, B_2, \ldots \in \hat{\mathcal{S}}$ of S, and define

$$\xi_{n,k} = (1_{B_n}\xi) 1\{\xi B_n \in (k-1,k]\}, \quad n,k \in \mathbb{N}.$$

Then $\xi = \sum_{n,k} \xi_{n,k}$, and so by (i) and monotone convergence

$$E\xi = E\sum_{n,k} \xi_{n,k} = \sum_{n,k} E\xi_{n,k},$$

[3]Recall that $\mu f = \int f\, d\mu$, whereas $f \cdot \mu$ denotes the measure $\nu \ll \mu$ with μ-density f.

whereas

$$\begin{aligned}\|E\xi_{n,k}\| &\le E(\xi B_n;\, \xi B_n \le k)\\ &\le k < \infty, \qquad n,k \in \mathsf{N}.\end{aligned}$$

(ii) Part (i) does not require ξ to be locally finite, and so it applies equally to $f \cdot \xi$. Using (i), we get for any measurable function $g \ge 0$ on S

$$\begin{aligned}\{E(f\cdot\xi)\}g &= E\{(f\cdot\xi)g\}\\ &= E\{\xi(fg)\} = (E\xi)(fg)\\ &= (f\cdot E\xi)g.\end{aligned}$$

□

For any random measure ξ on S, we define a *supporting measure* of ξ as a measure ν on S with $\xi f = 0$ a.s. iff νf. We show that such a measure ν exists and can be chosen to be bounded.

Corollary 2.5 *(supporting measure) For any random measure ξ on S, there exists a bounded measure ν on S, such that for every measurable function $f \ge 0$ on S,*

$$\xi f = 0 \;\; a.s. \quad \Leftrightarrow \quad \nu f = 0.$$

Proof: By Lemma 2.4 we have $E\xi = \sum_n \nu_n$ for some measures ν_n with $0 < \|\nu_n\| < \infty$. Defining[4]

$$\nu = \sum\nolimits_n 2^{-n}\nu_n/\|\nu_n\|,$$

we note that $\|\nu\| < 1$, and for measurable $f \ge 0$ we have

$$\begin{aligned}\xi f = 0 \;\text{ a.s.} \quad &\Leftrightarrow \quad E\xi f = 0\\ &\Leftrightarrow \quad \nu_n f \equiv 0 \quad \Leftrightarrow \quad \nu f = 0.\end{aligned}$$

□

For any random measure ξ on S, we say that $s \in S$ is a *fixed atom* of ξ if $P\{\xi\{s\} > 0\} > 0$. We show that this holds for at most countably many s.

Corollary 2.6 *(fixed atoms) A random measure ξ on S has at most countably many fixed atoms.*

Proof: Let ν be a bounded supporting measure of ξ, and note that for any $s \in S$

$$\begin{aligned}P\{\xi\{s\} > 0\} > 0 \quad &\Leftrightarrow \quad E\xi\{s\} > 0\\ &\Leftrightarrow \quad \nu\{s\} > 0.\end{aligned}$$

Since ν has at most countably many atoms, the assertion follows. □

For any random measure ξ on S, the product measure $\xi^n = \xi^{\otimes n}$ is a random measure on S^n. The associated intensity $E\xi^n$ is called the *n-th order moment measure* of ξ. When ξ is a point process, we may also define the *factorial processes* $\xi^{(n)}$ as in Lemma 1.12, and consider the *factorial moment measures* $E\xi^{(n)}$. Already the second order moment measures $E\xi^2$ and $E\xi^{(2)}$ reveal some important information about ξ:

[4]Recall that $\|\mu\| = \mu S$ denotes the total mass of the measure μ.

Lemma 2.7 *(simplicity and diffuseness) Let D denote the main diagonal in S^2. Then*

(i) *a point process ξ on S is a.s. simple iff $E\xi^{(2)}D = 0$,*

(ii) *a random measure ξ on S is a.s. diffuse iff $E\xi^2 D = 0$.*

Proof: (i) Writing $\xi = \sum_i \delta_{\sigma_i}$, we get by Lemma 1.12 (i)

$$\xi^{(2)}D = \sum_{i \neq j} \delta_{\sigma_i, \sigma_j} D = \sum_{i \neq j} 1\{\sigma_i = \sigma_j\},$$

and so

$$\begin{aligned} E\xi^{(2)}D = 0 \quad &\Leftrightarrow \quad E\sum_{i \neq j} 1\{\sigma_i = \sigma_j\} = 0 \\ &\Leftrightarrow \quad \sum_{i \neq j} 1\{\sigma_i = \sigma_j\} = 0 \ \text{ a.s.} \\ &\Leftrightarrow \quad \sigma_i \neq \sigma_j \ \text{ for all } \ i \neq j \ \text{ a.s.} \\ &\Leftrightarrow \quad \xi \ \text{ simple a.s.} \end{aligned}$$

(ii) By Fubini's theorem,

$$\xi^2 D = \int \xi(ds)\, \xi\{s\} = \sum\nolimits_s (\xi\{s\})^2,$$

and so

$$\begin{aligned} E\xi^2 D = 0 \quad &\Leftrightarrow \quad E\sum\nolimits_s (\xi\{s\})^2 = 0 \\ &\Leftrightarrow \quad \sum\nolimits_s (\xi\{s\})^2 = 0 \ \text{ a.s.} \\ &\Leftrightarrow \quad \xi\{s\} \equiv 0 \ \text{ a.s.} \\ &\Leftrightarrow \quad \xi \ \text{ diffuse a.s.} \end{aligned}$$

□

For any $\mu \in \mathcal{M}_S$ and $a > 0$, we define the measures μ_a^* and μ_a' by

$$\begin{aligned} \mu_a^* B &= \sum_{s \in B} 1\{\mu\{s\} \geq a\}, \\ \mu_a' B &= \mu B - \sum_{s \in B} \mu\{s\}\, 1\{\mu\{s\} \geq a\}. \end{aligned}$$

Theorem 2.8 *(dissection limits) Let ξ be a random measure on S. Fix a dissection system (I_{nj}) in $\hat{\mathcal{S}}$, let $B \in \hat{\mathcal{S}}$ be arbitrary, and put $B_{nj} = I_{nj} \cap B$. Then as $n \to \infty$,*

(i) $\sum_j 1\{\xi B_{nj} \geq a\} \to \xi_a^* B, \quad a > 0,$

(ii) $\liminf_{n \to \infty} \sum_j P\{\xi B_{nj} \geq a\} \geq E\xi_a^* B, \quad a > 0.$

(iii) *When $E\xi_a' B < \infty$,*

$$\sum\nolimits_j P\{a \leq \xi B_{nj} < b\} \to E(\xi_a^* - \xi_b^*)B, \quad 0 < a < b \leq \infty.$$

(iv) *For point processes* ξ,

$$\textstyle\sum_j P\{\xi B_{nj} > 0\} \to E\xi^* B.$$

Proof: (i) Fix any $\omega \in \Omega$. By the dissecting property of $\mathcal{I}$, all atoms of ξ_a^* in B will ultimately lie in different sets I_{nj}, and so for large enough $n \in \mathsf{N}$,

$$\textstyle\sum_j 1\{\xi B_{nj} \ge a\} \ge \xi_a^* B. \tag{1}$$

Now define $U_n = \bigcup_j \{B_{nj};\, \xi_a' B_{nj} \ge a\}$, and assume that $U_n \downarrow C$. If strict inequality holds in (1) for infinitely many n, then $\xi_a' U_n \ge a$ for all n, and so by dominated convergence $\xi_a' C \ge a$, which implies $C \ne \emptyset$. Fixing an $s \in C$ and letting I_n denote the set I_{nj} containing s, we have $I_n \downarrow \{s\}$ by the dissecting property of $\mathcal{I}$. Since $\xi_a' I_n \ge a$ for all n, we get $\xi_a'\{s\} \ge a$ by dominated convergence. This contradicts the definition of ξ_a', and so equality holds in (1) for large enough n, proving the desired convergence.

(ii) Use (i) and Fatou's lemma.

(iii) Writing $\xi^*_{[a,b)} = \xi_a^* - \xi_b^*$, we get from (i)

$$\lim_{n\to\infty} \textstyle\sum_j 1\{a \le \xi B_{nj} < b\} = \xi^*_{[a,b)} B. \tag{2}$$

Then (iii) follows by Fatou's lemma when $E\xi^*_{[a,b)}B = \infty$. If instead $E\xi^*_{[a,b)}B < \infty$, then (iii) follows from (2) by dominated convergence, since

$$\begin{aligned}
\textstyle\sum_j 1\{a \le \xi B_{nj} < b\} &\le \textstyle\sum_j 1\{\xi^*_{[a,b)} B_{nj} \ge 1\} \vee 1\{\xi_a' B_{nj} \ge a\} \\
&\le \textstyle\sum_j 1\{\xi^*_{[a,b)} B_{nj} \ge 1\} + \sum_j 1\{\xi_a' B_{nj} \ge a\} \\
&\le \textstyle\sum_j \xi^*_{[a,b)} B_{nj} + \sum_j a^{-1} \xi_a' B_{nj} \\
&= \xi^*_{[a,b)} B + a^{-1} \xi_a' B,
\end{aligned}$$

which is integrable and independent of n.

(iv) Use (ii) with $a = 1$. □

The last result yields some simple criteria for a point process to be simple or a random measure to be diffuse.

Corollary 2.9 *(simplicity and diffuseness) Fix any dissection system* (I_{nj}) *in* $\hat{\mathcal{S}}$, *and write* $B_{nj} = I_{nj} \cap B$ *for* $B \in \hat{\mathcal{S}}$. *Then*

(i) *a point process* ξ *is a.s. simple whenever*

$$\textstyle\sum_j P\{\xi B_{nj} > 1\} \to 0, \quad B \in \hat{\mathcal{S}},$$

(ii) *a random measure* ξ *is a.s. diffuse whenever*

$$\textstyle\sum_j E(\xi B_{nj} \wedge 1) \to 0, \quad B \in \hat{\mathcal{S}}.$$

The converse statements are also true when $E\xi$ is locally finite.

Proof: (i) By Theorem 2.8 (ii) with $a = 2$, the stated condition implies $E\xi_2^* = 0$, which means that ξ is a.s. simple. If $E\xi$ is locally finite, then taking $a = 2$ and $b = \infty$ in Theorem 2.8 (iii) gives $\sum_j P\{\xi B_{nj} > 1\} \to E\xi_2^* B$, and so the stated condition is equivalent to a.s. simplicity.

(ii) For any $B \in \hat{\mathcal{S}}$ and $\varepsilon \in (0,1)$, the stated condition gives

$$\sum_j P\{\xi B_{nj} > \varepsilon\} \le \varepsilon^{-1} \sum_j E(\xi B_{nj} \wedge 1) \to 0,$$

and so by Theorem 2.8 we have $E\xi_\varepsilon^* = 0$ for every $\varepsilon > 0$, which shows that ξ is a.s. diffuse. Conversely, if ξ is a.s. diffuse, we get $\sum_j (\xi B_{nj} \wedge 1) \to 0$ a.s. by Theorem 2.8 (i). When $E\xi$ is locally finite, the stated condition then follows by dominated convergence. □

2.2 Absolute Continuity and Conditioning

Throughout this section, we take $\xi, \eta, \ldots$ to be random measures on a localized Borel space S. First we consider the existence of densities $X = d\xi/d\eta$ as measurable processes on S, and of integral processes $X \cdot \eta$ and conditional intensities $E(\xi \,|\mathcal{F})$ as random measures on S.

Lemma 2.10 *(densities and conditioning)*

(i) *For any random measures ξ and η on S, the a.s. relation $\xi \ll \eta$ is measurable and implies the existence of a measurable process $X \ge 0$ on S such that $\xi = X \cdot \eta$ a.s.*

(ii) *For any random measure η and measurable process $X \ge 0$ on S, the integral process $\xi = X \cdot \eta$ exists and is measurable on $\hat{\mathcal{S}}$. When ξ is locally finite, it is again a random measure on S.*

(iii) *For any random measure ξ on S and σ-field $\mathcal{F}$, the process $E(\xi B \,|\mathcal{F})$ on $\hat{\mathcal{S}}$ has a measure-valued version η. When the latter is locally finite, it is an $\mathcal{F}$-measurable random measure on S.*

Proof: (i) This is immediate from Theorem 1.28.

(ii) For any random measure η, measurable process $X \ge 0$, and measurable function $f \ge 0$ on S, the integral $\xi f = \eta(fX)$ in again measurable by Lemma 1.15 (i), hence a random variable. If $\xi = X \cdot \eta$ is locally finite, it is then a random measure on S.

(iii) Regarding ξ as a random element in the space $\mathcal{M}_S$, which is again Borel by Theorem 1.5, we see from FMP 6.3 that $\rho = \mathcal{L}(\xi \,|\mathcal{F})$ has a regular version, as an $\mathcal{F}$-measurable random measure on $\mathcal{M}_S$. For any $B \in \hat{\mathcal{S}}$, we may then choose versions $\eta B = \int \mu B \, \rho(d\mu) = \rho\pi_B$, which are again $\mathcal{F}$-measurable by Lemma 1.14 and combine into a measure-valued version of η.

The last assertion is now obvious. □

It may be surprising that the joint distribution of X and η in part (ii) may not determine the distribution of $\xi = X \cdot \eta$. For a simple counter-example, let $\xi = \delta_\tau$, where τ is $U(0,1)$, and define $X_t = 1\{t = \tau\}$ and $Y_t = 0$ for all $t \in [0,1]$. Then X and Y are measurable with $X_t = Y_t = 0$ a.s. for every t, so that $(\xi, X) \stackrel{d}{=} (\xi, Y)$, and yet $\xi X = 1 \neq 0 = \xi Y$. We show that the desired uniqueness does hold under an additional assumption.

Theorem 2.11 *(uniqueness) Consider some random measures ξ, η and measurable processes $X, Y \geq 0$ on a Borel space S, such that $(\xi, X) \stackrel{d}{=} (\eta, Y)$ and $\xi \ll E\xi$ a.s. Then*

$$(\xi,\, X,\, X \cdot \xi) \stackrel{d}{=} (\eta,\, Y,\, Y \cdot \eta).$$

Proof: Invoking Corollary 2.5, we may choose a bounded measure $\mu \sim E\xi$. By monotone convergence, we may further take X and Y to be bounded. Since $\xi \ll \mu$, and hence $X \cdot \xi \ll \mu$ a.s, Lemma 1.29 yields some measurable functions $g, h \geq 0$ on $S \times \mathcal{M}_S$, such that a.s.

$$\xi = g(\cdot, \xi) \cdot \mu, \qquad X \cdot \xi = h(\cdot, X \cdot \xi) \cdot \mu. \tag{3}$$

Since the density is a.e. unique, we have

$$X_s\, g(s, \xi) = h(s, X \cdot \xi), \quad s \in S \ \text{ a.e. } \mu, \ \text{ a.s. } P,$$

and so by Fubini's theorem,

$$X_s\, g(s, \xi) = h(s, X \cdot \xi) \ \text{ a.s. } P, \quad s \in S \ \text{ a.e. } \mu. \tag{4}$$

Since $(\xi, X) \stackrel{d}{=} (\eta, Y)$, and $\mathcal{M}_S$ is Borel by Theorem 1.5, Lemma 1.16 (ii) yields a random measure ζ on S with

$$(\xi, X, X \cdot \xi) \stackrel{d}{=} (\eta, Y, \zeta). \tag{5}$$

By (4) and (5) we get

$$Y_s g(s, \eta) = h(s, \zeta) \ \text{ a.s. } P, \quad s \in S \ \text{ a.e. } \mu,$$

and so by Fubini's theorem,

$$Y_s g(s, \eta) = h(s, \zeta), \quad s \in S \ \text{ a.e. } \mu, \ \text{ a.s. } P.$$

Since also $\eta = g(\cdot, \eta) \cdot \mu$ by (3) and (5), we obtain

$$\begin{aligned} Y \cdot \eta &= Y g(\cdot, \eta) \cdot \mu \\ &= h(\cdot, \zeta) \cdot \mu \quad \text{a.s.} \end{aligned} \tag{6}$$

Combining (3), (5), and (6) gives

$$\begin{aligned}(\xi, X, X\cdot\xi) &= (\xi, X, h(\cdot, X\cdot\xi)\cdot\mu)\\ &\overset{d}{=} (\eta, Y, h(\cdot,\zeta)\cdot\mu)\\ &= (\eta, Y, Y\cdot\eta).\end{aligned}$$

□

The previous condition $\xi \ll E\xi$ can be replaced by $\xi \ll \mu$, for any fixed, s-finite measure μ on S. More generally, we show how any a.s. relation $\xi \ll \eta$ extends to a broader class of random measures ξ or η.

Theorem 2.12 *(absolute continuity) Let ξ and η be random measures on a Borel space S such that $\xi \ll \eta$ a.s. Then for any σ-field $\mathcal{F} \supset \sigma(\eta)$,*

$$\xi \ll E(\xi\,|\mathcal{F}) \ll \eta \quad a.s.$$

In particular, the relation $\xi \ll \eta$ implies $\xi \ll E(\xi|\,\eta) \ll \eta$ a.s. Note that the conditional intensity $E(\xi\,|\mathcal{F})$ in Lemma 2.10 (iii) is a.s. s-finite, though it may fail to be σ-finite. Thus, the asserted relations are measurable by Corollary 1.29.

Proof: We may clearly assume that $E\xi$ is σ-finite. By Lemma 1.29 there exists a measurable function $h \geq 0$ on $S \times \mathcal{M}_S$, such that $\xi = h(\cdot,\xi,\eta)\cdot\eta = X\cdot\eta$ a.s., where $X_s = h(s,\xi,\eta)$. Since $\mathcal{M}_S$ is Borel, we may choose an $\mathcal{F}$-measurable random probability measure ν on $\mathcal{M}_S$ such that $\nu = \mathcal{L}(\xi|\mathcal{F})$ a.s. Using the disintegration theorem and Fubini's theorem, we get a.s. for any $f \in \mathcal{S}_+$

$$\begin{aligned}E(\,\xi f\,|\,\mathcal{F}) &= E\{\eta(fX)\,|\,\mathcal{F}\}\\ &= E\Big\{\int \eta(ds)\, f(s)\, h(s,\xi,\eta)\,\Big|\,\mathcal{F}\Big\}\\ &= \int \nu(dm)\int \eta(ds)\, f(s)\, h(s,m,\eta)\\ &= \int \eta(ds)\, f(s)\int \nu(dm)\, h(s,m,\eta)\\ &= \int \eta(ds)\, f(s)\, E\{h(s,\xi,\eta)\,|\,\mathcal{F}\}\\ &= \eta\,\{f\,E(X|\,\mathcal{F})\} = \{E(X|\,\mathcal{F})\cdot\eta\,\}f.\end{aligned}$$

Since f was arbitrary, we obtain $E(\xi\,|\,\mathcal{F}) = E(X|\,\mathcal{F})\cdot\eta$ a.s., which implies $E(\xi\,|\,\mathcal{F}) \ll \eta$ a.s.

The previous calculation shows that the remaining claim, $\xi \ll E(\xi\,|\,\mathcal{F})$ a.s., is equivalent to

$$h(\cdot,\xi,\eta)\cdot\eta \ll \int \nu(dm)\, h(\cdot,m,\eta)\cdot\eta \quad \text{a.s.}$$

Using the disintegration theorem again, we may write this as

$$E\,\nu\Big\{\mu \in \mathcal{M}_S;\; h(\cdot,\mu,\eta)\cdot\eta \ll \int \nu(dm)\, h(\cdot,m,\eta)\cdot\eta\Big\} = 1,$$

which reduces the assertion to $X \cdot \eta \ll EX \cdot \eta$ a.s., for any non-random measure η and measurable process $X \geq 0$ on S.

To prove this, we note that $EX_s = 0$ implies $X_s = 0$ a.s. Hence,

$$X_s \ll EX_s \text{ a.s. } P, \quad s \in S,$$

and so by Fubini's theorem,

$$X_s \ll EX_s, \quad s \in S \text{ a.e. } \eta, \quad \text{a.s. } P.$$

Now fix any $\omega \in \Omega$ outside the exceptional P-null set, and let $B \in \mathcal{S}$ be arbitrary with $\eta(EX; B) = 0$. Splitting B into the four subsets where

$$\begin{aligned} &X_s = EX_s = 0, \qquad && X_s > EX_s = 0, \\ &EX_s > X_s = 0, \qquad && X_s EX_s > 0, \end{aligned}$$

we may easily verify that even $\eta(X; B) = 0$. This shows that indeed $X \cdot \eta \ll EX \cdot \eta$ a.s. □

Given a random measure ξ on S, a dissection system (I_{nj}) in S, and a constant $p > 1$, we define the *L^p-intensity* $\|\xi\|_p$ of ξ as the set function

$$\|\xi\|_p B = \lim_{n\to\infty} \textstyle\sum_j \|\xi B_{nj}\|_p, \quad B \in \hat{\mathcal{S}},$$

where $B_{nj} = I_{nj} \cap B$. Note that the limit exists, since the sum is non-decreasing by Minkowski's inequality.

Theorem 2.13 *(L^p-intensity) Let ξ be a random measure on S with $E\xi \in \mathcal{M}_S$, and fix a dissection system $\mathcal{I} = (I_{nj})$ and a constant $p > 1$. Then $\|\xi\|_p$ is locally finite iff $\xi = X \cdot E\xi$ a.s. for some measurable process X such that $\|X\|_p \cdot E\xi$ is locally finite, in which case $\|\xi\|_p = \|X\|_p \cdot E\xi$.*

Proof: Putting $E\xi = \mu$, we first assume that $\xi = X \cdot \mu$. Letting $p^{-1} + q^{-1} = 1$, we get by Fubini's theorem and Hölder's inequality

$$\begin{aligned} \|\mu X\|_p^p &= E(\mu X)^p = E(\mu X)(\mu X)^{p-1} \\ &= E\mu\Big\{X(\mu X)^{p-1}\Big\} = \mu E\Big\{X(\mu X)^{p-1}\Big\} \\ &\leq \mu\Big\{\|X\|_p \, \|(\mu X)^{p-1}\|_q\Big\} \\ &= \mu\|X\|_p \, \|\mu X\|_p^{p/q}. \end{aligned}$$

If $0 < \|\mu X\|_p < \infty$, we may divide by the second factor on the right to get $\|\mu X\|_p \leq \mu\|X\|_p$. This extends by monotone convergence to infinite $\|\mu X\|_p$, and it is also trivially true when $\|\mu X\|_p = 0$. Replacing X by $1_B X$ for a $B \in \hat{\mathcal{S}}$ gives $\|\xi B\|_p \leq (\|X\|_p \cdot \mu)B$, and so $\|\xi\|_p B \leq (\|X\|_p \cdot \mu)B$.

To prove the reverse inequality, fix any $B \in \hat{\mathcal{S}}$, and define $X_n(s) = \xi B_n(s)/\mu B_n(s)$, where $B_n(s)$ denotes the set B_{nj} containing $s \in B$. For fixed $\omega \in \Omega$, the functions X_n form a martingale on (B, μ) with a.e. limit X, and so $X_n \to X$ a.e. $\mu \otimes P$ on $B \times \Omega$ by Fubini's theorem. Hence, by Fatou's lemma and the definition of $\|\xi\|_p$,

$$\begin{aligned}(\|X\|_p \cdot \mu)B &= \int_B \|X\|_p \, d\mu \le \int_B \liminf_{n\to\infty} \|X_n\|_p \, d\mu \\ &\le \liminf_{n\to\infty} \int_B \|X_n\|_p \, d\mu \\ &= \lim_{n\to\infty} \sum\nolimits_j \|\xi B_{nj}\|_p = \|\xi\|_p B.\end{aligned}$$

Combining with the previous inequality gives $\|\xi\|_p B = (\|X\|_p \cdot \mu)B$, and since B was arbitrary, we get $\|\xi\|_p = \|X\|_p \cdot \mu$. In particular, $\|\xi\|_p$ is locally finite when $\|X\|_p$ is locally μ-integrable.

Now assume instead that $\|\xi\|_p B < \infty$ for all $B \in \hat{\mathcal{S}}$. Fixing B and defining X_n and q as above, we get by Hölder's inequality

$$\begin{aligned}E\int_B (X_n)^{1+1/q} \, d\mu &= \sum\nolimits_j E\,\xi B_{nj}\,(\xi B_{nj}/\mu B_{nj})^{1/q} \\ &\le \sum\nolimits_j \|\xi B_{nj}\|_p \,(E\xi B_{nj}/\mu B_{nj})^{1/q} \\ &= \sum\nolimits_j \|\xi B_{nj}\|_p \le \|\xi\|_p B < \infty.\end{aligned}$$

Hence, by Doob's inequality,

$$\begin{aligned}E\int_B (\sup\nolimits_n X_n)^{1+1/q} \, d\mu \;&\lesssim\; E \sup_n \int_B (X_n)^{1+1/q} \, d\mu \\ &\le\; \|\xi\|_p B < \infty,\end{aligned}$$

and so the inner integral on the left is a.s. finite. In particular, (X_n) is uniformly integrable on B, and so $X_n \to X$ a.e. and in $L^1(\mu)$ on $B \times \Omega$ for some measurable process X, and we get a.s.

$$\begin{aligned}\int_B X \, d\mu &= \lim_{n\to\infty} \int_B X_n \, d\mu \\ &= \sum\nolimits_j \xi B_{nj} = \xi B.\end{aligned}$$

Hence, $\xi B = (X \cdot \mu)B$ a.s. for all $B \in \hat{\mathcal{S}}$, and so $\xi = X \cdot \mu$ a.s. by Lemma 2.1. As before, we get $\int_B \|X\|_p \, d\mu \le \|\xi\|_p B < \infty$ for all B, which shows that $\|X\|_p$ is locally μ-integrable. □

We conclude with a technical martingale property, needed in Chapter 12.

Lemma 2.14 *(martingale density) Let μ be a σ-finite measure on a metric space S, and fix any $x \in \operatorname{supp}\mu$. Consider a measure-valued martingale (ξ_t) on S with induced filtration $\mathcal{F}$, such that $\xi_t = M^t \cdot \mu$ a.s. for all $t \ge 0$, where M_s^t is $(\mathcal{F}_t \otimes \mathcal{S})$-measurable in s for each t and L^1-continuous in t for $s = x$. Then M_x^t is a martingale in t.*

Proof: Fix any times $s \le t$, and let $A \in \mathcal{F}_s$ and $B \in \hat{\mathcal{S}}$. Using Fubini's theorem and the martingale property of ξ, we get

$$\begin{aligned}\int_B E(M_y^s; A)\,\mu(dy) &= E\Big\{\int_B M_y^s\,\mu(dy); A\Big\}\\ &= E(\,\xi_s B; A) = E(\,\xi_t B; A)\\ &= E\Big\{\int_B M_y^t\,\mu(dy); A\Big\}\\ &= \int_B E(M_y^t; A)\,\mu(dy).\end{aligned}$$

Since B was arbitrary, we get $E(M_y^s; A) = E(M_y^t; A)$ for $y \in S$ a.e. μ, and since $x \in \operatorname{supp}\mu$, we may choose some $x_n \to x$ such that $E(M_{x_n}^s; A) = E(M_{x_n}^t; A)$ for all n. By the L^1-continuity on each side, the relation extends to $E(M_x^s; A) = E(M_x^t; A)$, and since A was arbitrary, we conclude that $M_x^t = E(M_x^t \,|\, \mathcal{F}_s)$ a.s. □

2.3 Additive and Maxitive Processes

Random measures are often obtained by regularization of suitable vector measures. Indeed, exceedingly weak conditions are required:

Theorem 2.15 *(additive processes and random measures, Harris) Given a process $\eta \ge 0$ on a generating ring $\mathcal{U} \subset \hat{\mathcal{S}}$, there exists a random measure ξ on S with $\xi U = \eta U$ a.s. for all $U \in \mathcal{U}$, iff*

(i) $\eta(A \cup B) = \eta A + \eta B$ *a.s.,* $A, B \in \mathcal{U}$ *disjoint,*

(ii) $\eta A_n \xrightarrow{P} 0$ *as* $A_n \downarrow \emptyset$ *along* $\mathcal{U}$.

In that case, ξ is a.s. unique.

Proof: Since the kernel property is preserved by increasing limits, and since S is covered by countably many sets in $\mathcal{U}$, we may assume that $S \in \mathcal{U}$ and $\eta S < \infty$ a.s. Since (i) and (ii) remain true for the process $(1 + \eta S)^{-1}\eta$ on $\mathcal{U}$, we may even assume that $\eta S \le 1$ a.s. Then define a function $\lambda \ge 0$ on $\mathcal{U}$ by $\lambda A = E\eta A$, and note that λ is countably additive with $\lambda\emptyset = 0$ by (i) and (ii). By Carathéodory's extension theorem (FMP 2.5) it extends to a measure on $\sigma(\mathcal{U}) = \mathcal{S}$.

Introduce the metric $\rho(A, B) = \lambda(A\Delta B)$ in $\mathcal{S}$, and note that $\mathcal{U}$ is dense in $(\mathcal{S}, \rho)$ by a monotone-class argument. By (i) we have for any $A, B \in \mathcal{U}$

$$\begin{aligned}E|\eta A - \eta B| &= E|\eta(A \setminus B) - \eta(B \setminus A)|\\ &\le E\eta(A \setminus B) + E\eta(B \setminus A)\\ &= \lambda(A\Delta B),\end{aligned}$$

and so η is uniformly continuous from $(\mathcal{U}, \rho)$ to $L^1(P)$, hence admitting a uniformly continuous extension to $\mathcal{S}$. Property (i) extends to $\mathcal{S}$ by a simple approximation, whereas (ii) holds on $\mathcal{S}$ by the countable additivity of λ and the continuity of η. We may then assume that $\mathcal{U} = \mathcal{S}$.

The Borel space S can be identified with a Borel set in R. Extending η to $\mathcal{B}$ by writing $\eta B = \eta(B \cap S)$, we may reduce to the case of $S = \mathsf{R}$. Putting $Y_r = \eta(-\infty, r]$, we see from (i) that $Y_r \le Y_s$ a.s. for all $r < s$ in Q, and from (ii) that a.s. $Y_r \to 0$ or $Y_r \to \eta S$, respectively, as $r \to \pm\infty$ along Q. Redefining Y on the common null set, we may assume that all those conditions hold identically.

Since the process $X_t = Y_{t+} = \inf\{Y_r;\, r \in \mathsf{Q},\, r > t\}$, $t \in \mathsf{R}$, is non-decreasing and right-continuous with $X_{-\infty} = 0$ and $X_\infty = \eta S$, there exists by FMP 2.14 a function $\xi : \Omega \to \hat{\mathcal{M}}_{\mathsf{R}}$ with $\xi(-\infty, t] = X_t$ for all $t \in \mathsf{R}$, and ξ is a random measure on R by Lemma 1.14. Approximating from the right along Q and using (i) and (ii), we get $\xi(-\infty, t] = \eta(-\infty, t]$ a.s. for every $t \in \mathsf{R}$, and so a monotone-class argument yields $\xi B = \eta B$ a.s. for every $B \in \mathcal{B}$. $\square$

The last theorem is easily translated into an existence criterion in terms of finite-dimensional distributions. Here *consistency* is defined in the sense of Kolmogorov (FMP 6.16).

Corollary 2.16 *(finite-dimensional distributions) Fix a generating ring $\mathcal{U} \subset \hat{\mathcal{S}}$ in S. Then the probability measures $\mu_{B_1,\dots,B_n}$ on R_+^n, $B_1, \dots, B_n \in \mathcal{U}$, $n \in \mathsf{N}$, are finite-dimensional distributions of some random measure ξ on S, iff they are consistent and satisfy*

(i) $\mu_{A,B,A\cup B}\{(x,y,z) : x + y = z\} = 1,\;\; A, B \in \mathcal{U}$ *disjoint,*

(ii) $\mu_A \xrightarrow{w} \delta_0$ *as* $A \downarrow \emptyset$ *along* $\mathcal{U}$.

The distribution of ξ is then unique.

Proof: Assume the stated conditions. Then Kolmogorov's existence theorem (FMP 6.16) yields a process $\eta \ge 0$ on $\mathcal{U}$ with finite-dimensional distributions $\mu_{B_1,\dots,B_n}$, and (i) and (ii) imply the corresponding conditions in Theorem 2.15. Hence, there exists a random measure ξ on S with $\xi U = \eta U$ a.s. for all $U \in \mathcal{U}$, and we note that ξ and η have the same finite-dimensional distributions. $\square$

For a simple application of the last theorem, we may regularize the conditional intensity $E(\xi \,|\, \mathcal{F})$ of a random measure ξ with respect to an arbitrary σ-field $\mathcal{F} \subset \mathcal{A}$.

Corollary 2.17 *(conditional intensity) For any random measure ξ on S and σ-field $\mathcal{F} \subset \mathcal{A}$, there exists an $\mathcal{F}$-measurable kernel $\eta = E(\xi \,|\, \mathcal{F})$ from Ω to S, such that $\eta B = E(\xi B \,|\, \mathcal{F})$ a.s. for all $B \in \hat{\mathcal{S}}$. If $E\xi$ is locally finite, then η is again a random measure on S.*

Proof: When $E\xi \in \mathcal{M}_S$, define $\zeta B = E(\xi B \mid \mathcal{F})$ a.s. for $B \in \hat{\mathcal{S}}$, and note that ζ satisfies (i) and (ii) in Theorem 2.15. Hence, there exists an $\mathcal{F}$-measurable random measure η on S with $\eta B = \zeta B$ a.s. for all $B \in \hat{\mathcal{S}}$. For general ξ, choose some random measures $\xi_n \uparrow \xi$ with $E\xi_n \in \mathcal{M}_S$, and construct some $\mathcal{F}$-measurable random measures η_n with $\eta_n B = E(\xi_n B \mid \mathcal{F})$ a.s. for all $B \in \hat{\mathcal{S}}$. Since $\eta_n B$ is a.s. non-decreasing for every $B \in \hat{\mathcal{S}}$, even η_n is a.s. non-decreasing by Lemma 2.1, and so $\eta_n \uparrow \eta$ a.s. for some $\mathcal{F}$-measurable kernel η. Finally, $\eta B = E(\xi B \mid \mathcal{F})$ a.s. for all $B \in \hat{\mathcal{S}}$ by monotone convergence. □

When $\xi = \delta_\sigma$ for some random element σ in S, the last result reduces to the existence of regular conditional distributions (FMP 6.3). We turn to a similar result for simple point processes, where the finite additivity in (i) is replaced by the *weak maxitivity* $\eta(A \cup B) = \eta(A) \vee \eta(B)$ a.s.

Theorem 2.18 *(maxitivity and simple point processes) Let η be a $\{0,1\}$-valued process on a generating ring $\mathcal{U} \subset \hat{\mathcal{S}}$. Then there exists a simple point process ξ on S with $\xi B \wedge 1 = \eta(B)$ a.s. for all $B \in \mathcal{U}$, iff*

(i) $\eta(A \cup B) = \eta(A) \vee \eta(B)$ *a.s.,* $A, B \in \mathcal{U}$,

(ii) $\eta(A_n) \xrightarrow{P} 0$ *as* $A_n \downarrow \emptyset$ *along* $\mathcal{U}$,

(iii) $\{\sum_{B\in\pi} \eta(B);\ \pi \in \mathcal{P}_A\}$ *is tight,* $A \in \mathcal{U}$,

where $\mathcal{P}_A$ denotes the class of finite partitions π of A into sets $B \in \mathcal{U}$. In that case, ξ is a.s. unique.

Our proof relies on the following lemma. Given a class Ξ of $\bar{\mathsf{R}}$-valued random variables, we say that a random variable α is an *a.s. upper/lower bound* of Ξ if $\alpha \geq \xi$ [respectively $\alpha \leq \xi$] a.s. for every $\xi \in \Xi$. An a.s. upper/lower bound α is called an *a.s. maximum/minimum* of Ξ if $\alpha \in \Xi$. Writing A for the set of a.s. upper/lower bounds of Ξ, we say that α is an *essential supremum/infimum* of ξ if it is an a.s. minimum/maximum of A.

Lemma 2.19 *(essential supremum and maximum) Any set Ξ of $\bar{\mathsf{R}}$-valued random variables has an a.s. unique essential supremum α. If Ξ is non-empty and closed under finite maxima and increasing limits, then α is also the a.s. unique essential maximum. Similar statements hold for the essential infimum and minimum.*

Proof: First assume that Ξ is non-empty and closed under finite maxima and increasing limits. By a simple transformation, we may assume that all random variables take values in $[0,1]$. Define $c = \sup\{E\xi;\ \xi \in \Xi\}$, and choose $\xi_1, \xi_2, \ldots \in \Xi$ with $E\xi_n \to c$. The assumptions on Ξ give $\alpha \equiv \sup_n \xi_n \in \Xi$. For any $\xi \in \Xi$, we have even $\alpha \vee \xi \in \Xi$. Noting that $E\alpha = E(\alpha \vee \xi) = c$, we get $E(\alpha \vee \xi - \alpha) = 0$, and so $\xi \leq \alpha$ a.s., which shows that α is an a.s. upper bound of Ξ. If α' is another a.s. upper bound, then $\alpha \leq \alpha'$ a.s., which shows that α is an essential supremum and maximum.

For general Ξ, let A be the set of a.s. upper bounds of Ξ, and note that A is non-empty and closed under finite minima and decreasing limits. Applying the previous case to $-A$, we see that A has an essential minimum, which is then an essential supremum of Ξ. The a.s. uniqueness is obvious, and the results for the essential infimum or mimimum follow, as we apply the previous results to $-\Xi$. □

Proof of Theorem 2.18: For any $A \in \mathcal{U}$, let Ξ_A be the class of sums $\Sigma_\pi = \sum_{B\in\pi} \eta(B)$ with $\pi \in \mathcal{P}_A$, and define $\zeta A = \operatorname{ess\,sup} \Xi_A$. Then also $\zeta A = \operatorname{ess\,sup} \bar{\Xi}_A$ a.s., where $\bar{\Xi}_A$ denotes the closure of Ξ_A under finite maxima and increasing limits. By (i) we have $\Sigma_\pi \vee \Sigma_{\pi'} \leq \Sigma_{\pi\vee\pi'}$, where $\pi \vee \pi'$ denotes the smallest common refinement of π and π', and so the tightness in (iii) extends to $\bar{\Xi}_A$. Since $\zeta A \in \bar{\Xi}_A$ a.s. by Lemma 2.19, we get $\zeta A < \infty$ a.s.

Since Σ_π is non-decreasing under refinements of π, we may restrict π in the definition of ζA to refinements of some fixed partition π'. For disjoint $A, B \in \mathcal{U}$, we get a.s.

$$\begin{aligned}\zeta(A \cup B) &= \operatorname{ess\,sup}\left\{\Sigma_\pi + \Sigma_{\pi'};\ \pi \in \mathcal{P}_A,\ \pi' \in \mathcal{P}_B\right\}\\ &= \operatorname{ess\,sup} \Xi_A + \operatorname{ess\,sup} \Xi_B\\ &= \zeta A + \zeta B.\end{aligned}$$

Since $\eta(A) = 0$ implies $\zeta A = 0$ a.s. by (i), condition (ii) gives $\zeta A_n \xrightarrow{P} 0$ as $A_n \downarrow \emptyset$ along $\mathcal{U}$. Thus, ζ satisfies the conditions in Theorem 2.15, and so there exists a random measure ξ on S with $\xi A = \zeta A$ a.s. for all $A \in \mathcal{U}$. Since ζ is Z_+-valued, ξ is a point process by Lemma 2.1 (iii), and (i) yields $\xi A \wedge 1 = \zeta A \wedge 1 = \eta(A)$ a.s. Writing ξ^* for the simple point process on the same support, we get $\Xi_A \leq \xi^* A$ for all $A \in \mathcal{U}$, and so $\xi A \leq \xi^* A$ a.s. on $\mathcal{U}$, which implies $\xi \leq \xi^*$ by Lemma 2.1 (i). Hence, ξ is a.s. simple. □

We may restate the last theorem as a criterion for a function h on $\mathcal{U}$ to be the *avoidance function* $P\{\xi U = 0\}$ of a simple point process ξ on S. Then define recursively the difference operators $\Delta_{B_1,\dots,B_n}$ by

$$\begin{aligned}\Delta_B h(A) &= h(A \cup B) - h(A),\\ \Delta_{B_1,\dots,B_n} h(A) &= \Delta_{B_1} \Delta_{B_2,\dots,B_n} h(A), \quad n \geq 2,\end{aligned} \tag{7}$$

where the last difference is with respect to A for fixed $B_2, \dots, B_n$. Note that the operators Δ_B commute, so that $\Delta_{B_1,\dots,B_n}$ is independent of the order of the sets B_k. For consistency, we may take $\Delta_{B_1,\dots,B_n} h(A) = h(A)$ when $n = 0$, so that (7) remains true for $n = 1$.

A function $h \geq 0$ on $\mathcal{U}$ is said to be *completely monotone* if

$$(-1)^n \Delta_{B_1,\dots,B_n} h(A) \geq 0, \quad A, B_1, \dots, B_n \in \mathcal{U},\ n \in \mathsf{N}.$$

We show how this property is related to maxitivity:

Lemma 2.20 *(maxitivity and complete monotonicity) Let f be a $[0,1]$-valued function on a class of sets $\mathcal{U}$, closed under finite unions. Then these conditions are equivalent:*

(i) *h is completely monotone,*

(ii) *$h(A) = P\{\eta(A) = 0\}$, $A \in \mathcal{U}$, for some weakly maxitive, $\{0,1\}$-valued process η on $\mathcal{U}$.*

The distribution of η is then unique.

Proof: First assume (ii). We claim that

$$(-1)^n \Delta_{B_1,\dots,B_n} h(A) = P\Big\{\eta(A) = 0,\, \eta(B_1)\cdots\eta(B_n) = 1\Big\}, \tag{8}$$

for every $n \in \mathsf{Z}_+$. This is clearly true for $n = 0$. Proceeding by induction, suppose that (8) holds for $n-1$. Using the induction hypothesis and the maxitivity of η, we get

$$\begin{aligned}
(-1)^n \Delta_{B_1,\dots,B_n} h(A) &= -\Delta_{B_1} P\Big\{\eta(A) = 0,\, \eta_{B_2}\cdots\eta_{B_n} = 1\Big\} \\
&= P\Big\{\eta(A) = 0,\, \eta_{A\cup B_1}\eta_{B_2}\cdots\eta_{B_n} = 1\Big\} \\
&= P\Big\{\eta(A) = 0,\, \eta_{B_1}\cdots\eta_{B_n} = 1\Big\},
\end{aligned}$$

as required. Condition (i) is obvious from (8).

Now assume (i). Then (8) suggests that we define

$$\begin{aligned}
&P\Big\{\eta(A_1) = \cdots = \eta(A_m) = 0,\, \eta(B_1)\cdots\eta(B_n) = 1\Big\} \\
&\qquad = (-1)^n \Delta_{B_1,\dots,B_n} h(A_1 \cup \cdots \cup A_m), \quad m, n \ge 0.
\end{aligned} \tag{9}$$

Since set union is commutative and the difference operators commute, the right-hand side is independent of the order of events. We also note that

$$\begin{aligned}
&(-1)^{n+1} \Delta_{B_1,\dots,B_n,C}\, h(A) + (-1)^n \Delta_{B_1,\dots,B_n} h(A \cup C) \\
&\qquad = (-1)^n \Delta_{B_1,\dots,B_n} h(A),
\end{aligned}$$

which proves that the measures in (9) are consistent. When $h(\emptyset) = 1$, Kolmogorov's existence theorem (FMP 6.16) yields a process η on $\mathcal{U}$ satisfying (9) for all $m, n \ge 0$. If instead $h(\emptyset) < 1$, we may take $P\{\eta \equiv 1\} = 1 - h(\emptyset)$, and use (9) to define the probabilities on the set $\{\eta \not\equiv 1\}$.

From (9) we get for any $A, B \in \mathcal{U}$

$$\begin{aligned}
P\Big\{\eta(A \cup B) = 0,\, \eta(B) = 1\Big\} &= -\Delta_B\, h(A \cup B) \\
&= h(A \cup B) - h(A \cup B) = 0,
\end{aligned}$$

and so a.s.

$$\begin{aligned}
\{\eta(A \cup B) = 0\} &\subset \{\eta(B) = 0\}, \\
\{\eta(B) = 1\} &\subset \{\eta(A \cup B) = 1\}.
\end{aligned}$$

Combining with the same relations with A and B interchanged, we see that η is weakly maxitive. By iteration, we get for any $A_1, \ldots, A_m \in \mathcal{U}$, $m \in \mathsf{N}$,

$$\Big\{\eta(A_1 \cup \cdots \cup A_m) = 0\Big\} = \Big\{\eta(A_1) = \cdots = \eta(A_m) = 0\Big\} \text{ a.s.},$$

and so (8) implies (9), which proves the asserted uniqueness. □

We may now translate Theorem 2.18 into a characterization of avoidance functions of simple point processes.

Corollary 2.21 *(avoidance function, Karbe, Kurtz) Let h be a $[0,1]$-valued function on a generating ring $\mathcal{U} \subset \hat{\mathcal{S}}$. Then $h(B) = P\{\xi B = 0\}$, $B \in \mathcal{U}$, for some simple point process ξ on S iff*

(i) *h is completely monotone,*

(ii) *$h(B) \to 1$ as $B \downarrow \emptyset$ along $\mathcal{U}$,*

(iii) $\displaystyle \lim_{n\to\infty} \sup_{\pi \in \mathcal{P}_A} \sum_{m \ge n} \sum_{B_1,\ldots,B_m \in \pi} (-1)^m \Delta_{B_1,\ldots,B_m} h\Big(A \setminus \bigcup\nolimits_i B_i\Big) = 0, \quad A \in \mathcal{U}.$

The distribution of ξ is then unique.

Proof: With h and η related as in Lemma 2.20, we have $\eta(B) \xrightarrow{P} 0$ iff $h(B) \to 1$. Furthermore, (iii) can be written as

$$\lim_{n\to\infty} \sup_{\pi \in \mathcal{P}_A} P\Big\{\sum\nolimits_{B\in\pi} \eta(B) \ge n\Big\} = 0, \quad A \in \mathcal{U},$$

which expresses the tightness of $\{\sum_{B\in\pi} \eta(B);\, \pi \in \mathcal{P}_A\}$ for every $A \in \mathcal{U}$. Hence, the assertion is equivalent to Theorem 2.18. □

We proceed to show how a weakly maxitive process can be regularized into the hitting process of a closed random set. This requires some topological conditions on S, here assumed to be lcscH. Say that η is *inner continuous* if $\eta(G_n) \xrightarrow{P} \eta(G)$ as $G_n \uparrow G$ in $\mathcal{G}$.

Theorem 2.22 *(random sets) Let S be lcscH with class $\mathcal{G}$ of open sets.*

(i) *(Choquet) A function $h\colon \mathcal{G} \to [0,1]$ satisfies $h(G) = P\{\varphi \cap G = \emptyset\}$, $G \in \mathcal{G}$, for some closed random set φ in S, iff it is completely monotone and inner continuous with $h(\emptyset) = 1$. Then $\mathcal{L}(\varphi)$ is unique.*

(ii) *A process $\eta\colon \mathcal{G} \times \Omega \to \{0,1\}$ satisfies $\eta(G) = 1\{\varphi \cap G \neq \emptyset\}$ a.s., $G \in \mathcal{G}$, for some closed random set φ in S, iff it is weakly maxitive and inner continuous with $\eta(\emptyset) = 0$ a.s. Then φ is a.s. unique.*

Our proof relies on a simple topological fact:

Lemma 2.23 *(separating basis) Let S be lcscH with class $\mathcal{G}$ of open sets and sub-class $\hat{\mathcal{G}}$ of relatively compact sets. Then there exist some $B_1, B_2, \ldots \in \hat{\mathcal{G}}$ with*

$$G = \bigcup\nolimits_n \{B_n;\, \bar{B}_n \subset G\}, \quad G \in \mathcal{G}.$$

Proof: Choose a metrization ρ of the topology, such that (S, ρ) becomes separable and the ρ-bounded sets are relatively compact. Fixing a dense set $s_1, s_2, \ldots \in S$, enumerate the open ρ-balls in S with centers s_k and rational radii as $B_1, B_2, \ldots$. Given $G \in \mathcal{G}$, we may choose $n_s \in \mathsf{N}$, $s \in G$, with

$$s \in B_{n_s} \subset \bar{B}_{n_s} \subset G, \quad s \in G.$$

Then $G = \bigcup_s B_{n_s} = \bigcup_s \bar{B}_{n_s}$, and the result follows. □

Proof of Theorem 2.22: Let φ be a closed random set in S, and define $\eta(G) = 1\{\varphi \cap G \neq \emptyset\}$ and $h = 1 - E\eta$. For any $A, B \in \mathcal{G}$, we get

$$\begin{aligned} \eta(A \cup B) &= 1\Big\{\varphi \cap (A \cup B) \neq \emptyset\Big\} \\ &= 1\Big\{(\varphi \cap A) \cup (\varphi \cap B) \neq \emptyset\Big\} \\ &= 1\{\varphi \cap A \neq \emptyset\} \vee 1\{\varphi \cap B \neq \emptyset\} \\ &= \eta(A) \vee \eta(B), \end{aligned}$$

and so η is maxitive, whereas h is completely monotone by Lemma 2.20. Since $\eta(\emptyset) = 1\{\varphi \cap \emptyset \neq \emptyset\} \equiv 0$, we also have $h(\emptyset) = 1$. Finally, $G_n \uparrow G$ implies $\varphi \cap G_n \uparrow \varphi \cap G$, and so $\eta(G_n) \uparrow \eta(G)$ and $h(G_n) \downarrow h(G)$. This proves the necessity in (i) and (ii). We turn to the sufficiency:

(i) Let h be completely monotone and inner continuous with $h(\emptyset) = 1$. Fix a separating base $B_1, B_2, \ldots \in \mathcal{G}$ as in Lemma 2.23, let $\mathcal{U}_n$ consist of all unions formed by the sets $B_1, \ldots, B_n$, $n \in \mathsf{N}$, and put $\mathcal{U}_\infty = \bigcup_n \mathcal{U}_n$. Then Lemma 2.20 yields a strictly maxitive process $\eta_n : \mathcal{U}_n \times \Omega \to \{0, 1\}$ with $P\{\eta_n(U) = 0\} = h(U)$ for all $U \in \mathcal{U}_n$. For any $s_k \in B_k$, $k \in \mathsf{N}$, we introduce the closed random sets

$$\varphi_n = \{s_k;\ k \leq n,\ \eta_n(B_k) = 1\}, \quad n \in \mathsf{N}.$$

Using the maxitivity of η_n, we get for any $U \in \mathcal{U}_n$

$$\begin{aligned} 1\{\varphi_n \cap U \neq \emptyset\} &= \max_{k \leq n,\, B_k \subset U} 1\{\varphi_n \cap B_k \neq \emptyset\} \\ &= \max_{k \leq n,\, B_k \subset U} \eta_n(B_k) = \eta_n(U), \end{aligned}$$

and so $P\{\varphi_n \cap U = \emptyset\} = h(U)$ for all $U \in \mathcal{U}_n$. Since $\mathcal{F}_S$ is compact in the Fell topology (FMP A2.5), the sequence (φ_n) is trivially tight, and Prohorov's theorem (FMP 16.3 or Lemma 4.4 below) yields $\varphi_n \xrightarrow{d} \varphi$ along a sub-sequence $N \subset \mathsf{N}$ for some closed random set φ.

Now define $h_n(B) = P\{\varphi_n \cap B \neq \emptyset\}$ and $\tilde{h}(B) = P\{\varphi \cap B \neq \emptyset\}$ for all $B \in \mathcal{G} \cup \mathcal{F}$. Since the Fell topology is generated by the sets $\{F;\ F \cap G \neq \emptyset\}$ with $G \in \mathcal{G}$ and $\{F;\ F \cap K = \emptyset\}$ with $K \in \mathcal{K}$, the Portmanteau theorem (FMP 4.25) gives

$$\begin{aligned} \tilde{h}(G) &\geq \limsup_{n \in N,\, n \to \infty} h_n(G) \\ &\geq \liminf_{n \in N,\, n \to \infty} h_n(\bar{G}) \geq \tilde{h}(\bar{G}), \quad G \in \hat{\mathcal{G}}. \end{aligned} \tag{10}$$

For any $G \in \mathcal{G}$, we may choose $U_1, U_2, \ldots \in \mathcal{U}_\infty$ with $U_n \uparrow G$ and $\bar{U}_n \subset G$. Then (10) yields

$$\tilde{h}(U_n) \geq h(U_n) \geq \tilde{h}(\bar{U}_n) \geq \tilde{h}(G), \quad n \in \mathsf{N}.$$

Since h and $\tilde{h}$ are both inner continuous on $\mathcal{G}$, we can let $n \to \infty$ to obtain $\tilde{h} = h$ on $\mathcal{G}$.

(ii) Let the process $\eta\colon \mathcal{G} \times \Omega \to \{0, 1\}$ be weakly maxitive and inner continuous with $\eta(\emptyset) = 0$. Then the function $h = 1 - E\eta$ is completely monotone by Lemma 2.20 and also inner continuous with $h(\emptyset) = 1$. By (i) there exists a closed random set ψ with $P\{\psi \cap G = \emptyset\} = h(G)$ for all $G \in \mathcal{G}$. Writing $\zeta(G) = 1\{\psi \cap G \neq \emptyset\}$, we get $E\zeta = 1 - h = E\eta$ on $\mathcal{G}$, and a monotone-class argument gives $\zeta \stackrel{d}{=} \eta$. Hence, Lemma 1.16 yields $(\psi, \zeta) \stackrel{d}{=} (\varphi, \eta)$ for some closed random set φ in S. In particular, $\eta(G) = 1\{\varphi \cap G \neq \emptyset\}$ a.s. for all $G \in \mathcal{G}$. $\square$

The proof shows that (ii) is an easy consequence of (i). The two statements are in fact equivalent:

Proof of (ii)$\Rightarrow$(i) *in Theorem 2.22:* Let $h : \mathcal{G} \to [0, 1]$ be completely monotone and inner continuous with $h(\emptyset) = 1$. Then Lemma 2.20 gives $h(G) = P\{\eta(G) = 0\}$, $G \in \mathcal{G}$, for some weakly maxitive process $\eta\colon \mathcal{G} \times \Omega \to \{0, 1\}$, which is again inner continuous with $\eta(\emptyset) = 0$ a.s. Hence, (ii) yields $\eta(G) = 1\{\varphi \cap G \neq \emptyset\}$ a.s. for some closed random set φ in S, and so $h(G) = P\{\varphi \cap G = 0\}$, $G \in \mathcal{G}$. $\square$

Chapter 3

Poisson and Related Processes

The simplest point process is just a unit mass δ_σ at a random point σ in S. A *binomial process* is a finite sum $\xi = \delta_{\sigma_1} + \cdots + \delta_{\sigma_n}$ of such unit masses, where $\sigma_1, \ldots, \sigma_n$ are i.i.d. random elements in S. The term is suggested by the fact that ξB is binomially distributed for any $B \in \mathcal{S}$. Replacing n by an integer-valued random variable κ independent of the i.i.d. sequence $\sigma_1, \sigma_2, \ldots$, we obtain a *mixed binomial process*, and if we specialize to a Poisson distributed random variable κ, we obtain a *Poisson process*, whose *increments* ξB are independent and Poisson distributed for any disjoint sets B. A general Poisson process can be constructed by patching together independent mixed binomial processes of this type.

Just as Brownian motion plays a fundamental role in martingale theory and stochastic calculus, the general Poisson process is fundamental within the theory of random measures, for about the same reasons. Indeed, it is a non-trivial fact that any stochastically continuous, marked point process with independent increments is Poisson. We will also see in Chapter 9 how every sufficiently regular point process on R_+ can be reduced to Poisson through a suitable random change of time.

Many basic properties of Poisson processes are well known, such as the preservation of the Poisson property under measurable mappings, random thinnings, randomizations, and displacements. Their study leads inevitably to the richer class of *Cox processes*[1], which also turn up in a variety of contexts throughout random measure theory. For example, Cox processes often appear as limiting processes in various ergodic theorems for particle systems, they may be used as in Chapter 9 to approximate quite general point processes, and the basic stationary line and flat processes will be seen in Chapter 11 to be Cox.

To define a Cox process, we note that the distribution of a Poisson process ξ is specified by its *intensity measure* $E\xi = \lambda$, which is a locally finite measure on S. Just as in the previous case of mixing, we may now randomize by replacing the fixed measure λ by a random measure η on S. More precisely, we may choose ξ to be conditionally Poisson given η, with intensity $E(\xi \mid \eta) = \eta$ a.s. To motivate the term "doubly stochastic," we may think of first choosing η, and then, with η regarded as fixed, taking ξ to be a Poisson

[1]formerly called *doubly stochastic Poisson* processes

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_3

process with intensity η.

The *Cox transformation*, leading from a random measure η to an associated Cox process ξ, has many useful applications. For an important example, recall from the previous chapter that the distribution of a simple point process ξ on S is determined by the set of avoidance probabilities $P\{\xi B = 0\}$. Applying this result to the Cox transform of a diffuse random measure ξ, we see that the distribution of such a ξ is determined by the class of one-dimensional distributions $\mathcal{L}(\xi B)$, for arbitrary $B \in \hat{\mathcal{S}}$. In fact, it is enough to know the expected values $Ee^{-c\xi B}$ for a fixed $c > 0$. In this way, many results for simple point processes have natural counterparts for diffuse random measures. The indicated device becomes especially important in the context of convergence in distribution.

The theory of Poisson processes yields an easy access to infinitely divisible random measures. Just as for random variables, we say that a random measure ξ is *infinitely divisible*, if for every $n \in \mathsf{N}$ there exist some i.i.d. random measures $\xi_1, \dots, \xi_n$ satisfying $\xi \stackrel{d}{=} \xi_1 + \cdots + \xi_n$, where $\stackrel{d}{=}$ denotes equality in distribution. The definition for point processes is similar. In analogy with the classical Lévy–Itô representation, a random measure ξ on S is infinitely divisible iff it has a *cluster representation* $\xi = \alpha + \int \mu\, \eta(d\mu)$, where α is a non-random measure on S and η is a Poisson process on the associated measure space $\mathcal{M}_S$. Thus, the distribution of ξ is uniquely determined by the pair (α, λ), where $\lambda = E\eta$ is known as the *Lévy measure* of ξ. For point processes ξ, the only difference is that $\alpha = 0$ and η is a Poisson process on the subspace $\mathcal{N}_S$ of integer-valued measures on S. If the space S is Borel, then so are $\mathcal{M}_S$ and $\mathcal{N}_S$, and the basic theory applies even to η. Cluster representations play a fundamental role in the context of branching and superprocesses, which constitute the subject of Chapter 13.

Random thinnings may be regarded as special random transformations, where the individual points of a point process ξ are moved independently according to some probability kernel on the underlying space S. Such motions describe the simplest particle systems, where the individual particles are subject to independent displacements. In the easiest case, we may just attach some independent, uniformly distributed random variables to the points of ξ, to create a new process $\tilde{\xi}$ on the product space $S \times [0, 1]$, referred to as a *uniform randomization* of ξ. Such randomizations are used to represent *exchangeable* random measures, which form a richer class than the measures with stationary independent increments.

Many elementary propositions show how the classes of Poisson and Cox processes are affected by various randomizations and transformations. Here our basic technical tool is the Laplace transform, which often enables us to give two-line proofs of statements that would otherwise take a page or two of awkward computations[2].

[2] We don't hesitate to use possibly unpopular tools, whenever they lead to shorter or more transparent proofs.

It is well known how the discontinuous component of a general Lévy process can be represented in terms of a suitably compensated Poisson integral. Similar representations hold for general processes with independent increments, except that the Poisson process is then replaced by a more general point process with independent increments. Processes of the latter type have many properties in common with the Poisson family, and in particular it is natural to study integration with respect to processes belonging to this richer class. Here the main integrals are the positive, symmetric, or compensated ones. In each case, we give criteria for existence, convergence in probability to 0, and tightness, along with some partial conditions for divergence to ∞.

A basic result is the characterization of a stationary Poisson process on R_+ as a renewal process with exponentially distributed spacing variables. In Theorem 3.34 we prove the related fact that the spacing variables of a stationary binomial process on $[0,1]$ are exchangeable. Similarly we show that a stationary, simple point process on R_+ is mixed Poisson iff its associated sequence of spacing variables is exchangeable. We further show in Theorem 3.35 that a simple point process on R_+ or $[0,1]$ has exchangeable increments iff it is a mixed Poisson or binomial process, whereas a diffuse random measure on the same space is exchangeable iff it is a.s. invariant. A much more subtle fact, already alluded to above, is the representation in Theorem 3.37 of a general random measure with exchangeable increments.

We may finally mention the remarkable characterizations in Theorem 3.12 of suitable mixed Poisson or binomial processes in terms of birth or death processes. In particular, we prove in part (iii) of the cited theorem that a simple point process ξ on R_+ is mixed Poisson with random rate $\rho\, e^t$, for some exponentially distributed random variable ρ with mean 1, iff the counting process $N_t = 1 + \xi[0,t]$ is a pure birth process in N with rates $c_n = n$. The latter process, known as a *Yule process*, will be seen in Chapter 13 to arise naturally in the context of branching and superprocesses.

3.1 Basic Processes and Uniqueness Criteria

Given a random element σ in S, we see from Lemma 1.15 (i) that $\xi = \delta_\sigma$ is a simple point process on S. More generally, if $\sigma_1, \sigma_2, \ldots$ are i.i.d. random elements in S with a common distribution μ, then for every $n \in \mathsf{N}$ we may form a point process $\xi = \sum_{k\le n} \delta_{\sigma_k}$ on S, called a *binomial process* based on n and μ. The term is suggested by the fact that ξB has a binomial distribution with parameters n and μB for every $B \in \mathcal{S}$. Replacing n by a Z_+-valued random variable $\kappa \perp\!\!\!\perp (\sigma_k)$, we obtain a *mixed binomial process* based on μ and κ.

A point process ξ on S is said to be *Poisson* with *intensity* $\mu \in \mathcal{M}_S$, if for any disjoint sets $B_1, \ldots, B_n \in \hat{\mathcal{S}}$, $n \in \mathsf{N}$, the random variables $\xi B_1, \ldots, \xi B_n$ are independent and Poisson distributed with means $\mu B_1, \ldots, \mu B_n$. Given a measure $\mu \in \mathcal{M}_S$ and a random variable $\alpha \ge 0$, we may also define a *mixed*

Poisson process ξ based μ and α, by requiring ξ to be conditionally Poisson with intensity $\alpha\mu$ given α. More generally, given a random measure η on S, we say that ξ is a *Cox process directed by* η, if it becomes Poisson with intensity η under conditioning on η.

Now fix any $\mu \in \mathcal{N}_S$, along with a probability kernel ν from S to a Borel space T. Assuming $\mu = \sum_{k\le\kappa}\delta_{s_k}$, we may choose some independent random elements τ_k with distributions ν_{s_k}, $k \le \kappa$. When the point process $\zeta = \sum_{k\le\kappa}\delta_{\tau_k}$ on T is locally finite, it is called a ν-*transform* of μ. Replacing μ by a point process ξ on S, we say that ζ is a ν-transform of ξ, provided it is locally finite, and the previous property holds conditionally given ξ. When $\nu_s = \delta_s \otimes \rho_s$ it is called a ρ-*randomization* of ξ, and it is called a *uniform randomization* of ξ when ν_s equals Lebesgue measure on $[0,1]$. Given a measurable function $p\colon S \to [0,1]$, we may then form a p-*thinning* ξ_p of ξ by setting $\xi_p f = \zeta f_p$, where $f_p(s,t) = f(s)1\{t \le p(s)\}$.

Computations involving the mentioned processes are often simplified by the use of Laplace functionals.

Lemma 3.1 *(Laplace functionals) For any $f \in \hat{\mathcal{S}}_+$, we have a.s.*

(i) *when ξ is a mixed binomial process based on μ and κ,*

$$E(e^{-\xi f}\,|\,\kappa) = (\mu e^{-f})^\kappa,$$

(ii) *when ξ is a Cox process directed by η,*

$$E(e^{-\xi f}\,|\,\eta) = \exp\Big\{-\eta(1-e^{-f})\Big\},$$

(iii) *when ξ is a ν-transform of η,*

$$E(e^{-\xi f}\,|\,\eta) = \exp\Big(\eta \log \nu e^{-f}\Big),$$

(iv) *when ξ is a p-thinning of η,*

$$E(e^{-\xi f}\,|\,\eta) = \exp\Big(\eta \log\Big\{1 - p\,(1-e^{-f})\Big\}\Big).$$

Here and below, we are using the compact integral notation explained in previous chapters. For example, in a more traditional notation, the formula in part (iii) becomes

$$E\Big(\exp\Big\{-\int f(s)\,\xi(ds)\Big\}\Big|\,\eta\Big) = \exp\Big(\int \log\Big\{\int e^{-f(t)}\,\nu_s(dt)\Big\}\,\eta(ds)\Big).$$

In the sequel, such transliterations are usually left to the reader[3].

[3]They will hardly be needed much, since the reader will soon get used to and appreciate our simplified and transparent notation.

Proof: (i) Let $\xi = \sum_{k\le n} \delta_{\sigma_k}$, where $\sigma_1, \dots, \sigma_n$ are i.i.d. μ and $n \in \mathsf{Z}_+$. Then for any $f \in \mathcal{S}_+$,

$$\begin{aligned} Ee^{-\xi f} &= E \exp \sum_{k\le n} \{-f(\sigma_k)\} \\ &= \prod_{k\le n} Ee^{-f(\sigma_k)} = (\mu e^{-f})^n. \end{aligned}$$

The general result follows by conditioning on κ.

(ii) For a Poisson random variable κ with mean c, we have

$$Es^\kappa = e^{-c} \sum_{k\ge 0} \frac{c^k}{k!} s^k = e^{-c} e^{cs} = e^{-c(1-s)}, \quad s \in (0,1]. \tag{1}$$

Hence, if ξ is a Poisson process with intensity μ, we get for any function $f = \sum_{k\le n} c_k 1_{B_k}$, with disjoint sets $B_1, \dots, B_n \in \hat{\mathcal{S}}$ and constants $c_1, \dots, c_n \ge 0$,

$$\begin{aligned} Ee^{-\xi f} &= E \exp \sum_{k\le n} (-c_k \xi B_k) \\ &= \prod_{k\le n} E(e^{-c_k})^{\xi B_k} \\ &= \prod_{k\le n} \exp\{-\mu B_k (1 - e^{-c_k})\} \\ &= \exp \sum_{k\le n} \{-\mu B_k (1 - e^{-c_k})\} \\ &= \exp\Big\{-\mu(1 - e^{-f})\Big\}, \end{aligned}$$

which extends by monotone and dominated convergence to arbitrary $f \in \mathcal{S}_+$. The general result follows by conditioning on η.

(iii) First let $\eta = \sum_k \delta_{s_k}$ be non-random in $\mathcal{N}_S$. Choosing some independent random elements τ_k in T with distributions ν_{s_k}, we get

$$\begin{aligned} Ee^{-\xi f} &= E \exp \textstyle\sum_k \{-f(\tau_k)\} \\ &= \textstyle\prod_k Ee^{-f(\tau_k)} = \prod_k \nu_{s_k} e^{-f} \\ &= \exp \textstyle\sum_k \log \nu_{s_k} e^{-f} \\ &= \exp\Big(\eta \log \nu e^{-f}\Big). \end{aligned}$$

The general result follows by conditioning on η.

(iv) This is easily deduced from (iii). For a direct proof, let $\eta = \sum_k \delta_{s_k}$ as before, and put $\xi = \sum_k \vartheta_k \delta_{s_k}$ for some independent random variables ϑ_k in $\{0,1\}$ with $E\vartheta_k = p(s_k)$. Then

$$\begin{aligned} Ee^{-\xi f} &= E \exp \textstyle\sum_k \{-\vartheta_k f(s_k)\} \\ &= \textstyle\prod_k Ee^{-\vartheta_k f(s_k)} \end{aligned}$$

$$= \prod_k \Big\{1 - p(s_k)\Big(1 - e^{-f(s_k)}\Big)\Big\}$$
$$= \exp \sum_k \log\Big\{1 - p(s_k)\Big(1 - e^{-f(s_k)}\Big)\Big\}$$
$$= \exp\Big(\eta \log\Big\{1 - p\,(1 - e^{-f})\Big\}\Big),$$

which extends by conditioning to general η. □

For a first application, we prove some basic closure properties for Cox processes and point process transforms. In particular, the Poisson property is preserved by randomizations.

Theorem 3.2 *(mapping and marking) Fix some probability kernels $\mu\colon S \to T$ and $\nu\colon T \to U$.*

(i) *Let ξ be a Cox process directed by some random measure η on S, such that the μ-transform $\zeta \perp\!\!\!\perp_\xi \eta$ of ξ is locally finite. Then ζ is a Cox process on U directed by $\eta\mu$.*

(ii) *Let ξ be a μ-transform of some point process η on S, such that the ν-transform $\zeta \perp\!\!\!\perp_\xi \eta$ of ξ is locally finite. Then ζ is a $\mu\nu$-transform of η.*

Recall that the kernel $\mu\nu\colon S \to U$ is given by

$$(\mu\nu)_s f = \int (\mu\nu)_s(du)\, f(u) = \int \mu_s(dt) \int \nu_t(du)\, f(u).$$

Proof: (i) By the conditional independence and Lemma 3.1 (ii) and (iii), we have

$$\begin{aligned} E(e^{-\zeta f}\,|\,\eta) &= E\Big\{E(e^{-\zeta f}\,|\,\xi,\eta)\,\Big|\,\eta\Big\} \\ &= E\Big\{E(e^{-\zeta f}\,|\,\xi)\,\Big|\,\eta\Big\} \\ &= E\Big\{\exp(\,\xi \log \mu e^{-f})\,\Big|\,\eta\Big\} \\ &= \exp\Big\{-\eta(1 - \mu e^{-f})\Big\} \\ &= \exp\Big\{-\eta\mu(1 - e^{-f})\Big\}. \end{aligned}$$

Now use Lemmas 2.2 (ii) and 3.1 (iii).

(ii) Using Lemma 3.1 (iii) twice, we get

$$\begin{aligned} E(e^{-\zeta f}\,|\,\eta) &= E\Big\{E(e^{-\zeta f}\,|\,\xi)\,\Big|\,\eta\Big\} \\ &= E\Big\{\exp(\,\xi \log \nu e^{-f})\,\Big|\,\eta\Big\} \\ &= \exp(\,\eta \log \mu\nu e^{-f}\,). \end{aligned}$$

Now apply Lemmas 2.2 (ii) and 3.1 (iii). □

Next, we show that the Cox and thinning operations are invertible. Here the relation $\xi \overset{d}{\sim} \xi'$ between two point processes ξ and ξ' means that ξ is distributed as a p-thinning of ξ', or vice versa, for some $p \in (0,1]$. Similarly, for any random measures η and η', we write $\eta \overset{d}{\simeq} \eta'$ to mean that $\eta \overset{d}{=} c\eta'$ for some constant $c > 0$.

Theorem 3.3 *(Cox and thinning uniqueness)*

(i) *Let ξ be a Cox process directed by η or a p-thinning of a point process η, where $p \in (0,1]$. Then $\mathcal{L}(\xi)$ and $\mathcal{L}(\eta)$ determine each other uniquely.*

(ii) *Both $\overset{d}{\sim}$ and $\overset{d}{\simeq}$ are equivalence relations, and for any Cox processes ξ and ξ' directed by η and η', we have $\xi \overset{d}{\sim} \xi'$ iff $\eta \overset{d}{\simeq} \eta'$.*

Proof: (i) Inverting Lemma 3.1 (ii), we get for $f \in \hat{\mathcal{S}}_+$

$$Ee^{-\eta f} = E\exp\{\,\xi\log(1-f)\}, \quad f < 1,$$

whereas in Lemma 3.1 (iv) we have instead

$$Ee^{-\eta f} = E\exp\Big(\,\xi\log\Big\{1 - p^{-1}(1-e^{-f})\Big\}\Big), \quad f < -\log(1-p). \tag{2}$$

The assertion now follows by Lemma 2.2 (ii).

(ii) For the first assertion, it suffices to prove the transitivity of the relation $\overset{d}{\sim}$, so assume that $\xi \overset{d}{\sim} \eta \overset{d}{\sim} \zeta$. Then by Lemma 3.1 (iv) and (2) there exist some constants $p, q > 0$, such that

$$\begin{aligned} Ee^{-\xi f} &= E\exp\Big(\,\eta\log\Big\{1 - p\,(1-e^{-f})\Big\}\Big), \\ Ee^{-\eta g} &= E\exp\Big(\,\zeta\log\Big\{1 - q\,(1-e^{-g})\Big\}\Big), \end{aligned} \tag{3}$$

for all $f, g \in \mathcal{S}_+$ satisfying

$$p\,(1-e^{-f}) \vee q\,(1-e^{-g}) < 1.$$

For small enough f, we may choose $g = -\log\{1 - p\,(1-e^{-f})\}$, so that

$$\log\Big\{1 - q\,(1-e^{-g})\Big\} = \log\Big\{1 - p\,q\,(1-e^{-f})\Big\}.$$

Substituting into (3) gives

$$Ee^{-\xi f} = E\exp\Big(\zeta\log\Big\{1 - p\,q\,(1-e^{-f})\Big\}\Big),$$

and $\xi \overset{d}{\sim} \zeta$ follows as before by Lemma 2.2 (ii).

The last assertion follows easily from Theorem 3.2 (i). The required conditional independence is clearly irrelevant here, since we are only concerned with the marginal distributions of ξ and ξ'. $\square$

We turn to a basic relationship between mixed Poisson and binomial processes. Write $1_B\mu = 1_B \cdot \mu$ for the restriction of μ to B.

Theorem 3.4 *(mixed Poisson and binomial processes) Let ξ be a mixed Poisson or binomial process on S based on μ, and fix any $B \in \hat{\mathcal{S}}$. Then $1_B\xi$ is a mixed binomial process based on $1_B\mu$, and $1_B\xi \perp\!\!\!\perp_{\xi B} 1_{B^c}\xi$.*

Proof: First let ξ be a mixed Poisson process based on μ and ρ. Since $1_B\xi$ is again mixed Poisson and based on $1_B\mu$ and ρ, we may assume that $B = S$ and $\|\mu\| = 1$. Now consider a mixed binomial process η based on μ and κ, where κ is conditionally Poisson given ρ with $E(\kappa|\,\rho) = \rho$. Using (1) and Lemma 3.1 (i)−(ii), we get for any $f \in \mathcal{S}_+$

$$\begin{aligned} Ee^{-\eta f} &= E(\mu e^{-f})^\kappa = EE\Big\{(\mu e^{-f})^\kappa \Big|\, \rho\Big\} \\ &= E\exp\Big\{-\rho(1-\mu e^{-f})\Big\} \\ &= E\exp\Big\{-\rho\mu(1-e^{-f})\Big\} = Ee^{-\xi f}, \end{aligned}$$

and so $\xi \stackrel{d}{=} \eta$ by Lemma 2.2.

Next, let ξ be a mixed binomial process based on μ and κ. Fix any $B \in \mathcal{S}$ with $\mu B > 0$, and consider a mixed binomial process η based on $\hat{\mu}_B = 1_B\mu/\mu B$ and β, where β is conditionally binomially distributed given κ, with parameters κ and μB. Letting $f \in \mathcal{S}_+$ be supported by B and using Lemma 3.1 (i) twice, we get

$$\begin{aligned} Ee^{-\eta f} &= E(\hat{\mu}_B e^{-f})^\beta = EE\Big\{(\hat{\mu}_B e^{-f})^\beta \Big|\, \kappa\Big\} \\ &= E\Big\{1-\mu B(1-\hat{\mu}_B e^{-f})\Big\}^\kappa \\ &= E(\mu e^{-f})^\kappa = Ee^{-\xi f}, \end{aligned}$$

and so $1_B\xi \stackrel{d}{=} \eta$ by Lemma 2.2.

To prove the stated conditional independence, we may assume that $\|\mu\| < \infty$, since the case of $\|\mu\| = \infty$ will then follow by a martingale or monotone-class argument. Then consider any mixed binomial process $\xi = \sum_{k\le\kappa}\delta_{\sigma_k}$, where the σ_k are i.i.d. μ and independent of κ. For any $B \in \mathcal{S}$, we note that $1_B\xi$ and $1_{B^c}\xi$ are conditionally independent binomial processes based on $1_B\mu$ and $1_{B^c}\mu$, respectively, given the variables κ and $\vartheta_k = 1\{k \le \kappa,\ \sigma_k \in B\}$, $k \in \mathsf{N}$. Thus, $1_B\xi$ is conditionally a binomial process based on μ_B and ξB, given $1_{B^c}\xi$, ξB, and $\vartheta_1, \vartheta_2, \ldots$. Since the conditional distribution depends only on ξB, we obtain $1_B\xi \perp\!\!\!\perp_{\xi B} (1_{B^c}\xi, \vartheta_1, \vartheta_2, \ldots)$, and the assertion follows. $\square$

It is now easy to prove the existence of the mentioned processes.

Theorem 3.5 *(Cox and transform existence)*

(i) *For any random measure η on S, there exists a Cox process ξ directed by η.*

(ii) *For any point process ξ on S and probability kernel $\nu : S \to T$, there exists a ν-transform ζ of ξ.*

Proof: By Corollary 1.16 and Lemmas 1.18 and 3.1, it is enough to prove the existence of Poisson processes. Then let $\mu \in \mathcal{M}_S$ be arbitrary, and fix a partition $B_1, B_2, \ldots \in \hat{\mathcal{S}}$ of S. By Theorem 3.4 there exist some independent Poisson processes $\xi_1, \xi_2, \ldots$ on S with intensities $\mu_n = 1_{B^c}\mu$, $n \in \mathbb{N}$. Putting $\xi = \sum_k \xi_k$ and using monotone and dominated convergence, we get for any $f \in \mathcal{S}_+$

$$\begin{aligned} Ee^{-\xi f} &= E \exp \textstyle\sum_n (-\xi_n f) = \prod_n Ee^{-\xi_n f} \\ &= \textstyle\prod_n \exp\big\{-\mu_n(1-e^{-f})\big\} \\ &= \exp\big\{-\mu(1-e^{-f})\big\}, \end{aligned}$$

and so ξ is Poisson with intensity λ by Lemmas 2.2 and 3.1. □

We proceed with some basic properties of Poisson and Cox processes.

Lemma 3.6 *(Cox simplicity and integrability) Let ξ be a Cox process directed by some random measure η on S. Then*

(i) *ξ is a.s. simple iff η is a.s. diffuse,*

(ii) *$1\{\xi f < \infty\} = 1\{\eta(f \wedge 1) < \infty\}$ a.s. for all $f \in \mathcal{S}_+$.*

Proof: (i) First reduce by conditioning to the case where ξ is a Poisson process with intensity μ. Since both properties are local, we may also assume that $\|\mu\| \in (0, \infty)$. By a further conditioning based on Theorem 3.4, we may then take ξ to be a binomial process based on μ, say $\xi = \sum_{k \le n} \delta_{\sigma_k}$, where the σ_k are i.i.d. μ. By Fubini's theorem

$$\begin{aligned} P\{\sigma_i = \sigma_j\} &= \int \mu\{s\}\, \mu(ds) \\ &= \textstyle\sum_s (\mu\{s\})^2, \quad i \neq j, \end{aligned}$$

which shows that the σ_k are a.s. distinct iff μ is diffuse.

(ii) Conditioning on η and using the disintegration theorem, we may again reduce to the case where ξ is Poisson with intensity μ. For any $f \in \mathcal{S}_+$, we get as $0 < r \to 0$

$$\exp\big\{-\mu(1-e^{-rf})\big\} = Ee^{-r\xi f} \to P\{\xi f < \infty\}.$$

Since $1 - e^{-f} \asymp f \wedge 1$, we have $\mu(1 - e^{-rf}) \equiv \infty$ when $\mu(f \wedge 1) = \infty$, whereas $\mu(1 - e^{-rf}) \to 0$ by dominated convergence when $\mu(f \wedge 1) < \infty$. □

We turn to a basic extension property for mixed binomial processes. A more general version is given in Corollary 3.38 below.

Theorem 3.7 *(binomial extension) Let ξ be a point process on S, and fix any $B_1, B_2, \ldots \in \hat{\mathcal{S}}$ with $B_n \uparrow S$. Then ξ is a mixed Poisson or binomial process, iff $1_{B_n}\xi$ is a mixed binomial process for every $n \in \mathbb{N}$.*

Proof: Let the restrictions $1_{B_n}\xi$ be mixed binomial processes based on some probability measures μ_n. Since the μ_n are unique (unless $\xi B_n = 0$ a.s.), Theorem 3.4 yields $\mu_m = 1_{B_m}\mu_n/\mu_n B_m$ whenever $m \le n$, and so there exists a measure $\mu \in \mathcal{M}_S$ with $\mu_n = 1_{B_n}\mu/\mu B_n$ for all $n \in \mathsf{N}$.

When $\|\mu\| < \infty$, we may first normalize to $\|\mu\| = 1$. Since $s_k^{n_k} \to s^n$ as $s_k \to s \in (0,1)$ and $n_k \to n \in \bar{\mathsf{N}}$, we get for any $f \in \hat{\mathcal{S}}_+$ with $\mu f > 0$, by Lemma 3.1 (i) and monotone and dominated convergence,

$$\begin{aligned} Ee^{-\xi f} \leftarrow Ee^{-\xi_n f} &= E(\mu_n e^{-f})^{\xi B_n} \\ &\to E(\mu e^{-f})^{\xi S}. \end{aligned}$$

Choosing $f = \varepsilon 1_{B_n}$ and letting $\varepsilon \downarrow 0$, we get $\xi S < \infty$ a.s., and so the previous calculation applies to arbitrary $f \in \hat{\mathcal{S}}_+$, which shows that ξ is a mixed binomial process based on μ.

Next let $\|\mu\| = \infty$. If $f \in \hat{\mathcal{S}}_+$ is supported by B_m for a fixed $m \in \mathsf{N}$, we have for any $n \ge m$

$$Ee^{-\xi f} = E(\mu_n e^{-f})^{\xi B_n} = E\Big\{1 - \frac{\mu(1-e^{-f})}{\mu B_n}\Big\}^{\mu B_n \alpha_n},$$

where $\alpha_n = \xi B_n/\mu B_n$. By Helly's selection theorem (FMP 5.19 or Lemma 4.4 below), we have convergence $\alpha_n \xrightarrow{d} \alpha$ in $[0,\infty]$ along a sub-sequence. Noting that $(1-m_n^{-1})^{m_n x_n} \to e^{-x}$ as $m_n \to \infty$ and $0 \le x_n \to x \in [0,\infty]$, we get by extended continuous mapping (FMP 4.27)

$$Ee^{-\xi f} = E\exp\big\{-\alpha\mu(1-e^{-f})\big\}.$$

Arguing as before gives $\alpha < \infty$ a.s., and so by monotone and dominated convergence, we may extend the displayed formula to arbitrary $f \in \hat{\mathcal{S}}_+$. Hence, by Lemmas 2.2 and 3.1 (ii), ξ is distributed as a mixed Poisson process based on μ and α. □

Using the Cox and thinning transforms, we may next extend the general uniqueness criteria of Theorem 2.2.

Theorem 3.8 *(simple and diffuse uniqueness) Fix any generating ring $\mathcal{U} \subset \hat{\mathcal{S}}$ and constant $c > 0$. Then*

(i) *for any point processes ξ and η on S, we have $\xi^* \stackrel{d}{=} \eta^*$ iff*

$$P\{\xi U = 0\} = P\{\eta U = 0\}, \quad U \in \mathcal{U},$$

(ii) *for any simple point processes or diffuse random measures ξ and η on S, we have $\xi \stackrel{d}{=} \eta$ iff*

$$Ee^{-c\xi U} = Ee^{-c\eta U}, \quad U \in \mathcal{U},$$

(iii) *for any simple point process or diffuse random measure* ξ *and a general random measure* η *on* S, *we have* $\xi \stackrel{d}{=} \eta$ *iff*

$$\xi U \stackrel{d}{=} \eta U, \quad U \in \mathcal{U}.$$

Proof: (i) See Theorem 2.2 (ii).

(ii) First let ξ and η be diffuse. Letting $\tilde{\xi}$ and $\tilde{\eta}$ be Cox processes directed by $c\xi$ and $c\eta$, respectively, we get by conditioning and hypothesis

$$\begin{aligned} P\{\tilde{\xi}U = 0\} &= Ee^{-c\xi U} = Ee^{-c\eta U} \\ &= P\{\tilde{\eta}U = 0\}, \quad U \in \mathcal{U}. \end{aligned}$$

Since $\tilde{\xi}$ and $\tilde{\eta}$ are a.s. simple by Lemma 3.6 (i), part (i) yields $\tilde{\xi} \stackrel{d}{=} \tilde{\eta}$, and so $\xi \stackrel{d}{=} \eta$ by Lemma 3.3 (i).

Next let ξ and η be simple point processes. For $p = 1 - e^{-c}$, let $\tilde{\xi}$ and $\tilde{\eta}$ be p-thinnings of ξ and η, respectively. Since Lemma 3.1 (iv) remains true when $0 \le f \le \infty$, we may take $f = \infty \cdot 1_U$ for any $U \in \mathcal{U}$ to get

$$\begin{aligned} P\{\tilde{\xi}U = 0\} &= E\exp\{\,\xi \log(1 - p\,1_U)\} \\ &= E\exp\{\,\xi U \log(1-p)\} = Ee^{-c\xi}, \end{aligned}$$

and similarly for $\tilde{\eta}$ in terms of η. Hence, the simple point processes $\tilde{\xi}$ and $\tilde{\eta}$ are such as in (i), and so as before $\xi \stackrel{d}{=} \eta$.

(iii) First let ξ be a simple point process. The stated condition yields $\eta U \in \mathsf{Z}_+$ a.s. for every $U \in \mathcal{U}$, and so η is a point process by Lemma 2.1. By (i) we get $\xi \stackrel{d}{=} \eta^*$, and in particular $\eta U \stackrel{d}{=} \xi U \stackrel{d}{=} \eta^* U$ for all $U \in \mathcal{U}$, which implies $\eta = \eta^*$ since $\mathcal{U}$ is covering.

Next let η be a.s. diffuse. Letting $\tilde{\xi}$ and $\tilde{\eta}$ be Cox processes directed by ξ and η, respectively, we get $\tilde{\xi}U \stackrel{d}{=} \tilde{\eta}U$ for any $U \in \mathcal{U}$ by Lemma 3.1 (ii). Since $\tilde{\xi}$ is simple by Lemma 3.6 (i), we conclude as before that $\tilde{\xi} \stackrel{d}{=} \tilde{\eta}$, and so $\xi \stackrel{d}{=} \eta$ by Lemma 3.3 (i). □

For a simple application, we consider a classical Poisson criterion. Further applications are given in the next section.

Corollary 3.9 *(Poisson criterion, Rényi) Let* ξ *be a random measure on* S *with* $E\xi\{s\} \equiv 0$, *and fix a generating ring* $\mathcal{U} \subset \hat{\mathcal{S}}$. *Then* ξ *is a Poisson process iff* ξU *is Poisson distributed for every* $U \in \mathcal{U}$, *in which case* $E\xi \in \mathcal{M}_S^*$.

Proof: Assume the stated condition. Since $\lambda = E\xi \in \mathcal{M}_S$, Theorem 3.5 yields a Poisson process η on S with $E\eta = \lambda$. Here $\xi U \stackrel{d}{=} \eta U$ for all $U \in \mathcal{U}$, and since η is simple by Lemma 3.6 (i), we get $\xi \stackrel{d}{=} \eta$ by Theorem 3.8 (iii). □

We conclude with an elementary estimate needed in Chapter 13. Say that the space S is *additive*, if it is endowed with an associative, commutative, and measurable operation "+" with a unique neutral element 0. Define $l(s) \equiv s$, and put $\xi l = \int_S s\,\xi(ds)$, interpreted as 0 when $\xi = 0$.

Lemma 3.10 *(Poisson estimate) Consider a Poisson process ξ and a measurable function $f \geq 0$ on an additive Borel space S, where f and $E\xi$ are bounded and $f(0) = 0$. Then*

$$|Ef(\xi l) - E\xi f| \leq \|f\| \, \|E\xi\|^2.$$

Proof: Since $f(0) = 0$, we have

$$\begin{aligned} Ef(\xi l) - E\xi f &= E\{f(\xi l);\ \|\xi\| > 1\} \\ &\quad - \Big(E\xi f - E\{f(\xi l);\ \|\xi\| = 1\}\Big), \end{aligned}$$

where both terms on the right are ≥ 0. Writing $p = \|E\xi\|$, we obtain

$$\begin{aligned} E\{f(\xi l);\ \|\xi\| > 1\} &\leq P\{\|\xi\| > 1\} \, \|f\| \\ &= \{1 - (1-p)e^{-p}\} \, \|f\| \\ &\leq \{1 - (1+p)(1-p)\} \, \|f\| = p^2 \, \|f\|, \end{aligned}$$

Furthermore, we get by Theorem 3.4

$$E\{f(\xi l);\ \|\xi\| = 1\} = P\{\|\xi\| = 1\} \, \frac{E\xi f}{\|E\xi\|} = e^{-p} \, E\xi f,$$

and so

$$E\xi f - E\{f(\xi l);\ \|\xi\| = 1\} = (1 - e^{-p}) \, E\xi f \leq p^2 \, \|f\|.$$

The assertion now follows by combination. □

3.2 Linear Poisson and Binomial Processes

For a point process $\xi = \sum_k \delta_{\tau_k}$ on R_+ or $[0,1]$, we may list the points in order $\tau_1 \leq \tau_2 \leq \cdots$, where $\tau_n = \infty$ or 1, respectively, when $\|\xi\| < n$. Note that ξ is simple iff $\tau_n < \tau_{n+1}$ on $\|\xi\| \geq n$ for all $n \in \mathsf{N}$. The following result gives the basic relationship between mixed Poisson and binomial processes on the random intervals $[0, \tau_n]$. Similar results for fixed intervals were given in Theorems 3.4 and 3.7. For processes on $[0,1]$, we define stationarity in the cyclic sense, as invariance in distribution modulo 1.

Theorem 3.11 *(binomial scaling) Let $\tau_1 < \tau_2 < \cdots$ form a simple point process ξ on $I = \mathsf{R}_+$ or $[0,1]$, and put $\rho_k^n = \tau_k/\tau_n$ when $\tau_n < \infty$.*

(i) *If ξ is a stationary, mixed Poisson or binomial process on R_+ or $[0,1]$, respectively, then conditionally on $\|\xi\| \geq n$, the ratios $\rho_1^n, \ldots, \rho_{n-1}^n$ form a stationary binomial process on $[0,1]$, independent of $\tau_n, \tau_{n+1}, \ldots$.*

(ii) *When* $I = \mathsf{R}_+$ *and* $\|\xi\| = \infty$ *a.s.,* ξ *is a stationary, mixed Poisson process, iff for every* $n \in \mathsf{N}$ *the ratios* $\rho_1^n, \dots, \rho_{n-1}^n$ *form a stationary binomial process on* $[0,1]$ *independent of* τ_n*.*

Proof: (i) It is enough to condition on the event $\tau_n < t$ for a fixed $t \in I$. By Theorem 3.4 and a further scaling, we may then take ξ to be a mixed binomial process on $[0,1]$. By a further conditioning, we may assume ξ to be a stationary binomial process on $[0,1]$ with $\|\xi\| = m \geq n$, so that $\tau_1, \dots, \tau_m$ have joint distribution $m!\,\lambda^m$ on the set $\Delta_m = \{0 < t_1 < \cdots < t_m < 1\}$. Then conditionally on $\tau_n, \dots, \tau_m$, the remaining variables $\tau_1, \dots, \tau_{n-1}$ have joint distribution $(n-1)!\,\lambda^{n-1}/\tau_n^{n-1}$ on $\tau_n \Delta_{n-1}$. By the disintegration theorem (FMP 6.4) it follows that the ratios $\rho_1^n, \dots, \rho_{n-1}^n$ have joint conditional distribution $(n-1)!\,\lambda^{n-1}$, independently of $\tau_n, \dots, \tau_m$.

(ii) The necessity is clear from (i). Now assume that for every $n \in \mathsf{N}$, the variables $\rho_1^n, \dots, \rho_{n-1}^n$ form a stationary binomial process on $[0,1]$ independent of τ_n. Equivalently, $\tau_1, \dots, \tau_{n-1}$ form a stationary binomial process on $[0, \tau_n]$, conditionally on τ_n. Then for fixed $t > 0$, Theorem 3.4 shows that ξ is a stationary, mixed binomial process on $[0,t]$, conditionally on $\tau_n > t$. Since $\tau_n \to \infty$ a.s. by the local finiteness of ξ, the same property holds unconditionally, and so by Theorem 3.7 the process ξ is stationary, mixed Poisson. $\Box$

An integer-valued process N on R_+ is called a *birth process* with *rates* c_n, if it is a continuous-time Markov process, increasing by unit jumps with transition probabilities $p_{n,n+1}(t) = c_n t + o(t)$. Similarly, it is called a *death process* with *rates* c_n, if it is Markov and decreases by unit jumps with transition probabilities $p_{n,n-1}(t) = c_n t + o(t)$. In either case, the holding times in state n are exponentially distributed with mean c_n^{-1}. We prove some classical characterizations in terms of birth or death processes.

Theorem 3.12 *(birth & death processes) Let* ξ *be a simple point process on* R_+*, and put* $\mu_\pm(dt) = e^{\pm t}dt$*. Then*

(i) ξ *is a Poisson process with* $E\xi = \lambda$*, iff* $N_t = \xi[0,t]$ *is a birth process on* Z_+ *with rates* $c_n = 1$*,*

(ii) *(Rényi)* ξ *is a mixed binomial process based on* μ_-*, iff* $N_t = \xi(t, \infty)$ *is a death process on* Z_+ *with rates* $c_n = n$*,*

(iii) *(Kendall)* ξ *is a mixed Poisson process based on* μ_+*, along with a random variable* $\rho > 0$ *with distribution* μ_-*, iff* $N_t = \xi[0,t] + 1$ *is a birth process on* N *with rates* $c_n = n$*.*

The process in (iii), known as a *Yule process*, arises naturally in the context of branching and super-processes, as well as in the study of Brownian excursions.

Proof: Since a point process ξ on R_+ and its set of points $\tau_1 \le \tau_2 \le \cdots$ determine each other uniquely by measurable mappings, the associated distributions are related by a $1-1$ correspondence. Hence, it suffices to prove each implication in only one direction.

(i) Let ξ be a Poisson process with intensity $E\xi = \lambda$ and points $\tau_1 < \tau_2 < \cdots$. Define $\tau_k^n = \inf\{j2^{-n} > \tau_k;\ j \in \mathsf{N}\}$ for any $k, n \in \mathsf{N}$, and write θ_t for the shift operators on $\mathcal{N}_{\mathsf{R}_+}$. Since ξ has stationary, independent increments, we get for any $A \in \sigma(\tau_1, \dots, \tau_k)$ and $B \in \mathcal{N}$

$$P\Big(A \cap \{\theta_{\tau_k^n}\xi \in B\}\Big) = P(A)\, P\{\xi \in B\}, \quad n \in \mathsf{N}, \tag{4}$$

by decomposing A into the subsets where $\tau_k \in I_{nj} \equiv 2^{-n}[j-1, j)$. Considering sets B of the form $\bigcap_{i \le m}\{\mu;\ \sigma_i(\mu) \le x_i\}$ and extending by a monotone-class argument, we see that (4) remains valid with τ_k^n replaced by τ_k, which shows that $\gamma_1, \gamma_2, \dots$ are i.i.d. Finally, $\mathcal{L}(\gamma_1) = \mu_-$ since $P\{\gamma_1 > t\} = P\{\xi[0,t] = 0\} = e^{-t}$ for all t. Thus, N is a birth process with rates $c_n = 1$.

(ii) We may take $\|\xi\| = n$ to be non-random, since the general result will then follow by a conditioning argument. First let ξ be a stationary binomial process on $[0,1]$ with points $\sigma_1 < \cdots < \sigma_n$. Then by Theorem 3.11 (i) we have on $[\sigma_{n-k}, 1]^k \cap \Delta_k$

$$\mathcal{L}(\, \sigma_{n-k+1}, \dots, \sigma_n \,|\, \sigma_1, \dots, \sigma_{n-k}\,) = \frac{k!\,\lambda^k}{(1-\sigma_{n-k})^k}\,.$$

If instead $\tau_1 < \cdots < \tau_n$ form a binomial process based on μ_-, we get by the scaling (loss of memory) property of μ_-

$$\mathcal{L}\Big\{(\tau_{n-k+1}, \dots, \tau_n) - \tau_{n-k} \,\Big|\, \tau_1, \dots, \tau_{n-k}\Big\} = k!\,\mu_-^k \ \text{ on } \ \Delta_k\,,$$

which agrees with the distribution for a binomial process based on μ_- and k. Since the minimum of k independent, exponential random variables with mean 1 is again exponentially distributed with mean k^{-1}, we get

$$P(\,\tau_{n-k+1} - \tau_{n-k} > t \,|\, \tau_1, \dots, \tau_{n-k}\,) = e^{-nt}, \quad t \ge 0,$$

which shows that N is a death process with rates $c_k = k$.

(iii) Let N be a birth process on N with rates $c_n = n$. Then the joint distribution of the spacing variables $\gamma_1, \dots, \gamma_n$ has density

$$n!\, \exp \sum_{k \le n} (-k s_k), \quad s_1, \dots, s_n > 0,$$

and so the joint density of $\tau_1, \dots, \tau_n$ equals

$$n!\, \exp\{(t_1 + \cdots + t_{n-1}) - n t_n\}, \quad 0 < t_1 < \cdots < t_n.$$

Thus, $P\{\,(\tau_1, \dots, \tau_{n-1}) \in \cdot \,|\, \tau_n\}$ is proportional to μ_+^n on $[0, \tau_n]^{n-1} \cap \Delta_n$, and so, conditionally on τ_n, the variables $\tau_1, \dots, \tau_{n-1}$ form a binomial process on

$[0, \tau_n]$ based on μ_+. By Lemma 3.11 (ii), ξ is then a mixed Poisson process based on μ_+ and some random variable $\rho \geq 0$. In particular,

$$\begin{aligned} e^{-t} &= P\{\tau_1 > t\} = Ee^{-\rho\,\mu_+[0,t]} \\ &= Ee^{-\rho(e^t-1)}, \quad t \geq 0. \end{aligned}$$

Since $Ee^{-\alpha r} = (r+1)^{-1}$ when $\mathcal{L}(\alpha) = \mu_-$, we get $\rho \overset{d}{=} \alpha$ by Lemma 2.2. □

We conclude with some simple properties of stationary Poisson and binomial processes, needed in Chapter 13, beginning with the following sampling property of stationary binomial processes.

Lemma 3.13 *(binomial sampling) Let $\tau_1 < \cdots < \tau_n$ form a stationary binomial process on $[0,1]$, and consider an independent, uniformly distributed subset $\varphi \subset \{1, \ldots, n\}$ of fixed cardinality $|\varphi| = k$. Then the sets $\{\tau_j;\ j \in \varphi\}$ and $\{\tau_j;\ j \in \varphi^c\}$ form independent, stationary binomial processes on $[0,1]$ of orders k and $n-k$.*

Proof: We may assume that $\varphi = \{\pi_1, \ldots, \pi_k\}$, where $\pi_1, \ldots, \pi_n$ form an exchangeable permutation of $1, \ldots, n$, independent of $\tau_1, \ldots, \tau_n$. The random variables $\sigma_j = \tau \circ \pi_j$, $j = 1, \ldots, n$, are then i.i.d. $U(0,1)$, and we have

$$\begin{aligned} \{\tau_j;\ j \in \varphi\} &= \{\sigma_1, \ldots, \sigma_k\}, \\ \{\tau_j;\ j \notin \varphi\} &= \{\sigma_{k+1}, \ldots, \sigma_n\}. \end{aligned} \qquad \Box$$

This leads to the following domination estimate:

Lemma 3.14 *(binomial domination) For each $n \in \mathsf{N}$, let ξ_n be a stationary binomial process on $[0,1]$ with $\|\xi_n\| = n$. Then for any point process $\eta \leq \xi_n$ with non-random $\|\eta\| = k \leq n$, we have*

$$\mathcal{L}(\eta) \leq \binom{n}{k} \mathcal{L}(\xi_k).$$

Proof: Let $\tau_1 < \cdots < \tau_n$ be the points of ξ_n. Writing $\xi_n^J = \sum_{j \in J} \delta_{\tau_j}$ for subsets $J \subset \{1, \ldots, n\}$, we have $\eta = \xi_n^\varphi$ for some random subset $\varphi \subset \{1, \ldots, n\}$ with $|\varphi| = k$ a.s. Choosing $\psi \subset \{1, \ldots, n\}$ to be independent of ξ_n and uniformly distributed with $|\psi| = k$, we get

$$\begin{aligned} \mathcal{L}(\eta) &= \mathcal{L}(\xi_n^\varphi) = \textstyle\sum_J \mathcal{L}(\xi_n^J;\ \varphi = J) \\ &\leq \textstyle\sum_J \mathcal{L}(\xi_n^J) = \binom{n}{k} \mathcal{L}(\xi_n^\psi) \\ &= \binom{n}{k} \mathcal{L}(\xi_k), \end{aligned}$$

with summations over subsets J of cardinality k, where the last equality holds by Lemma 3.13. □

We turn to the distributions of the points τ_k and the joint conditional distribution of the remaining points.

Lemma 3.15 *(binomial density and conditioning) Let $\tau_1 < \cdots < \tau_n$ form a stationary binomial process on $[0,t]$, and fix any $k \in \{1,\dots,n\}$. Then*

(i) *the distrinution of τ_k has density*

$$p_k(s) = k\binom{n}{k} s^{k-1}(t-s)^{n-k}\, t^{-n}, \quad s \in [0,t],$$

(ii) *conditionally on τ_k, the sets $\{\tau_j;\, j<k\}$ and $\{\tau_j;\, j>k\}$ form independent, stationary binomial processes on $[0,\tau_k]$ and $[\tau_k,t]$, respectively, of cardinalities $k-1$ and $n-k$.*

Proof: Part (i) is elementary and classical. As for part (ii), we see from Theorem 3.11 (i) that the times $\tau_1,\dots,\tau_{k-1}$ form a stationary binomial process on $[0,\tau_k]$, conditionally on $\tau_k,\dots,\tau_n$. By symmetry, the times $\tau_{k+1},\dots,\tau_n$ form a stationary binomial process on $[\tau_k,1]$, conditionally on $\tau_1,\dots,\tau_k$. The conditional independence follows by FMP 6.6 from the fact that those conditional distributions depend only on τ_k. □

We further need an identity for stationary Poisson processes on R_+, stated in terms of the tetrahedral sets

$$t\Delta_n = \Big\{s \in \mathsf{R}^n;\ s_1 < \cdots < s_n < t\Big\}, \quad t>0,\ n \in \mathsf{N}.$$

Lemma 3.16 *(Poisson identity) Let $\tau_1 < \tau_2 < \cdots$ form a stationary Poisson process on R_+ with rate $c>0$, and fix any $n \in \mathsf{N}$. Then for any measurable function $f \geq 0$ on R_+^{n+1}, we have*

$$Ef(\tau_1,\dots,\tau_{n+1}) = c^n E\int\cdots\int_{\tau_1\Delta_n} f(s_1,\dots,s_n,\tau_1)\, ds_1\cdots ds_n. \tag{5}$$

Proof: Since

$$E\tau_1^n = \int_0^\infty t^n c\, e^{-ct}\, dt = n!\, c^{-n},$$

the right-hand side of (5) defines the joint distribution of some random variables $\sigma_1,\dots,\sigma_{n+1}$. Noting that $\mathcal{L}(\tau_{n+1})$ has density

$$g_{n+1}(s) = \frac{c^{n+1} s^n\, e^{-cs}}{n!}, \quad s \geq 0,$$

we get for any measurable function $f \geq 0$ on R_+

$$\begin{aligned} Ef(\tau_{n+1}) &= \int_0^\infty f(s)\, g_{n+1}(s)\, ds \\ &= \frac{c^{n+1}}{n!}\int_0^\infty s^n f(s)\, e^{-cs}\, ds \\ &= \frac{c^n}{n!}\, E\,\tau_1^n f(\tau_1) \\ &= c^n E\int\cdots\int_{\tau_1\Delta_n} f(\tau_1)\, ds_1\cdots ds_n, \end{aligned}$$

and so $\sigma_{n+1} \stackrel{d}{=} \tau_{n+1}$. Furthermore, the expression in (5) shows that $\sigma_1, \ldots, \sigma_n$ form a stationary binomial process on $[0, \sigma_n]$, conditionally on σ_{n+1}. Since the corresponding property holds for $\tau_1, \ldots, \tau_{n+1}$ by Theorem 3.11 (i), we obtain

$$(\sigma_1, \ldots, \sigma_{n+1}) \stackrel{d}{=} (\tau_1, \ldots, \tau_{n+1}),$$

which implies (5). □

3.3 Independence and Infinite Divisibility

We begin with some basic characterizations of simple point processes or diffuse random measures with independent increments.

Theorem 3.17 *(simple and diffuse independence, Erlang, Lévy, Wiener)*

(i) *Let ξ be a simple point process on S with $E\xi\{s\} \equiv 0$. Then ξ has pairwise independent increments iff it is Poisson.*

(ii) *Let ξ be a diffuse random measure on S. Then ξ has pairwise independent increments iff it is a.s. non-random.*

Proof: Define a set function ρ on $\hat{\mathcal{S}}$ by

$$\rho B = -\log E e^{-\xi B}, \quad B \in \hat{\mathcal{S}},$$

and note that $\rho B \in \mathsf{R}_+$ for all $B \in \hat{\mathcal{S}}$, since $\xi B \in \mathsf{R}_+$ a.s. If ξ has pairwise independent increments, then for disjoint $B, C \in \hat{\mathcal{S}}$,

$$\begin{aligned} \rho(B \cup C) &= -\log E e^{-\xi B - \xi C} \\ &= -\log\left\{E e^{-\xi B} E e^{-\xi C}\right\} \\ &= \rho B + \rho C, \end{aligned}$$

which shows that ρ is finitely additive. In fact, it is even countably additive, since $B_n \uparrow B$ in $\hat{\mathcal{S}}$ implies $\rho B_n \uparrow \rho B$ by monotone and dominated convergence. Finally, $\rho\{s\} = -\log E e^{-\xi\{s\}} = 0$ for all $s \in S$, which means that ρ is diffuse.

(i) By Theorem 3.5 there exists a Poisson process η on S with $E\eta = c^{-1}\rho$, where $c = 1 - e^{-1}$. For any $B \in \hat{\mathcal{S}}$ we get by Lemma 3.1 (ii)

$$\begin{aligned} E e^{-\eta B} &= \exp\left\{-c^{-1}\rho(1 - e^{-1_B})\right\} \\ &= e^{-\rho B} = E e^{-\xi B}. \end{aligned}$$

Since ξ and η are simple by hypothesis and Lemma 3.6 (i), respectively, Theorem 3.8 (ii) yields $\xi \stackrel{d}{=} \eta$.

(ii) Taking $\eta \equiv \rho$ gives $E e^{-\xi B} = e^{-\rho B} = E e^{-\eta B}$ for all $B \in \hat{\mathcal{S}}$. Since ξ and η are both diffuse, Theorem 3.8 (ii) yields $\xi \stackrel{d}{=} \eta = \rho$, and so $\xi = \rho$ a.s.

by Lemma 2.1. □

We proceed to characterize the class of random measures ξ on a product space $S \times T$ with independent S-increments, beginning with the special case of marked point processes. Here we say that ξ has *independent S-increments*, if for any disjoint sets $B_1, \ldots, B_n \in \hat{\mathcal{S}}$, $n \in \mathsf{N}$, the random measures $\xi(B_1 \times \cdot), \ldots, \xi(B_n \times \cdot)$ on T are independent. Define $\mathcal{M}'_T = \mathcal{M}_T \setminus \{0\}$ and $T^\Delta = T \cup \{\Delta\}$, where the point $\Delta \notin T$ is arbitrary.

Theorem 3.18 *(marked independence, Itô, Kingman) A T-marked point process ξ on S has independent S-increments iff*

$$\xi = \eta + \sum\nolimits_k (\delta_{s_k} \otimes \delta_{\tau_k}) \quad a.s.,$$

for some Poisson process η on $S \times T$ with $E\eta(\{s\} \times T) \equiv 0$, some distinct points $s_1, s_2, \ldots \in S$, and some independent random elements $\tau_1, \tau_2, \ldots$ in T^Δ with distributions $\nu_1, \nu_2, \ldots$, such that the measure

$$\rho = E\eta + \sum\nolimits_k (\delta_{s_k} \otimes \nu_k) \tag{6}$$

on $S \times T$ is locally finite. Here ρ is unique and determines $\mathcal{L}(\xi)$, and any measure $\rho \in \mathcal{M}_{S \times T}$ with $\sup_s \rho(\{s\} \times T) \leq 1$ *may occur.*

Proof: Clearly $E\xi(\{s\} \times T) > 0$ for at most countably many $s \in S$, say $s_1, s_2, \ldots$, and we may separate the corresponding sum $\sum_k (\delta_{s_k} \otimes \delta_{\tau_k})$, where the random elements τ_k in T^Δ are mutually independent and independent of the remaining process ξ'.

Now fix any $B \in \hat{\mathcal{S}} \otimes \hat{\mathcal{T}}$, and note that the simple point process $\xi_B = 1_B \xi'(\cdot \times T)$ on S has independent increments and satisfies $E\xi_B\{s\} \equiv 0$. Hence, ξ_B is Poisson by Theorem 3.17 (i), and in particular $\xi' B = \xi_B S$ is a Poisson random variable. Since B was arbitrary, ξ' is again Poisson by Corollary 3.9. The last assertions are obvious. □

Theorem 3.19 *(independent increments) A random measure ξ on $S \times T$ has independent S-increments iff a.s.*

$$\xi = \alpha + \sum\nolimits_k (\delta_{s_k} \otimes \beta_k) + \iint (\delta_s \otimes \mu)\, \eta(ds\, d\mu), \tag{7}$$

for some non-random measure $\alpha \in \mathcal{M}_{S \times T}$ with $\alpha(\{s\} \times T) \equiv 0$, a Poisson process η on $S \times \mathcal{M}'_T$ with $E\eta(\{s\} \times \mathcal{M}'_T) \equiv 0$, some distinct points $s_1, s_2, \ldots \in S$, and some independent random measures $\beta_1, \beta_2, \ldots$ on T with distributions $\nu_1, \nu_2, \ldots$, such that the measure ρ in (6) satisfies

$$\int (\mu C \wedge 1)\, \rho(B \times d\mu) < \infty, \quad B \in \hat{\mathcal{S}},\ C \in \hat{\mathcal{M}}_T. \tag{8}$$

The pair (α, ρ) is unique and determines $\mathcal{L}(\xi)$, and any α and ρ with the stated properties may occur.

Proof: Again, we have $E\xi(\{s\}\times T) \neq 0$ for at most countably many $s \in S$, and we may separate the corresponding sum $\sum_k(\delta_{s_k} \otimes \beta_k)$, for some distinct points $s_1, s_2, \ldots \in S$ and associated random measures β_k on T, where the latter are mutually independent and independent of the remaining process ξ'. The remaining discontinuities may be encoded by a marked point process η on $S \times \mathcal{M}'_T$ given by

$$\eta = \sum\nolimits_k (\delta_{\sigma_k} \otimes \delta_{\gamma_k}),$$

for some a.s. distinct random elements $\sigma_1, \sigma_2, \ldots$ in S and associated random measures $\gamma_k = \xi'(\{\sigma_k\} \times \cdot)$ on T. Here η depends measurably on ξ' by Lemmas 1.6 and 1.15 (iii), and it inherits from ξ' the independence of the S-increments. Since η has no fixed discontinuities, it is Poisson by Lemma 3.18, and the corresponding component of ξ' may be represented as the integral term in (7). Subtracting even this part, we end up with a random measure α on $S \times T$ with independent increments satisfying $\alpha(\{s\} \times T) \equiv 0$ a.s. Proceeding as in the previous proof and using Lemma 3.17 (ii), we see that αB is a.s. non-random for every $B \in \hat{\mathcal{S}} \otimes \hat{\mathcal{T}}$, and so α is a.s. non-random by Lemma 2.1.

For any $B \in \hat{\mathcal{S}}$ and $C \in \hat{\mathcal{T}}$, we see from (7) that

$$\sum_{s_k \in B} \beta_k C + \int (\mu C)\, \eta(B \times d\mu) < \infty \text{ a.s.},$$

which is equivalent to (8) by Lemma 3.1 (i) and FMP 4.17. The uniqueness of α and ρ is obvious, since α, η, and the β_k are a.s. unique, measurable functions of ξ. The last assertion follows by the same argument, once we note that the β_k are now $\mathcal{N}_T$-valued, and any measure $\alpha \in \mathcal{N}_{S\times T}$ with $\alpha(\{s\} \times T) \equiv 0$ equals 0. □

A random measure ξ on S is said to be *infinitely divisible*, if for every $n \in \mathbb{N}$ there exist some i.i.d. random measures $\xi_{n1}, \ldots, \xi_{nn}$ such that $\xi \overset{d}{=} \sum_{j\le n} \xi_{nj}$. For point processes ξ, the ξ_{nj} are again required to be point processes. The distinction is important, since every non-random measure $\mu \in \mathcal{N}_S$ is trivially infinitely divisible as a random measure, but not as a point process. The general characterization is the following:

Theorem 3.20 *(infinitely divisible cluster representation, Matthes, Jiřina, Lee) A random measure ξ on S is infinitely divisible iff*

$$\xi = \alpha + \int \mu\, \eta(d\mu) \ \ a.s.,$$

for some non-random measure $\alpha \in \mathcal{M}_S$ and a Poisson process η on $\mathcal{M}'_S$ with intensity λ satisfying

$$\int (\mu B \wedge 1)\, \lambda(d\mu) < \infty, \quad B \in \hat{\mathcal{S}}.$$

The pair (α, λ) is unique and determines $\mathcal{L}(\xi)$, and any α and λ with the stated properties may occur. For point processes ξ, the result remains valid with $\alpha = 0$ and with λ restricted to $\mathcal{N}'_S$.

Here λ is called the *Lévy measure* of ξ, and we say that a random measure or point process ξ of the stated form is *directed by* (α, λ) or λ, respectively. In terms of Laplace transforms, the representation becomes

$$-\log Ee^{-\xi f} = \alpha f + \int (1 - e^{-\mu f})\,\lambda(d\mu), \quad f \in \mathcal{S}_+. \tag{9}$$

Proof: The sufficiency is obvious, since both α and η are trivially infinitely divisible. Now let ξ be infinitely divisible. Then for every $n \in \mathsf{Z}_+$, we have $\xi \overset{d}{=} \sum_{j \le 2^n} \xi_{nj}$ for some i.i.d. random measures ξ_{nj}, $1 \le j \le 2^n$. Since

$$Ee^{-\xi_{nj} f} = (Ee^{-\xi f})^{2^{-n}}, \quad f \in \hat{\mathcal{S}}_+,\ j \le 2^n,\ n \in \mathsf{Z}_+,$$

the ξ_{nj} are uniquely distributed by Lemma 2.2, and so

$$\xi_{n+1,\,2j-1} + \xi_{n+1,\,2j} \overset{d}{=} \xi_{nj}, \quad j \le 2^n,\ n \in \mathsf{Z}_+. \tag{10}$$

Using Lemma 1.16 recursively, and modifying the resulting ξ_{nj} on a common null set, we may assume that (10) holds identically. We may then define, on the dyadic rationals $j2^{-n}$ in $[0,1]$, a measure-valued process X by

$$X_t = \sum_{i \le j} \xi_{ni}, \quad t = j2^{-n},\ j \le 2^n,\ n \in \mathsf{Z}_+.$$

Since X is non-decreasing, we may next form the right-hand limits $X_t^+ = X_{t+} = \inf_{s>t} X_t$ for $t \in [0,1)$ and put $X_1^+ = X_1$, to define a right-continuous, non-decreasing, measure-valued process X^+ on $[0,1]$, where the measure property holds for every t by dominated convergence. Since the sequence ξ_{nj} is stationary for each n, the increments of X^+ are stationary under dyadic shifts, which extends by right continuity to arbitrary shifts. In particular, $X_t^+ = X_t$ a.s. for dyadic times t, which shows that the dyadic increments are also independent. Even the latter property extends by right continuity to arbitrary increments.

By an elementary construction (cf. FMP, p. 59), we may form a random measure $\tilde{\xi}$ on $[0,1] \times S$ with $\tilde{\xi}([0,t] \times \cdot) = X_t$ for all $t \in [0,1]$. Since $\tilde{\xi}$ has again stationary, independent increments along $[0,1]$, Theorem 3.19 yields an a.s. representation as in (7), in terms of a non-random measure $\tilde{\alpha}$ on $[0,1] \times S$ and a Poisson process $\tilde{\eta}$ on $[0,1] \times \mathcal{M}'_S$ with intensity $\tilde{\lambda}$ as in (8). Projecting onto S yields the desired representation of ξ itself. A similar argument based on the last assertion in Theorem 3.19 yields the corresponding representation of infinitely divisible point processes.

For any other pair (α', λ'), we may construct an associated random measure $\tilde{\xi}'$ on $[0,1] \times S$ with stationary, independent increments and $\tilde{\xi}'([0,1] \times \cdot) = \xi$ a.s. Here $\tilde{\xi}' \overset{d}{=} \tilde{\xi}$ by the uniqueness of the distributions $\mathcal{L}(\xi_{nj})$ noted earlier. Writing η' for the Poisson process on $\mathcal{M}'_S$ associated with $\tilde{\xi}'$, we obtain $(\alpha', \eta') \overset{d}{=} (\alpha, \eta)$, since (α, η) and (α', η') are measurably determined by $\tilde{\xi}$ and $\tilde{\xi}'$, respectively, through the same construction. Taking expected values gives

$\alpha' = \alpha$ and $\lambda' = \lambda$, which proves the asserted uniqueness. □

For the next result, measures of the form $r\delta_s$ are said to be *degenerate.*

Corollary 3.21 *(independent increments, Lee) An infinitely divisible random measure on S has independent increments iff its Lévy measure is supported by the set of degenerate measures on S. Thus, a random measure ξ is infinitely divisible with independent increments iff*

$$\xi = \alpha + \iint r\delta_s \, \eta(ds\, dr) \quad a.s.,$$

for some non-random measure $\alpha \in \mathcal{M}_S$ and a Poisson process η on $S \times (0, \infty)$ with intensity ρ satisfying

$$\int (r \wedge 1)\, \rho(B \times dr) < \infty, \quad B \in \hat{\mathcal{S}}.$$

For point processes ξ we have $\alpha = 0$, and ρ is restricted to $S \times \mathsf{N}$.

Proof: The sufficiency is obvious. Now suppose that ξ is infinitely divisible with independent increments. In particular, $\xi\{s\}$ is infinitely divisible for every $s \in S$. Treating separately the countably many points s with $E\xi\{s\} > 0$, and applying the last two theorems, we obtain the required representation. Letting ζ be the image of η under the mapping $(s, r) \mapsto r\delta_s$, we get a representation as in Theorem 3.20, with $\lambda = E\zeta$ restricted to the set of degenerate measures. The necessity now follows by the uniqueness of λ. □

The infinite divisibility of a random measure is preserved by both the Cox transformation and by arbitrary randomizations. We proceed to identify the associated L'evy measures. As before, we write $\psi(x) = 1 - e^{-x}$.

Corollary 3.22 *(infinite divisibility preservation)*

(i) *Let ξ be a Cox process directed by an infinitely divisible random measure η on S with directing pair (α, λ). Then ξ is again infinitely divisible with a Lévy measure $\tilde{\lambda}$, given for $f \in \mathcal{S}_+$ and $g = \psi \circ f$ by*

$$\int (1 - e^{-\mu f})\, \tilde{\lambda}(d\mu) = \alpha g + \int (1 - e^{-\mu g})\, \lambda(d\mu).$$

(ii) *Given a probability kernel $\nu : S \to T$, let ξ be a ν-transform of an infinitely divisible point process η on S with Lévy measure λ. Then ξ is again infinitely divisible with a Lévy measure $\tilde{\lambda}$, given as in* (i) *with $\alpha = 0$ and $g = -\log \nu e^{-f}$.*

When λ is a probability measure on $\mathcal{M}_S$ or $\mathcal{N}_S$, respectively, Lemma 3.1 shows that $\tilde{\lambda}$ is the distribution of a corresponding Cox process or ν-randomization, restricted to $\mathcal{N}'_S$. The Cox transformation or randomization extends to any σ-finite measure λ. In particular, it applies to the Lévy measure above, which yields a measure $\tilde{\lambda}$ as in (i) with $\alpha = 0$. Writing

$$\begin{aligned} \alpha g = \alpha(1 - e^{-f}) &= \int (1 - e^{-\delta_s f})\, \alpha(ds) \\ &= \int (1 - e^{-\mu f})\, \alpha\delta(d\mu), \end{aligned}$$

we see that the independent Poisson component generated by α gives rise to the component $\alpha\delta = \int \alpha(ds)\, \delta_s$ of $\tilde{\lambda}$.

Proof: (i) For fixed $n \in \mathsf{N}$, let $\eta \overset{d}{=} \eta_1 + \cdots + \eta_n$ for some i.i.d. random measures η_k, and choose some conditionally independent Cox processes $\xi_1, \ldots, \xi_n$ directed by $\eta_1, \ldots, \eta_n$, so that $\xi_1, \ldots, \xi_n$ are again i.i.d. Writing $g = 1 - e^{-f}$ for arbitrary $f \in \mathcal{S}_+$ and letting $\tilde{\xi} = \xi_1 + \cdots + \xi_n$, we get by Lemma 3.1

$$\begin{aligned} Ee^{-\tilde{\xi} f} &= \prod\nolimits_k Ee^{-\xi_k f} = \prod\nolimits_k Ee^{-\eta_k g} \\ &= Ee^{-\eta g} = Ee^{-\xi f}, \end{aligned}$$

which shows that $\xi \overset{d}{=} \tilde{\xi}$. Since n was arbitrary, this shows that ξ is again infinitely divisible. Letting $\tilde{\lambda}$ be the Lévy measure of ξ, we get by (9) and Lemma 3.1

$$\begin{aligned} \int (1 - e^{-\mu f})\, \tilde{\lambda}(d\mu) &= -\log Ee^{-\xi f} \\ &= -\log Ee^{-\eta g} \\ &= \alpha g + \int (1 - e^{-\mu g})\, \lambda(d\mu). \end{aligned}$$

(ii) Here a similar argument applies with $g = -\log \nu e^{-f}$. □

Next, we consider the atomic properties of an infinitely divisible random measure and its distribution. Given a measure μ on S, we define its *atomic support* as the set $\mathrm{supp}_+ \mu = \{s \in S;\, \mu\{s\} > 0\}$. For any subset A of an additive semi-group S, we write $A^{\oplus n}$ for the *Minkowski sum* of n copies of A, consisting of all sums $x_1 + \cdots + x_n$ with $x_1, \ldots, x_n \in A$.

Theorem 3.23 *(simplicity and diffuseness) Let ξ be an infinitely divisible random measure or point process on S directed by (α, λ) or λ. Then*

(i) *ξ is a.s. diffuse iff $\alpha \in \mathcal{M}_S^*$ and $\lambda(\mathcal{M}_S^*)^c = 0$,*

(ii) *when ξ is a point process, it is a.s. simple iff $\sup_s \lambda\{\mu\{s\} > 0\} = 0$ and $\lambda(\mathcal{N}_S^*)^c = 0$,*

(iii) $$\mathrm{supp}_+ \mathcal{L}(\xi) = \begin{cases} \emptyset, & \|\lambda\| = \infty, \\ \alpha + \bigcup_{n \ge 0} (\mathrm{supp}_+ \lambda)^{\oplus n}, & \|\lambda\| < \infty. \end{cases}$$

Proof: (i) By Theorem 3.20 we have $\xi = \alpha + \sum_k \xi_k$, where $\xi_1, \xi_2, \ldots$ form a Poisson process on $\mathcal{M}_S$ with intensity λ. Here ξ is diffuse iff α and all ξ_k are diffuse. By Theorem 3.4 the latter holds a.s. iff λ is supported by $\mathcal{M}_S^*$.

(ii) Here the representation simplifies to $\xi = \sum_k \xi_k$, and ξ is simple iff $\sup_s \sum_k \xi_k\{s\} \le 1$, which holds iff the ξ_k are simple with disjoint sets of atoms. By Theorem 3.4 we may take the ξ_k to be i.i.d. with distribution λ. They are then a.s. simple iff λ is supported by $\mathcal{N}_S^*$, in which case Fubini's theorem yields

$$\begin{aligned} E\sum\nolimits_s \Big(\xi_j\{s\} \wedge \xi_k\{s\}\Big) &= E\int \xi_j\{s\}\,\xi_k(ds) \\ &= \int E\xi_j\{s\}\,E\xi_k(ds) \\ &= \sum\nolimits_s (E\xi_k\{s\})^2 \\ &= \sum\nolimits_s (\lambda\{\mu\{s\} > 0\})^2, \end{aligned}$$

which vanishes iff the second condition is fulfilled.

(iii) When $\|\lambda\| < \infty$, the number of terms ξ_k is finite and Poisson distributed, and the stated formula follows. If instead $\|\lambda\| = \infty$, we may fix a partition $B_1, B_2, \ldots \in \hat{\mathcal{S}}$ of S, choose $c_1, c_2, \ldots > 0$ with $P\{\xi B_k > c_k\} \le 2^{-k}$ for all k, and put $f = \sum_k 2^{-k} c_k^{-1} 1_{B_k}$. Then $\xi f < \infty$ a.s. by the Borel–Cantelli lemma, and it suffices to show that $\mathcal{L}(\xi f - \alpha f)$ is diffuse, which reduces the discussion to the case of infinitely divisible random variables $\xi \ge 0$ directed by $(0, \lambda)$.

For every $n \in \mathsf{N}$, we may choose a measure $\lambda_n \le \lambda$ on $(0, n^{-1})$ with $\|\lambda_n\| = \log 2$. Let ξ_n and η_n be independent and infinitely divisible with Lévy measures $\lambda - \lambda_n$ and λ_n, respectively. Then $\xi \overset{d}{=} \xi_n + \eta_n$, and so for any $x \ge 0$ and $\varepsilon > 0$,

$$\begin{aligned} P\{\xi = x\} \le\; & P\{\xi_n = x,\, \eta_n = 0\} \\ & + P\Big\{\xi_n \in [x - \varepsilon, x),\, \eta_n > 0\Big\} + P\{\eta_n > \varepsilon\} \\ =\; & \tfrac{1}{2}\, P\{\xi_n \in [x - \varepsilon, x]\} + P\{\eta_n > \varepsilon\}. \end{aligned}$$

Letting $n \to \infty$ and then $\varepsilon \to 0$, and noting that $\xi_n \xrightarrow{d} \xi$ and $\eta_n \xrightarrow{P} 0$, we obtain $P\{\xi = x\} \le \frac{1}{2} P\{\xi = x\}$, which implies $\xi \ne x$ a.s. □

We conclude with some finite-dimensional criteria for infinite divisibility.

Theorem 3.24 *(finite infinite divisibility) For a generating semiring $\mathcal{I} \subset \hat{\mathcal{S}}$,*

(i) *a random measure ξ on S is infinitely divisible, iff $(\xi I_1, \ldots, \xi I_n)$ is so for any $I_1, \ldots, I_n \in \mathcal{I}$, $n \in \mathsf{N}$,*

(ii) *a point process ξ on S is infinitely divisible as a random measure, iff ξf is infinitely divisible for every $f \in \mathcal{I}_+$.*

Proof: Since the necessity is obvious in both cases, we need to prove only the sufficiency.

(i) Suppose that ξ satisfies the stated condition. By considering sums of components ξI_k, we may extend the condition to the ring $\mathcal{U}$ generated by $\mathcal{I}$. Now fix any $m \in \mathsf{N}$. By hypothesis, we may choose some probability measures $\mu_{B_1,\dots,B_n}$ on R_+^n satisfying

$$(\mu_{B_1,\dots,B_n})^{*m} = \mathcal{L}(\xi B_1, \dots, \xi B_n), \quad B_1, \dots, B_n \in \mathcal{U}, \ n \in \mathsf{N}, \tag{11}$$

whose uniqueness is clear from the uniqueness theorem for Laplace transforms. They are further consistent in the sense of Kolmogorov, and so there exists a process $\eta \geq 0$ on $\mathcal{U}$ with those measures as finite-dimensional distributions. Here (11) yields $\eta(A \cup B) = \eta A + \eta B$ for disjoint $A, B \in \mathcal{U}$. Furthermore, we may combine the continuity $\xi A \to 0$ as $A \downarrow \emptyset$ along $\hat{\mathcal{S}}$ with the obvious stochastic boundedness to see that $\eta A \xrightarrow{P} 0$ as $A \downarrow \emptyset$ along $\mathcal{U}$. Hence, Theorem 2.15 yields a random measure ζ on S with $\zeta B = \eta B$ a.s. for all $B \in \mathcal{U}$, and so $\{\mathcal{L}(\zeta)\}^{*m} = \mathcal{L}(\xi)$ by (11) and Lemma 2.2. Since m was arbitrary, this shows that ξ is infinitely divisible.

(ii) Let ξ be a point process satisfying the stated condition. For any $f = \sum_{k \leq n} t_j 1_{I_j}$ with $I_1, \dots, I_n \in \mathcal{I}$ and $t = (t_1, \dots, t_n) \in (0, \infty)^n$, the random variable ξf is infinitely divisible with vanishing constant component, and so by Theorem 3.20

$$\xi f = \sum_{j \leq n} t_j \, \xi I_j = \alpha_t + \int r \, \eta_t(dr) \ \text{ a.s.}, \tag{12}$$

for some constant $\alpha_t \geq 0$ and Poisson process η_t on $(0, \infty)$ with bounded intensity. Since ξf is a.s. restricted to the semigroup $H_t = \{ \sum_j t_j z_j; \, z_1, \dots, z_n \in \mathsf{Z}_+\}$, the same thing is true for α_t and η_t by Theorem 3.23. For rationally independent $t_1, \dots, t_n > 0$, the mapping $\varphi_t \colon (z_1, \dots, z_n) \mapsto \sum_j t_j z_j$ is bijective from Z_+^n to H_t, and (12) becomes equivalent to

$$(\xi I_1, \dots, \xi I_n) = \alpha + \int (z_1, \dots z_n) \, \eta(dz_1 \cdots dz_n) \ \text{ a.s.},$$

where α and η denote the images of α_t and η_t under the inverse map φ_t^{-1}, so that $\alpha \in \mathsf{Z}_+$, and η is a Poisson process on $\mathsf{Z}_+^n \setminus \{0\}$. In particular, $(\xi I_1, \dots, \xi I_n)$ is infinitely divisible, and the assertion follows from (i) since $I_1, \dots, I_n \in \mathcal{I}$ were arbitrary. □

3.4 Poisson and Related Integrals

Here we derive some existence, convergence, and tightness criteria for Poisson and related integrals ξf, along with their symmetric or compensated versions $\tilde{\xi} f$ and $(\xi - \nu) f$. The general setting of independent increments, which covers both continuous and discrete time, is needed for extensions to more general processes in subsequent chapters.

We begin with some simple formulas for the integrals ξf with respect to an independent-increment process ξ. To simplify the statements in this section, we assume that all sets B and functions f on $S \times T$ are measurable. We also define

$$\psi(x) = 1 - e^{-x}, \quad x \ge 0.$$

Lemma 3.25 *(moment relations) Let ξ be a T-marked point process on S with independent increments and $E\xi = \nu$. Then for any sets B and functions $f, g \ge 0$ on $S \times T$, we have*

(i) $E\psi(\xi f) \ge \psi(\nu(\psi \circ f))$,

(ii) $P\{\xi B > 0\} \ge \psi(\nu B)$,

(iii) $\mathrm{Cov}(\xi f, \xi g) \le \nu(fg)$,

where in (iii) *we assume that $f, g \in L^2(\nu)$. All relations hold with equality when ξ is Poisson on S with intensity $E\xi = \nu$.*

Proof: (i) When ξ is Poisson ν, this holds with equality by Lemma 3.1 (ii). For marked point processes ξ with independent increments, we have by Lemma 3.18

$$\begin{aligned} \nu &= \alpha + \textstyle\sum_k (\delta_{s_k} \otimes \beta_k) = \alpha + \sum_k \nu_k, \\ \xi &= \eta + \textstyle\sum_k (\delta_{s_k} \otimes \delta_{\tau_k}) \text{ a.s.}, \end{aligned}$$

where η is a Poisson process with $E\eta = \alpha$, the points $s_1, s_2, \ldots \in S$ are distinct, and $\tau_1, \tau_2, \ldots$ are independent random elements in T^Δ with distributions $\beta_1, \beta_2, \ldots$. Using the Poisson result along with the independence of increments, and noting that $\log(1-x) \le -x$ for $x \in [0,1)$, we get

$$\begin{aligned} Ee^{-\xi f} &= Ee^{-\eta f} \textstyle\prod_k Ee^{-f(s_k, \tau_k)} \\ &= e^{-\alpha(\psi \circ f)} \textstyle\prod_k \{1 - \nu_k(\psi \circ f)\} \\ &= e^{-\alpha(\psi \circ f)} \exp \textstyle\sum_k \log\{1 - \nu_k(\psi \circ f)\} \\ &\le e^{-\alpha(\psi \circ f)} \exp \textstyle\sum_k \{-\nu_k(\psi \circ f)\} = e^{-\nu(\psi \circ f)}. \end{aligned}$$

(ii) Taking $f = t1_B$ in (i) gives

$$E(1 - e^{-t\xi B}) \ge \psi\{(1 - e^{-t})\nu B\}, \quad t \ge 0,$$

with equality when ξ is Poisson ν, and the assertion follows by dominated convergence as $t \to \infty$.

(iii) In the Poisson case, Theorem 3.2 allows us to assume that ν is diffuse, so that ξ is simple by Corollary 3.6 (i). Define $D = \{(s,s);\ s \in S\}$, and let ν_D be the image of ν under the mapping $s \mapsto (s,s)$. Starting from non-diagonal rectangles and using the independence property of ξ, we get $E\xi^2 = \nu^2$ on

$S^2 \setminus D$. Since the restriction of ξ^2 to D equals the projection of ξ onto D, we further note that $E\xi^2 = \nu_D$ on D. Hence, by combination, $E\xi^2 = \nu^2 + \nu_D$, and so $E(\xi f \cdot \xi g) = \nu f \cdot \nu g + \nu(fg)$, which yields the desired formula.

Now consider a general process ξ with independent increments. Using the decomposition in (i), the Poisson result, the independence of $\eta, \tau_1, \tau_2, \ldots$, and monotone convergence, we get for any $f, g \in L^2(\nu)$

$$\begin{aligned} \mathrm{Cov}(\xi f, \xi g) &= \mathrm{Cov}(\eta f, \eta g) + \textstyle\sum_k \mathrm{Cov}\Big\{ f(s_k, \tau_k), g(s_k, \tau_k) \Big\} \\ &= \alpha(fg) + \textstyle\sum_k \Big\{ \nu_k(fg) - \nu_k(f)\nu_k(g) \Big\} \\ &\le \alpha(fg) + \textstyle\sum_k \nu_k(fg) = \nu(fg). \end{aligned}$$

□

The last lemma leads immediately to criteria for $\xi_n f_n \xrightarrow{P} 0$ or ∞, when the ξ_n are Poisson processes on S with arbitrary intensities $E\xi_n = \nu_n$. For later needs, we consider the more general case where the ξ_n are marked point processes with independent increments.

Theorem 3.26 *(positive integrals) Let $\xi, \xi_1, \xi_2, \ldots$ be T-marked point processes on S with independent S-increments and intensities $\nu, \nu_1, \nu_2, \ldots$. Then for any functions $f, f_1, f_2, \ldots \ge 0$ on $S \times T$,*

(i) $P\{\xi f < \infty\} = 1\{\nu(f \wedge 1) < \infty\}$,

(ii) $\xi_n f_n \xrightarrow{P} 0 \iff \nu_n(f_n \wedge 1) \to 0$,

(iii) $(\xi_n f_n)$ *is tight* $\iff \nu_n(t_n f_n \wedge 1) \to 0$ *as* $0 < t_n \to 0$,

(iv) $$\begin{aligned} \nu_n(f_n \wedge 1) \to \infty \quad &\Rightarrow \quad \xi_n f_n \xrightarrow{P} \infty \\ &\Rightarrow \quad \nu_n f_n \to \infty, \end{aligned}$$

(v) *the first implication in* (iv) *is reversible when*

$$\limsup_{n\to\infty} \Big(\|\Delta\nu_n\| \wedge \|\psi \circ f_n\| \Big) < 1.$$

The tightness condition in (iii) is clearly equivalent to a combination of the conditions

$$\sup_{n\ge 1} \nu_n(f_n \wedge 1) < \infty, \qquad \lim_{t\to\infty} \limsup_{n\to\infty} \nu_n\{f_n > t\} = 0.$$

A similar remark applies to part (iii) of Theorem 3.27 below.

Proof: It suffices to prove (ii) and (iv), since (i) and (iii) are easy consequences of (ii). Indeed, $\xi f < \infty$ a.s. iff $t\xi f \xrightarrow{P} 0$ as $t \to 0$, which holds by (ii) iff $\nu(tf \wedge 1) \to 0$, and this is in turn equivalent to $\nu(f \wedge 1) < \infty$ by dominated convergence. Now (i) follows by Kolmogorov's $0-1$ law. Similarly, the sequence $(\xi_n f_n)$ is tight iff $t_n \xi_n f_n \xrightarrow{P} 0$ whenever $t_n \to 0$, which holds by (ii) iff $\nu_n(t_n f_n \wedge 1) \to 0$.

Now consider the atomic decompositions

$$\nu_n = \alpha_n + \textstyle\sum_j (\delta_{s_{nj}} \otimes \beta_{nj}), \quad n \in \mathsf{N},$$

where α_n denotes the S-continuous component of ν_n. Writing $f_s = f(s,\cdot)$, we get by Lemmas 3.1 and 3.18 for every $n \in \mathbb{N}$

$$-\log Ee^{-\xi_n f_n} = \alpha_n(\psi \circ f_n) - \sum_k \log\Big(1 - \beta_k\{\psi \circ f_n(s_k)\}\Big). \qquad (13)$$

Noting that

$$1 \le -\frac{\log(1-x)}{x} \le -\frac{\log(1-p)}{p} \equiv C_p, \quad 1 < x \le p \le 1,$$

and assuming $\|\Delta\nu_n\| \le p \le 1$, we obtain

$$\begin{aligned} \nu_n(\psi \circ f_n) &\le -\log Ee^{-\xi_n f_n} \\ &\le C_p\, \nu_n(\psi \circ f_n), \quad n \in \mathbb{N}. \end{aligned} \qquad (14)$$

(ii) Decompose ξ_n into complementary $\frac{1}{2}$-thinnings ξ_n' and ξ_n'', and note that by symmetry $\xi_n f_n \xrightarrow{P} 0$ iff $\xi_n' f_n \xrightarrow{P} 0$. Since also $E\xi_n' = \frac{1}{2}\,\nu_n$, it is enough to prove the assertion with the ξ_n replaced by ξ_n'. It remains to notice that (14) applies to ξ_n' with $p = \frac{1}{2}$.

(iv) Use the lower bound in (14) and the fact that $E\xi_n f_n = \nu_n f_n$.

(v) We may assume that $\|\Delta\nu_n\| \wedge \|\psi \circ f_n\| \le p < 1$. Write $\nu_n = \nu_n' + \nu_n''$, where $\|\Delta\nu_n'\| \le p$, and the S-projection of ν_n'' is purely atomic with point masses $\|\Delta\nu_n''\| > p$. By the choice of p, we have $\psi \circ f_n < p$ on the support S_n of ν_n'', and so on S_n we may replace f_n by $f_n \wedge c$ with $c = -\log(1-p) < \infty$.

Now introduce the corresponding decomposition $\xi_n = \xi_n' + \xi_n''$. Using the independence of ξ_n' and ξ_n'', the upper bound in (14), and Jensen's inequality, we get

$$\begin{aligned} -\log Ee^{-\xi_n f_n} &= -\log Ee^{-\xi_n' f_n} - \log Ee^{-\xi_n'' f_n} \\ &\le C_p\, \nu_n'(\psi \circ f_n) + \nu_n''(f_n \wedge c) \\ &\lesssim \nu_n(f_n \wedge 1). \end{aligned}$$

Letting $n \to \infty$, we see that in this case $\xi_n f_n \xrightarrow{P} \infty$ implies $\nu_n(f_n \wedge 1) \to \infty$, as asserted. □

We turn to the corresponding criteria for symmetric point process integrals. Given a simple point process ξ with intensity ν, we may use Theorem 3.2 (i) to write $\xi = \xi' + \xi''$, where ξ' and ξ'' are complementary $\frac{1}{2}$-thinnings of ξ with intensity $\nu/2$. The symmetric version $\tilde{\xi}$ of ξ may then be defined as $\xi' - \xi''$. Equivalently, we may form $\tilde{\xi}$ by attaching independent, symmetric signs to the atoms of ξ.

The integral $\tilde{\xi} f$ is elementary when the support of f has finite ν-measure. In general, it may be defined by the condition $\tilde{\xi} f_n \xrightarrow{P} \tilde{\xi} f$, for any functions $f_n \to f$ with bounded supports and $|f_n| \le |f|$, whenever the limit exists and is a.s. independent of the choice of sequence (f_n).

Theorem 3.27 *(symmetric integrals) Let $\tilde\xi, \tilde\xi_1, \tilde\xi_2, \ldots$ be symmetric, T-marked point processes on S with independent S-increments and intensities $\nu, \nu_1, \nu_2, \ldots$. Then for any functions $f, f_1, f_2, \ldots$ on $S \times T$,*

(i) *$\tilde\xi f$ exists $\Leftrightarrow$ $\nu(f^2 \wedge 1) < \infty$, and when $S = \mathsf{R}_+$ we have*

$$P\{(f \cdot \tilde\xi)^* < \infty\} = 1\{\nu(f^2 \wedge 1) < \infty\},$$

(ii) $\tilde\xi_n f_n \xrightarrow{P} 0 \;\Leftrightarrow\; \nu_n(f_n^2 \wedge 1) \to 0,$

(iii) $(\tilde\xi_n f_n)$ *is tight* $\Leftrightarrow$ $\nu_n(t_n f_n^2 \wedge 1) \to 0$ *as* $0 < t_n \to 0,$

(iv)
$$\begin{aligned} \nu_n(f_n^2 \wedge 1) \to \infty \quad &\Rightarrow \quad |\tilde\xi_n f_n| \xrightarrow{P} \infty \\ &\Rightarrow \quad \nu_n f_n^2 \to \infty, \end{aligned}$$

(v) *the first implication in* (iv) *is reversible when*

$$\limsup_{n\to\infty} \Big(\|\Delta\nu_n\| \wedge \|\psi \circ f_n^2\| \Big) < 1. \tag{15}$$

Proof: Our proof of (i) will be based on criterion (ii′), defined as the restriction of (ii) to functions f_n with bounded support. Once (i) is established, the remaining statements make sense for arbitrary f_n, and (ii) will follow by the same proof as (ii′).

(ii′)−(ii): For any triangular array of independent, symmetric random variables α_{nj}, we have by FMP 5.11

$$\sum\nolimits_j \alpha_{nj} \xrightarrow{P} 0 \quad \Leftrightarrow \quad \sum\nolimits_j \alpha_{nj}^2 \xrightarrow{P} 0. \tag{16}$$

Applying this to the integrals $\tilde\xi_n f_n$ gives

$$E\Big(|\tilde\xi_n f_n| \wedge 1 \,\Big|\, \xi_n \Big) \to 0 \;\text{ a.s.} \quad \Leftrightarrow \quad \xi_n f_n^2 \to 0 \;\text{ a.s.}$$

Using a sub-sequence argument, dominated convergence, and Theorem 3.26 (ii), we obtain

$$\begin{aligned} \tilde\xi_n f_n \xrightarrow{P} 0 \quad &\Leftrightarrow \quad \xi_n f_n^2 \xrightarrow{P} 0 \\ &\Leftrightarrow \quad E\xi_n(f_n^2 \wedge 1) \to 0. \end{aligned}$$

(i) If $\nu(f^2 \wedge 1) < \infty$, we may choose some functions $f_n \to f$ with bounded supports and $|f_n| \le |f|$. Since $|f_m - f_n| \le 2|f|$, we have $(f_m - f_n)^2 \wedge 1 \le 4(f^2 \wedge 1)$, and so by dominated convergence $\nu\{(f_m - f_n)^2 \wedge 1\} \to 0$, which implies $\tilde\xi f_m - \tilde\xi f_n \xrightarrow{P} 0$ by (ii′), so that $\tilde\xi f_n \xrightarrow{P}$ some γ by completeness. If also $f_n' \to f$ with $|f_n'| \le |f|$, we may apply the same argument to the alternating sequence $f_1, f_1', f_2, f_2', \ldots$, to see that the limit γ is a.s. independent of the choice of sequence (f_n). This proves the existence of the integral $\tilde\xi f$.

Conversely, assume $\nu(f^2 \wedge 1) = \infty$. Choosing functions $f_n \to f$ with bounded supports and $|f_n| \le |f|$ such that $\nu(f_n^2 \wedge 1) \to \infty$, we get

$$\nu(f_n^2 \wedge 1) \le 4\nu(f_m^2 \wedge 1) + 4\nu\Big\{(f_m - f_n)^2 \wedge 1\Big\},$$

and so $\nu\{(f_m - f_n)^2 \wedge 1\} \to \infty$ as $n \to \infty$ for fixed m, which implies $\nu\{(f_m - f_n)^2 \wedge 1\} \not\to 0$ as $m, n \to \infty$. Then $\tilde\xi f_m - \tilde\xi f_n \overset{P}{\not\to} 0$ by (ii′), and so $\tilde\xi f_n$ fails to converge in probability.

(iii) This follows from (ii), since the sequence $(\tilde\xi_n f_n)$ is tight iff $t_n\tilde\xi_n f_n \xrightarrow{P} 0$ for every sequence $t_n \to 0$.

(iv)−(v): By Theorem 3.26 it is enough to prove that $|\tilde\xi_n f_n| \xrightarrow{P} \infty$ iff $\xi_n f_n^2 \xrightarrow{P} \infty$. Assuming the latter condition, we get by Theorem A1.6 for any $t > 0$

$$P\Big(\Big|\tilde\xi_n f_n| \le t \,\Big|\, \xi_n\Big) \lesssim t\,(\xi_n f_n^2)^{-1/2} \xrightarrow{P} 0.$$

Hence, $P\{|\tilde\xi_n f_n| \le t\} \to 0$ by dominated convergence, and $|\tilde\xi_n f_n| \xrightarrow{P} \infty$ follows since t was arbitrary.

Conversely, let $|\tilde\xi_n f_n| \xrightarrow{P} \infty$. Then for any sub-sequence $N' \subset \mathsf{N}$, we have $|\tilde\xi_n f_n| \to \infty$ a.s. along a further sub-sequence N'', and so

$$\xi_n f_n^2 = E\Big\{(\tilde\xi_n f_n)^2 \Big|\, \xi_n\Big\} \to \infty \;\text{ a.s. along } N''.$$

Since N' was arbitrary, it follows that $\xi_n f_n^2 \xrightarrow{P} \infty$. □

Given a marked point process ξ with independent increments and intensity ν, we may next consider the compensated integrals $(\xi - \nu)f$, defined by the condition $(\xi - \nu)f_n \xrightarrow{P} (\xi - \nu)f$, for any bounded functions $f_n \to f$ with bounded supports and $|f_n| \le |f|$, whenever the limit exists and is a.s. independent of approximating sequence (f_n). To avoid near cancellations, we need to assume that

$$\sup_{s\in S} \nu\{s\} < 1, \qquad \sup_{n\ge 1}\sup_{s\in S} \nu_n\{s\} < 1. \tag{17}$$

Theorem 3.28 *(compensated integrals) Let $\xi, \xi_1, \xi_2, \ldots$ be simple point processes on S with independent increments and intensities $\nu, \nu_1, \nu_2, \ldots$ satisfying* (17)*. Then for any functions $f, f_1, f_2, \ldots \ge 0$ on S,*

(i) *$(\xi-\nu)f$ exists $\Leftrightarrow$ $\nu(f^2 \wedge f) < \infty$, and when $S = \mathsf{R}_+$ and f is bounded, we have*

$$P\Big\{(f\cdot(\xi - \nu))^* < \infty\Big\} = 1\{\nu f^2 < \infty\},$$

(ii) $(\xi_n - \nu_n)f_n \xrightarrow{P} 0 \;\Leftrightarrow\; \nu_n(f_n^2 \wedge f_n) \to 0,$

(iii) $\{(\xi_n - \nu_n)f_n\}$ *is tight* $\Leftrightarrow\; \limsup_n \nu_n(f_n^2 \wedge f_n) < \infty,$

(iv)
$$\begin{aligned} \nu_n(f_n^2 \wedge 1) \to \infty \quad &\Rightarrow \quad |(\xi_n - \nu_n)f_n| \xrightarrow{P} \infty \\ &\Rightarrow \quad \nu_n(f_n^2 \wedge f_n) \to \infty. \end{aligned}$$

The implications in (iv) *are irreversible, even when the ξ_n are Poisson.*

Proof: (i) Write $f = g + h$, where $g = f1\{f \le 1\}$ and $h = f1\{f > 1\}$. First assume $\nu(f^2 \wedge f) < \infty$, so that $\nu g^2 < \infty$ and $\nu h < \infty$. Since $E\{(\xi - \nu)f\}^2 = \nu f^2$, we see as in the previous proof that $(\xi - \nu)g$ exists. Furthermore, $(\xi - \nu)h = \xi h - \nu h$ exists by Theorem 3.26 (i).

Conversely, suppose that $(\xi - \nu)f$ exists. Taking differences, we see that even $(\xi - \xi')f$ exists, and so $\nu(f^2 \wedge 1) < \infty$ by Theorem 3.27 (i). Hence, ξh is just a finite sum, and so even $\nu h = \xi h - (\xi - \nu)h$ exists. This gives $\nu(f^2 \wedge f) \le \nu(f^2 \wedge 1) + \nu h < \infty$.

(ii) Assume $\nu_n(f_n^2 \wedge f_n) \to 0$, and write $f_n = g_n + h_n$ as before. Then $E\{(\xi_n - \nu_n)g_n\}^2 = \nu_n g_n^2 \to 0$, and so $(\xi_n - \nu_n)g_n \xrightarrow{P} 0$. Furthermore, $E\xi_n h_n = \nu_n h_n \to 0$, and so $(\xi_n - \nu_n)h_n = \xi_n h_n - \nu_n h_n \xrightarrow{P} 0$. Hence, by combination, $\xi_n f_n \xrightarrow{P} 0$.

Conversely, let $(\xi_n - \nu_n)f_n \xrightarrow{P} 0$. Then even $(\xi_n - \xi_n')f_n \xrightarrow{P} 0$, and so $\nu_n(f_n^2 \wedge 1) \to 0$ by Theorem 3.27 (ii). Here $(\xi_n - \nu_n)g_n \xrightarrow{P} 0$ as before, and so even $(\xi_n - \nu_n)h_n \xrightarrow{P} 0$. Since also $\nu_n\{h_n > 0\} \to 0$, we get $\xi_n\{h_n > 0\} \xrightarrow{P} 0$, which implies $\xi_n h_n \xrightarrow{P} 0$. Then $\nu_n h_n = \xi_n h_n - (\xi_n - \nu_n)h_n \xrightarrow{P} 0$, and so by combination $\nu_n(f_n^2 \wedge f_n) \to 0$.

(iii) Recall that $\{(\xi_n - \nu_n)f_n\}$ is tight iff $t_n(\xi_n - \nu_n)f_n \xrightarrow{P} 0$ as $t_n \to 0$. By (ii) this is equivalent to $\nu_n(t_n^2 f_n^2 \wedge t_n f_n) \to 0$ as $t_n \downarrow 0$, which holds iff $\limsup_n \nu_n(f_n^2 \wedge f_n) < \infty$.

(iv) If $\nu_n(f_n^2 \wedge 1) \to \infty$, then Theorem 3.26 (iv) yields $\xi_n f_n^2 \xrightarrow{P} \infty$, and so as before $|(\xi_n - \nu_n)f_n| \xrightarrow{P} \infty$ by Theorem A1.6. If instead $\nu_n(f_n^2 \wedge f) \not\to \infty$, then $\nu_n(f_n^2 \wedge f)$ is bounded along a sub-sequence N', and so by (iii) $\{(\xi_n - \nu_n)f_n\}$ is tight along N', which excludes the possibility of $(\xi_n - \nu_n)f_n \xrightarrow{P} \infty$.

The last assertion requires two counter-examples. Here we take $\xi_n \equiv \xi$ to be Poisson on S with intensity ν. First let $\|\nu\| = 1$ and $1 \le f_n \equiv c_n \to \infty$, so that $\nu(f_n^2 \wedge f_n) = c_n \to \infty$. Then the random variable ξS is Poisson with mean 1, and so

$$P\{(\xi - \nu)f_n = 0\} = P\{\xi S = 1\} = e^{-1} > 0, \quad n \in \mathsf{N},$$

which implies $|(\xi - \nu)f_n| \stackrel{P}{\not\to} \infty$. This shows that the second implication in (iv) is not reversible.

Next let ν be Lebesgue measure on $S = [0, 1]$, and choose $f_n(x) \equiv 2nx$ for $x \in S$ and $n \in \mathsf{N}$. Then Theorem 3.4 yields $\xi f_n = 2n \sum_{j \le \kappa} \vartheta_j$, where $\vartheta_1, \vartheta_2, \dots$ are i.i.d. $U(0, 1)$ and $\kappa = \xi S$ is an independent Poisson random variable with mean 1. Since also $\nu f_n = n$, we have $|(\xi - \nu)f_n| = n \to \infty$ on $\{\kappa = 0\}$, whereas for fixed $k, t > 0$

$$\begin{aligned} P\Big(|(\xi - \nu)f_n| \le t \,\Big|\, \kappa = k\Big) &\le \sup_r P\Big(|\xi f_n - r| \le t \,\Big|\, \kappa = k\Big) \\ &\le \sup_r P\{\,|2n\vartheta_1 - r| \le t\} \\ &= P\Big\{\,|\vartheta_1 - \tfrac{1}{2}| \le t/2n\Big\} \le t/n \to 0. \end{aligned}$$

Thus, $|(\xi-\nu)f_n| \xrightarrow{P} \infty$. Since also $\nu(f_n^2 \wedge 1) \le 1 < \infty$, the first implication in (iv) is not reversible either. □

We also consider an intermediate case between the symmetric and compensated integrals in Theorems 3.27 and 3.28.

Corollary 3.29 *(centered integrals) Let $\xi^\pm, \xi_1^\pm, \xi_2^\pm, \ldots$ be pairwise independent, simple point processes on S with independent increments and intensities $\nu^\pm, \nu_1^\pm, \nu_2^\pm, \ldots$, satisfying conditions as in* (17). *Put $\xi = \xi^+ - \xi^-$, $\nu = \nu^+ - \nu^-$, and $|\nu| = \nu^+ + \nu^-$, and similarly for ξ_n, ν_n, and $|\nu_n|$. Then for any uniformly bounded functions $f, f_1, f_2, \ldots$ on S,*

(i) $(\xi-\nu)f$ *exists* $\Leftrightarrow$ $|\nu|f^2 < \infty$, *and when $S = \mathsf{R}_+$ we have*

$$P\Big\{(f\cdot(\xi-\nu))^* < \infty\Big\} = 1\{|\nu|f^2 < \infty\},$$

(ii) $(\xi_n-\nu_n)f_n \xrightarrow{P} 0 \;\Leftrightarrow\; |\nu_n|f_n^2 \to 0$,

(iii) $\{(\xi_n-\nu_n)f_n\}$ *is tight* $\Leftrightarrow$ $\limsup_n |\nu_n|f_n^2 < \infty$,

(iv) $|(\xi_n-\nu_n)f_n| \xrightarrow{P} \infty \;\Leftrightarrow\; |\nu_n|f_n^2 \to \infty$.

Proof: (ii) If $(\xi_n-\nu_n)f_n \xrightarrow{P} 0$, then $\tilde\xi_n f_n \xrightarrow{P} 0$ by subtraction, and so $|\nu_n|f_n^2 \to 0$ by Theorem 3.27 (ii). Conversely, $|\nu_n|f_n^2 \to 0$ yields $\nu_n^\pm f_n^2 \to 0$, and since the f_n are uniformly bounded, we get $(\xi_n^\pm - \nu_n^\pm)f_n \xrightarrow{P} 0$ by Theorem 3.28 (ii), which implies $(\xi_n-\nu_n)f_n \xrightarrow{P} 0$ by subtraction.

(i) & (iii): Use (ii) as before.

(iv) If $|\nu_n|f_n^2 \to \infty$, then $|\xi_n|f_n^2 \xrightarrow{P} \infty$ by Theorem 3.26 (iv), and so $|(\xi_n-\nu_n)f_n| \xrightarrow{P} \infty$ by Theorem A1.6. Conversely, $|\nu_n|f_n^2 \not\to \infty$ implies that $|\nu_n|f_n^2$ is bounded along a sub-sequence $N' \subset \mathsf{N}$. Then by (iii) the sequence $\{(\xi_n-\nu_n)f_n\}$ is tight along N', which contradicts $|(\xi_n-\nu_n)f_n| \xrightarrow{P} \infty$. □

The independent-increment processes of the previous results can be replaced by Poisson processes. In particular, this yields a simple comparison between random sums and Poisson integrals.

Corollary 3.30 *(Poisson reduction) The criteria in Theorems 3.26–3.29 remain valid for Poisson processes $\xi, \xi_1, \xi_2, \ldots$ on S with arbitrary intensities $\nu, \nu_1, \nu_2, \ldots$.*

Proof: In proving the Poisson cases of the stated results, we never needed the diffuseness of $\nu, \nu_1, \nu_2, \ldots$. □

For later needs, we show how, in the previous results, we may add some independent Gaussian random variables to the symmetric or compensated point process integrals $\tilde\xi_n f_n$ or $(\xi_n - \nu_n)f_n$. We state the resulting extensions only for Theorem 3.27 (ii) and (iv), the remaining cases being similar.

Corollary 3.31 *(Gaussian extensions) For $n \in \mathbb{N}$, let $\tilde{\xi}_n$ be symmetric, T-marked point processes on S with independent S-increments and intensities ν_n, and let γ be an independent $N(0,1)$ random variable. Then for any functions f_n on $S \times T$ and constants $\sigma_n \in \mathbb{R}$,*

(i) $\tilde{\xi}_n f_n + \sigma_n \gamma \xrightarrow{P} 0 \;\Leftrightarrow\; \nu_n(f_n^2 \wedge 1) + \sigma_n^2 \to 0$,

(ii) $\nu_n(f_n^2 \wedge 1) + \sigma_n^2 \to \infty \;\Rightarrow\; |\tilde{\xi}_n f_n + \sigma_n \gamma| \xrightarrow{P} \infty$,
with equivalence under condition (15).

Similar extensions hold for Theorems 3.28 and 3.29.

Proof: (i) Using the equivalence (16), we get

$$\tilde{\xi}_n f_n + \sigma_n \gamma \xrightarrow{P} 0 \quad \Leftrightarrow \quad \tilde{\xi}_n f_n \xrightarrow{P} 0, \;\; \sigma_n \gamma \xrightarrow{P} 0.$$

Since $\sigma_n \gamma \xrightarrow{P} 0$ iff $\sigma_n^2 \to 0$, this reduces the assertion to the previous case.

(ii) Assuming $\nu_n(f_n^2 \wedge 1) + \sigma_n^2 \to \infty$, we see from Theorem 3.27 that $\xi_n f_n^2 + \sigma_n^2 \gamma^2 \xrightarrow{P} \infty$, and so by Theorem A1.6 we have for any $t > 0$

$$P\Big(|\tilde{\xi}_n f_n + \sigma_n \gamma| \le t \,\Big|\, \xi, \gamma^2 \Big) \lesssim t \Big(\xi_n f_n^2 + \sigma_n^2 \Big)^{-1/2} \xrightarrow{P} 0,$$

which implies $|\tilde{\xi}_n f_n + \sigma_n \gamma| \xrightarrow{P} \infty$.

Conversely, let $|\tilde{\xi}_n f_n + \sigma_n \gamma| \xrightarrow{P} \infty$. Turning to a.s. convergent sub-sequences, we get

$$\xi_n f_n^2 + \sigma_n^2 \gamma^2 = E\Big\{ (\tilde{\xi}_n f_n + \sigma_n \gamma)^2 \Big|\, \xi_n, \gamma^2 \Big\} \to \infty \;\text{ a.s.} \tag{18}$$

To prove that $\nu_n(f_n^2 \wedge 1) + \sigma_n^2 \to \infty$, assume the opposite. Then σ_n^2 is bounded along a further sub-sequence $N' \subset \mathbb{N}$, and so $\sigma_n^2 \gamma^2$ is a.s. bounded as well. Then (18) yields $\xi_n f_n^2 \to \infty$ a.s. along N', and so as before $\nu_n(f_n^2 \wedge 1) \to \infty$ along N', a contradiction proving our claim. □

For certain purposes in Chapter 9, we need to strengthen the previous results to criteria for uniform boundedness and convergence, in the special case of processes on $\mathbb{R}_+$.

Corollary 3.32 *(uniform boundedness and convergence) Let $X, X_1, X_2, \ldots$ be centered processes with independent increments and symmetric, positive, or uniformly bounded jumps. Then*

(i) $X^* < \infty$ *a.s.* $\Leftrightarrow$ $|X_\infty| < \infty$ *a.s.*,

(ii) $X_n^* \xrightarrow{P} 0 \;\Leftrightarrow\; X_n(\infty) \xrightarrow{P} 0$.

Proof: (i) Write $h(x) \equiv x$, let σ^2 be the Gaussian variance of X_∞, and put $\nu = E\xi_\infty$, where ξ denotes the jump component of X. Using Theorems 3.26 – 3.29 along with Corollary 3.31, we see that the a.s. condition $|X_\infty| < \infty$ is equivalent, in the three cases, to respectively

(B_1) $\sigma^2 + \nu(h^2 \wedge 1) < \infty$,

(B_2) $\sigma^2 + \nu(h^2 \wedge h) < \infty$,

(B_3) $\sigma^2 + \nu h^2 < \infty$.

By Lemma 3.25, we conclude from (B_3) that X is an L^2-bounded martingale, in which case $X^* < \infty$ a.s. by Doob's maximum inequality. Subtracting the Gaussian component, along with all compensated jumps smaller than 1, we see from (B_2) that the remaining compensated jumps form a process of integrable variation, which implies $X^* < \infty$ a.s. Similarly, (B_1) implies that X has a.s. finitely many jumps of size > 1. By the symmetry in this case, the corresponding component of X is simply the sum of those large jumps, and the a.s. condition $X^* < \infty$ is trivially fulfilled.

(ii) As before, we note that $X_n(\infty) \xrightarrow{P} 0$ is equivalent, in the three cases, to respectively

(C_1) $\sigma_n^2 + \nu_n(h^2 \wedge 1) \to 0$,

(C_2) $\sigma_n^2 + \nu_n(h^2 \wedge h) \to 0$,

(C_3) $\sigma_n^2 + \nu_n h^2 \to 0$.

Here (C_3) yields $E(X_n^*)^2 \to 0$ by Doob's inequality, which implies $X_n^* \xrightarrow{P} 0$. Subtracting the Gaussian component from X_n, along with the sum of compensated jumps ≤ 1, we see from (C_2) that the total variation of the remaining process tends to 0 in L^1, which again implies $X_n^* \xrightarrow{P} 0$. Finally, (C_1) shows that the expected number of jumps > 1 tends to 0. Since there is no compensation in this case, we have again $X_n^* \xrightarrow{P} 0$. □

Finally, we show how the previous results lead to similar statements for thinnings of general point processes. This will be useful in Chapter 10.

Corollary 3.33 *(thinning equivalence) Let $\xi_1, \xi_2, \ldots$ be arbitrary point processes on S with symmetrizations $\tilde{\xi}_n$ and associated p-thinnings ξ_n^p and $\tilde{\xi}_n^p$. Then for any functions f_n on S and constants $p, q \in (0,1)$, all these conditions are equivalent:*

(i) $\xi_n^p f_n^2 \xrightarrow{P} \infty$, (i′) $\xi_n^q f_n^2 \xrightarrow{P} \infty$,

(ii) $|\tilde{\xi}_n^p f_n| \xrightarrow{P} \infty$, (ii′) $|\tilde{\xi}_n^q f_n| \xrightarrow{P} \infty$.

Proof: First assume (i). By a sub-sequence argument, we may assume that $\xi_n^p f_n^2 \to \infty$ a.s., and by conditioning we may take the ξ_n to be non-random. Turning to uniform randomizations, if necessary, we may further assume the ξ_n to be simple. Then Theorem 3.26 yields $\xi_n(f_n^2 \wedge 1) \to \infty$, which implies (i′), (ii), and (ii′) by Theorems 3.26 and 3.27. The remaining proofs are similar. □

3.5 Symmetric Sequences and Processes

Here we consider some symmetry properties of Poisson and related processes, and derive the general representations of exchangeable random measures. Given a simple point process ξ on $[0,1]$ with atoms at $\tau_1 < \cdots < \tau_n$, we define the associated *spacing variables* $\gamma_1, \ldots, \gamma_{n+1}$ by $\gamma_k = \tau_k - \tau_{k-1}$, $1 \le k \le n+1$, where $\tau_0 = 0$ and $\tau_{n+1} = 1$. Similarly, for a simple point process on R_+ with points $\tau_1 < \tau_2 < \cdots$, we define the spacing variables $\gamma_1, \gamma_2, \ldots$ by $\gamma_k = \tau_k - \tau_{k-1}$, $k \ge 1$, with $\tau_0 = 0$.

Recall that a random sequence is said to be *exchangeable* if all permutations have the same distribution, and *contractable*[4] if the same property holds for all sub-sequences. For infinite sequences (but not in general!), the two notions are known to be equivalent.

Theorem 3.34 *(spacing symmetries)*

(i) *Let ξ be a binomial process on* $[0,1]$ *based on λ and n. Then the spacing variables $\gamma_1, \ldots, \gamma_{n+1}$ of ξ are exchangeable.*

(ii) *Let ξ be a simple point process on R_+ with $\|\xi\| = \infty$ a.s. Then ξ is a mixed Poisson process based on λ, iff it is stationary with contractable spacing variables $\gamma_1, \gamma_2, \ldots$.*

Proof: (i) It is enough to prove the invariance under transpositions of adjacent intervals. Applying Theorem 3.11 (i) twice, we may reduce to the case where $\|\xi\| = 1$. But then $\xi = \delta_\tau$ where τ is $U(0,1)$, and the desired invariance holds since $1 - \tau \stackrel{d}{=} \tau$.

(ii) The necessity follows from (i) and Theorem 3.4 by an elementary conditioning argument. Now let ξ be stationary with contractable spacing variables $\gamma_1, \gamma_2, \ldots$. By the de Finetti/Ryll-Nardzewski theorem (FMP 11.10), the γ_k are then conditionally i.i.d., given the invariant σ-field $\mathcal{I}_\xi$. Since conditioning on $\mathcal{I}_\xi$ preserves the stationarity of ξ, we may henceforth take ξ to be a stationary renewal process. Writing $f(t) = P\{\gamma_1 > t\}$ and $c = (E\gamma_1)^{-1}$, we see from FMP 9.18 that $f'(t) = -cf(t)$ for $t > 0$, and since clearly $f(0+) = 1$, we get $f(t) = e^{-ct}$. Hence, Theorem 3.12 shows that ξ is a Poisson process based on $c\lambda$. □

We proceed to characterize exchangeable and related random measures ξ on a Borel space S. Given a diffuse measure $\lambda \in \mathcal{M}_S$, we say that ξ is *λ-symmetric*, if[5] $\xi \circ f^{-1} \stackrel{d}{=} \xi$ for any measurable function $f: S \to S$ with $\lambda \circ f^{-1} = \lambda$. Similarly, when λ is Lebesgue measure on $S = [0,1]$ or R_+, we define the contraction maps $f_B : B \to [0, \lambda B]$ by $f_B(t) = \lambda([0,t] \cap B)$, and say that ξ is *contractable* if $\xi \circ f_B^{-1} \stackrel{d}{=} 1_{[0,\lambda B]}\xi$ for every $B \in \hat{\mathcal{S}}$. In both cases, it is enough to consider countable classes of transformations.

[4] short for *contraction invariant in distribution*

[5] Recall that the measure $\mu \circ f^{-1}$ is given by $(\mu \circ f^{-1})B = \mu(f^{-1}B)$ or $(\mu \circ f^{-1})g = \mu(g \circ f)$.

Theorem 3.35 *(simple and diffuse symmetry) Fix any diffuse measure $\lambda \in \mathcal{M}_S$, where S is Borel.*

(i) *A simple point process on S is λ-symmetric, iff it is a mixed Poisson or binomial process based on λ and some random variable ρ in R_+ or κ in Z_+, respectively.*

(ii) *A diffuse random measure on S is λ-symmetric, iff $\xi = \alpha\lambda$ a.s. for some random variable $\alpha \geq 0$.*

When $S = \mathsf{R}_+$ or $[0,1]$ with Lebesgue measure λ, it is equivalent that ξ be contractable.

Proof: The necessity is obvious. Conversely, when ξ is contractable, we get in both cases

$$Ee^{-\xi B} = \varphi(\lambda B), \quad B \in \mathcal{B},$$

for some function $\varphi\colon S \to [0,1]$. If $\eta = \xi \circ f^{-1}$ for some λ-preserving function f on S, then for any $B \in \mathcal{B}$, we have

$$\begin{aligned} Ee^{-\eta B} &= Ee^{-\xi \circ f^{-1} B} \\ &= \varphi(\lambda \circ f^{-1} B) \\ &= \varphi(\lambda B) = Ee^{-\xi B}, \end{aligned}$$

and so $\xi \stackrel{d}{=} \eta$ by Theorem 3.8 (ii), which shows that ξ is even exchangeable.

(i) When $S = [0,1]$, the variable $\|\xi\|$ is invariant under λ-preserving transformations, and so ξ remains conditionally exchangeable, given $\|\xi\|$. We may then reduce by conditioning to the case where $\|\xi\| = n$ is a constant, so that $\xi = \sum_{k \leq n} \delta_{\sigma_k}$ for some random variables $\sigma_1 < \cdots < \sigma_n$ in $[0,1]$. Now consider any sub-intervals $I_1 < \cdots < I_n$ of $[0,1]$, where $I < J$ means that $s < t$ for all $s \in I$ and $t \in J$. By contractability, the probability

$$P \bigcap\nolimits_k \{\sigma_k \in I_k\} = P \bigcap\nolimits_k \{\xi I_k > 0\}$$

is invariant under individual shifts of $I_1, \ldots, I_n$, subject to the mentioned restriction. Thus, $\mathcal{L}(\sigma_1, \ldots, \sigma_n)$ is shift invariant on the set $\Delta_n = \{s_1 < \cdots < s_n\} \subset [0,1]^n$ and hence equals $n!\,\lambda^n$ there, which means that ξ is a binomial process on $[0,1]$ based on λ and n.

When $S = \mathsf{R}_+$, the restrictions $1_{[0,t]}\xi$ are mixed binomial processes based on $1_{[0,t]}\lambda$, and so ξ is mixed Poisson on R_+ by Theorem 3.7.

(ii) It is enough to take $S = [0,1]$. As in (i), we may next reduce by conditioning to the case where $\|\xi\| = 1$ a.s. Then by contractability we have $E\xi = \lambda$, and for any intervals $I_1 < I_2$ in I, the moment $E\xi^2(I_1 \times I_2)$ is invariant under individual shifts of I_1 and I_2, so that $E\xi^2$ is proportional to λ^2 on Δ_2. Since ξ is diffuse, we have also $E\xi^2 = 0$ on the diagonal $\{s_1 = s_2\}$ in $[0,1]^2$. Hence, by symmetry and normalization, $E\xi^2 = \lambda^2$ on $[0,1]^2$, and so

$$\begin{aligned}\mathrm{Var}(\xi B) &= E(\xi B)^2 - (E\xi B)^2\\ &= \lambda^2 B^2 - (\lambda B)^2 = 0, \quad B \in \mathcal{B}_{[0,1]}.\end{aligned}$$

This gives $\xi B = \lambda B$ a.s. for every B, and so $\xi = \lambda$ a.s. by Lemma 2.1. □

We turn to the case of marked point processes.

Theorem 3.36 *(marked symmetry) Let ξ be a T-marked point process on S, and fix any $\lambda \in \mathcal{M}_S^*$ with $\|\lambda\| = 1$ or ∞. Then ξ is λ-symmetric iff*

(i) *for $\|\lambda\| = 1$, it is a λ-randomization of some point process β on T,*

(ii) *for $\|\lambda\| = \infty$, it is a Cox process directed by $\lambda \otimes \nu$, for some random measure ν on T.*

Proof: The necessity is again obvious. Now let ξ be λ-symmetric. Since S is Borel, we may assume that $S = [0,1]$ or R_+. By a simple transformation, we may further reduce to the case where λ equals Lebesgue measure on S.

(i) Here we may take λ to be Lebesgue measure on $I = [0,1]$. Since $\beta = \xi(I \times \cdot)$ is invariant under λ-preserving transformations of I, we may reduce by conditioning to the case where β is non-random. It suffices to derive the stated representation on $I \times C$ for any $C \in \hat{\mathcal{T}}$, and so we may assume that $\|\beta\| < \infty$. Now we have

$$\beta = \sum_{i \le n} \kappa_i \, \delta_{\tau_i}, \qquad \xi = \sum_{i \le n} \sum_{j \le \kappa_i} \delta_{\sigma_{ij}, \tau_i},$$

for some constants $\kappa_1, \ldots, \kappa_n \in \mathsf{N}$ and distinct elements $\tau_1, \ldots, \tau_n \in T$, along with some random variables σ_{ij} with $\sigma_{i,1} < \cdots < \sigma_{i,\kappa_i}$ for each i. For any disjoint intervals $I_{ij} \subset I$, $j \le \kappa_i$, $i \le n$, such that $I_{i,1} < \cdots < I_{i,\kappa_i}$ for each i, the probability

$$P \bigcap_{i,j} \{\sigma_{ij} \in I_{ij}\} = P \bigcap_{i,j} \Big\{ \xi(I_{ij} \times \{\tau_i\}) > 0 \Big\}$$

is invariant under individual shifts of the intervals I_{ij}, subject to the stated restrictions. The joint distribution of the σ_{ij} is then shift invariant on the non-diagonal part of $\Delta_{\kappa_1} \times \cdots \times \Delta_{\kappa_n}$, and is therefore proportional to Lebesgue measure there. Since the σ_{ij} are a.s. distinct, the contributions to the diagonal spaces vanish. Thus, the point processes $\xi_i = \xi(\cdot \times \{\tau_i\})$ are independent binomial processes based on λ and κ_i, which means that ξ is a λ-randomization of β.

(ii) Here we may take λ to be Lebesgue measure on R_+. Fixing a dissection system (I_{nj}) in T, we note that the projections $\xi_{nj} = \xi(\cdot \times I_{nj})$ are exchangeable, simple point processes on R_+. Hence, by Theorem 3.35 (ii) they are mixed Poisson processes based on λ and some random variables $\rho_{nj} \ge 0$. By the law of large numbers, the latter are measurable with respect to the tail σ-field in R_+, and so the exchangeability is preserved under conditioning on the ρ_{nj}, which may then be assumed to be non-random.

By a simple limiting argument based on (i), we note that for any fixed n and j, the process ξ_{nj} remains conditionally exchangeable, given the family $\{\xi_{ni};\, i \neq j\}$. Under the same conditioning, it then remains a Poisson process with intensity $\rho_{nj}\lambda$, which shows that the ξ_{nj} are independent in j for fixed n. Letting $\mathcal{U}$ denote the ring in $S \times T$ generated by the measurable rectangles $B \times I_{nj}$ with $B \in \hat{\mathcal{S}}$, we conclude that the variable ξU is Poisson distributed for every $U \in \mathcal{U}$. Since $\mathcal{U}$ generates $\mathcal{S} \otimes \mathcal{T}$, Corollary 3.9 shows that ξ is a Poisson process on $S \times T$. The λ-invariance of $E\xi$ is clear from the exchangeability of ξ, and so $E\xi = \lambda \otimes \nu$ for some $\nu \in \mathcal{M}_T$ by FMP 2.6. □

Using the previous results, we may easily derive a representation of general λ-symmetric random measures on $S \times T$.

Theorem 3.37 *(symmetric random measures) Fix any $\lambda \in \mathcal{M}_S^*$ with $\|\lambda\| = 1$ or ∞. Then a random measure ξ on $S \times T$ is λ-symmetric iff*

$$\xi = \lambda \otimes \alpha + \iint (\delta_s \otimes \mu)\, \eta(ds\, d\mu) \quad a.s.,$$

where α is a random measure on T, and

(i) *for $\|\lambda\| = 1$, η is a λ-randomization of some point process β on $\mathcal{M}_T \setminus \{0\}$ with $\eta \perp\!\!\!\perp_\beta \alpha$ and $\int \beta(d\mu)\, \mu B < \infty$ a.s. for all $B \in \hat{\mathcal{T}}$,*

(ii) *for $\|\lambda\| = \infty$, η is a Cox process on $S \times \mathcal{M}_T$ directed by $\lambda \otimes \nu$, for some random measure ν on $\mathcal{M}_T \setminus \{0\}$, such that $\eta \perp\!\!\!\perp_\nu \alpha$ and $\int (\mu B \wedge 1)\, \nu(d\mu) < \infty$ a.s. for all $B \in \hat{\mathcal{T}}$.*

The distribution of (α, β) or (α, ν) is then unique, and any such pair with the stated properties may occur.

Proof: First let ξ be λ-symmetric with $\xi(\{s\} \times \cdot) \equiv 0$ a.s., where $\|\lambda\| = 1$. Writing $\alpha = \xi(S \times \cdot)$, we see from Theorem 3.35 that $\xi(\cdot \times C) = \alpha(C)\lambda$ a.s. for every $C \in \hat{\mathcal{T}}$, and so $\xi = \lambda \otimes \alpha$ a.s. by Lemma 2.1. The result extends immediately to the case where $\|\lambda\| = \infty$.

As in the proof of Theorem 3.19, we may write

$$\xi = \xi_c + \iint (\delta_s \otimes \mu)\, \eta(ds\, d\mu) \quad \text{a.s.},$$

for some point process η on $S \times (\mathcal{M}_T \setminus \{0\})$ and a random measure ξ_c on $S \times T$ with $\xi_c(\{s\} \times \cdot) \equiv 0$. By the uniqueness of the representation, η and ξ_c inherit the exchangeability from ξ. As before, we get $\xi_c = \lambda \otimes \alpha$ a.s. for some random measure α on T, and Lemma 3.36 shows that η is a λ-randomization of some point process β on $\mathcal{M}_T \setminus \{0\}$, respectively a Cox process directed by $\lambda \otimes \nu$ for some random measure ν on $\mathcal{M}_T \setminus \{0\}$. Since α and β or ν are invariant under any λ-preserving transformations of ξ, the distributional properties of η remain conditionally true, given (α, β) or (α, ν), respectively. Since the conditional distribution depends only on β or ν, the asserted conditional independence follows. □

The point process version of Theorem 3.36 leads to an interesting extension of Theorem 3.7, which may also be proved by a simple but less intuitive compactness argument.

Corollary 3.38 *(Cox processes and thinnings, Mecke) A point process ξ on S is Cox, iff for every $p \in (0,1)$ it is a p-thinning of some point process ξ_p on S. In that case, the processes ξ_p are again Cox, and if ξ is directed by the random measure η on S, then the ξ_p are directed by the measures $p^{-1}\eta$.*

Proof: Assume the stated property. For every $n \in \mathsf{N}$, let ξ be an n^{-1}-thinning of β_n. Then introduce a λ/n-randomization η_n of β_n on $[0,n] \times S$, and note that $\xi \overset{d}{=} \eta_n([0,1] \times \cdot)$. Since ξ is distributed as an n^{-1}-thinning of both β_n and $\eta_{n+1}([0,n] \times \cdot)$, Lemma 3.3 (ii) yields $\beta_n \overset{d}{=} \eta_{n+1}([0,n] \times \cdot)$, which implies $\eta_n \overset{d}{=} \eta_{n+1}$ on $[0,n] \times S$, since both processes are λ/n-randomizations with equally distributed S-projections. Using Lemma 1.16 recursively, we may redefine the η_n for $n > 1$, such that $\eta_n = \eta_{n+1}$ a.s. on $[0,n] \times S$ for all $n \in \mathsf{N}$. Introducing a point process η on $\mathsf{R}_+ \times S$ with $\eta = \eta_n$ a.s. for all n, we note that η is λ-symmetric, since this holds for all η_n. By Lemma 3.36 (ii), η is then a Cox process directed by some random measure of the form $\lambda \otimes \nu$, and so $\xi = \eta([0,1] \times \cdot)$ is Cox and directed by ν. The converse assertion is obvious, by a similar embedding argument. □

Theorem 3.35 is essentially equivalent to Bernstein's Theorem A6.1. Here we first prove the implication in one direction:

Proof of Theorem 3.35 from Theorem A6.1: Let ξ be an exchangeable simple point process on $I = [0,1]$ or R_+. The process $\eta(B) = 1\{\xi B > 0\}$ is maxitive, and so the function $h = 1 - E\eta$ is completely monotone with $h(\emptyset) = 1$. The exchangeability gives $h(B) = f(\lambda B)$ for some function f on I, which is again completely monotone with $f(0) = f(0+) = 1$. Hence, Theorem A6.1 yields $f(t) = Ee^{-\rho t}$ when $I = \mathsf{R}_+$ and $f(t) = E(1-t)^\kappa$ when $I = [0,1]$, and so $P\{\xi B = 0\} = Ee^{-\rho\lambda B}$ or $P\{\xi B = 0\} = E(1 - \lambda B)^\kappa$, respectively. By Theorem 3.8, ξ is then a mixed Poisson or binomial process based on (λ, ρ) or (λ, κ), respectively. □

To extend the latter implication to an equivalence, we need a characterization of symmetric random sets. Here we say that a random set in R_+ or $[0,1]$ is *exchangeable*, if its distribution is invariant under permutations of intervals.

Lemma 3.39 *(exchangeable random sets) For any closed exchangeable random set φ in $I = [0,1]$ or R_+, we have a.s. either $\varphi \in \hat{\mathcal{F}}$ or $\varphi = I$.*

Proof: It is enough to take $I = [0,1]$. We need to show that $P(\varphi = I \mid \kappa = \infty) = 1$, where κ denotes the cardinality of φ. Since κ is invariant under permutations of intervals, the exchangeability of φ is preserved by conditioning on κ, which allows us to assume that $\kappa = \infty$ a.s. Now define $I_{nj} = 2^{-n}(j-1, j]$, $j = 1, \dots, 2^n$, and put $\kappa_n = \sum_j 1\{\varphi \cap I_{nj} \neq \emptyset\}$, $n \in \mathsf{N}$. For every $n \in \mathsf{N}$, consider an exchangeable permutation π_{nj}, $j \leq 2^n$, of $1, \dots, 2^n$, and let η_{nj}, $j \leq 2^n$, be i.i.d. and $U\{1, \dots, 2^n\}$. Then for any $k \leq n$

$$\begin{aligned} P \bigcup_{i \leq 2^k} \{\varphi \cap I_{ki} = \emptyset\} &\leq 2^k P\{\varphi \cap I_{k1} = \emptyset\} \\ &= 2^k\, P \bigcap_{j \leq \kappa_n} \{\pi_{nj} > 2^{n-k}\} \\ &\leq 2^k\, P \bigcap_{j \leq \kappa_n} \{\eta_{nj} > 2^{n-k}\} \\ &= 2^k E \Big(P\{\eta_{n1} > 2^{n-k}\} \Big)^{\kappa_n} \\ &= 2^k E (1 - 2^{-k})^{\kappa_n} \\ &\leq 2^k E \exp(-2^{-k} \kappa_n), \end{aligned}$$

which tends to 0 as $n \to \infty$ for fixed k, since $\kappa_n \to \kappa = \infty$ a.s. Thus, the 2^{-k}-neighborhood of φ equals I a.s. Since k was arbitrary, φ is a.s. dense in I, and we get $\varphi = I$ a.s. since φ is closed. □

Proof of Theorem A6.1 from Theorem 3.35: Let f be completely monotone on I with $f(0) = f(0+) = 1$. Then the set function $h(B) = f(\lambda B)$ is again completely monotone with $h(\emptyset) = 1$. By Lemma 2.20 there exists a weakly maxitive process η on I with $P\{\eta(B) = 0\} = h(B)$. Define

$$\varphi_n = \bigcup\nolimits_j \Big\{ \bar{I}_{nj};\, \eta(I_{nj}) = 1 \Big\}, \quad n \in \mathsf{N},$$

where $I_{nj} = 2^{-n}(j-1, j]$, and note that the sequence $\varphi_1, \varphi_2, \dots$ is a.s. nondecreasing. The intersection $\varphi = \bigcap_n \varphi_n$ is a closed random set, exchangeable under dyadic permutations of intervals, and satisfying $P\{\varphi \cap U = \emptyset\} = h(U)$ for dyadic interval unions U. Since $f(0+) = 1$, Lemma 3.39 shows that φ is a.s. locally finite. By Lemma 1.7 it then supports a simple point process ξ on I with $P\{\xi U = 0\} = h(U) = f(\lambda U)$. Since ξ is again exchangeable, it is a mixed Poisson or binomial process by Theorem 3.35, and so $f(\lambda U) = E e^{-\rho \lambda U}$ when $I = \mathsf{R}_+$ and $f(\lambda U) = E(1 - \lambda U)^\kappa$ when $I = [0, 1]$, for some random variables $\rho \geq 0$ and κ in Z_+. This gives the desired representation for dyadic $t \in I$, and the general result follows by the monotonicity of f. □

Chapter 4

Convergence and Approximation

For any random elements ξ and $\xi_1, \xi_2, \ldots$ in a topological space S, *convergence in distribution*, $\xi_n \xrightarrow{d} \xi$, means that $Ef(\xi_n) \to Ef(\xi)$ for every bounded continuous function f on S. This may also be written as $\mathcal{L}(\xi_n) \xrightarrow{w} \mathcal{L}(\xi)$, where for any bounded measures μ_n and μ, the *weak convergence* $\mu_n \xrightarrow{w} \mu$ means that $\mu_n f \to \mu f$ for all functions f as above[1]. Since random measures on S can be regarded as random elements in the measure space $\mathcal{M}_S$, their distributional convergence depends on the choice of topology on $\mathcal{M}_S$, which in turn depends on the topology on S.

To get a streamlined theory, yet broad enough to cover most applications, we take S to be a separable and complete metric space. We may then introduce the *vague topology* on $\mathcal{M}_S$, generated by the *integration maps* $\pi_f : \mu \mapsto \mu f$, for all bounded continuous functions f on S with bounded support, which makes even $\mathcal{M}_S$ a Polish space. On the subspace $\hat{\mathcal{M}}_S$ of bounded measures, we may also introduce the *weak topology*, generated by the maps π_f for all bounded continuous functions f. The corresponding modes of convergence in distribution, written as $\xi_n \xrightarrow{vd} \xi$ and $\xi_n \xrightarrow{wd} \xi$, respectively, are both equivalent to weak convergence of the associated distributions, but for different choices of underlying topology on $\mathcal{M}_S$ or $\hat{\mathcal{M}}_S$.

The stated definitions cover even the case of point processes $\xi, \xi_1, \xi_2, \ldots$ on S. In particular, if $\xi = \delta_\sigma$ and $\xi_n = \delta_{\sigma_n}$ for some random elements σ and σ_n in S, then $\xi_n \xrightarrow{vd} \xi$ and $\xi_n \xrightarrow{wd} \xi$ are both equivalent to $\sigma_n \xrightarrow{d} \sigma$. For the vague topology we may also have $\xi_n \xrightarrow{vd} 0$, which holds iff $\sigma_n \xrightarrow{P} \infty$, in the sense that $\rho(\sigma_n, s) \xrightarrow{P} \infty$ for any fixed point $s \in S$, where ρ denotes the metric on S.

With the appropriate topologies in place, we are ready to characterize the convergence $\xi_n \xrightarrow{vd} \xi$, for arbitrary random measures ξ_n and ξ on S. In Theorem 4.11 we prove that $\xi_n \xrightarrow{vd} \xi$ iff $\xi_n f \xrightarrow{d} \xi f$, for any bounded continuous function f on S with bounded support. In view of the well-known complications arising in the standard weak convergence theory for processes on $[0,1]$ or R_+, it is quite remarkable that, in the present context, no extra tightness condition is needed for convergence. The mentioned criterion is equivalent to the finite-dimensional convergence $(\xi_n B_1, \ldots, \xi_n B_m) \xrightarrow{d} (\xi B_1, \ldots, \xi B_m)$,

[1] Recall the integral notation $\mu f = \int f \, d\mu$.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_4

for any ξ-continuity sets $B_1, \ldots, B_m \in \hat{\mathcal{S}}$, where the ξ-*continuity* means that $\xi \partial B_k = 0$ a.s. for all k.

Since the distribution of a simple point process or diffuse random measure is determined by all one-dimensional distributions, we may expect the corresponding convergence criteria to simplify accordingly. Indeed, when $\xi, \xi_1, \xi_2, \ldots$ are either simple point processes or diffuse random measures, the convergence $\xi_n \xrightarrow{vd} \xi$ becomes equivalent to $\xi_n B \xrightarrow{d} \xi B$ for all ξ-continuity sets $B \in \hat{\mathcal{S}}$. Several weaker and more refined criteria are given in Theorems 4.15, 4.16, and 4.18. Another basic result is Theorem 4.19, which clarifies the relationship between the notions of distributional convergence for the weak and vague topologies.

An important special case is when $\xi_n = \sum_j \xi_{nj}$ for each $n \in \mathsf{N}$, where the ξ_{nj} form a *null array* of random measures, in the sense that the ξ_{nj} are independent in j for fixed n and satisfy $\xi_{nj} \xrightarrow{vd} 0$ as $n \to \infty$, uniformly in j. Here the classical Corollary 4.25 provides necessary and sufficient conditions for convergence to a Poisson process ξ with intensity ρ, when the ξ_{nj} form a null array of point processes on S. In the simplest case, the conditions for convergence become

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

for any ρ-continuity sets $B \in \hat{\mathcal{S}}$. For general sums $\sum_j \xi_{nj}$, the possible limits are infinitely divisible, and the conditions for convergence can be expressed in terms of the Lévy measure λ of ξ. In fact, there is a striking analogy between the criteria for $\xi_n \xrightarrow{vd} \xi$ and $\sum_j \xi_{nj} \xrightarrow{vd} \xi$.

There is also a stronger, non-topological notion of convergence, which is of equally great importance. Starting with the case of non-random measures μ_n and μ, we may write $\mu_n \xrightarrow{u} \mu$ for the total-variation convergence $\|\mu_n - \mu\| \to 0$. We also consider the local version $\mu_n \xrightarrow{ul} \mu$, defined by $1_B \mu_n \xrightarrow{u} 1_B \mu$ for all $B \in \hat{\mathcal{S}}$, where $1_B \mu = 1_B \cdot \mu$ denotes the restriction of μ to the set B. For random measures ξ_n and ξ, the corresponding global notion of convergence, $\xi_n \xrightarrow{ud} \xi$, is defined by $\mathcal{L}(\xi_n) \xrightarrow{u} \mathcal{L}(\xi)$. Finally, the associated local version is $\xi_n \xrightarrow{uld} \xi$, meaning that $1_B \xi_n \xrightarrow{ud} 1_B \xi$ for all $B \in \hat{\mathcal{S}}$. Note that the latter modes of convergence involve only the local structure on S, but no topology.

Among the many powerful results involving strong convergence, we quote at this point only the deep Theorem 4.38. Letting ξ and $\xi_1, \xi_2, \ldots$ be infinitely divisible point processes on some bounded space S with Lévy measures λ and $\lambda_1, \lambda_2, \ldots$, we prove that $\xi_n \xrightarrow{ud} \xi$ iff $\lambda_n \xrightarrow{u} \lambda$. This implies a similar equivalence for random measures on a localized space S. To emphasize the analogy with results in the topological setting such as Corollary 4.39, we note that for any null array of point processes ξ_{nj}, we have the striking equivalences

$$\sum\nolimits_j \xi_{nj} \xrightarrow{vd} \xi \quad \Leftrightarrow \quad \sum\nolimits_j \lambda_{nj} \xrightarrow{vw} \lambda,$$

$$\sum_j \xi_{nj} \xrightarrow{uld} \xi \quad \Leftrightarrow \quad \sum_j \lambda_{nj} \xrightarrow{ul} \lambda,$$

where the precise definitions are postponed until later sections.

The mentioned results have many applications, some of which appear in the next chapter. Here we comment specifically on Theorem 4.40, where general criteria are given for the convergence $\xi_n \xrightarrow{d} \xi$, when the ξ_n are ν_n-transforms of some point processes η_n on S, for some probability kernels ν_n on S. When the kernels ν_n are *dissipative*, in the sense that $\nu_n \xrightarrow{v} 0$ uniformly on S, all limiting processes ξ turn out to be Cox. Writing ζ for the associated directing random measure, we have $\xi_n \xrightarrow{vd} \xi$ iff $\eta_n \nu_n \xrightarrow{vd} \zeta$. This applies in particular to the case where each ξ_n is a p_n-thinning of η_n for some constants $p_n \to 0$, in which case the latter condition reduces to $p_n \eta_n \xrightarrow{vd} \zeta$.

4.1 Weak and Vague Topologies

Let (S, d) be a complete, separable metric space with classes $\hat{\mathcal{S}}$ of bounded Borel sets and $\mathcal{K}$ of compact sets. Write C_S for the class of bounded, continuous functions $f: S \to \mathsf{R}_+$, and let $\hat{C}_S$ denote the subclass of functions with bounded support. If S is locally compact, we may choose a metrization d of S, such that a set $B \subset S$ is relatively compact iff it is bounded.

On $\mathcal{M}_S$ we introduce the *vague topology*, generated by the maps π_f : $\mu \mapsto \mu f$ for all $f \in \hat{C}_S$, i.e., the coarsest topology making all π_f continuous. On $\hat{\mathcal{M}}_S$ we also consider the *weak topology*, generated by the maps π_f with $f \in C_S$. This agrees with the vague topology on $\hat{\mathcal{M}}_S$, when d is replaced by the metric $\hat{d} = d \wedge 1$. We state some basic facts about the vague topology on $\mathcal{M}_S$. The associated notions of *vague* and *weak convergence*, denoted by $\mu_n \xrightarrow{v} \mu$ and $\mu_n \xrightarrow{w} \mu$, respectively, are defined by the conditions $\mu_n f \to \mu f$ for all $f \in \hat{C}_S$ or C_S.

Recall that $\hat{\mathcal{S}}_\mu$ denotes the class of sets $B \in \hat{\mathcal{S}}$ with $\mu \partial B = 0$. Say that $\mathcal{C} \subset \hat{C}_+$ is an *approximating class* of $\mathcal{I} \subset \hat{\mathcal{S}}_\mu$, if for any $B \in \mathcal{I}$ there exist some functions $f_n, g_n \in \mathcal{C}$ with $f_n \uparrow 1_{B^o}$ and $g_n \downarrow 1_{\bar{B}}$.

Lemma 4.1 *(vague convergence) Let $\mu, \mu_1, \mu_2, \ldots \in \mathcal{M}_S$, and fix any dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}_\mu$ with approximating class $\mathcal{C} \subset \hat{C}_+$. Then these conditions are equivalent:*

(i) $\mu_n \xrightarrow{v} \mu$,

(ii) $\mu_n f \to \mu f$ *for all* $f \in \mathcal{C}$,

(iii) $\mu_n I \to \mu I$ *for all* $I \in \mathcal{I}$,

(iv) $\mu B^o \leq \liminf_{n\to\infty} \mu_n B \leq \limsup_{n\to\infty} \mu_n B \leq \mu \bar{B}, \quad B \in \hat{\mathcal{S}}$.

On product spaces $S = S' \times S''$, it suffices in (ii)−(iv) *to consider the corresponding tensor products $f = f' \otimes f''$ or product sets $B = B' \times B''$.*

Proof, (i) ⇒ (ii): This holds by the definition of vague convergence.

(ii) ⇒ (iii): Letting $0 \le f \le 1_I \le g \le 1$ with $I \in \mathcal{I}$ and $f, g \in \mathcal{C}$, we get by (ii)

$$\mu f \le \liminf_{n\to\infty} \mu_n I \le \limsup_{n\to\infty} \mu_n I \le \mu g,$$

and (iv) follows with $B = I$, by dominated convergence as $f \uparrow 1_{I^o}$ and $g \downarrow 1_{\bar{I}}$. Since $\mu\partial I = 0$, we have $\mu I^o = \mu\bar{I} = \mu I$, and (iii) follows.

(iii) ⇒ (iv): Write $\mathcal{U}$ for the ring generated by $\mathcal{I}$. Letting $B \in \hat{\mathcal{S}}$ and $U, V \in \mathcal{U}$ with $U \subset B^o \subset \bar{B} \subset V$, we get by (iii)

$$\mu U \le \liminf_{n\to\infty} \mu_n B \le \limsup_{n\to\infty} \mu_n B \le \mu V,$$

and (iv) follows by dominated convergence, as $U \uparrow B^o$ and $V \downarrow \bar{B}$.

(iv) ⇒ (i): The relations in (iv) yield $\mu_n B \to \mu B$ for all $B \in \hat{\mathcal{S}}_\mu$. Now fix any $f \in \hat{C}_S$, and note that

$$\{f > t\} \equiv \{s \in S;\, f(s) > t\} \in \hat{\mathcal{S}}, \quad t \ge 0.$$

Since

$$\partial\{s \in S;\, f(s) > t\} \subset \{s \in S;\, f(s) = t\}, \quad t > 0,$$

we have in fact $\{f > t\} \in \hat{\mathcal{S}}_\mu$ for all but countably many $t \ge 0$. Clearly, $\{f > t\} = \emptyset$ for $t > \|f\|$, and by (iv) we have

$$\begin{aligned}\limsup_{n\to\infty} \mu_n\{f > 0\} &\le \mu\overline{\{f > 0\}} \\ &= \mu(\operatorname{supp} f) < \infty.\end{aligned}$$

Using Fubini's theorem and dominated convergence, we get from (iv)

$$\begin{aligned}\mu_n f &= \int_0^\infty \mu_n\{f > t\}dt \\ &\to \int_0^\infty \mu\{f > t\}dt = \mu f,\end{aligned}$$

which shows that $\mu_n f \to \mu f$. Since f was arbitrary, this proves (i).

The last assertion follows from (ii) and (iii) above, along with Lemma 1.9 (vi). □

We turn to the basic metrization property and compactness criteria.

Theorem 4.2 *(metrization and compactness, Prohorov) The vague topology on $\mathcal{M}_S$ is Polish with Borel σ-field $\mathcal{B}_{\mathcal{M}_S}$. Furthermore, a set $A \subset \mathcal{M}_S$ is vaguely relatively compact iff*

(i) $\sup_{\mu\in A} \mu B < \infty, \quad B \in \hat{\mathcal{S}},$

(ii) $\inf_{K\in\mathcal{K}} \sup_{\mu\in A} \mu(B \setminus K) = 0, \quad B \in \hat{\mathcal{S}}.$

In particular, a set $A \subset \hat{\mathcal{M}}_s$ is weakly relatively compact, iff (i) *and* (ii) *hold with $B = S$.*

The result will be proved in steps, beginning with the special case of the weak topology on $\hat{\mathcal{M}}_S$. For any $B \subset S$ and $\varepsilon > 0$, put $B^\varepsilon = \{s \in S;\ d(s,B) < \varepsilon\}$, where $d(s,B) = \inf\{d(s,t);\ t \in B\}$. Given any $\mu, \nu \in \hat{\mathcal{M}}_S$, we define

$$\rho(\mu,\nu) = \inf\Big\{\varepsilon > 0;\ \mu B \le \nu B^\varepsilon + \varepsilon,\ \nu B \le \mu B^\varepsilon + \varepsilon,\ B \in \mathcal{S}\Big\}.$$

Lemma 4.3 *(Prohorov metric) The function ρ is a metric on $\hat{\mathcal{M}}_S$ inducing the weak topology, so that $\mu_n \xrightarrow{w} \mu$ in $\hat{\mathcal{M}}_S$ iff $\rho(\mu,\mu_n) \to 0$.*

Proof: The see that ρ is a metric, it suffices to prove the triangle inequality. Letting $\rho(\lambda,\mu) < \delta$ and $\rho(\mu,\nu) < \varepsilon$, we get for any $B \in \mathcal{S}$

$$\begin{aligned}\lambda B &\le \mu B^\delta + \delta \\ &\le \nu (B^\delta)^\varepsilon + \delta + \varepsilon \\ &\le \nu B^{\delta+\varepsilon} + \delta + \varepsilon,\end{aligned}$$

and similarly with λ and ν interchanged, which gives $\rho(\lambda,\nu) \le \delta + \varepsilon$. Taking the infimum over all such δ and ε gives

$$\rho(\lambda,\nu) \le \rho(\lambda,\mu) + \rho(\mu,\nu).$$

Now assume that $\rho(\mu_n,\mu) \to 0$ for some $\mu, \mu_1, \mu_2, \ldots \in \hat{\mathcal{M}}_S$. For any closed set $F \subset S$, we get

$$\limsup_{n\to\infty} \mu_n F \le \inf_{\varepsilon>0} \mu F^\varepsilon = \mu \bigcap_{\varepsilon>0} F^\varepsilon = \mu F.$$

Since also $\liminf_{n\to\infty} \mu_n S \ge \mu S$, we have $\mu_n S \to \mu S$, and so for any open set $G \subset S$,

$$\begin{aligned}\liminf_{n\to\infty} \mu_n G &= \mu S - \limsup_{n\to\infty} \mu_n G^c \\ &\ge \mu S - \mu G^c = \mu G.\end{aligned}$$

Hence, Lemma 4.1 yields $\mu_n \xrightarrow{w} \mu$.

Conversely, assume $\mu_n \xrightarrow{w} \mu$, and let $\varepsilon > 0$ be arbitrary. Since S is separable, it can be covered by some sets $I_1, I_2, \ldots \in \mathcal{S}_\mu$ of diameter $< \varepsilon$. Since μ is bounded, we may choose $m \in \mathbb{N}$ so large that $\mu U_m^c < \varepsilon$, where $U_m = \bigcup_{k\le m} I_k$. Let $\mathcal{U}_m$ consist of all unions of sets in $\{I_1, \ldots, I_m\}$. Since $\mathcal{U}_m$ is finite, Lemma 4.1 yields an $n_0 \in \mathbb{N}$ such that $|\mu_n U - \mu U| < \varepsilon$ for all $U \in \mathcal{U}_m$ and $n > n_0$. We may also require $|\mu_n U_m^c - \mu U_m^c| < \varepsilon$ for all $n > n_0$, so that $\mu_n U_m^c < 2\varepsilon$ for all such n. For any $B \in \mathcal{S}$, let $\tilde{B}$ be the union of all sets $I_1, \ldots, I_m$ intersecting B. Then

$$\begin{aligned}\mu B &\le \mu \tilde{B} + \varepsilon \le \mu_n \tilde{B} + 2\varepsilon \\ &\le \mu_n B^\varepsilon + 2\varepsilon, \\ \mu_n B &\le \mu_n \tilde{B} + 2\varepsilon \le \mu \tilde{B} + 3\varepsilon \\ &\le \mu B^\varepsilon + 3\varepsilon,\end{aligned}$$

which shows that $\rho(\mu,\mu_n) \le 3\varepsilon$ for all $n > n_0$. Since ε was arbitrary, we obtain $\rho(\mu,\mu_n) \to 0$. □

The following lemma characterizes compactness in $\hat{\mathcal{M}}_S$.

Lemma 4.4 *(weak compactness in $\hat{\mathcal{M}}_S$) A set $A \subset \hat{\mathcal{M}}_S$ is weakly relatively compact iff*

(i) $\sup_{\mu\in A} \mu S < \infty$,

(ii) $\inf_{K\in\mathcal{K}} \sup_{\mu\in A} \mu K^c = 0.$

The sufficiency of (i) and (ii) is actually true under weaker conditions on S (cf. FMP 16.3). However, assuming S to be separable and complete simplifies the proof.

Proof (necessity): Condition (i) holds since the mapping π_S is weakly continuous on $\hat{\mathcal{M}}_S$. To prove (ii), fix any $\varepsilon > 0$. Since S is separable, it is covered by some open ε-balls $B_1, B_2, \ldots$. Putting $G_n = \bigcup_{k\le n} B_k$, we claim that

$$\lim_{n\to\infty} \sup_{\mu\in A} \mu G_n^c = 0. \tag{1}$$

Indeed, we may otherwise choose some $\mu_1,\mu_2,\ldots \in A$ with $\inf_n \mu_n G_n^c = c > 0$. Since A is relatively compact, $\mu_n \xrightarrow{w} \mu \in A$ along a sub-sequence N', and Lemma 4.1 gives $\mu G_m^c \ge \limsup_{n\in N'} \mu_n G_m^c \ge c$ for every $m \in \mathsf{N}$, which yields the contradiction $c \le 0$, by dominated convergence as $m \to \infty$.

If $\varepsilon > 0$ and $k \in \mathsf{N}$, then by (1) we may choose a finite union F_k of closed 2^{-k}-balls with $\sup_{\mu\in A} \mu F_k^c < \varepsilon 2^{-k}$. The set $K = \bigcap_k F_k$ is totally bounded and complete, hence compact, and $\sup_{\mu\in A} \mu K^c < \sum_k \varepsilon 2^{-k} = \varepsilon$. Since ε was arbitrary, condition (ii) follows.

(Sufficiency): By Lemma 4.3 it suffices to show that A is sequentially weakly relatively compact, in the sense that every sequence $\mu_1,\mu_2,\ldots \in A$ has a weakly convergent sub-sequence. Since S is separable, it is homeomorphic to a subset of R_+^∞, and so we may assume that $S \subset \mathsf{R}_+^\infty$. Writing π_d for the projection of R_+^∞ onto the first d coordinates, we note that (i) and (ii) extend by continuity to the projections $\mu_n \circ \pi_d^{-1}$. If the assertion holds in R_+^d for every d, then a diagonal argument yields convergence $\mu_n \circ \pi_d^{-1} \xrightarrow{w} \nu_d$ for every $d \in \mathsf{N}$, along some sub-sequence $N' \subset \mathsf{N}$. The sequence of limits ν_d is clearly projective, and so the Daniell–Kolmogorov theorem (FMP 6.14) yields a bounded measure μ on R_+^∞ with $\mu \circ \pi_d^{-1} = \nu_d$ for all d, and we get $\mu_n \xrightarrow{w} \mu$ along N'. To see that $\mu S^c = 0$, use (ii) to choose some compact sets $K_m \subset S$ with $\sup_n \mu_n K_m^c < 2^{-m}$. Since the K_m remain compact in R_+^∞, Lemma 4.1 yields

$$\begin{aligned}\mu K_m^c &\leq \liminf_{n\in N'} \mu_n K_m^c \\ &\leq \sup_n \mu_n K_m^c < 2^{-m},\end{aligned}$$

and so $\mu S^c \leq \mu \bigcap_m K_m^c = 0$, as required.

It remains to take $S = \mathsf{R}_+^d$ for a fixed $d \in \mathsf{N}$. Then define $F_n(x) = \mu_n[0,x]$ for $x \in \mathsf{R}_+^d$, where $[0,x] = \bigotimes_i [0,x_i]$. Since the F_n are bounded by (i), a diagonal argument yields convergence $F_n(r) \to G(r)$ along a sub-sequence $N' \subset \mathsf{N}$, for all $r \in \mathsf{Q}_+^d$. The function $F(x) = \inf\{G(r);\, r > x\}$ is clearly bounded and right-continuous on R_+^d, with $F_n(x) \to F(x)$ along N' for almost all $x \geq 0$. Proceeding as in FMP 3.25, we may construct a bounded measure μ on R_+^d with $\mu[0,x] = F(x)$ for all $x \in \mathsf{R}_+^d$. Then $\mu_n(x,y] \to \mu(x,y]$ along N' for all $x < y$ in R_+^d, and a simple approximation yields condition (ii) in Lemma 4.1. By a further approximation based on (ii), the condition extends to unbounded Borel sets $B \subset \mathsf{R}^d$, and so the lemma gives $\mu_n \xrightarrow{w} \mu$ along N'. $\square$

We proceed to show that $\hat{\mathcal{M}}_S$ is again Polish for the weak topology:

Lemma 4.5 *(Polishness of $\hat{\mathcal{M}}_S$) The space $\hat{\mathcal{M}}_S$ is again Polish, and a complete metrization is given by the ρ of Lemma 4.3.*

Proof: Choose a dense set $D = \{s_1, s_2, \ldots\}$ in S, and let M denote the countable set of measures $\mu = \sum_{k\leq m} r_k \delta_{s_k}$ with $m \in \mathsf{N}$ and $r_1, \ldots, r_m \in \mathsf{Q}_+$. For any $\varepsilon > 0$, we may choose a countable partition of S into sets $B_k \in \mathcal{S}$, each within distance ε of some point $t_k \in D$. For any $\mu \in \hat{\mathcal{M}}_s$, the measure $\mu' = \sum_k \mu B_k\, \delta_{t_k}$ satisfies $\rho(\mu, \mu') \leq \varepsilon$. By a suitable truncation and adjustment of coefficients, we may next choose a measure $\mu'' \in M$ with $\|\mu' - \mu''\| < \varepsilon$. Then $\rho(\mu, \mu'') < 2\varepsilon$, which shows that M is dense.

Next, let $\mu_1, \mu_2, \ldots \in \hat{\mathcal{M}}_S$ with $\rho(\mu_m, \mu_n) \to 0$ as $m, n \to \infty$, and note that $\sup_n \mu_n S < \infty$. For any $\varepsilon > 0$, let $m \in \mathsf{N}$ be such that $\rho(\mu_m, \mu_n) < \varepsilon$ for all $n > m$. Since the μ_n are individually tight, there exists a compact set $K_\varepsilon \subset S$ with $\mu_n K_\varepsilon^c < \varepsilon$ for all $n \leq m$. By compactness, we may cover K_ε by a finite union U_ε of ε-balls. By the definition of ρ, we have for any $n > m$

$$\begin{aligned}\mu_n (U_\varepsilon^\varepsilon)^c &= \mu_n S - \mu_n U_\varepsilon^\varepsilon \\ &\leq \mu S - \mu U_\varepsilon + 2\varepsilon \\ &= \mu U_\varepsilon^c + 2\varepsilon \\ &\leq \mu K_\varepsilon^c + 2\varepsilon \leq 3\varepsilon,\end{aligned}$$

which clearly remains true for $n \leq m$. Now put $\varepsilon_k = \varepsilon 2^{-k}$ for $k \in \mathsf{N}$, and define $K = \bigcap_k \overline{U_{\varepsilon_k}^{\varepsilon_k}}$. Then K is totally bounded and hence compact, and for every $n \in \mathsf{N}$

$$\begin{aligned}\mu_n K^c &\leq \mu_n \bigcup\nolimits_k (U_{\varepsilon_k}^{\varepsilon_k})^c \\ &\leq \sum\nolimits_k \mu_n (U_{\varepsilon_k}^{\varepsilon_k})^c \\ &\leq 3\varepsilon \sum\nolimits_k 2^{-k} = 3\varepsilon.\end{aligned}$$

Thus, (μ_n) is tight, and Lemma 4.4 yields weak convergence along a sub-sequence. Since (μ_n) is Cauchy, the convergence extends to the full sequence, which shows that $\hat{\mathcal{M}}_S$ is complete for the metric ρ. $\square$

The previous results for $\hat{\mathcal{M}}_S$ with the weak topology may now be extended to $\mathcal{M}_S$ with the vague topology:

Lemma 4.6 *(metrization and compactness in $\mathcal{M}_S$) The metrization and compactness properties of $\hat{\mathcal{M}}_S$ extend to $\mathcal{M}_S$ with the vague topology.*

Proof: Fix a point $s_0 \in S$, and define $B_n = \{s \in S;\, d(s, s_0) < n\}$. For every $k \in \mathsf{N}$, we may choose an $f_k \in \hat{C}_S$ with $1_{B_k} \le f_k \le 1_{B_{k+1}}$, so that $\mu_n \xrightarrow{v} \mu$ in $\mathcal{M}_S$ iff $f_k \cdot \mu_n \xrightarrow{w} f_k \cdot \mu$ for all k. Writing $\hat{\rho}$ for the Prohorov metric on $\hat{\mathcal{M}}_S$, we may then choose the metrization

$$\rho(\mu, \mu') = \sum\nolimits_k 2^{-k} \Big\{ \hat{\rho}(f_k \cdot \mu,\, f_k \cdot \mu') \wedge 1 \Big\}, \quad \mu, \mu' \in \mathcal{M}_S,$$

of the vague topology on $\mathcal{M}_S$. For any countable, weakly dense set $D \subset \hat{\mathcal{M}}$, the countable set $\bigcup_n (D^n \times \{0\}^\infty)$ is dense in $(\hat{\mathcal{M}}_S)^\infty$ with the product topology, which implies the vague separability of $\mathcal{M}_S$, regarded as a subset of $(\hat{\mathcal{M}}_S)^\infty$.

To prove that ρ is complete, let $\mu_1, \mu_2, \ldots \in \mathcal{M}_S$ with $\rho(\mu_m, \mu_n) \to 0$ as $m, n \to \infty$. Then $\hat{\rho}(f_k \cdot \mu_m,\, f_k \cdot \mu_n) \to 0$ for each $k \in \mathsf{N}$. By the completeness of $\hat{\rho}$ and a diagonal argument, we have convergence $f_k \cdot \mu_n \xrightarrow{w} \nu_k$ for all k, along a sub-sequence $N' \subset \mathsf{N}$. Noting that $\nu_k = \nu_l$ on B_k for all $k < l$ and putting $\mu = \sup_k (1_{B_k} \nu_k)$, we get $\mu_n \xrightarrow{v} \mu$ along N', as required.

Now let $A \subset \mathcal{M}_S$ be vaguely relatively compact. Since the maps $\varphi_n \colon \mu \mapsto f_n \cdot \mu$ are continuous from $\mathcal{M}_S$ to $\hat{\mathcal{M}}_S$, the images $A_n = \varphi_n(A)$ are weakly relatively compact, and so by Lemma 4.4

$$\begin{aligned} \sup_{\mu \in A} \mu B_n &\le \sup_{\mu \in A} (f_n \cdot \mu) B_n < \infty, \\ \inf_{K \in \mathcal{K}} \sup_{\mu \in A} \mu(B_n \setminus K) &\le \inf_{K \in \mathcal{K}} \sup_{\mu \in A} (f_n \cdot \mu) K^c = 0, \end{aligned}$$

which implies (i) and (ii) of Theorem 4.2, since n was arbitrary.

Conversely, assuming (i) and (ii), we get for any $k \in \mathsf{N}$

$$\begin{aligned} \sup_{\mu \in A} (f_k \cdot \mu) S &\le \sup_{\mu \in A} \mu B_{k+1} < \infty, \\ \inf_{K \in \mathcal{K}} \sup_{\mu \in A} (f_k \cdot \mu) K^c &\le \inf_{K \in \mathcal{K}} \sup_{\mu \in A} \mu(B_{k+1} \setminus K) = 0, \end{aligned}$$

and so by Lemma 4.4 the sets A_k are weakly relatively compact. Hence, for any $\mu_1, \mu_2, \ldots \in A$, a diagonal argument yields convergence $f_k \cdot \mu_n \xrightarrow{w} \nu_k$ for every $k \in \mathsf{N}$, along a sub-sequence $N' \in \mathsf{N}$, and so as before $\mu_n \xrightarrow{v} \mu \in \mathcal{M}_S$ along N'. Thus, A is vaguely relatively sequentially compact. Since the vague topology is metrizable, the vague compactness follows. $\square$

It remains to identify the Borel sets in $\mathcal{M}_S$.

Lemma 4.7 *(Borel σ-field) The following families of sets or maps generate the same σ-field in $\mathcal{M}_S$:*

(i) *the vaguely open sets in $\mathcal{M}_S$,*

(ii) *the maps π_f with $f \in \hat{C}_S$,*

(iii) *the maps π_I with sets I in a dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}$.*

Proof: Let the generated σ-fields be Σ_i, $i = 1, 2, 3$. Since the sets $\{\mu;\, \mu f < r\}$ are vaguely open for all $f \in \hat{C}_S$ and $r > 0$, we have $\Sigma_2 \subset \Sigma_1$. Conversely, since $\mathcal{M}_S$ is vaguely separable, any open set in $\mathcal{M}_S$ is a countable union of basis sets $\{\mu;\, |\mu f - r| < \varepsilon\}$ with $f \in \hat{C}_S$ and $r, \varepsilon > 0$, and so $\Sigma_1 \subset \Sigma_2$.

For any open set $G \in \hat{\mathcal{S}}$, choose $f_1, f_2, \ldots \in \hat{C}_S$ with $f_n \uparrow 1_G$. Then $\mu f_n \uparrow \mu G$ by monotone convergence, and so the maps π_G are Σ_2-measurable. Fixing any open set $G_0 \in \hat{\mathcal{S}}$, let $\mathcal{G}_0$ be the class of open sets in G_0, and write $\mathcal{D}$ for the class of Borel sets $B \subset G_0$, such that π_B is Σ_2-measurable. Then $\mathcal{D}$ is a λ-system containing the π-system $\mathcal{G}_0$, and so a monotone-class argument yields $\mathcal{D} = \mathcal{S} \cap G_0$. Hence, π_B is Σ_2-measurable for every $B \in \hat{\mathcal{S}}$, which implies $\Sigma_3 \subset \Sigma_2$.

As for the converse relation, a similar monotone-class argument allows us to choose $\mathcal{I} = \hat{\mathcal{S}}$. For any $f \in \hat{C}_S$, let $f_1, f_2, \ldots$ be simple, $\hat{\mathcal{S}}$-measurable functions with $0 \le f_n \uparrow f$. Then $\mu f_n \uparrow \mu f$ by monotone convergence, which shows that $\Sigma_2 \subset \Sigma_3$. □

We proceed with some criteria for convergence of random measures ξ_n toward a limit ξ. Here the condition $\xi_n \xrightarrow{v} \xi$ a.s. means of course that $\xi_n(\omega) \xrightarrow{v} \xi(\omega)$ for all $\omega \in \Omega$ outside a P-null set. Similarly, $\xi_n \xrightarrow{vP} \xi$ means that $\xi_n f \xrightarrow{P} \xi f$ for all $f \in \hat{C}_S$, whereas $\xi_n \xrightarrow{v} \xi$ in L^1 is defined by $\xi_n f \to \xi f$ in L^1 for every $f \in \hat{C}_S$.

Lemma 4.8 *(convergence of random measures) Let ξ and $\xi_1, \xi_2, \ldots$ be random measures on S. Then*

(i) *$\xi_n \xrightarrow{v} \xi$ a.s. $\Leftrightarrow$ $\xi_n f \to \xi f$ a.s. for all $f \in \hat{C}_S$,*

(ii) *$\xi_n \xrightarrow{v} \xi$ in L^1 $\Leftrightarrow$ $E\xi_n \xrightarrow{v} E\xi \in \mathcal{M}_S$ and $\xi_n \xrightarrow{vP} \xi$,*

(iii) *$\xi_n \xrightarrow{vP} \xi$ $\Leftrightarrow$ for any sub-sequence $N' \subset \mathsf{N}$, we have $\xi_n \xrightarrow{v} \xi$ a.s. along a further sub-sequence $N'' \subset N'$.*

Proof: (i) If $\xi_n \xrightarrow{v} \xi$ a.s., then $\xi_n f \to \xi f$ holds a.s. for all $f \in \hat{C}_S$, by the definition of vague convergence. Conversely, Lemma 1.9 yields the existence of a countable dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}_\xi$, to which we can associate a countable approximating class $\mathcal{C} \subset \hat{C}_S$. The stated condition gives $\xi_n f \to \xi f$ for all $f \in \mathcal{C}$ outside a common P-null set, and so $\xi_n \xrightarrow{v} \xi$ a.s. by Lemma 4.1.

(ii) For any $f \in \hat{C}_S$, we see from FMP 4.12 that $\xi_n f \to \xi f$ in L^1 iff $E\xi_n f \to E\xi f < \infty$ and $\xi_n f \xrightarrow{P} \xi f$. The stated equivalence now follows by the definitions of vague convergence in $\mathcal{M}_S$ and convergence in probability or in L^1 for random measures on S.

(iii) Assume the stated condition. Since $\xi_n \xrightarrow{v} \xi$ a.s. implies $\xi_n f \to \xi f$ a.s. for every $f \in \hat{C}_S$, we get $\xi_n f \xrightarrow{P} \xi f$ for all $f \in \hat{C}_S$ by FMP 4.2, which means that $\xi_n \xrightarrow{vP} \xi$. Conversely, assuming $\xi_n \xrightarrow{vP} \xi$, fix any sub-sequence $N' \subset \mathsf{N}$. Choose a countable approximating class $\mathcal{C} \subset \hat{C}_S$, as in part (i). By FMP 4.2 there exists a further sub-sequence $N'' \subset N'$, such that $\xi_n f \to \xi f$ along N'' for all $f \in \mathcal{C}$, outside a common P-null set. Hence, $\xi_n \xrightarrow{v} \xi$ a.s. along N'' by Lemma 4.1. $\Box$

It is sometimes useful to characterize convergence of random measures ξ_n in terms of the convergence of $\xi_n f$ or $\xi_n B$, for suitable sub-classes of functions f or sets B.

Corollary 4.9 *(sub-class criteria) Let ξ and $\xi_1, \xi_2, \ldots$ be random measures on S, and fix a dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}_\xi$ with approximating class $\mathcal{C}$. Then*

(i) $\xi_n \xrightarrow{v} \xi$ *a.s.* $\Leftrightarrow$ $\xi_n f \to \xi f$ *a.s. for all* $f \in \mathcal{C}$,

(ii) $\xi_n \xrightarrow{v} \xi$ *in* L^1 $\Leftrightarrow$ $\xi_n f \to \xi f$ *in* L^1 *for all* $f \in \mathcal{C}$,

(iii) $\xi_n \xrightarrow{vP} \xi$ $\Leftrightarrow$ $\xi_n f \xrightarrow{P} \xi f$ *for all* $f \in \mathcal{C}$,

All statements remain true with the functions $f \in \mathcal{C}$ replaced by sets $I \in \mathcal{I}_\xi$.

Proof: The necessity in (i)–(iii) is immediate from the definitions. To prove the sufficiency, we may invoke Lemma 1.9 to reduce to some countable classes $\mathcal{I}$ and $\mathcal{C}$. We consider only the case of functions $f \in \mathcal{C}$, the case of sets $I \in \mathcal{I}$ being similar.

(i) Let $\xi_n f \to \xi f$ a.s. for all $f \in \mathcal{C}$. Since $\mathcal{I} \subset \mathcal{S}_\xi$ remains a.s. true, Lemma 4.1 yields $\xi_n \xrightarrow{v} \xi$ a.s.

(iii) Let $\xi_n f \xrightarrow{P} \xi f$ for all $f \in \mathcal{C}$. Then for any sub-sequence $N' \subset \mathsf{N}$, we have a.s. $\xi_n f \to \xi f$ for all $f \in \mathcal{C}$ along a further sub-sequence $N'' \subset N'$, and so by (i) we get $\xi_n \xrightarrow{v} \xi$ a.s. along N''. Hence, $\xi_n \xrightarrow{P} \xi$ by Lemma 4.8 (iii).

(ii) Let $\xi_n f \to \xi f$ in L^1 for all $f \in \mathcal{C}$. In particular, $E\xi_n f \to E\xi f$ for all $f \in \mathcal{C}$, and so $E\xi_n \xrightarrow{v} E\xi$ by Lemma 4.1. Furthermore, $\xi_n f \xrightarrow{P} \xi f$ for all $f \in \mathcal{C}$, and so $\xi_n \xrightarrow{vP} \xi$ by part (iii). The convergence $\xi_n \xrightarrow{v} \xi$ in L^1 now follows by Lemma 4.8 (ii). $\Box$

4.2 Convergence in Distribution

The compactness criteria of Theorem 4.2 yield tightness criteria for random measures on S, subject to the vague or weak topologies on $\mathcal{M}_S$ or $\mathcal{N}_S$.

Theorem 4.10 *(tightness of random measures) A collection Ξ of random measures on S is vaguely tight iff*

(i) $\lim_{r\to\infty} \sup_{\xi\in\Xi} P\{\xi B > r\} = 0, \quad B \in \hat{\mathcal{S}},$

(ii) $\inf_{K\in\mathcal{K}} \sup_{\xi\in\Xi} E[\,\xi(B \setminus K) \wedge 1] = 0, \quad B \in \hat{\mathcal{S}}.$

In particular, a family Ξ of a.s. bounded random measures on S is weakly tight iff (i) *and* (ii) *hold with $B = S$.*

Proof: Let Ξ be vaguely tight. Then for every $\varepsilon > 0$ there exists a vaguely compact set $A \subset \mathcal{M}_S$, such that $P\{\xi \notin A\} < \varepsilon$ for all $\xi \in \Xi$. For fixed $B \in \hat{\mathcal{S}}$, Theorem 4.2 yields $r = \sup_{\mu\in A} \mu B < \infty$, and so $\sup_{\xi\in\Xi} P\{\xi B > r\} < \varepsilon$. The same theorem yields a $K \in \mathcal{K}$ with $\sup_{\mu\in A} \mu(B \setminus K) < \varepsilon$, which implies $\sup_{\xi\in\Xi} E\{\xi(B\setminus K)\wedge 1\} \le 2\varepsilon$. Since ε was arbitrary, this proves the necessity of (i) and (ii).

Now assume (i) and (ii). Fix any $s_0 \in S$, and put $B_n = \{s \in S;\, d(s, s_0) < n\}$. Given any $\varepsilon > 0$, we can choose some constants $r_1, r_2, \ldots > 0$ and sets $K_1, K_2, \ldots \in \mathcal{K}$, such that for all $\xi \in \Xi$ and $n \in \mathsf{N}$,

$$P\{\xi B_n > r_n\} < 2^{-n}\varepsilon, \qquad E\{\,\xi(B_n \setminus K_n) \wedge 1\,\} < 2^{-2n}\varepsilon. \tag{2}$$

Writing A for the set of measures $\mu \in \mathcal{M}_S$ with

$$\mu B_n \le r_n, \quad \mu(B_n \setminus K_n) \le 2^{-n}, \qquad n \in \mathsf{N},$$

we see from Theorem 4.2 that A is vaguely relatively compact. Noting that

$$P\{\xi(B_n \setminus K_n) > 2^{-n}\} \le 2^n E\{\,\xi(B_n \setminus K_n) \wedge 1\,\}, \quad n \in \mathsf{N},$$

we get from (2) for any $\xi \in \Xi$

$$P\{\xi \notin A\} \le \sum\nolimits_n P\{\xi B_n > r_n\} + \sum\nolimits_n 2^n E\{\,\xi(B_n \setminus K_n) \wedge 1\,\} < 2\varepsilon,$$

and since ε was arbitrary, we conclude that Ξ is tight. $\square$

We turn to some basic criteria for convergence in distribution of random measures with respect to the vague topology, here denoted by $\xrightarrow{vd}$. For any random measure ξ, we write $\hat{\mathcal{S}}_\xi = \hat{\mathcal{S}}_{E\xi}$, for convenience. The phrase $\xi_n \xrightarrow{d}$ some ξ means that ξ_n converges in distribution toward an unspecified limit ξ. For any class of subsets $\mathcal{I}$ of S, let $\hat{\mathcal{I}}_+$ denote the class of simple, $\mathcal{I}$-measurable functions $f \ge 0$ on S.

Theorem 4.11 *(convergence of random measures, Harris, Matthes et al.) Let $\xi, \xi_1, \xi_2, \ldots$ be random measures on S, and fix a dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}_\xi$. Then these conditions are equivalent:*

(i) $\xi_n \xrightarrow{vd} \xi$,

(ii) $\xi_n f \xrightarrow{d} \xi f$ *for all* $f \in \hat{C}_S$ *or* $\hat{\mathcal{I}}_+$,

(iii) $Ee^{-\xi_n f} \to Ee^{-\xi f}$ *for all* $f \in \hat{C}_S$ *or* $\hat{\mathcal{I}}_+$ *with* $f \le 1$.

The necessity of (ii) and (iii) is a consequence of the following result. For any function f on a topological space, let D_f denote the set of discontinuity points of f. Note in particular that $D_f = \partial B$ when $f = 1_B$.

Lemma 4.12 *(limits of integrals) Let $\xi, \xi_1, \xi_2, \ldots$ be random measures on S with $\xi_n \xrightarrow{d} \xi$, and let $f \in \hat{\mathcal{S}}_+$ with $\xi D_f = 0$ a.s. Then $\xi_n f \xrightarrow{d} \xi f$.*

Proof: By the continuous mapping theorem (FMP 4.27), it is enough to prove that if $\mu_n \xrightarrow{v} \mu$ with $\mu D_f = 0$, then $\mu_n f \to \mu f$. By a suitable truncation and normalization, we may take μ and all μ_n to be probability measures, in which case the same theorem yields $\mu_n \circ f^{-1} \xrightarrow{w} \mu \circ f^{-1}$. This implies $\mu_n f \to \mu f$ since f is bounded. □

The sufficiency in Theorem 4.11 requires yet another lemma.

Lemma 4.13 *(continuity sets) Let $\xi, \eta, \xi_1, \xi_2, \ldots$ be random measures on S with $\xi_n \xrightarrow{d} \eta$, and fix a dissecting ring $\mathcal{U}$ and a constant $t > 0$. Then $\hat{\mathcal{S}}_\eta \supset \hat{\mathcal{S}}_\xi$ whenever*

$$\liminf_{n\to\infty} Ee^{-t\xi_n U} \geq Ee^{-t\xi U}, \quad U \in \mathcal{U}. \tag{3}$$

For point processes $\xi, \xi_1, \xi_2, \ldots$, we may assume instead that

$$\liminf_{n\to\infty} P\{\xi_n U = 0\} \geq P\{\xi U = 0\}, \quad U \in \mathcal{U}.$$

Proof: Fix any $B \in \hat{\mathcal{S}}_\xi$. Let $\partial B \subset F \subset U$ with $F \in \hat{\mathcal{S}}_\eta$ and $U \in \mathcal{U}$, where F is closed. Then by (3) and Lemma 4.12,

$$\begin{aligned} Ee^{-t\eta\partial B} \geq Ee^{-t\eta F} &= \lim_{n\to\infty} Ee^{-t\xi_n F} \\ &\geq \liminf_{n\to\infty} Ee^{-t\xi_n U} \\ &\geq Ee^{-t\xi U}. \end{aligned}$$

By Lemma 1.9 (iii) and (iv), we may let $U \downarrow F$ and then $F \downarrow \partial B$ to get $Ee^{-t\eta\partial B} \geq Ee^{-t\xi\partial B} = 1$ by dominated convergence. Hence, $\eta\partial B = 0$ a.s., which means that $B \in \hat{\mathcal{S}}_\eta$. The proof of the last assertion is similar. □

Proof of Theorem 4.11: The implication (i) ⇒ (ii) holds by Lemma 4.12, whereas (ii) ⇔ (iii) by the continuity theorem for Laplace transforms. Thus, it remains to prove that (ii) ⇒ (i). Then suppose that $\xi_n f \xrightarrow{d} \xi f$ for all $f \in \hat{C}_S$ or $\hat{\mathcal{I}}_+$, respectively. For any $B \in \hat{\mathcal{S}}$ there exists an f with $1_B \leq f$, and so $\xi_n B \leq \xi_n f$ is tight by the convergence of $\xi_n f$. Thus, (ξ_n) satisfies condition (i) of Theorem 4.10.

To prove condition (ii) of the same result, we may assume that $\xi\partial B = 0$ a.s. A simple approximation yields $(1_B\xi_n)f \xrightarrow{d} (1_B\xi)f$ for all $f \in \hat{C}_S$ or $\hat{\mathcal{I}}_+$, and so we may take $S = B$ and let the measures ξ_n and ξ be a.s. bounded with $\xi_n f \xrightarrow{d} \xi f$ for all $f \in C_S$ or $\mathcal{I}_+$. Here the Cramér–Wold theorem yields

$(\xi_n f, \xi_n S) \xrightarrow{d} (\xi f, \xi S)$ for all such f, and so $\xi_n f/(\xi_n S \vee 1) \xrightarrow{d} \xi f/(\xi S \vee 1)$, which allows us to assume $\|\xi_n\| \le 1$ for all n. Then $\xi_n f \xrightarrow{d} \xi f$ implies $E\xi_n f \to E\xi f$ for all f as above, and so $E\xi_n \xrightarrow{w} E\xi$ by Lemma 4.1. Here Theorem 4.2 yields $\inf_{K\in\mathcal{K}} \sup_n E\xi_n K^c = 0$, which proves condition (ii) of Theorem 4.10 for the sequence (ξ_n).

The latter theorem shows that (ξ_n) is tight. If $\xi_n \xrightarrow{d} \eta$ along a sub-sequence for some random measure η on S, then by Lemma 4.12 we get $\xi_n f \xrightarrow{d} \eta f$ for all $f \in \hat{C}_S$. In the former case, $\eta f \overset{d}{=} \xi f$ for all $f \in C_S$, which implies $\eta \overset{d}{=} \xi$ by Lemma 2.2. In the latter case, Lemma 4.13 yields $\hat{\mathcal{S}}_\eta \supset \hat{\mathcal{S}}_\xi \supset \mathcal{I}$, and so $\xi_n f \xrightarrow{d} \eta f$ for all $f \in \hat{\mathcal{I}}_+$, which again implies $\eta \overset{d}{=} \xi$. Since the sub-sequence was arbitrary, Theorem 4.2 yields $\xi_n \xrightarrow{vd} \xi$. □

Even stronger results are available when S is locally compact:

Corollary 4.14 *(existence of limit) Let $\xi_1, \xi_2, \dots$ be random measures on an lcscH space S, such that $\xi_n f \xrightarrow{d}$ some γ_f for every $f \in \hat{C}_S$. Then $\xi_n \xrightarrow{vd} \xi$ for some random measure ξ on S with $\xi f \overset{d}{=} \gamma_f$ for all $f \in \hat{C}_S$.*

Proof: Here condition (ii) of Theorem 4.10 is void, and the tightness of (ξ_n) follows already from the tightness of $(\xi_n f)$ for all $f \in \hat{C}_S$. Hence, Theorem 4.2 yields $\xi_n \xrightarrow{vd} \xi$ along a sub-sequence $N' \subset \mathsf{N}$, for some random measure ξ on S, and we get $\xi_n f \xrightarrow{d} \xi f$ along N' for every $f \in \hat{C}_S$. The latter convergence extends by hypothesis to N, and so $\xi_n \xrightarrow{vd} \xi$ by Theorem 4.11. □

When $\xi, \xi_1, \xi_2, \dots$ are point processes and ξ is simple, $\xi_n \xrightarrow{vd} \xi$ follows already from the one-dimensional convergence $\xi_n U \xrightarrow{d} \xi U$, with U restricted to a dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\xi$. In fact, we have the following stronger result.

Theorem 4.15 *(convergence of point processes) Let $\xi, \xi_1, \xi_2, \dots$ be point processes on S, where ξ is simple, and fix a dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\xi$ and semi-ring $\mathcal{I} \subset \mathcal{U}$. Then $\xi_n \xrightarrow{vd} \xi$ iff*

(i) $P\{\xi_n U = 0\} \to P\{\xi U = 0\}, \quad U \in \mathcal{U},$

(ii) $\limsup_{n\to\infty} P\{\xi_n I > 1\} \le P\{\xi I > 1\}, \quad I \in \mathcal{I}.$

Proof: The necessity is clear from Lemma 4.12. Now assume (i) and (ii). By Lemma 1.9 (i), we may assume $\mathcal{U}$ and $\mathcal{I}$ to be countable. Define $\eta(U) = \xi U \wedge 1$ and $\eta_n(U) = \xi_n U \wedge 1$ for $U \in \mathcal{U}$. Taking successive differences in (i), we get $\eta_n \xrightarrow{d} \eta$ with respect to the product topology on $\mathsf{R}_+^{\mathcal{U}}$. Hence, by Skorohod coupling (FMP 4.30), there exist some processes $\tilde{\eta} \overset{d}{=} \eta$ and $\tilde{\eta}_n \overset{d}{=} \eta_n$, such that a.s. $\tilde{\eta}_n(U) \to \tilde{\eta}(U)$ for all $U \in \mathcal{U}$. By Lemma 1.16 we may choose $\tilde{\xi}_n \overset{d}{=} \xi_n$ and $\tilde{\xi} \overset{d}{=} \xi$, such that a.s. $\tilde{\eta}_n(U) = \tilde{\xi}_n U \wedge 1$ and $\tilde{\eta}(U) = \tilde{\xi} U$ for all $U \in \mathcal{U}$. We may then assume $\eta_n(U) \to \eta(U)$ for all $U \in \mathcal{U}$ a.s.

Now fix an $\omega \in \Omega$ such that $\eta_n(U) \to \eta(U)$ for all $U \in \mathcal{U}$, and let $U \in \mathcal{U}$ be arbitrary. By Lemma 1.9 (ii), we may choose a finite partition of U into sets $I_1, \dots, I_m \in \mathcal{I}$ with $\xi I_k \le 1$ for all $k \le m$. Then

$$\begin{aligned} \eta_n(U) \to \eta(U) &\le \xi U \\ &= \textstyle\sum_j \xi I_j = \sum_j \eta(I_j) \\ &\leftarrow \textstyle\sum_j \eta_n(I_j) \\ &\le \textstyle\sum_j \xi_n I_j = \xi_n U, \end{aligned}$$

which shows that a.s.

$$\limsup_{n\to\infty} (\xi_n U \wedge 1) \le \xi U \le \liminf_{n\to\infty} \xi_n U, \quad U \in \mathcal{U}. \tag{4}$$

Next we note that, for any $h, k \in \mathsf{Z}_+$,

$$\begin{aligned} \{k \le h \le 1\}^c &= \{h > 1\} \cup \{h < k \wedge 2\} \\ &= \{k > 1\} \cup \{h = 0,\, k = 1\} \cup \{h > 1 \ge k\}, \end{aligned}$$

where all unions are disjoint. Substituting $h = \xi I$ and $b = \xi_n I$, we get from (ii) and (4)

$$\lim_{n\to\infty} P\{\xi I < \xi_n I \wedge 2\} = 0, \quad I \in \mathcal{I}. \tag{5}$$

Now fix any $U \in \mathcal{U}$. For any partition of U into sets $I_1, \dots, I_m \in \mathcal{I}$, we have by (5)

$$\begin{aligned} \limsup_{n\to\infty} P\{\xi_n U > \xi U\} &\le \limsup_{n\to\infty} P \textstyle\bigcup_j \{\xi_n I_j > \xi I_j\} \\ &\le P \textstyle\bigcup_j \{\xi I_j > 1\}. \end{aligned}$$

By Lemma 1.9 (ii) and dominated convergence, we can make the right-hand side arbitrarily small, and so $P\{\xi_n U > \xi U\} \to 0$. Combining with (4) gives $P\{\xi_n U \ne \xi U\} \to 0$, which means that $\xi_n U \xrightarrow{P} \xi U$. Hence, $\xi_n f \xrightarrow{P} \xi f$ for every $f \in \mathcal{U}_+$, and Theorem 4.11 (i) yields $\xi_n \xrightarrow{vd} \xi$. □

The one-dimensional convergence criterion for point processes may be extended to general random measures $\xi_1, \xi_2, \dots$ with diffuse limits ξ. Here we prove a stronger result involving exponential moments, which also admits a version for point processes.

Theorem 4.16 *(exponential criteria) Let $\xi, \xi_1, \xi_2, \dots$ be random measures or point processes on S, where ξ is diffuse or simple, respectively. Fix any $t > s > 0$, along with a dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\xi$ and semi-ring $\mathcal{I} \subset \mathcal{U}$. Then $\xi_n \xrightarrow{vd} \xi$ iff*

(i) $Ee^{-t\xi_n U} \to Ee^{-t\xi U}, \quad U \in \mathcal{U}$,

(ii) $\liminf_{n\to\infty} Ee^{-s\xi_n I} \ge Ee^{-s\xi I}, \quad I \in \mathcal{I}$.

Our proof relies on the following weak bi-continuity of the Cox and thinning transforms of Lemma 3.3. A similar property of strong continuity is established in Theorem 4.33.

Lemma 4.17 *(Cox and thinning continuity) Let $\xi_1, \xi_2, \ldots$ be Cox processes on S directed by some random measures $\eta_1, \eta_2, \ldots$. Then $\xi_n \xrightarrow{vd}$ some ξ iff $\eta_n \xrightarrow{vd}$ some η, in which case ξ is distributed as a Cox process directed by η. The corresponding statement holds when the ξ_n are p-thinnings of some point processes η_n on S, for any fixed $p \in (0,1]$.*

Proof: Beginning with the case of Cox processes, suppose that $\eta_n \xrightarrow{vd} \eta$ for some random measure η on S. Then Lemma 3.1 yields $Ee^{-\xi_n f} \to Ee^{-\xi f}$ for any $f \in \hat{\mathcal{S}}_+$, where ξ is a Cox process directed by η, and so by Theorem 4.11 we have $\xi_n \xrightarrow{vd} \xi$. Conversely, suppose that $\xi_n \xrightarrow{vd} \xi$ for some point process ξ on S. Then the sequence (ξ_n) is vaguely tight, and so by Theorem 4.10 we have for any $B \in \hat{\mathcal{S}}$ and $u > 0$

$$\lim_{t \to 0} \inf_{n \geq 0} Ee^{-t\xi_n B} = \sup_{K \in \mathcal{K}} \inf_{n \geq 1} Ee^{-u\xi_n (B \setminus K)} = 1.$$

By Lemma 3.1 the same relations hold for (η_n), with t and u replaced by $t' = 1 - e^{-t}$ and $u' = 1 - e^{-u}$, and so even (η_n) is vaguely tight by Theorem 4.10, which implies that (η_n) is vaguely relatively compact in distribution. If $\eta_n \xrightarrow{vd} \eta$ along a sub-sequence, then $\xi_n \xrightarrow{vd} \xi'$ as before, where ξ' is a Cox process directed by η. Since $\xi' \stackrel{d}{=} \xi$, the distribution of η is unique by Lemma 3.3, and so the convergence $\eta_n \xrightarrow{vd} \eta$ extends to the entire sequence. The proof for p-thinnings is similar. □

Proof of Theorem 4.16: The necessity is clear from Lemma 4.12. Now assume (i) and (ii). By Lemma 1.9 (i), we may take $\mathcal{U}$ to be countable. In the point process case, define

$$p = 1 - e^{-t}, \qquad r = -\log\left\{1 - \frac{1 - e^{-s}}{1 - e^{-t}}\right\},$$

and let $\eta, \eta_1, \eta_2, \ldots$ be p-thinnings of $\xi, \xi_1, \xi_2, \ldots$, respectively. Then η is again simple, and Lemma 3.1 (iv) shows that (i) and (ii) are equivalent to

$$P\{\eta_n U = 0\} \to P\{\eta U = 0\}, \quad U \in \mathcal{U}, \tag{6}$$

$$\liminf_{n \to \infty} Ee^{-r\eta_n I} \geq Ee^{-r\eta I}, \quad I \in \mathcal{I}. \tag{7}$$

Using (6) and arguing as in the preceding proof, we may assume that a.s.

$$\liminf_{n \to \infty} \eta_n U \geq \eta U, \quad U \in \mathcal{U}. \tag{8}$$

Next, we have for any $I \in \mathcal{I}$

$$\begin{aligned} 0 &\le (1-e^{-r})\, E\Big(e^{-r\eta I};\, \eta_n I > \eta I\Big) \\ &\le E\Big(e^{-r\eta I} - e^{-r\eta_n I};\, \eta_n I > \eta I\Big) \\ &= Ee^{-r\eta I} - Ee^{-r\eta_n I} - E\Big(e^{-r\eta I} - e^{-r\eta_n I};\, \eta_n I < \eta I\Big). \end{aligned}$$

By (7) and (8) we obtain $E(e^{-r\eta I};\, \eta_n I > \eta I) \to 0$, which yields $P\{\eta_n I > \eta I\} \to 0$ since $e^{-r\eta I} > 0$. Combining with (8) gives $\eta_n I \xrightarrow{P} \eta I$ for all $\mathcal{I}$, and so $\eta_n \xrightarrow{vd} \eta$ by Theorem 4.11 (i). Hence, $\xi_n \xrightarrow{vd} \xi$ by Lemma 4.17 (ii).

In the random measure case, we may take $\eta, \eta_1, \eta_2, \ldots$ to be Cox processes directed by $t\xi, t\xi_1, t\xi_2, \ldots$ respectively. Then η is simple by Lemma 3.6 (i), and (6) and (7) remain valid with $r = -\log(1 - s/t)$. Hence, $\eta_n \xrightarrow{vd} \eta$ as before, and $\xi_n \xrightarrow{vd} \xi$ follows by Lemma 4.17 (i). $\square$

We proceed with some sufficient conditions for convergence. Given a semi-ring $\mathcal{I}$ and a set $J \in \mathcal{I}$, let $\mathcal{I}_J$ denote the class of finite partitions π of J into $\mathcal{I}$-sets I.

Theorem 4.18 *(sufficient conditions) Let ξ and $\xi_1, \xi_2, \ldots$ be point processes on S, where ξ is simple, and fix any dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\xi$ and semi-ring $\mathcal{I} \subset \mathcal{U}$. Then $\xi_n \xrightarrow{vd} \xi$ whenever*

(i) $P\{\xi_n U = 0\} \to P\{\xi U = 0\}, \quad U \in \mathcal{U},$

(ii) $\limsup_{n\to\infty} E\xi_n I \le E\xi I < \infty, \quad I \in \mathcal{I}.$

For any fixed $t > 0$, we may replace (i) *by*

(i′) $Ee^{-t\xi_n U} \to Ee^{-t\xi U}, \quad U \in \mathcal{U}.$

In the latter form, the statement remains true for any random measures ξ and $\xi_1, \xi_2, \ldots$ with ξ diffuse. The previous statements also hold, in the simple and diffuse cases, with (ii) *replaced by respectively*

(ii′) $\inf_{\pi \in \mathcal{I}_J} \limsup_{n\to\infty} \sum_{I\in\pi} P\{\xi_n I > 1\} = 0, \quad J \in \mathcal{I},$

(ii″) $\inf_{\pi \in \mathcal{I}_J} \limsup_{n\to\infty} \sum_{I\in\pi} E\{(\xi_n I)^2 \wedge 1\} = 0, \quad J \in \mathcal{I}.$

Proof: First assume (i) and (ii). By the proof of Theorem 4.15, we may assume that (4) holds a.s. with $\mathcal{U}$ countable. Writing

$$E|\,\xi_n I - \xi I| = E\xi_n - E\xi I + 2E\{\,\xi I;\, \xi_n I < \xi I\,\}, \quad I \in \mathcal{I},$$

we see from (ii) and (4) that $E|\,\xi_n I - \xi I| \to 0$, and so $\xi_n I \xrightarrow{P} \xi I$, which implies $\xi_n \xrightarrow{vd} \xi$ by Theorem 4.11 (i). Replacing (i) by (i′) and putting $p = 1 - e^{-t}$, we see that (i) and (ii) hold for some point processes $\eta, \eta_1, \eta_2, \ldots$, defined as p-thinnings of $\xi, \xi_1, \xi_2, \ldots$ or Cox processes directed by $t\xi, t\xi_1, t\xi_2, \ldots,$

respectively. As before, we get $\eta_n \xrightarrow{vd} \eta$, and $\xi_n \xrightarrow{vd} \xi$ follows by Lemma 4.17.

Now assume (i) and (ii′), which also allows us to assume (4). Fix any $J \in \mathcal{I}$ with partition $\pi \in \mathcal{I}_J$. Putting $A_n = \{\max_{I\in\pi} \xi_n I \le 1\}$, we see from (4) that

$$\begin{aligned}\limsup_{n\to\infty} 1_{A_n}\xi_n J &\le \limsup_{n\to\infty} \sum_{I\in\pi}(\xi_n I \wedge 1)\\ &\le \sum_{I\in\pi} \limsup_{n\to\infty}(\xi_n I \wedge 1)\\ &\le \sum_{I\in\pi} \xi I = \xi J.\end{aligned}$$

Writing

$$P\{\xi_n J \ne \xi J\} \le P\{\xi_n J < \xi J\} + P\{\,\xi_n J > \xi J;\, A_n\} + P(A_n^c),$$

we conclude that

$$\begin{aligned}\limsup_{n\to\infty} P\{\xi_n J \ne \xi J\} &\le \limsup_{n\to\infty} P(A_n^c)\\ &\le \limsup_{n\to\infty} \sum_{I\in\pi} P\{\xi_n I > 1\}.\end{aligned}$$

Taking the infimum over $\pi \in \mathcal{I}_J$ gives $\xi_n J \xrightarrow{P} \xi J$ by (ii′), and since J was arbitrary, we obtain $\xi_n \xrightarrow{vd} \xi$ by Theorem 4.11 (i).

Finally, let $\xi, \xi_1, \xi_2, \ldots$ be random measures with ξ diffuse, satisfying (i′) and (ii″). Then (i) holds for the Cox processes $\eta, \eta_1, \eta_2, \ldots$ directed by $t\xi, t\xi_1, t\xi_2, \ldots$, and

$$\begin{aligned}P\{\eta_n I > 1\} &= EP(\,\eta_n I > 1 \,|\, \xi_n\,)\\ &\asymp E\Big\{(t\xi_n I)^2 \wedge 1\Big\}\\ &\asymp E\Big\{(\xi_n I)^2 \wedge 1\Big\}.\end{aligned} \tag{9}$$

Hence, (ii″) is equivalent to (ii) for the sequence (η_n), and so $\eta_n \xrightarrow{vd} \eta$ as before, which implies $\xi_n \xrightarrow{vd} \xi$ by Lemma 4.17. □

For a.s. bounded random measures $\xi, \xi_1, \xi_2, \ldots$ on S, we now examine how the notions of convergence in distribution are related, for the weak and vague topologies on $\mathcal{M}_S$. Note that the weak topology on $\mathcal{M}_S$ for the metric d agrees with the vague topology on $\mathcal{M}_S$ for the bounded metric $\hat{d} = d \wedge 1$. Fixing any $s_0 \in S$, we may define an equivalent, complete metric on $S' = S \setminus \{s_0\}$, by setting

$$d'(s,t) = \hat{d}(s,t) \vee \left|\frac{1}{\hat{d}(s,s_0)} - \frac{1}{\hat{d}(t,s_0)}\right|, \quad s,t \in S',$$

so that a subset of S' is d'-bounded iff it is bounded away from s_0. Convergence in $\mathcal{M}_{S'}$ is defined with respect to the associated vague topology.

Theorem 4.19 *(weak and vague topologies) Let $\xi, \xi_1, \xi_2, \dots$ be a.s. bounded random measures on S, fix any $s_0 \in S$, and put $S' = S \setminus \{s_0\}$. Then these conditions are equivalent:*

(i) $\xi_n \xrightarrow{wd} \xi$,

(ii) $\xi_n \xrightarrow{vd} \xi$ *and* $\xi_n S \xrightarrow{d} \xi S$,

(iii) $\xi_n \xrightarrow{vd} \xi$ *and* $\inf_{B \in \hat{\mathcal{S}}} \limsup_{n\to\infty} E(\xi_n B^c \wedge 1) = 0$,

(iv) $(\xi_n, \xi_n S) \xrightarrow{vd} (\xi, \xi S)$ *in* $\mathcal{M}_{S'} \times \mathsf{R}_+$.

Proof, (i) $\Rightarrow$ (ii): This is obvious since $1 \in C_S$.

(ii) $\Rightarrow$ (iii): If (iii) fails, then a diagonal argument yields a sub-sequence $N' \subset \mathsf{N}$, such that

$$\inf_{B \in \hat{\mathcal{S}}} \liminf_{n \in N'} E(\xi_n B^c \wedge 1) > 0. \tag{10}$$

Under (ii) the sequences (ξ_n) and $(\xi_n S)$ are vaguely tight, and so the same thing is true for the sequence of pairs $(\xi_n, \xi_n S)$. Hence, we have convergence $(\xi_n, \xi_n S) \xrightarrow{vd} (\tilde{\xi}, \alpha)$ along a further sub-sequence $N'' \subset N'$. Clearly $\tilde{\xi} \overset{d}{=} \xi$, and so by Lemma 1.16 we may take $\tilde{\xi} = \xi$. Then for any $B \in \hat{\mathcal{S}}_\xi$, we get along N''

$$0 \le \xi_n B^c = \xi_n S - \xi_n B \xrightarrow{d} \alpha - \xi B,$$

and so $\xi B \le \alpha$ a.s., which implies $\xi S \le \alpha$ a.s., since B was arbitrary. Since also $\alpha \overset{d}{=} \xi S$, we get

$$\begin{aligned} E|\, e^{-\xi S} - e^{-\alpha}| &= Ee^{-\xi S} - Ee^{-\alpha} \\ &= Ee^{-\xi S} - Ee^{-\xi S} = 0, \end{aligned}$$

and so in fact $\alpha = \xi S$ a.s. Hence, for any $B \in \hat{\mathcal{S}}_\xi$, we have along N''

$$\begin{aligned} E(\xi_n B^c \wedge 1) &= E\{(\xi_n S - \xi_n B) \wedge 1\} \\ &\to E\{(\xi S - \xi B) \wedge 1\} \\ &= E(\,\xi B^c \wedge 1\,), \end{aligned}$$

which tends to 0 as $B \uparrow S$ by dominated convergence. This contradicts (10), and (iii) follows.

(iii) $\Rightarrow$ (i): Writing $K^c \subset B^c \cup (B \setminus K)$ for $K \in \mathcal{K}$ and $B \in \hat{\mathcal{S}}$, and using the additivity of P and ξ_n and the sub-additivity of $x \wedge 1$ for $x \ge 0$, we get

$$\begin{aligned} &\limsup_{n\to\infty} E(\xi_n K^c \wedge 1) \\ &\qquad \le \limsup_{n\to\infty} \Big(E(\xi_n B^c \wedge 1) + E\{\xi_n (B \setminus K) \wedge 1\} \Big) \\ &\qquad \le \limsup_{n\to\infty} E(\xi_n B^c \wedge 1) + \limsup_{n\to\infty} E\{\xi_n (B \setminus K) \wedge 1\}. \end{aligned}$$

Taking the infimum over $K \in \mathcal{K}$ and then letting $B \uparrow S$, we get by Theorem 4.10 and (iii)

$$\inf_{K \in \mathcal{K}} \limsup_{n\to\infty} E(\xi_n K^c \wedge 1) = 0. \tag{11}$$

Similarly, for any $r > 1$,

$$\begin{aligned}\limsup_{n\to\infty} & P\{\xi_n S > r\} \\ &\le \limsup_{n\to\infty} \Big(P\{\xi_n B > r-1\} + P\{\xi_n B^c > 1\} \Big) \\ &\le \limsup_{n\to\infty} P\{\xi_n B > r-1\} + \limsup_{n\to\infty} E(\xi_n B^c \wedge 1).\end{aligned}$$

Letting $r \to \infty$ and then $B \uparrow S$, we get by Theorem 4.10 and (iii)

$$\lim_{r\to\infty} \liminf_{n\to\infty} P\{\xi_n S > r\} = 0. \tag{12}$$

By Theorem 4.10, we see from (11) and (12) that (ξ_n) is weakly tight. Assuming $\xi_n \xrightarrow{wd} \eta$ along a sub-sequence, we have also $\xi_n \xrightarrow{vd} \eta$, and so $\eta \stackrel{d}{=} \xi$. Since the sub-sequence was arbitrary, (i) follows by Theorem 4.2.

(i) $\Rightarrow$ (iv): This holds by continuity.

(iv) $\Rightarrow$ (i): By continuous mapping, we may take $\xi, \xi_1, \xi_2, \ldots$ to be non-random. Fix any $f \in C_S$, and write $f = f(s_0) + f_+ - f_-$, where $f_\pm \in C_S$ with $f_\pm(s_0) = 0$. Letting $g_\pm \in \hat{C}_{S'}$ and using (iv), we get

$$\limsup_{n\to\infty} |\, \xi_n f - \xi f| \le 2\Big(\|f_+ - g_+\| + \|f_- - g_-\| \Big)\, \xi S,$$

which can be made arbitrarily small, by suitable choices of $g_\pm$. □

For random measures on R_+, we define $\xi_n \xrightarrow{Sd} \xi$ by $X_n \xrightarrow{d} X$ for the Skorohod J_1-topology (FMP A2.2), where $X, X_1, X_2, \ldots$ are the cumulative processes on R_+, given by $X(t) = \xi[0,t]$ and $X_n(t) = \xi_n[0,t]$ for $t \ge 0$. The following result relates the notions of convergence in distribution, for the vague and Skorohod topologies.

Theorem 4.20 *(vague and Skorohod convergence) Let $\xi, \xi_1, \xi_2, \ldots$ be random measures on R_+. Then $\xi_n \xrightarrow{Sd} \xi$ implies $\xi_n \xrightarrow{vd} \xi$, and the two statements are equivalent under each of these conditions:*

(i) *$\xi, \xi_1, \xi_2, \ldots$ are point processes, and ξ is a.s. simple with $\xi\{0\} = 0$,*

(ii) *ξ is diffuse.*

Note that the converse assertion fails without additional assumptions on ξ or $\xi_1, \xi_2, \ldots$. For example, choosing $\xi = 2\delta_1$ and $\xi_n = \delta_1 + \delta_{1+1/n}$, we have $\xi_n \xrightarrow{vd} \xi$, whereas $\xi_n \stackrel{Sd}{\not\to} \xi$.

Proof: By continuous mapping (FMP 4.27), we may take the measures $\mu = \xi$ and $\mu_n = \xi_n$ to be non-random. First assume the Skorohod convergence $\mu_n \xrightarrow{S} \mu$. To prove $\mu_n \xrightarrow{v} \mu$, we need to show that $\mu_n f \to \mu f$ for any continuous function $f : \mathsf{R}_+ \to \mathsf{R}_+$ with bounded support. The convergence $\mu_n \xrightarrow{S} \mu$ yields $\mu_n f_n \to \mu f$, where $f_n = f \circ h_n$ for some increasing bijections h_n on R_+ with $h_n(t) \to t$ for each $t > 0$. Here the f_n have uniformly bounded supports, and $\|f_n - f\| \to 0$ by uniform continuity and convergence. Hence,

$$|\mu_n f - \mu f| \le \mu_n |f - f_n| + |\mu_n f - \mu f| \to 0.$$

Conversely, let $\mu, \mu_1, \mu_2, \ldots$ be point measures with $\mu_n \xrightarrow{v} \mu$, and assume μ to be simple with $\mu\{0\} = 0$. Writing $\mu = \sum_k \delta_{\tau_k}$ and $\mu_n = \sum_k \delta_{\tau_{nk}}$ for some finite or infinite, non-decreasing sequences (τ_k) and (τ_{nk}), $n \in \mathsf{N}$, we obtain $\tau_{nk} \to \tau_k$, as $n \to \infty$ for fixed k. In particular, the μ_n are eventually simple on every fixed, bounded interval. For each $n \in \mathsf{N}$, define $h_n(\tau_k) = \tau_{nk}$ for all $k \ge 0$, where $\tau_0 = \tau_{n0} = 0$, and extend h_n to R_+ by linear interpolation. (When $\|\mu\| < \infty$, we may choose $h'_n = 1$ after the last point τ_k.) Then $\mu_n = \mu \circ h_n^{-1}$, and since $h_n(t) \to t$ for all $t > 0$, we get $\mu_n \xrightarrow{S} \mu$.

Finally, if $\mu_n \xrightarrow{v} \mu$ with μ diffuse, then $\mu_n[0,t] \to \mu[0,t]$ for all $t \ge 0$. Since the limit is continuous in t, the convergence is uniform on bounded intervals, and so $\mu_n \xrightarrow{S} \mu$. □

For special purposes in Chapter 12, we also need to consider weak convergence of distributions on the Skorohod space $D = D(\mathsf{R}_+, S)$ with respect to the *uniform topology*, where S is a separable and complete metric space. Since D is not separable in the uniform topology, the standard weak convergence theory does not apply to this case. Note the distinction from the uniform convergence of distributions on D, considered in Section 4.4.

For suitable subsets $U \subset \mathsf{R}_+$, write $\|\cdot\|_U$ for the supremum or total variation on U. A function $f : D \to \mathsf{R}$ is said to be continuous for the uniform topology on D, if $\|x_n - x\| \to 0$ in D implies $f(x_n) \to f(x)$ in R. For probability measures μ_n and μ on D, the weak convergence $\mu_n \xrightarrow{uw} \mu$ means that $\mu_n f \to \mu f$ for every such function.

Lemma 4.21 *(weak convergence for the uniform topology) Let $X, X_1, X_2,$ $\ldots$ be rcll processes with distributions $\mu, \mu_1, \mu_2, \ldots,$ in a separable and complete metric space (S, ρ). Fix an element $a \in S$ and a class $\mathcal{U}$ of finite interval unions in R_+. Then $\mu_n \xrightarrow{uw} \mu$ holds under these conditions:*

(i) $\lim_{n\to\infty} \|\mu_n - \mu\|_U = 0, \quad U \in \mathcal{U},$

(ii) $\inf_{U \in \mathcal{U}} \limsup_{n\to\infty} E\|\rho(X_n, a) \wedge 1\|_{U^c} = 0,$

(iii) $\inf_{U \in \mathcal{U}} E\|\rho(X, a) \wedge 1\|_{U^c} = 0.$

Proof: Assume (i)−(iii), take $\rho \le 1$ for simplicity, and write $\rho(x) = \rho(x,s)$. For any $\varepsilon > 0$, we may choose $U \in \mathcal{U}$ such that

$$\limsup_{n\to\infty} E\|\rho(X_n)\|_U < \varepsilon, \qquad E\|\rho(X)\| < \varepsilon.$$

Keeping U fixed, we may next choose $m \in \mathsf{N}$ so large that

$$\|\mu_n - \mu\|_U < \varepsilon, \quad E\|\rho(X_n)\|_{U^c} < \varepsilon, \qquad n > m.$$

By an elementary construction, there exist some processes $Y \stackrel{d}{=} \pi_U X$ and $Y_n \stackrel{d}{=} \pi_U X_n$ on U, such that $P\{Y = Y_n\} > 1 - \varepsilon$ for all $n > m$, where π_U denotes restriction to U. By a transfer argument we may finally choose some processes $X_n' \stackrel{d}{=} X_n$ with $\pi_U(X, X_n') \stackrel{d}{=} (Y, Y_n)$. Using the triangle inequality for both the supremum norm and the metric ρ, we get for any $n > m$

$$E\|\rho(X, X_n')\| \le E\|\rho(X, X_n')\|_U + E\|\rho(X)\|_{U^c} + E\|\rho(X_n')\|_{U^c} < 3\varepsilon.$$

Repeating the construction for an arbitrary sequence $\varepsilon_n \to 0$, we obtain some processes $\tilde{X}_n \stackrel{d}{=} X_n$ satisfying $E\|\rho(X, \tilde{X}_n)\| \to 0$, which implies $\|\rho(X, \tilde{X}_n)\| \xrightarrow{P} 0$.

Now consider any bounded, measurable function f on D, which is continuous with respect to the uniform topology. Turning to a.s. convergent sub-sequences, we get by continuity and dominated convergence

$$\mu_n f = Ef(\tilde{X}_n) \to Ef(X) = \mu f. \qquad \square$$

4.3 Null Arrays and Infinite Divisibility

Given a complete, separable metric space S, an infinitely divisible random measure or point process ξ on S is said to be *directed* by (α, λ) or λ, respectively, if it can be represented as in Theorem 3.20. In that case, we define $\hat{\mathcal{S}}_\lambda$ as the class of sets $B \in \hat{\mathcal{S}}$ with $\lambda\{\mu;\, \mu\partial B > 0\} = 0$. A similar terminology and notation is used for random measures with independent increments. The random measures ξ_{nj} on S, $n, j \in \mathsf{N}$, are said to form a *null array*, if they are independent in j for fixed n and satisfy

$$\lim_{n\to\infty} \sup_{j\in\mathsf{N}} E(\xi_{nj}B \wedge 1) = 0, \quad B \in \hat{\mathcal{S}}.$$

In the point process case, we have clearly $E(\xi_{nj}B \wedge 1) = P\{\xi_{nj}B > 0\}$. For any $\mathbf{I} = (I_1, \dots, I_m)$, write $\pi_{\mathbf{I}}\mu = \mu\mathbf{I} = (\mu I_1, \dots, \mu I_m)$.

We begin with some general convergence criteria for null arrays of random measures. In the point process case, a similar criterion for convergence in total variation is given in Corollary 4.39.

Theorem 4.22 *(convergence in null arrays) Let (ξ_{nj}) be a null array of random measures on S. If $\sum_j \xi_{nj} \xrightarrow{vd} \xi$, then ξ is infinitely divisible, and convergence holds with ξ directed by (α, λ) iff*

(i) $\sum_j (1 - Ee^{-\xi_{nj} f}) \to \alpha f + \int (1 - e^{-\mu f})\, \lambda(d\mu), \quad f \in \hat{C}_S.$

For any dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}_\alpha \cap \hat{\mathcal{S}}_\lambda$, this holds iff

(ii) $\sum_j \mathcal{L}(\xi_{nj}\mathbf{I}) \xrightarrow{v} \lambda \circ \pi_{\mathbf{I}}^{-1}$ *on* $\overline{\mathsf{R}^m} \setminus \{0\}, \quad \mathbf{I} \in \mathcal{I}^m,\ m \in \mathsf{N},$

(iii) $\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \left| \alpha I - \sum_j E(\xi_{nj} I;\, \xi_{nj} I < \varepsilon) \right| = 0, \quad I \in \mathcal{I}.$

For point processes ξ_{nj}, it is also equivalent that

(ii′) $\sum_j \mathcal{L}(\xi_{nj}\mathbf{I}) \xrightarrow{w} \lambda \circ \pi_{\mathbf{I}}^{-1}$ *on* $\mathsf{Z}^m \setminus \{0\}, \quad \mathbf{I} \in \mathcal{I}^m,\ m \in \mathsf{N}.$

For the proof, we first reduce the statement to a limit theorem for infinitely divisible distributions. Given any random measure ξ on S, we may construct an *associated compound Poisson* random measure $\tilde{\xi} = \xi_1 + \cdots + \xi_\kappa$, where $\xi_1, \xi_2, \ldots$ are i.i.d. random measures with distribution $\mathcal{L}(\xi)$, and κ is an independent Poisson random variable with $E\kappa = 1$, so that $\tilde{\xi}$ becomes infinitely divisible with directing pair $\{0, \mathcal{L}(\xi)\}$. Similarly, given a null array (ξ_{nj}) of random measures on S, we may form an associated compound Poisson array $(\tilde{\xi}_{nj})$, where the associated compound Poisson random measures $\tilde{\xi}_{nj}$ are chosen to be independent in j for each n. For any random measures $\xi_1, \xi_2, \ldots$ and $\eta_1, \eta_2, \ldots$, the relation $\xi_n \overset{vd}{\sim} \eta_n$ means that ξ_n and η_n converge simultaneously in distribution to a common limit $\xi \overset{d}{=} \eta$.

The following result gives the desired equivalence, in the sense of weak convergence. For null arrays of point processes, a similar strong equivalence is given by Lemma 4.35.

Lemma 4.23 *(compound Poisson approximation) Let (ξ_{nj}) be a null array of random measures on S, with an associated compound Poisson array $(\tilde{\xi}_{nj})$. Then $\sum_j \xi_{nj} \overset{vd}{\sim} \sum_j \tilde{\xi}_{nj}$.*

Proof: Write $\xi_n = \sum_j \xi_{nj}$ and $\tilde{\xi}_n = \sum_j \tilde{\xi}_{nj}$. By Theorem 4.11 it is enough to show that $Ee^{-\xi_n f} \sim Ee^{-\tilde{\xi}_n f}$ for every $f \in \hat{C}_S$, in the sense that convergence of either side implies that both sides converge to the same limit. Now Lemma 3.1 (ii) yields

$$Ee^{-\tilde{\xi}_n f} = \prod_j Ee^{-\tilde{\xi}_{nj} f} = \prod_j \exp\Big(Ee^{-\xi_{nj} f} - 1\Big).$$

Writing $p_{nj} = 1 - Ee^{-\xi_{nj} f}$, we need to show that $\prod_j e^{-p_{nj}} \sim \prod_j (1 - p_{nj})$. Since $\sup_j p_{nj} \to 0$ by the "null" property of $(\xi_{nj} f)$ (FMP 5.6), this holds by a first order Taylor expansion of $\sum_j \log(1 - p_{nj})$. □

The last lemma reduces the proof of Theorem 4.22 to the corresponding criteria for infinitely divisible random measures. For point processes, a similar strong result will appear as Theorem 4.38.

Lemma 4.24 *(infinitely divisible random measures) For every $n \in \mathsf{N}$, let ξ_n be an infinitely divisible random measure on S directed by (α_n, λ_n). If $\xi_n \xrightarrow{vd} \xi$, then ξ is again infinitely divisible, and convergence holds with ξ directed by (α, λ) iff*

$$\text{(i)} \quad \alpha_n f + \int (1 - e^{-\mu f})\, \lambda_n(d\mu) \to \alpha f + \int (1 - e^{-\mu f})\, \lambda(d\mu), \quad f \in \hat{C}_S.$$

For any dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}_\alpha \cap \hat{\mathcal{S}}_\lambda$, this holds iff

$$\text{(ii)} \quad \lambda_n \circ \pi_{\mathbf{I}}^{-1} \xrightarrow{v} \lambda \circ \pi_{\mathbf{I}}^{-1} \ \text{on } \overline{\mathsf{R}^m} \setminus \{0\}, \quad \mathbf{I} \in \mathcal{I}^m,\ m \in \mathsf{N},$$

$$\text{(iii)} \quad \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \left| \alpha I - \alpha_n I - \int_{[0,\varepsilon)} r\, \lambda_n\{\mu I \in dr\} \right| = 0, \quad I \in \mathcal{I}.$$

For point processes ξ_n, it is also equivalent that

$$\text{(ii}'\text{)} \quad \lambda_n \circ \pi_{\mathbf{I}}^{-1} \xrightarrow{w} \lambda \circ \pi_{\mathbf{I}}^{-1} \ \text{on } \mathsf{Z}^m \setminus \{0\}, \quad \mathbf{I} \in \mathcal{I}^m,\ m \in \mathsf{N}.$$

Proof: Suppose that $\xi_n \xrightarrow{vd} \xi$. Then (ξ_n) is tight by Theorem 4.2, and so it satisfies the conditions in Theorem 4.10. Now fix any $m \in \mathsf{N}$. Since the ξ_n are infinitely divisible, we have $\xi_n \stackrel{d}{=} \xi_{n1} + \cdots + \xi_{nm}$ for each n, where the ξ_{nj} are i.i.d. in j for fixed n. From Theorem 4.10 we see that (ξ_{nj}) is again tight in n for fixed j. The tightness extends immediately to the sequence of m-tuples $(\xi_{n1}, \ldots, \xi_{nm})$, and so by Theorem 4.2 we have convergence $(\xi_{n1}, \ldots, \xi_{nm}) \xrightarrow{vd} (\eta_1, \ldots, \eta_m)$ along a sub-sequence, where the η_j are again i.i.d. Hence, $\xi_n \stackrel{d}{=} \sum_j \xi_{nj} \xrightarrow{vd} \sum_j \eta_j$, which implies $\xi \stackrel{d}{=} \sum_j \eta_j$. The infinite divisibility of ξ now follows since m was arbitrary.

If ξ is directed by (α, λ), then by Theorem 3.20 and Lemma 3.1 (ii),

$$-\log E e^{-\xi f} = \alpha f + \int (1 - e^{-\mu f})\, \lambda(d\mu), \quad f \in \hat{\mathcal{S}}_+, \tag{13}$$

and similarly for each ξ_n in terms of (α_n, λ_n). By Theorem 4.11 we have $\xi_n \xrightarrow{vd} \xi$ iff $\xi_n f \xrightarrow{d} \xi f$ for all $f \in \hat{C}_S$ or all $f \in \hat{\mathcal{I}}_+$, which is equivalent to (i) for functions f in either class, by the continuity theorem for Laplace transforms, applied to the random variables $\xi_n f$ and ξf with f fixed.

Now assume (i) on $\mathcal{I}_+$. Fix any $\mathbf{I} = (I_1, \ldots, I_m) \in \mathcal{I}^m$, $m \in \mathsf{N}$, and let $s, t \in \mathsf{R}_+^m$ be arbitrary. Taking differences for the functions $f = \sum_k (s_k + t_k) 1_{I_k}$ and $g = \sum_k t_k 1_{I_k}$ and writing $s\mu\mathbf{I} = \sum_k s_k \mu I_k$ and $t\mu\mathbf{I} = \sum_k t_k \mu I_k$, we get

$$\begin{aligned} s\alpha_n\mathbf{I} + \int e^{-t\mu\mathbf{I}}\, (1 - e^{-s\mu\mathbf{I}})\, \lambda_n(d\mu) \\ \to s\alpha\mathbf{I} + \int e^{-t\mu\mathbf{I}}\, (1 - e^{-s\mu\mathbf{I}})\, \lambda(d\mu). \end{aligned}$$

By the continuity theorem for multivariate Laplace transforms (FMP 5.3), we obtain on R_+^m

$$\begin{aligned} (s\alpha_n\mathbf{I})\, \delta_0 + (1 - e^{-\pi_s}) \cdot (\lambda_n \circ \pi_{\mathbf{I}}^{-1}) \\ \xrightarrow{w} (s\alpha\mathbf{I})\, \delta_0 + (1 - e^{-\pi_s}) \cdot (\lambda \circ \pi_{\mathbf{I}}^{-1}), \end{aligned} \tag{14}$$

where $\pi_s x = sx$ for $s, x \in \mathsf{R}_+^m$. Taking $s = (1, \ldots, 1)$ gives (ii), whereas (iii) is obtained for $m = s = 1$. Conversely, (ii) and (iii) imply (14) for any $\mathbf{I}$ and s as above, and (i) follows for $f = \sum_k s_k 1_{I_k}$. In the point process case we have $\alpha_n = \alpha = 0$, and (14) simplifies to (ii′). $\Box$

Specializing to Poisson limits gives the following classical result. Here a strong counterpart will appear in Theorem 4.34.

Corollary 4.25 *(Poisson convergence, Grigelionis) Let (ξ_{nj}) be a null array of point processes on S, let ξ be a Poisson process on S with $E\xi = \rho$, and fix any dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\rho$ and semi-ring $\mathcal{I} \subset \mathcal{U}$. Then $\sum_j \xi_{nj} \xrightarrow{vd} \xi$ iff*

(i) $\sum_j P\{\xi_{nj} I = 1\} \to \rho I, \quad I \in \mathcal{I},$

(ii) $\sum_j P\{\xi_{nj} U > 1\} \to 0, \quad U \in \mathcal{U}.$

Proof: Fix any disjoint sets $I_1, \ldots, I_m \in \mathcal{I}$, and put $U = \bigcup_k I_k$. Writing $\lambda_n = \sum \mathcal{L}(\xi_{nj})$, we get for any $k \le m$

$$\begin{aligned}\lambda_n\{\mu I_k = \mu U = 1\} &= \lambda_n\{\mu I_k = 1\} - \lambda_n\{\mu I_k = 1 < \mu U\} \to \rho I_k,\\ \lambda_n\Big\{\textstyle\sum_k \mu I_k > 1\Big\} &= \lambda_n\{\mu U > 1\} \to 0,\end{aligned}$$

which proves (ii′) of Theorem 4.22 with $\lambda = \rho\{s;\, \delta_s \in \cdot\}$. $\Box$

In the special case of limits with independent increments, Theorem 4.22 simplifies as follows. For any classes $\mathcal{A} \subset \mathcal{B}$, we mean by $\mathcal{A} \ll \mathcal{B}$ that every $A \in \mathcal{A}$ is contained in some $B \in \mathcal{B}$.

Theorem 4.26 *(limits with independent increments) Let (ξ_{nj}) be a null array of random measures on S, let ξ be an infinitely divisible random measure with independent increments directed by (α, ρ), and fix any dissecting semi-ring $\mathcal{I} \subset \hat{\mathcal{S}}_\alpha \cap \hat{\mathcal{S}}_\rho$ with $\hat{\mathcal{S}} \ll \mathcal{I}$. Then $\sum_j \xi_{nj} \xrightarrow{vd} \xi$ iff $\sum_j \xi_{nj} I \xrightarrow{d} \xi I$ for all $I \in \mathcal{I}$, and also iff*

(i) $\sum_j \mathcal{L}(\xi_{nj} I) \xrightarrow{v} \rho(I \times \cdot)$ *on* $(0, \infty], \quad I \in \mathcal{I},$

(ii) $\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \Big| \alpha I - \sum_j E(\xi_{nj} I;\, \xi_{nj} I < \varepsilon) \Big| = 0, \quad I \in \mathcal{I}.$

For point processes ξ_{nj} and ξ, it is further equivalent that

(i′) $\sum_j \mathcal{L}(\xi_{nj} I) \xrightarrow{w} \rho(I \times \cdot)$ *on* $\mathsf{N}, \quad I \in \mathcal{I}.$

Proof: If $\sum_j \xi_{nj} \xrightarrow{vd} \xi$, then Lemma 4.12 yields $\sum_j \xi_{nj} I \xrightarrow{vd} \xi I$ for all $I \in \mathcal{I}$, which in turn implies (i) and (ii) by Theorem 4.22 for singleton sets S. By the same theorem, it remains to show that (i) and (ii) imply condition

(ii) of Theorem 4.22, with $\lambda = \rho\{(s,r);\, r\delta_s \in \cdot\}$. By (i) it is then enough to show that, for any disjoint sets $I_1, I_2 \in \mathcal{I}$,

$$\lambda_n\{\mu;\, \mu I_1 \wedge \mu I_2 \ge \varepsilon\} \to 0, \quad \varepsilon > 0, \tag{15}$$

where $\lambda_n = \sum_j \mathcal{L}(\xi_{nj})$. Since $\mathcal{I} \gg \hat{\mathcal{S}}$, we may choose an $I \in \mathcal{I}$ with $I_1, I_2 \subset I$, and since $\mathcal{I}$ is a semi-ring, we may write $I = \bigcup_k I_k$ with $I_1, \ldots, I_m \in \mathcal{I}$ disjoint. Put $T = \{t > 0;\, \rho(I \times \{t\}) = 0\}$, and note that T is dense in $(0, \infty)$. When $\mu I_1 \wedge \mu I_2 > 0$, we have $\max_k \mu I_k < \mu I$, and we may choose a $t \in T$ with $\max_k \mu I_k < t < \mu I$. By compactness in $\overline{\mathsf{R}}_+^m \setminus \{0\}$, the μ-set in (15) is covered by finitely many sets $M_t = \{\mu;\, \max_k \mu I_k < t < \mu I\}$, and it suffices to show that $\lambda_n M_t \to 0$, as $n \to \infty$ for fixed t. Then note that

$$\begin{aligned}
t\,\lambda_n M_t &\le \int \Big(\mu I;\, \max_k \mu I_k < t,\, \mu I \ge t\Big)\, \lambda_n(d\mu)\\
&\le \textstyle\sum_h \displaystyle\int \Big(\mu I_h;\, \mu I_h < t,\, \mu I \ge t\Big)\, \lambda_n(d\mu)\\
&= \textstyle\sum_h \displaystyle\int \Big\{(\mu I_h;\, \mu I_h < t) - (\mu I_h;\, \mu I < t)\Big\}\, \lambda_n(d\mu)\\
&= \textstyle\sum_h \displaystyle\int (\mu I_h;\, \mu I_h < t)\, \lambda_n(d\mu) - \int (\mu I;\, \mu I < t)\, \lambda_n(d\mu)\\
&\to \textstyle\sum_h \displaystyle\Big\{\alpha I_h + \int_0^t r\,\rho(I_h \times dr)\Big\} - \alpha I - \int_0^t r\,\rho(I \times dr) = 0,
\end{aligned}$$

where the convergence follows from (i) and (ii). In the point process case, (ii) is void and (i) reduces to (i′). □

We turn to a version of Theorem 4.15 for null arrays.

Corollary 4.27 *(point process case) Let (ξ_{nj}) be a null array of point processes on S, let ξ be an infinitely divisible point process on S with Lévy measure λ restricted to $\mathcal{N}_S^*$, and fix any dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\lambda$ and semi-ring $\mathcal{I} \subset \mathcal{U}$. Then $\sum_j \xi_{nj} \xrightarrow{vd} \xi$ iff*

(i) $\sum_j P\{\xi_{nj} U > 0\} \to \lambda\{\mu;\, \mu U > 0\}, \quad U \in \mathcal{U}$,

(ii) $\limsup_{n\to\infty} \sum_j P\{\xi_{nj} I > 1\} \le \lambda\{\mu;\, \mu I > 1\}, \quad I \in \mathcal{I}$.

Proof: Assume (i) and (ii), and put $\lambda_n = \sum_j \mathcal{L}(\xi_{nj})$. By Theorem 4.22, we need to show that

$$\lambda_n \circ \pi_{\mathbf{I}}^{-1} \xrightarrow{w} \lambda \circ \pi_{\mathbf{I}}^{-1} \text{ on } \mathsf{Z}^m \setminus \{0\}, \quad \mathbf{I} \in \mathcal{I}^m,\ m \in \mathsf{N}. \tag{16}$$

When considering subsets $I_1, \ldots, I_m \in \mathcal{I}$ of some fixed set $U \in \mathcal{U}$, we may assume that ξ and all ξ_{nj} are restricted to U. Then $\lambda, \lambda_1, \lambda_2, \ldots$ become uniformly bounded on $\mathcal{N}_S \setminus \{0\}$, hence allowing extensions to bounded measures on $\mathcal{N}_S$ with a common total mass. Thus, (16) holds by Theorem 4.15. □

We may now prove a version for null arrays of Theorem 4.16:

Theorem 4.28 *(exponential criteria) Let (ξ_{nj}) be a null array of random measures or point processes on S, and let ξ be an infinitely divisible random measure or point process on S directed by (α,λ) or λ, where λ is restricted to $\mathcal{M}_S^*$ or $\mathcal{N}_S^*$, respectively. Fix any $t > s > 0$, along with a dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\lambda \cap \hat{\mathcal{S}}_\alpha$ and semi-ring $\mathcal{I} \subset \mathcal{U}$. Then $\sum_j \xi_{nj} \xrightarrow{vd} \xi$ iff*

(i) $\sum_j (1 - Ee^{-t\xi_{nj}U}) \to t\alpha U + \int (1 - e^{-t\mu U})\, \lambda(d\mu), \quad U \in \mathcal{U},$

(ii) $\limsup_{n\to\infty} \sum_j (1 - Ee^{-s\xi_{nj}I}) \le s\alpha I + \int (1 - e^{-s\mu I})\, \lambda(d\mu), \quad I \in \mathcal{I}.$

Proof: As before, we may consider the restrictions to a fixed set $U \in \mathcal{U}$. In the point process case, we have

$$\begin{aligned}(1 - e^{-t})\lambda\{\mu U > 0\} &\le \int (1 - e^{-t\mu U})\, \lambda(d\mu)\\ &\le \lambda\{\mu U > 0\},\end{aligned}$$

and similarly with λ replaced by $\lambda_n = \sum_n \mathcal{L}(\xi_{nj})$. By (i) the measures λ and λ_n are then uniformly bounded, and the assertion follows as before from Theorem 4.16.

For general random measures, it suffices by Lemma 4.23 to show that (i) and (ii) imply $\xi_n \xrightarrow{vd} \xi$, where the ξ_n are infinitely divisible and directed by the pairs $(0, \lambda_n)$. Then fix any $u \in (t, \infty)$, let $\tilde{\xi}$ and $\tilde{\xi}_n$ be Cox processes directed by $u\xi$ and $u\xi_n$, respectively, and let $\tilde{\lambda}$ and $\tilde{\lambda}_n$ denote the associated Lévy measures. Writing $\tilde{s} = -\log(1 - s/u)$ and $\tilde{t} = -\log(1 - t/u)$, and using (13) and Lemma 3.1, we get by Corollary 3.22 for any $B \in \hat{\mathcal{S}}$

$$\int (1 - e^{-\tilde{s}\mu B})\, \tilde{\lambda}(d\mu) = s\,\alpha B + \int (1 - e^{-s\mu B})\, \lambda(d\mu),$$

and similarly with $\tilde{s}$ replaced by $\tilde{t}$ and/or $\tilde{\lambda}$ replaced by $\tilde{\lambda}_n$. Hence, $\tilde{\xi}, \tilde{\xi}_1, \tilde{\xi}_2, \ldots$ satisfy suitable counterparts of (i) and (ii), and so the point process version yields $\tilde{\xi}_n \xrightarrow{vd} \tilde{\xi}$, which implies $\xi_n \xrightarrow{vd} \xi$ by Lemma 4.17 (i). □

We can also give a version of Theorem 4.18 for null arrays:

Theorem 4.29 *(sufficient conditions) Let (ξ_{nj}) be a null array of point processes on S, let ξ be an infinitely divisible point process on S with Lévy measure λ restricted to $\mathcal{N}_S^*$, and fix any dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}_\xi$ and semi-ring $\mathcal{I} \subset \mathcal{U}$. Then $\sum_j \xi_{nj} \xrightarrow{vd} \xi$ whenever*

(i) $\sum_j P\{\xi_{nj}U > 0\} \to \lambda\{\mu;\, \mu U > 0\}, \quad U \in \mathcal{U},$

(ii) $\limsup_{n\to\infty} \sum_j E\xi_{nj} I \le \int \mu(I)\, \lambda(d\mu), \quad I \in \mathcal{I}.$

For any $t > 0$, we may replace (i) *by*

(i′) $\sum_j (1 - Ee^{-t\xi_{nj}U}) \to \int (1 - e^{-t\mu U})\, \lambda(d\mu), \quad U \in \mathcal{U}.$

For general random measures ξ_{nj}*, and for* ξ *directed by* (α, λ) *with* λ *restricted to* $\mathcal{M}_S^*$*, the latter statement holds with* αI *and* $t\alpha U$ *added on the right of* (ii) *and* (i′)*, respectively. The previous statements remain true in the simple and diffuse cases, with* (ii) *replaced by respectively*

(ii′) $\displaystyle \inf_{\pi \in \mathcal{I}_J} \limsup_{n\to\infty} \sum_{I \in \pi} \sum_{j \in \mathbb{N}} P\{\xi_{nj} I > 1\} = 0, \quad J \in \mathcal{I},$

(ii″) $\displaystyle \inf_{\pi \in \mathcal{I}_J} \limsup_{n\to\infty} \sum_{I \in \pi} \sum_{j \in \mathbb{N}} E[(\xi_{nj} I)^2 \wedge 1] = 0, \quad J \in \mathcal{I}.$

Proof: Putting $\lambda_n = \sum_j \mathcal{L}(\xi_{nj})$, we may rewrite the displayed conditions in the form

(i) $\lambda_n\{\mu U > 0\} \to \lambda\{\mu U > 0\}, \quad U \in \mathcal{U},$

(ii) $\displaystyle \limsup_{n\to\infty} \lambda_n \pi_I \le \lambda \pi_I, \quad I \in \mathcal{I},$

(i′) $\lambda_n(1 - e^{-t\pi_U}) \to \lambda(1 - e^{-t\pi_U}), \quad U \in \mathcal{U},$

(ii′) $\displaystyle \inf_{\pi \in \mathcal{I}_J} \limsup_{n\to\infty} \sum_{I \in \pi} \lambda_n\{\mu I > 1\} = 0, \quad J \in \mathcal{I},$

(ii″) $\displaystyle \inf_{\pi \in \mathcal{I}_J} \limsup_{n\to\infty} \sum_{I \in \pi} \lambda_n(\pi_I^2 \wedge 1) = 0, \quad J \in \mathcal{I}.$

In the point process case, we may assume again that λ and all λ_n are uniformly bounded, which allows extensions to bounded measures on $\mathcal{N}_S$ with a common total mass. The desired convergence then follows by Theorems 4.18 and 4.22.

For general random measures, let $\xi_1, \xi_2, \ldots$ be infinitely divisible random measures directed by the pairs $(0, \lambda_n)$, and choose some Cox processes $\tilde\xi, \tilde\xi_1, \tilde\xi_2, \ldots$ directed by $t\xi, t\xi_1, t\xi_2, \ldots$ respectively, with Lévy measures $\tilde\lambda, \tilde\lambda_1, \tilde\lambda_2, \ldots$. Then Corollary 3.22 yields

$$\tilde\lambda\{\mu U > 0\} = t\alpha U + \lambda(1 - e^{-t\pi_U}), \quad U \in \mathcal{U},$$

and similarly for $\tilde\lambda_n$ with $\alpha_n = 0$. Hence, (i′) turns into condition (i) for the measures $\tilde\lambda$ and $\tilde\lambda_n$. Since also $\tilde\lambda \pi_I = t\lambda\pi_I$ and $\tilde\lambda_n \pi_I = t\lambda_n \pi_I$, we conclude as before that $\tilde\xi_n \xrightarrow{vd} \tilde\xi$, which implies $\xi_n \xrightarrow{vd} \xi$. Finally, the same Cox transformation converts (ii″) into condition (ii′) for the measures $\tilde\lambda_n$, by Corollary 3.22 and the estimate (9) applied to each λ_n. □

We turn to a relationship between the supports of an infinitely divisible distribution and its Lévy measure. A similar result for the atomic support was considered in Theorem 3.23.

Theorem 4.30 *(vague support) Let* ξ *be an infinitely divisible random measure on* S *directed by* (α, λ)*. Then*

$$\operatorname{supp} \mathcal{L}(\xi) = \alpha + \overline{\bigcup_{n \ge 0} (\operatorname{supp} \lambda)^{\oplus n}}.$$

Proof: Since the mapping $\mu \mapsto \mu + \alpha$ is vaguely bi-continuous, we have

$$\operatorname{supp} \mathcal{L}(\xi) = \alpha + \operatorname{supp} \mathcal{L}(\xi - \alpha),$$

and we may henceforth assume that $\alpha = 0$. For the remaining relation, we begin with the case where $0 < \|\lambda\| < \infty$.

First let μ belong to the closure on the right. Since supports are closed, in order to prove $\mu \in \operatorname{supp} \mathcal{L}(\xi)$, we may assume that $\mu \in (\operatorname{supp} \lambda)^{\oplus n}$ for some $n \in \mathsf{Z}_+$. Thus, let $\mu = \mu_1 + \cdots + \mu_n$, where $\mu_1, \ldots, \mu_n \in \operatorname{supp} \lambda$. Let $\eta_1, \eta_2, \ldots$ be i.i.d. random measures with distribution $\lambda/\|\lambda\|$, and let κ be an independent Poisson random variable with mean $\|\lambda\|$, so that $\xi \overset{d}{=} S_\kappa$ by Theorem 3.4, where $S_k = \eta_1 + \cdots + \eta_k$. Suppose that $\mu \in G$ for some vaguely open set $G \subset \mathcal{M}_S$. Since addition in $\mathcal{M}_S$ is continuous, there exist some open sets $G_1, \ldots G_n \subset \mathcal{M}_S$ with $\mu_k \in G_k$ for all k and $\bigoplus_k G_k \subset G$. Then

$$\begin{aligned} P\{\xi \in G\} &\geq P\Big(\bigcap\nolimits_k \{\eta_k \in G_k\};\, \kappa = n\Big) \\ &= P\{\kappa = n\} \prod\nolimits_k P\{\eta_k \in G_k\} > 0, \end{aligned}$$

which implies $\mu \in \operatorname{supp} \mathcal{L}(\xi)$, since G was arbitrary.

Conversely, let $\mu \in \operatorname{supp} \mathcal{L}(\xi)$, so that $P\{\xi \in G\} > 0$ for every open set $G \subset \mathcal{M}_S$ containing μ. With S_n and κ as before, let $\xi_n = S_{\kappa \wedge n}$, and note that $\xi_n \xrightarrow{d} \xi$. Hence, Lemma 4.1 yields $P\{\xi_n \in G\} > 0$ for large enough n. Since G was arbitrary, we may choose some $\mu_n \in \operatorname{supp} \mathcal{L}(\xi_n)$ with $\mu_n \xrightarrow{v} \mu$, and it suffices to show that $\mu_n \in \bigcup_k (\operatorname{supp} \lambda)^{\oplus k}$ for all k. Noting that $\operatorname{supp} \mathcal{L}(\xi_n) \subset \bigcup_{k \leq n} \operatorname{supp} \mathcal{L}(S_k)$, we may finally reduce the assertion to $\operatorname{supp} \mathcal{L}(S_n) \subset (\operatorname{supp} \lambda)^{\oplus n}$.

Fixing any $\mu \in \operatorname{supp} \mathcal{L}(S_n)$, consider an open set G containing μ, and let $G_1, \ldots, G_n$ be open with $P\{\xi_k \in G_k\} > 0$ and $\bigoplus_k G_k \subset G$. Finally, choose some measures $\mu_k^G \in G_k \cap \operatorname{supp} \lambda$, so that $\sum_k \mu_k^G \in G$. Letting $G \downarrow \{\mu\}$ along a sequence, we obtain some measures $\mu^G = \sum_k \mu_k^G$ with $\mu^G \xrightarrow{v} \mu$. For fixed k, Theorem 4.2 shows that the sequence (μ_k^G) is relatively compact in the vague topology, which yields convergence $\mu_k^G \xrightarrow{v} \mu_k$ along a sub-sequence, for all $k \leq n$. Then $\sum_k \mu_k = \mu$ by continuity, and $\mu_k \in \operatorname{supp} \lambda$ for all k, since the support is closed. Thus, $\mu \in (\operatorname{supp} \lambda)^{\oplus n}$.

When $\|\lambda\| = \infty$, we may choose some $\lambda_n \uparrow \lambda$ with $\lambda_n \sim \lambda$ and $\|\lambda_n\| < \infty$. Next, let ξ_n and ξ'_n be independent, infinitely divisible random measures directed by λ_n and $\lambda - \lambda_n$, respectively, so that $\xi_n + \xi'_n \overset{d}{=} \xi$. Note that $\xi_n \xrightarrow{vd} \xi$ and $\xi'_n \xrightarrow{vd} 0$, by the representation in Theorem 3.20 and dominated convergence. If $\mu \in \operatorname{supp} \mathcal{L}(\xi)$, we may choose $\mu_n \in \operatorname{supp} \mathcal{L}(\xi_n)$ as before with $\mu_n \xrightarrow{v} \mu$. Then

$$\mu_n \in \overline{\bigcup\nolimits_k (\operatorname{supp} \lambda_n)^{\oplus k}} \subset \overline{\bigcup\nolimits_k (\operatorname{supp} \lambda)^{\oplus k}}, \quad n \in \mathsf{N},$$

which implies that even μ lies in the closure on the right.

Conversely, assume $\mu \in (\operatorname{supp} \lambda)^{\oplus k}$ for some k, so that $\mu \in \operatorname{supp} \mathcal{L}(\xi_n)$ for every n. Fix any open set G around μ. Since addition in $\mathcal{M}_S$ is continuous, we may choose some open sets B and B' containing μ and 0, respectively, such that $m \in B$ and $m' \in B'$ imply $m + m' \in G$. Since $\xi'_n \xrightarrow{vd} 0$, we may next choose n so large that $P\{\xi'_n \in B'\} > 0$. Then

$$\begin{aligned} P\{\xi \in G\} &\geq P\{\xi_n \in B,\, \xi'_n \in B'\} \\ &= P\{\xi_n \in B\}\, P\{\xi'_n \in B'\} > 0, \end{aligned}$$

and $\mu \in \operatorname{supp} \mathcal{L}(\xi)$ follows since G was arbitrary. □

We conclude with a simple approximation, needed in a later chapter. For any random measures ξ and $\xi_1, \xi_2, \ldots$, we mean by $\xi_n \xrightarrow{vP} \xi$ that, for any sub-sequence $N' \subset \mathbb{N}$, the a.s. convergence $\xi_n \xrightarrow{v} \xi$ holds along a further sub-sequence $N'' \subset N'$.

Corollary 4.31 *(Cox approximation) For any random measure ξ on S, let η_n be a Cox process directed by $n\xi$, $n \in \mathbb{N}$. Then $\eta_n/n \xrightarrow{vP} \xi$. In particular, ξ is infinitely divisible iff this is true for every η_n.*

Proof: By conditioning, we may assume that ξ is non-random. Then by monotone convergence,

$$Ee^{-\eta_n f/n} = \exp\Big\{-n\xi(1 - e^{-f/n})\Big\} \to e^{-\xi f}, \quad f \in \hat{C}_S,$$

and so $\eta_n/n \xrightarrow{vd} \xi$ by Theorem 4.11. Since the limit is non-random and $\mathcal{M}_S$ is vaguely separable by Lemma 4.6, this is equivalent to $\eta_n/n \xrightarrow{vP} \xi$. The last assertion follows from Lemma 4.24. □

4.4 Strong Approximation and Convergence

For measures μ_n and μ on a Borel space S, write $\mu_n \xrightarrow{u} \mu$ for convergence in total variation, and $\mu_n \xrightarrow{ul} \mu$ for the corresponding local version, defined by $1_B\mu_n \xrightarrow{u} 1_B\mu$ for every $B \in \hat{\mathcal{S}}$. When ξ_n and ξ are random measures on S, write $\xi_n \xrightarrow{ud} \xi$ for the convergence $\mathcal{L}(\xi_n) \xrightarrow{u} \mathcal{L}(\xi)$, and $\xi_n \xrightarrow{uld} \xi$ for the corresponding local version where $1_B\xi_n \xrightarrow{ud} 1_B\xi$ for all $B \in \hat{\mathcal{S}}$. Define $\xi_n \overset{ud}{\sim} \eta_n$ by $\|\mathcal{L}(\xi_n) - \mathcal{L}(\eta_n)\| \to 0$, and similarly for the local relation $\overset{uld}{\sim}$.

We begin with a simple probabilistic interpretation of the strong convergence $\xi_n \xrightarrow{ud} \xi$ and associated equivalence $\xi_n \overset{ud}{\sim} \zeta_n$.

Lemma 4.32 *(coupling) Let μ and ν be probability measures on a Borel space S. Then*

$$\|\mu - \nu\| = 2 \inf P\{\xi \neq \eta\},$$

where the infimum extends over all pairs of random elements ξ and η in S with marginal distributions μ and ν.

Proof: For any ξ and η as stated, put $A = \{\xi = \eta\}$ and define

$$\mu' = \mathcal{L}(\xi;\, A^c), \qquad \nu' = \mathcal{L}(\eta;\, A^c).$$

Then

$$\begin{aligned}\|\mu - \nu\| &= \|\mu' - \nu'\| \le \|\mu'\| + \|\nu'\| \\ &= 2\,P(A^c) = 2\,P\{\xi \ne \eta\}.\end{aligned}$$

Conversely, define

$$\hat{\mu} = \mu \wedge \nu, \qquad \mu' = \mu - \hat{\mu}, \qquad \nu' = \nu - \hat{\mu}.$$

and choose a $B \in \mathcal{S}$ with

$$\hat{\mu} B^c = \mu' B = \nu' B = 0.$$

For any ξ and η as stated, we have

$$\mathcal{L}(1_B \xi) = \mathcal{L}(1_B \eta) = \hat{\mu},$$

and so the transfer theorem (FMP 6.10) yields a random element $\tilde{\eta} \overset{d}{=} \eta$ with

$$1_B \xi = 1_B \tilde{\eta} \text{ a.s.}$$

Here clearly

$$P\{\xi = \tilde{\eta}\} \ge P\{\xi \in B\} = \|\hat{\mu}\|,$$

and so

$$P\{\xi \ne \tilde{\eta}\} \le 1 - \|\hat{\mu}\| = \tfrac{1}{2}\|\mu - \nu\|.$$

The asserted relation now follows by combination of the two bounds. □

The following strong continuity theorem for Poisson and Cox processes may be compared with the weak version in Lemma 4.17. By $\eta_n \xrightarrow{ulP} \eta$ we mean that $\|\eta_n - \eta\|_B \xrightarrow{P} 0$ for all $B \in \hat{\mathcal{S}}$, and similarly for the global version $\eta_n \xrightarrow{uP} \eta$.

Theorem 4.33 *(Poisson and Cox continuity) Let $\xi, \xi_1, \xi_2, \ldots$ be Cox processes directed by some random measures $\eta, \eta_1, \eta_2, \ldots$ on S. Then*

$$\eta_n \xrightarrow{ulP} \eta \quad \Rightarrow \quad \xi_n \xrightarrow{uld} \xi,$$

with equivalence when η and all η_n are non-random.

Proof: For both parts, it is clearly enough to consider bounded S. First let $\eta = \lambda$ and all $\eta_n = \lambda_n$ be non-random. For every $n \in \mathbb{N}$, define

$$\hat{\lambda}_n = \lambda \wedge \lambda_n, \qquad \lambda'_n = \lambda - \hat{\lambda}_n, \qquad \lambda''_n = \lambda_n - \hat{\lambda}_n,$$

so that

$$\lambda = \hat{\lambda}_n + \lambda'_n, \qquad \lambda_n = \hat{\lambda}_n + \lambda''_n, \qquad \|\lambda_n - \lambda\| = \|\lambda'_n\| + \|\lambda''_n\|.$$

Letting $\hat{\xi}_n$, ξ'_n, and ξ''_n be independent Poisson processes with intensities $\hat{\lambda}$, λ'_n, and λ''_n, respectively, we note that

$$\xi \stackrel{d}{=} \hat{\xi}_n + \xi'_n, \qquad \xi_n \stackrel{d}{=} \hat{\xi}_n + \xi''_n.$$

Assuming $\lambda_n \stackrel{u}{\to} \lambda$, we get

$$\begin{aligned} \|\mathcal{L}(\xi) - \mathcal{L}(\xi_n)\| &\le \|\mathcal{L}(\xi) - \mathcal{L}(\hat{\xi}_n)\| + \|\mathcal{L}(\xi_n) - \mathcal{L}(\hat{\xi}_n)\| \\ &\lesssim P\{\xi'_n \neq 0\} + P\{\xi''_n \neq 0\} \\ &= (1 - e^{-\|\lambda'_n\|}) + (1 - e^{-\|\lambda''_n\|}) \\ &\le \|\lambda'_n\| + \|\lambda''_n\| = \|\lambda - \lambda_n\| \to 0, \end{aligned}$$

which shows that $\xi_n \xrightarrow{ud} \xi$.

Conversely, let $\xi_n \xrightarrow{ud} \xi$. Then for any $B \in \mathcal{S}$ and $n \in \mathsf{N}$,

$$\begin{aligned} P\{\xi_n B = 0\} &\ge P\{\xi_n = 0\} \\ &\to P\{\xi = 0\} = e^{-\|\lambda\|} > 0. \end{aligned}$$

Since $(\log x)' = x^{-1} \lesssim 1$ on every interval $[\varepsilon, 1]$ with $\varepsilon > 0$, we get

$$\begin{aligned} \|\lambda - \lambda_n\| &\lesssim \sup_B |\lambda B - \lambda_n B| \\ &= \sup_B \Big| \log P\{\xi B = 0\} - \log P\{\xi_n B = 0\} \Big| \\ &\lesssim \sup_B \Big| P\{\xi B = 0\} - P\{\xi_n B = 0\} \Big| \\ &\le \|\mathcal{L}(\xi) - \mathcal{L}(\xi_n)\| \to 0, \end{aligned}$$

which shows that $\lambda_n \stackrel{u}{\to} \lambda$.

Turning to the case of random η and η_n, suppose that $\eta_n \xrightarrow{uP} \eta$. To prove $\xi_n \xrightarrow{ud} \xi$, it is enough to show that, for any sub-sequence $N' \subset \mathsf{N}$, the desired convergence holds along a further sub-sequence N''. By the sub-sequence criterion for convergence in probability, we may then assume that $\eta_n \stackrel{u}{\to} \eta$ a.s. Assuming ξ and the ξ_n to be conditionally independent Poisson processes with intensities η and η_n, respectively, we conclude from the previous case that

$$\Big\| \mathcal{L}(\xi_n | \eta_n) - \mathcal{L}(\xi | \eta) \Big\| \to 0 \ \text{ a.s.},$$

and so by dominated convergence

$$\begin{aligned} \| \mathcal{L}(\xi_n) - \mathcal{L}(\xi) \| &= \Big\| E\Big\{ \mathcal{L}(\xi_n | \eta_n) - \mathcal{L}(\xi | \eta) \Big\} \Big\| \\ &= \sup_{|f| \le 1} \Big| E\Big(E\{f(\xi_n) | \eta_n\} - E\{f(\xi) | \eta\} \Big) \Big| \\ &\le \sup_{|f| \le 1} E \Big| E\{f(\xi_n) | \eta_n\} - E\{f(\xi) | \eta\} \Big| \\ &\le E \Big\| \mathcal{L}(\xi_n | \eta_n) - \mathcal{L}(\xi | \eta) \Big\| \to 0, \end{aligned}$$

which shows that $\xi_n \xrightarrow{ud} \xi$. □

We turn to a total-variation version of Corollary 4.25.

Theorem 4.34 *(strong Poisson convergence) Let* (ξ_{nj}) *be a null array of point processes on* S*, and let* η *be a Poisson process on* S *with* $E\eta = \lambda$*. Then for any covering class* $\mathcal{C} \subset \hat{\mathcal{S}}$*, we have* (i) $\Leftrightarrow$ (ii)$-$(iii)*, where*

(i) $\sum_j \xi_{nj} \xrightarrow{uld} \eta$,

(ii) $\left\| \sum_j E(\xi_{nj};\, \xi_{nj}B = 1) - \lambda \right\|_B \to 0, \quad B \in \mathcal{C}$,

(iii) $\sum_j P\{\xi_{nj}B > 1\} \to 0, \quad B \in \mathcal{C}$.

Under the stronger requirement

(iii′) $\sum_j E(\xi_{nj}B;\, \xi_{nj}B > 1) \to 0, \quad B \in \mathcal{C}$,

condition (i) *becomes equivalent to*

(ii′) $\sum_j E\xi_{nj} \xrightarrow{ul} \lambda$.

Our proof relies on the following preliminary result, which allows us to replace the sums $\sum_j \xi_{nj}$ by suitable Poisson processes η_n. This may be compared with the weak version in Lemma 4.23.

Lemma 4.35 *(Poisson approximation) Let* (ξ_{nj}) *be a null array of point processes on a bounded space* S*, and choose some Poisson processes* η_n *on* S *with intensities*

$$\lambda_n = \sum\nolimits_j E(\xi_{nj};\, \|\xi_{nj}\| = 1), \quad n \in \mathsf{N}.$$

Then the conditions

$$\sum\nolimits_j P\{\|\xi_{nj}\| > 1\} \to 0, \qquad \limsup_{n\to\infty} \|\lambda_n\| < \infty,$$

imply

$$\sum\nolimits_j \xi_{nj} \overset{ud}{\sim} \eta_n.$$

Proof: Putting $\lambda_{nj} = E(\xi_{nj};\, \|\xi_{nj}\| = 1)$, we may choose some independent Poisson processes η_{nj} on S with $E\eta_{nj} = \lambda_{nj}$, so that $\eta_n = \sum_j \eta_{nj}$ is again Poisson with $E\eta_n = \lambda_n$. Clearly

$$\begin{aligned} P\{\|\eta_{nj}\| = 1\} &= \|\lambda_{nj}\|\, e^{-\|\lambda_{nj}\|} \\ &= \|\lambda_{nj}\| + O(\|\lambda_{nj}\|^2), \\ P\{\|\eta_{nj}\| > 1\} &= 1 - (1 + \|\lambda_{nj}\|)\, e^{-\|\lambda_{nj}\|} \\ &= O(\|\lambda_{nj}\|^2), \end{aligned}$$

and by Theorem 3.4,

$$E(\eta_{nj} \mid \|\eta_{nj}\| = 1) = E(\xi_{nj} \mid \|\xi_{nj}\| = 1).$$

Hence,

$$\begin{aligned} \|\mathcal{L}(\xi_{nj}) - \mathcal{L}(\eta_{nj})\| &\lesssim |P\{\|\xi_{nj}\| = 1\} - P\{\|\eta_{nj}\| = 1\}| \\ &\qquad + P\{\|\xi_{nj}\| > 1\} + P\{\|\eta_{nj}\| > 1\} \\ &\lesssim P\{\|\xi_{nj}\| > 1\} + \|\lambda_{nj}\|^2, \end{aligned}$$

and so by independence

$$\begin{aligned}\Big\|\mathcal{L}\Big(\textstyle\sum_j\xi_{nj}\Big)-\mathcal{L}(\eta_n)\Big\| &\le \textstyle\sum_j\|\mathcal{L}(\xi_{nj})-\mathcal{L}(\eta_{nj})\|\\ &\lesssim \textstyle\sum_j\Big(P\{\|\xi_{nj}\|>1\}+\|\lambda_{nj}\|^2\Big)\\ &\le \textstyle\sum_j P\{\|\xi_{nj}\|>1\}+\|\lambda_n\|\sup_j P\{\xi_{nj}\ne 0\},\end{aligned}$$

which tends to 0 as $n\to\infty$ by the stated hypotheses. □

Proof of Theorem 4.34: First let S be bounded, and take $\mathcal{C}=\{S\}$. Consider some Poisson processes η_n with intensities

$$\lambda_n=\textstyle\sum_j E(\xi_{nj};\,\|\xi_{nj}\|=1),\quad n\in\mathsf{N}.$$

Assuming (ii) and (iii), we may write

$$\Big\|\mathcal{L}\Big(\textstyle\sum_j\xi_{nj}\Big)-\mathcal{L}(\eta)\Big\|\le\Big\|\mathcal{L}\Big(\textstyle\sum_j\xi_{nj}\Big)-\mathcal{L}(\eta_n)\Big\|+\|\mathcal{L}(\eta_n)-\mathcal{L}(\eta)\|,$$

which tends to 0 as $n\to\infty$ by Lemma 4.35 and Theorem 4.33.

Conversely, (i) implies $\sum_j\|\xi_{nj}\|\xrightarrow{d}\|\eta\|$, and so by Corollary 4.25 (or FMP 5.7) we obtain (iii), along with the convergence

$$\|\lambda_n\|=\textstyle\sum_j P\{\|\xi_{nj}\|=1\}\to\|\lambda\|,$$

which shows that $\|\lambda_n\|$ is bounded. Defining η_n as before, we may write

$$\|\mathcal{L}(\eta_n)-\mathcal{L}(\eta)\|\le\Big\|\mathcal{L}(\eta_n)-\mathcal{L}\Big(\textstyle\sum_j\xi_{nj}\Big)\Big\|+\Big\|\mathcal{L}\Big(\textstyle\sum_j\xi_{nj}\Big)-\mathcal{L}(\eta)\Big\|,$$

which tends to 0 as $n\to\infty$ by Lemma 4.35 and (i). Hence, (ii) follows by Theorem 4.33.

Now assume (iii′). This clearly implies (iii), and furthermore

$$\begin{aligned}\Big\|\textstyle\sum_j E\xi_{nj}-\sum_j E(\xi_{nj};\,\|\xi_{nj}\|=1)\Big\| &= \Big\|\textstyle\sum_j E(\xi_{nj};\,\|\xi_{nj}\|>1)\Big\|\\ &= \textstyle\sum_j E(\xi_{nj};\,\|\xi_{nj}\|>1)\to 0,\end{aligned}$$

which shows that (ii) and (ii′) are equivalent. This proves the last assertion, and hence completes the proof for bounded S.

For general S, it suffices to apply the previous version to the restrictions of all ξ_{nj} and η to an arbitrary set $B\in\mathcal{C}$. □

For any Borel spaces S and T, the kernels $\nu_1,\nu_2,\ldots\colon S\to T$ are said to be *dissipative* if

$$\lim_{n\to\infty}\sup_{x\in T}\nu_n(x,B)=0,\quad B\in\hat{\mathcal{S}}.$$

In a non-topological setting, we further say that the random measures ξ_n on S are *tight*, if the sequence ξ_nB is tight in R_+ for every $B\in\hat{\mathcal{S}}$. The following Cox approximation may be compared with the weak version in Theorem 4.40 below.

Theorem 4.36 *(Cox approximation) Consider a dissipative sequence of probability kernels $\nu_n: T \to S$, where S and T are Borel. For every $n \in \mathsf{N}$, let ξ_n be a ν_n-transform of some point process η_n on T, and let ζ_n be a Cox process directed by $\eta_n \nu_n$. Further, let ζ be a Cox process on S directed by some random measure ρ. Then*

(i) *(ξ_n), (ζ_n), and $(\eta_n \nu_n)$ are simultaneously tight,*

(ii) *$\xi_n \overset{uld}{\sim} \zeta_n$ whenever either side is tight,*

(iii) *$\eta_n \nu_n \xrightarrow{ulP} \rho$ implies $\xi_n \xrightarrow{uld} \zeta$ and $\zeta_n \xrightarrow{uld} \zeta$.*

Note that $\eta_n \nu_n$ is defined in the sense of kernel multiplication, as in Section 1.3, hence as a kernel from Ω to S, or simply a random measure on S, so that for measurable functions $f \geq 0$ on S,

$$(\eta_n \nu_n) f = \int \eta_n(dt) \int \nu_n(t, ds)\, f(t).$$

Proof: (i) Let $t_n > 0$, and write $s_n = 1 - e^{-t_n}$, so that $t_n = -\log(1 - s_n)$. By Lemma 3.1, we have for any $B \in \hat{\mathcal{S}}$

$$\begin{aligned} E \exp(-t_n \zeta_n B) &= E \exp(-s_n \eta_n \nu_n B), \\ E \exp(-t_n \xi_n B) &= E \exp\Big\{\eta_n \log(1 - s_n \nu_n B)\Big\}. \end{aligned}$$

Since $x \leq -\log(1 - x) \leq 2x$ for small $x > 0$, the latter relation yields for small enough t_n

$$\begin{aligned} E \exp(-2 s_n \eta_n \nu_n B) &\leq E \exp(-t_n \xi_n B) \\ &= E \exp(-s_n \eta_n \nu_n B). \end{aligned}$$

Letting $t_n \to 0$, corresponding to $s_n \to 0$, we see that the conditions

$$t_n \xi_n B \xrightarrow{P} 0, \qquad t_n \zeta_n B \xrightarrow{P} 0, \qquad s_n \eta_n \nu_n B \xrightarrow{P} 0,$$

are equivalent. The assertion now follows by FMP 4.9.

(ii) Again, we may take S to be bounded, so that the sequence of random variables $\|\eta_n \nu_n\|$ is tight. To prove the equivalence $\xi_n \overset{ud}{\sim} \zeta_n$, it is enough, for any sub-sequence $N' \subset \mathsf{N}$, to establish the stated convergence along a further sub-sequence N''. By tightness we may then assume that $\|\eta_n \nu_n\|$ converges in distribution, and by Skorohod coupling and transfer (FMP 4.30, 6.10), we may even assume the convergence to hold a.s. In particular, $\|\eta_n \nu_n\|$ is then a.s. bounded.

Assuming the appropriate conditional independence, we may now apply Lemma 4.35 to the conditional distributions of the sequences (ξ_n) and (ζ_n), given (η_n), to obtain

$$\Big\| \mathcal{L}(\xi_n |\, \eta_n) - \mathcal{L}(\zeta_n |\, \eta_n) \Big\| \to 0 \ \text{ a.s.},$$

which implies $\xi_n \overset{ud}{\sim} \zeta_n$, as in the proof of Theorem 4.33.

(iii) The hypothesis implies that $(\eta_n \nu_n)$ is tight, and so by (ii) it suffices to show that $\zeta_n \xrightarrow{uld} \zeta$, which holds by Theorem 4.33. $\square$

The last result leads to the following invariance criterion[2].

Corollary 4.37 *(transform invariance) Fix a probability kernel ν on S with dissipative powers ν^n, and consider a point process ξ on S with ν-transform ξ_ν. Then $\xi_\nu \overset{d}{=} \xi$ iff ξ is a Cox process directed by some random measure η with $\eta\nu \overset{d}{=} \eta$.*

Proof: If ξ is Cox and directed by η, then by Theorem 3.2 (i) the ν-transform ξ_ν is again Cox and directed by $\eta\nu$, in which case $\xi_\nu \overset{d}{=} \xi$ iff $\eta\nu \overset{d}{=} \eta$ by Lemma 3.3 (i).

Conversely, suppose that $\xi_\nu \overset{d}{=} \xi$. For every $n \in \mathsf{N}$, let ξ_n be a ν^n-transform of ξ and let ζ_n be a Cox process directed by $\xi\nu^n$. Then Theorem 4.36 (ii) yields $\xi \overset{d}{=} \xi_n \overset{uld}{\sim} \zeta_n$, and so $\zeta_n \xrightarrow{uld} \xi$. Embedding any $B \in \hat{\mathcal{S}}$ into $[0,1]$, we see from Lemma 4.17 that ξ is Cox, and so as before the directing measure η satisfies $\eta\nu \overset{d}{=} \eta$. $\square$

We proceed with a strong continuity theorem for infinitely divisible point processes, here stated for simplicity only for bounded spaces S. A corresponding weak version appeared as part of Lemma 4.24.

Theorem 4.38 *(infinitely divisible point processes, Matthes et al.) Given a bounded Borel space S, let ξ and $\xi_1, \xi_2, \ldots$ be infinitely divisible point processes on S with Lévy measures λ and $\lambda_1, \lambda_2, \ldots$. Then $\xi_n \xrightarrow{ud} \xi$ iff $\lambda_n \xrightarrow{u} \lambda$.*

Proof: For any bounded, signed measure μ on $\mathcal{N}_S$, let μ^{*n} denote the n-th convolution power of μ, and put $\mu^{*0} = \delta_0$ for consistency, where δ_0 denotes the unit mass at $0 \in \mathcal{N}_S$. Define the bounded, signed measure $\exp \mu = e^\mu$ on $\mathcal{N}_S$ by the power series

$$\exp \mu = \sum_{n \ge 0} \frac{\mu^{*n}}{n!},$$

which converges in total variation. It is easy to verify that

$$e^{\mu+\nu} = e^\mu * e^\nu, \qquad \|e^\mu\| = e^{\|\mu\|}, \qquad e^{c\delta_0} = e^c \delta_0. \tag{17}$$

Now let ξ be an infinitely divisible point process on S with bounded Lévy measure λ, and conclude from Theorems 3.4 and 3.20 that

$$\mathcal{L}(\xi) = e^{-\|\lambda\|} \sum_{n \ge 0} \frac{\|\lambda\|^n}{n!} (\lambda / \|\lambda\|)^{*n}$$

[2]Note that $\eta\nu$ and ν^n are defined in the sense of kernel multiplication, as in Section 1.3. Hence, the ν^n form a Markov transition semigroup on S, and $\eta\nu$ is the "distribution" at time 1 of an associated Markov chain with initial distribution η.

$$\begin{aligned} &= e^{-\|\lambda\|} \sum_{n\geq 0} \frac{\lambda^{*n}}{n!} \\ &= e^{-\|\lambda\|}\, e^{\lambda} = e^{\lambda}/\|e^{\lambda}\|. \end{aligned} \tag{18}$$

Recall from Theorem 3.20 that λ is uniquely determined by $\mathcal{L}(\xi)$.

Next suppose that $e^{\mu} = e^{\nu}$, for some bounded, signed measures $\mu = \mu' - \mu''$ and $\nu = \nu' - \nu''$. Then by (17)

$$\begin{aligned} e^{\mu'+\nu''} &= e^{\mu} * e^{\mu''+\nu''} \\ &= e^{\nu} * e^{\mu''+\nu''} = e^{\mu''+\nu'}. \end{aligned}$$

Dividing by the norms $\|e^{\mu'+\nu''}\| = \|e^{\mu''+\nu'}\|$ and using the uniqueness in (18), we obtain $\mu' + \nu'' = \mu'' + \nu'$, and so $\mu = \nu$.

For any signed measure μ on $\mathcal{N}_S$ with $\|\mu - \delta_0\| < 1$, we define

$$\log \mu = -\sum_{n\geq 1} \frac{(\delta_0 - \mu)^{*n}}{n},$$

which again converges in total variation. Under the same condition, a standard computation, for absolutely convergent power series, yields $\exp(\log \mu) = \mu$. Assuming $\mu = e^{\lambda}$ and using the previous uniqueness, we conclude that also $\lambda = \log \mu$. Thus, $\mu = \exp \lambda$ iff $\lambda = \log \mu$, whenever $\|\mu - \delta_0\| < 1$.

For any bounded, signed measures μ and ν, we note that

$$\|\mu + \nu\| \leq \|\mu\| + \|\nu\|, \qquad \|\mu * \nu\| \leq \|\mu\|\,\|\nu\|, \tag{19}$$

and so for all $n \in \mathsf{N}$,

$$\|\mu^{*n} - \nu^{*n}\| \leq n\big(\|\mu\| \vee \|\nu\|\big)^{n-1} \|\mu - \nu\|.$$

Hence,

$$\begin{aligned} \|e^{\mu} - e^{\nu}\| &\leq \sum_{n\geq 1} \frac{\|\mu^{*n} - \nu^{*n}\|}{n!} \\ &\leq \|\mu - \nu\| \sum_{n\geq 1} \frac{\big(\|\mu\| \vee \|\nu\|\big)^{n-1}}{(n-1)\,!} \\ &= \|\mu - \nu\|\, \exp\big(\|\mu\| \vee \|\nu\|\big), \end{aligned}$$

which shows that the function $\exp \mu$ is continuous in total variation. Similarly, for any μ and ν with $\|\mu - \delta_0\| < 1$ and $\|\nu - \delta_0\| < 1$,

$$\begin{aligned} \|\log \mu - \log \nu\| &\leq \sum_{n\geq 1} n^{-1} \Big\|(\delta_0 - \mu)^{*n} - (\delta_0 - \nu)^{*n}\Big\| \\ &\leq \|\mu - \nu\| \sum_{n\geq 1} \big(\|\delta_0 - \mu\| \vee \|\delta_0 - \nu\|\big)^{n-1} \\ &= \|\mu - \nu\| \big(1 - \|\delta_0 - \mu\| \vee \|\delta_0 - \nu\|\big)^{-1}, \end{aligned}$$

which shows that even $\log \mu$ is continuous on the open ball where $\|\mu - \delta_0\| < 1$.

Now let ξ and $\xi_1, \xi_2, \ldots$ be infinitely divisible point processes on S with bounded Lévy measures λ and $\lambda_1, \lambda_2, \ldots$. Assuming $\lambda_n \xrightarrow{u} \lambda$, we have $e^{\lambda_n} \xrightarrow{u} e^{\lambda}$ and hence also $\|e^{\lambda_n}\| \to \|e^{\lambda}\|$, by the continuity of e^{λ}, and so $\xi_n \xrightarrow{ud} \xi$ by (18). Conversely, assume $\xi_n \xrightarrow{ud} \xi$, so that $e^{\lambda_n}/\|e^{\lambda_n}\| \xrightarrow{u} e^{\lambda}/\|e^{\lambda}\|$ by (18). Since

$$\begin{aligned} \|e^{\lambda_n}\|^{-1} &= e^{-\|\lambda_n\|} = P\{\xi_n = 0\} \\ &\to P\{\xi = 0\} = e^{-\|\lambda\|} = \|e^{\lambda}\|^{-1}, \end{aligned}$$

we get $e^{\lambda_n} \xrightarrow{u} e^{\lambda}$. Hence, by (17) and (19),

$$\begin{aligned} \|e^{\lambda_n - \lambda} - \delta_0\| &= \|e^{-\lambda}(e^{\lambda_n} - e^{\lambda})\| \\ &\le \|e^{-\lambda}\| \, \|e^{\lambda_n} - e^{\lambda}\| \to 0, \end{aligned}$$

and so, by the continuity of $\log \mu$, we get for large enough n

$$\|\lambda_n - \lambda\| = \Big\| \log e^{\lambda_n - \lambda} - \log \delta_0 \Big\| \to 0,$$

which means that $\lambda_n \xrightarrow{u} \lambda$. □

The last result yields a strong convergence criterion for null arrays of point processes. For comparison, the result is stated alongside the corresponding weak version[3]. We give the criteria in a striking but slightly simplified form, to be clarified after the formal statements.

Corollary 4.39 *(null arrays of point processes) Let (ξ_{nj}) be a null array of point processes on S with distributions λ_{nj}, and consider an infinitely divisible point process ξ on S with Lévy measure λ. Then*

(i) $\quad \sum_j \xi_{nj} \xrightarrow{vd} \xi \quad \Leftrightarrow \quad \sum_j \lambda_{nj} \xrightarrow{vw} \lambda,$

(ii) $\quad \sum_j \xi_{nj} \xrightarrow{uld} \xi \quad \Leftrightarrow \quad \sum_j \lambda_{nj} \xrightarrow{ul} \lambda.$

For bounded S, the conditions $\lambda_n \xrightarrow{vw} \lambda$ and $\lambda_n \xrightarrow{ul} \lambda$ reduce to the global versions $\lambda_n \xrightarrow{ww} \lambda$ and $\lambda_n \xrightarrow{u} \lambda$, respectively, where the former condition is defined by $\lambda_n f \to \lambda f$, for any bounded and weakly continuous function $f \ge 0$ on $\mathcal{N}_S \setminus \{0\}$, whereas the latter condition means that $\|\lambda_n - \lambda\| \to 0$ on the same space. The definitions extend to unbounded S by an obvious localization.

Proof: (i) This is equivalent to the point process case of Theorem 4.22.

(ii) We may clearly take S to be bounded. Since the array (ξ_{nj}) is null, we have $\sup_j \|\lambda_{nj}\| \to 0$. Now write $\xi_n = \sum_j \xi_{nj}$ and $\lambda_n = \sum_j \lambda_{nj}$, let η and

[3]For the latter, S is understood to be a separable and complete metric space.

η_n be a Poisson processes on $\mathcal{N}_S \setminus \{0\}$ with intensities λ and λ_n, respectively, and consider on the same space the point processes

$$\zeta_n = \sum\nolimits_j \delta_{\xi_{nj}}, \quad n \in \mathsf{N}.$$

Since either condition implies $\limsup_n \|\lambda_n\| < \infty$ by part (i), Lemma 4.35 yields $\zeta_n \overset{ud}{\sim} \eta_n$, and so $\xi_n \overset{ud}{\sim} \tilde{\xi}_n$ by Theorem 3.20, where the $\tilde{\xi}_n$ are infinitely divisible with Lévy measures λ_n. It remains to note that $\tilde{\xi}_n \xrightarrow{ud} \xi$ iff $\lambda_n \xrightarrow{u} \lambda$, by Theorem 4.38. $\square$

4.5 Transforms and Symmetries

Here we consider some weak limit theorems of special importance, beginning with a weak version of Theorem 4.36. Recall that the kernels $\nu_n : T \to S$ are said to be *dissipative* if $\sup_t \nu_n(t, B) \to 0$ for all $B \in \hat{\mathcal{S}}$.

Theorem 4.40 *(Cox convergence) Fix a Borel space T and a Polish space S. For every $n \in \mathsf{N}$, let ξ_n be a ν_n-transform of a point process η_n on T, for some dissipative probability kernels $\nu_n : T \to S$. Then $\xi_n \xrightarrow{vd}$ some ξ iff $\eta_n \nu_n \xrightarrow{vd}$ some ζ, in which case ξ is distributed as a Cox process directed by ζ.*

Proof: By Lemma 3.1 (iii), we have

$$Ee^{-\xi_n f} = E \exp\Big(\eta_n \log \nu_n e^{-f}\Big), \quad f \in \mathcal{S}_+, \; n \in \mathsf{N}.$$

Let the functions f be supported by a fixed set $B \in \hat{\mathcal{S}}$, and write $g = 1 - e^{-f}$. Noting that

$$t \le -\log(1-t) \le t + O(t^2), \quad t \in [0,1),$$

we get for every $n \in \mathsf{N}$

$$\begin{aligned} E \exp(-\rho_n \eta_n \nu_n g) &\le E \exp(-\xi_n f) \\ &\le E \exp(-\eta_n \nu_n g), \end{aligned} \tag{20}$$

where $1 \le \rho_n \to 1$ a.s. If $\eta_n \nu_n \xrightarrow{vd} \zeta$, then even $\rho_n \eta_n \nu_n \xrightarrow{vd} \zeta$, and so $Ee^{-\xi_n f} \to Ee^{-\zeta g}$. Since f was arbitrary, Theorem 4.11 yields $\xi_n \xrightarrow{vd} \xi$, where ξ is a Cox process directed by ζ.

Conversely, suppose that $\xi_n \xrightarrow{vd} \xi$ for some point process ξ on S. Then by Theorem 4.10 and FMP 5.2,

$$\lim_{t \to 0} \inf_n Ee^{-t\xi_n B} = \sup_{K \in \mathcal{K}} \inf_n Ee^{-\xi_n (B \setminus K)} = 1, \quad B \in \hat{\mathcal{S}}.$$

Since $Ee^{-t' \eta_n \nu_n B} \ge Ee^{-t \xi_n B}$ with $t' = 1 - e^{-t}$ by (20), similar conditions hold for the random measures $\eta_n \nu_n$, which are then tight by Theorem 4.10.

If $\eta_n \nu_n \xrightarrow{vd} \zeta$ along a sub-sequence $N' \subset \mathsf{N}$, the direct assertion yields $\xi_n \xrightarrow{vd} \xi'$ along N', where ξ' is a Cox process directed by ζ. Since $\xi' \overset{d}{=} \xi$, the distribution of ζ is unique by Lemma 3.3, and so the convergence $\eta_n \nu_n \xrightarrow{vd} \zeta$ extends to the original sequence. □

We can also give a short proof, based on strong approximation:

Second proof of Theorem 4.40: Either condition $\xi_n \xrightarrow{vd} \xi$ or $\eta_n \nu_n \xrightarrow{vd} \zeta$ implies that both sequences (ξ_n) and $(\eta_n \nu_n)$ are tight, in the non-topological sense of Theorem 4.36, and so the same result yields $\xi_n \overset{uld}{\sim} \zeta_n$, where the ζ_n are Cox and directed by $\eta_n \nu_n$. Thus, by FMP 4.28, we have $\xi_n \xrightarrow{vd} \xi$ iff $\zeta_n \xrightarrow{vd} \xi$, which holds by Lemma 4.17 iff $\eta_n \nu_n \xrightarrow{vd} \zeta$, and ξ is Cox with directing measure ζ. □

The last result extends with the same proof to the case of random probability kernels, i.e., to probability kernels $\nu_n : T \times \Omega \to S$. In particular, the notion of p-thinnings from Chapter 3 extends immediately to the case where $p = X$ is a product-measurable process on S. Extending the thinning rates X_n to suitable random probability kernels ν_n, we obtain a basic limit theorem for independent thinnings.

Corollary 4.41 *(thinning limits) For every $n \in \mathsf{N}$, let ξ_n be an X_n-thinning of a point process η_n on a Polish space S, where the processes X_n are product-measurable with $\sup_{s \in B} X_n(s) \to 0$ a.s. for all $B \in \hat{\mathcal{S}}$. Then $\xi_n \xrightarrow{vd}$ some ξ iff $X_n \cdot \eta_n \xrightarrow{vd}$ some ζ, in which case ξ is distributed as a Cox process directed by ζ.*

This yields a short proof of a Cox criterion from Chapter 3:

Second proof of Corollary 3.38: If ξ is a p-thinning for every $p \in (0,1]$, it is Cox by Theorem 4.41. Conversely, if ξ is a Cox process directed by η, and ξ_p is a p-thinning of a Cox process directed by $p^{-1}\eta$, then $\xi_p \overset{d}{=} \xi$ by Lemmas 2.2 and 3.1 (ii), (iv). □

We turn to some convergence criteria for λ-symmetric random measures ξ on $S \times I$, as in Theorem 3.37, where λ is taken to be Lebesgue measure on $I = [0,1]$ or R_+. Recall that, for $I = [0,1]$, we have a representation

$$\xi = \alpha \otimes \lambda + \textstyle\sum_j \beta_j \otimes \delta_{\tau_j} \quad \text{a.s.},$$

where α and $\beta_1, \beta_2, \ldots$ are random measures on S, and $\tau_1, \tau_2, \ldots$ is an independent sequence of i.i.d. $U(0,1)$ random variables. Assuming $\beta_1, \beta_2, \ldots \neq 0$ a.s. and introducing the point process $\beta = \sum_j \delta_{\beta_j}$ on $\mathcal{M}'_S = \mathcal{M}_S \setminus \{0\}$, we note that the distributions of ξ and (α, β) determine each other uniquely. We say that ξ is *directed* by (α, β).

For our present purposes, we also introduce the random measures

$$\hat{\beta} = \alpha \otimes \delta_0 + \sum\nolimits_j \beta_j \otimes \delta_{\beta_j}, \qquad \bar{\xi} = \xi(\cdot \times [0,1]), \tag{21}$$

on $S \times \mathcal{M}_S$ and S, respectively. Given any bounded, complete metrization ρ of $\mathcal{M}_S$, we may regard $\hat{\beta}$ as a random element in $\mathcal{M}_{S\times\mathcal{M}_S}$, endowed with the associated vague topology. We also regard β as a random element in $\mathcal{N}_{\mathcal{M}'_S}$ with the vague topology, where the underlying complete metric ρ' in $\mathcal{M}'_S$ is defined as in Theorem 4.19. Finally, we introduce the measure-valued process

$$X_t = \xi(\cdot \times [0,t]), \quad t \in [0,1], \tag{22}$$

regarded as a random element in the Skorohod space $D([0,1], \mathcal{M}_S)$.

Theorem 4.42 *(symmetric measures on $S \times [0,1]$) Let ξ and $\xi_1, \xi_2, \ldots$ be λ-symmetric random measures on $S \times [0,1]$, directed by some pairs (α, β) and (α_n, β_n), $n \in \mathsf{N}$, and define the random measures $\hat{\beta}, \hat{\beta}_1, \hat{\beta}_2, \ldots$ and processes $X, X_1, X_2, \ldots$ as in* (21) *and* (22)*. Then these conditions are equivalent:*

(i) $\xi_n \xrightarrow{vd} \xi$ *in* $\mathcal{M}_{S\times[0,1]}$,

(ii) $X_n \xrightarrow{Sd} X$ *in* $D([0,1], \mathcal{M}_S)$,

(iii) $(\beta_n, \bar{\xi}_n) \xrightarrow{vd} (\beta, \bar{\xi})$ *in* $\mathcal{N}_{\mathcal{M}'_S} \times \mathcal{M}_S$,

(iv) $\hat{\beta}_n \xrightarrow{vd} \hat{\beta}$ *in* $\mathcal{M}_{S\times\mathcal{M}_S}$.

For a singleton S, the convergence in (i)$-$(iv) is defined with respect to the weak topology in $\mathcal{M}_{[0,1]}$, the Skorohod topology in $D([0,1], \mathsf{R}_+)$, the vague topology in $\mathcal{N}_{(0,\infty]} \times \mathsf{R}_+$, and the weak topology in $\mathcal{M}_{[0,\infty)}$, respectively.

Proof, (ii) $\Rightarrow$ (i): This holds by Theorem 4.11, since (ii) implies convergence of the finite-dimensional distributions.

(i) $\Rightarrow$ (ii): For every $n \in \mathsf{N}$, consider an optional time τ_n with respect to X_n and a constant $h_n \geq 0$, such that $\tau_n + h_n \leq 1$ and $h_n \to 0$. Using (i) and the strong stationarity of X_n (FMP 11.12), we get for any $B \in \hat{\mathcal{S}}$ and $h \in (0,1]$

$$\begin{aligned} E\Big(\xi_n\{B \times [\tau_n, \tau_n + h_n]\} \wedge 1\Big) &= E\Big(\xi_n\{B \times [0, h_n]\} \wedge 1\Big) \\ &\leq E\Big(\xi_n\{B \times [0, h]\} \wedge 1\Big) \\ &\to E\Big(\xi\{B \times [0, h]\} \wedge 1\Big), \end{aligned}$$

which tends to 0 as $h \to 0$. Thus, $\xi_n(B \times [\tau_n, \tau_n + h_n]) \xrightarrow{P} 0$, which implies $\xi_n(\cdot \times [\tau_n, \tau_n + h_n]) \xrightarrow{P} 0$, since B was arbitrary. By Aldous' criterion (FMP 16.11), the sequence (X_n) is then tight in the Skorohod space $D([0,1], \mathcal{M}_S)$. If $X_n \xrightarrow{d} Y$ along a sub-sequence, the direct assertion yields $\xi_n \xrightarrow{vd} \eta$, where

η is the random measure induced by Y. But then $\eta \stackrel{d}{=} \xi$ by (i), and so $Y \stackrel{d}{=} X$. Since the sub-sequence was arbitrary, Theorem 4.2 yields (ii), even along the original sequence.

(ii) $\Rightarrow$ (iii): By the definition of weak convergence, it suffices to show that the mapping $Y \mapsto (\beta, \bar{\xi})$ is continuous with respect to the topologies in (ii) and (iii), which reduces the proof to the case of non-random α and β. Here $\bar{\xi} = Y(1)$ is continuous, by the definition of the J_1-topology. Writing $\beta = \sum_j \delta_{\beta^j}$ and $\beta_n = \sum_j \delta_{\beta_n^j}$ for the jump point processes associated with the exchangeable processes Y and Y_n, respectively, we see that also $Y_n \xrightarrow{S} Y$ implies $\sup_j \rho(\beta_n^j, \beta^j) \to 0$, for a suitable ordering of the atoms. Since ultimately any bounded continuity set contains a fixed, finite number of points β_n^j, we obtain $\beta_n \xrightarrow{v} \beta$ in $\mathcal{N}_{\mathcal{M}'_S}$.

(iii) $\Rightarrow$ (i): Since $\bar{\xi}_n \xrightarrow{vd} \bar{\xi}$ in $\mathcal{M}_S$, the sequence $(\bar{\xi}_n)$ satisfies the tightness conditions in Theorem 4.10, which imply the corresponding conditions for (ξ_n), with B and K replaced by $B \times [0,1]$ and $K \times [0,1]$, respectively. Since every bounded set in $S \times [0,1]$ is contained in a set $B \times [0,1]$, and since the sets $K \times [0,1]$ are again compact in $S \times [0,1]$, the same theorem shows that (ξ_n) is tight. If $\xi_n \xrightarrow{vd} \eta$ along a sub-sequence, then η is again λ-symmetric, and the implication (i) $\Rightarrow$ (iii) yields $(\beta_n, \bar{\xi}_n) \xrightarrow{vd} (\gamma, \bar{\eta})$, where γ denotes the β-process of η. But then $(\gamma, \bar{\eta}) \stackrel{d}{=} (\beta, \bar{\xi})$ by (iii), and so $\eta \stackrel{d}{=} \xi$, which implies $\xi_n \xrightarrow{vd} \xi$, even along N.

(iv) $\Rightarrow$ (iii): By continuous mapping, we may take the pairs (α, β) and (α_n, β_n) to be non-random. The map $\hat{\beta} \mapsto \bar{\xi}$ is just the projection of $\hat{\beta}$ onto S, which is continuous, since $f \in \hat{C}_S$ implies $f \otimes 1 \in \hat{C}_{S\times[0,1]}$. To prove the continuity of the map $\hat{\beta} \mapsto \beta$, fix any $f \in \hat{C}_{\mathcal{M}'_S}$. Since the support of f is bounded away from 0, we may choose a $g \in \hat{C}_S$ such that $\inf\{\mu g;\, f(\mu) > 0\} = \varepsilon > 0$. The function

$$h(s, \mu) = \frac{f(\mu)\, g(s)}{\mu g \vee \varepsilon}, \quad s \in S,\ \mu \in \mathcal{M}_S,$$

belongs to $\hat{C}_{S\times\mathcal{M}_S}$, and $\hat{\beta} h = \beta f$. Hence, $\hat{\beta}_n \xrightarrow{v} \hat{\beta}$ implies $\beta_n f \to \beta f$, and $\beta_n \xrightarrow{v} \beta$ follows since f was arbitrary.

(iii) $\Rightarrow$ (iv): Again, we may take all pairs (α, β) to be non-random. For any $f \in \hat{C}_{S\times\mathcal{M}'_S}$, the function $g(\mu) = \int f(s, \mu)\, \mu(ds)$ belongs to $\hat{C}_{\mathcal{M}'_S}$ by the extended continuous mapping theorem (FMP 4.27). Since $\hat{\beta} f = \beta g$, we conclude from (iii) that $\hat{\beta}_n \xrightarrow{v} \hat{\beta}$ in $\mathcal{M}_{S\times\mathcal{M}'_S}$. To extend the convergence to $\mathcal{M}_{S\times\mathcal{M}_S}$, it suffices by Lemma 4.1 to show that $\hat{\beta}_n f \to \hat{\beta} f$ for functions $f = g \otimes h$ with $g \in \hat{C}_S$ and $h \in C_{\mathcal{M}_S}$. Then write $h = h(0) + h_+ - h_-$, where $h_\pm \in C_{\mathcal{M}_S}$ with $h_\pm(0) = 0$. Approximating $h_\pm$ by functions $\tilde{h}_\pm \in \hat{C}_{\mathcal{M}'_S}$, and noting that

$$\hat{\beta}(g \otimes 1) = \bar{\xi} g, \qquad \hat{\beta}\big(g \otimes |h_\pm - \tilde{h}_\pm|\big) \le \|h_\pm - \tilde{h}_\pm\|\, \bar{\xi} g,$$

we get $\hat{\beta}_n f \to \hat{\beta} f$, as $n \to \infty$ and then $\|h_\pm - \tilde{h}_\pm\| \to 0$. □

Turning to the case of λ-symmetric random measures ξ on $S \times \mathsf{R}_+$, we assume for simplicity that the metric in S is bounded. Then ξ has a representation

$$\xi = \alpha \otimes \lambda + \iint (\mu \otimes \delta_t)\, \eta(d\mu\, dt) \ \text{ a.s.}, \tag{23}$$

where α is a random measure on S and η is a Cox process on $\mathcal{M}'_S \times \mathsf{R}_+$, directed by some random measure $\nu \otimes \lambda$ with $\alpha \perp\!\!\!\perp_\nu \eta$ and $\int (\mu S \wedge 1)\, \nu(d\mu) < \infty$ a.s. Here $\mathcal{L}(\xi)$ and $\mathcal{L}(\alpha, \nu)$ determine each other uniquely, and we say that ξ is directed by (α, ν). Writing $\hat{\mu} = \mu/(\mu S \vee 1)$, we introduce on $S \times \mathcal{M}_S$ the bounded random measure

$$\hat{\nu} = \alpha \otimes \delta_0 + \int (\hat{\mu} \otimes \delta_\mu)\, \nu(d\mu). \tag{24}$$

Theorem 4.43 *(symmetric measures on $S \times \mathsf{R}_+$) For bounded S, let ξ and $\xi_1, \xi_2, \ldots$ be λ-symmetric random measures on $S \times \mathsf{R}_+$, directed by some pairs (α, ν) and (α_n, ν_n), $n \in \mathsf{N}$, and define the processes $X, X_1, X_2, \ldots$ and random measures $\hat{\nu}, \hat{\nu}_1, \hat{\nu}_2, \ldots$ as in* (22) *and* (24)*. Then these conditions are equivalent:*

(i) $\xi_n \xrightarrow{vd} \xi$ *in* $\mathcal{M}_{S \times \mathsf{R}_+}$,

(ii) $X_n \xrightarrow{Sd} X$ *in* $D(\mathsf{R}_+, \mathcal{M}_S)$,

(iii) $\hat{\nu}_n \xrightarrow{wd} \hat{\nu}$ *in* $\mathcal{M}_{S \times \mathcal{M}_S}$.

When S is a singleton, so that $\xi, \xi_1, \xi_2, \ldots$ are λ-symmetric random measures on R_+, the space in (iii) reduces to $\mathcal{M}_{\mathsf{R}_+}$.

Proof, (i) $\Leftrightarrow$ (ii): This follows immediately from the corresponding equivalence in Theorem 4.42. It can also be proved directly by the same argument.

(iii) $\Rightarrow$ (i): By continuous mapping and dominated convergence, we may take the pairs (α, ν) and (α_n, ν_n) to be non-random. Then by (23) and Lemma 3.1

$$-t^{-1} \log E e^{-X_t f} = \alpha f + \int (1 - e^{-\mu f})\, \nu(d\mu) = \hat{\nu} g, \quad f \in C_S, \ t > 0,$$

and similarly for the processes $X^n = X_n$, where

$$g(s, \mu) = f(s)\, \frac{1 - e^{-\mu f}}{\mu f}\, (\mu S \wedge 1), \quad s \in S, \ \mu \in \mathcal{M}_S,$$

with $0/0$ interpreted as 1. Here g is again continuous, and when $\varepsilon \le f \le c$ it is also bounded, since in that case $g \le c$ for $\mu S \le 1$ and $g \le c/\varepsilon$ for $\mu S > 1$. Hence, (iii) then implies $X_t^n f \xrightarrow{d} X_t f$. In particular, this gives $X_t^n S \xrightarrow{d} X_t S$, and so, by approximation (FMP 4.28), $X_t^n f \xrightarrow{d} X_t f$ for arbitrary $f \in C_S$.

Here Theorem 4.11 yields $X_t^n \xrightarrow{wd} X_t$, which extends to the finite-dimensional distributions, since the ξ_n have independent increments. Condition (i) now follows by the same theorem.

(i) ⇒ (iii): Let the restriction of ξ to $S \times [0,1]$ be directed by (α,β), so that β is a Cox process directed by ν with $\alpha \perp\!\!\!\perp_\nu \beta$, and define $\hat\beta'$ as in (24) with ν replaced by β, and similarly for $\hat\beta'_1, \hat\beta'_2, \ldots$ in terms of $\xi_1, \xi_2, \ldots$. Assuming (i), we get $\hat\beta_n \xrightarrow{wd} \hat\beta$ by Theorem 4.42. Since the mapping $\mu \mapsto \hat\mu$ is bounded and continuous, $\hat\beta'$ is a continuous function of $\hat\beta$, and so $\hat\beta'_n \xrightarrow{wd} \hat\beta'$ by continuous mapping. Writing $\mathcal{K}' = \mathcal{K}_{S\times\mathcal{M}_S}$, we get by Theorem 4.10

$$\lim_{t\to 0} \inf_n E \exp(-t\|\hat\beta'_n\|) = \sup_{K\in\mathcal{K}'} \inf_n E \exp(-t'\hat\beta'_n K^c) = 1. \tag{25}$$

Now fix any $A \in \mathcal{S} \otimes \mathcal{M}_S$, and put $A_0\{s \in S;\, (s,0) \in A\}$, $h(\mu) = (\hat\mu \otimes \delta_\mu)A$, and $t' = 1 - e^{-t}$. Then Lemma 3.1 yields

$$\begin{aligned} Ee^{-t\hat\beta'_n A} &= E \exp(-t\alpha_n A_0 - t\beta_n h) \\ &= E \exp\Big\{-t\alpha_n A_0 - \nu_n(1 - e^{-th})\Big\} \\ &\le E \exp(-t'\alpha_n A_0 - t'\nu_n h) \\ &= Ee^{-t'\hat\nu_n A}, \end{aligned} \tag{26}$$

and so (25) remains true for the sequence $(\hat\nu_n)$, which is then weakly relatively compact by Theorem 4.10. If $\hat\nu_n \xrightarrow{wd} \rho$ along a sub-sequence $N' \subset \mathsf{N}$, for some random measure ρ on $S \times \mathcal{M}_S$, the direct assertion yields $\xi_n \xrightarrow{vd} \eta$ along N', for an associated λ-symmetric random measure η. But then $\eta \overset{d}{=} \xi$, and so $\rho \overset{d}{=} \hat\nu$ by the uniqueness of (α,ν), and (iii) follows, even along N. □

We finally consider convergence criteria for λ-symmetric random measures on some product spaces $S \times [0, t_n]$, where S is bounded and $t_n \to \infty$. Here the directing pairs (α_n, β_n) refer to the scaled random measures ξ'_n on $S \times [0,1]$, given by $\xi'_n B = \xi_n(t_n B)$. As in the previous proof, we need to modify the definition of $\hat\beta$ in (21), by writing $\hat\mu = \mu/(\mu S \vee 1)$ and setting

$$\hat\beta' = \alpha \otimes \delta_0 + \int (\hat\mu \otimes \delta_\mu)\, \beta(d\mu). \tag{27}$$

Theorem 4.44 *(symmetric measures on increasing sets) For bounded S, consider some λ-symmetric random measures ξ on $S\times\mathsf{R}_+$ and ξ_n on $S\times[0,t_n]$ with $t_n \to \infty$, directed by (α,ν) and (α_n,β_n), respectively, and define the random measures $\hat\nu$ and $\hat\beta'_1, \hat\beta'_2, \ldots$ and processes $X, X_1, X_2, \ldots$ as in* (24), (27)*, and* (22)*. Then these conditions are equivalent:*

(i) $\xi_n \xrightarrow{vd} \xi$ *in* $\mathcal{M}_{S\times[0,t]}$, $t > 0$,

(ii) $X_n \xrightarrow{Sd} X$ *in* $D([0,t], \mathcal{M}_S)$, $t > 0$,

(iii) $t_n^{-1}\hat\beta'_n \xrightarrow{wd} \hat\nu$ *in* $\mathcal{M}_{S\times\mathcal{M}_S}$.

As before, the space in (iii) reduces to $\mathcal{M}_{\mathsf{R}_+}$ when the ξ_n are λ-symmetric random measures on some increasing intervals $[0, t_n]$. Our proof is based on an elementary lemma, where $\|\cdot\|_t$ denotes total variation on the interval $[0, t]$.

Lemma 4.45 *(Poisson approximation) Let τ be a $U(0,1)$ random variable. Then for every $t \in [0,1]$, there exists a unit rate Poisson process η_t on $[0,1]$, satisfying*

$$E\|\delta_\tau - \eta_t\|_t^2 \le 3t^2/2, \quad t \in [0,1].$$

Proof: Let η be a unit rate Poisson process on $[0,1]$, and fix any $t \in [0,1]$. By Theorem 3.4 we may write $\eta = \sigma_1 + \cdots + \sigma_\kappa$, where $\sigma_1, \sigma_2, \dots$ are i.i.d. $U(0,t)$ and independent of $\kappa = \eta[0,t]$. Since $P\{\kappa > 0\} = 1 - e^{-t} \le t$, we may choose a $\kappa' \perp\!\!\!\perp \sigma_1$ in $\{0,1\}$ with $\kappa' \ge \kappa \wedge 1$ and $P\{\kappa' = 1\} = t$. Define $\sigma = \sigma_1 1\{\kappa' = 1\} + 1\{\kappa' = 0\}$, so that $\sigma \overset{d}{=} \tau$ on $[0,t]$, and note that

$$\begin{aligned}
E\|\delta_\sigma - \eta\|_t^2 &= E(\kappa' - \kappa)^2 \\
&= P\{\kappa < \kappa'\} + E\Big\{(\kappa - 1)^2;\, \kappa \ge 2\Big\} \\
&= t - (1 - e^{-t}) + \sum_{k \ge 2} (k-1)^2 \frac{t^k}{k!}\, e^{-t} \\
&\le \int_0^t (1 - e^{-s})\, ds + t^2 \sum_{k \ge 2} \frac{t^{k-2}}{(k-2)!}\, e^{-t} \le 3t^2/2.
\end{aligned}$$

Finally, Lemma 1.16 yields an $\eta_t \overset{d}{=} \eta$ with $(\tau, \eta_t) \overset{d}{=} (\sigma, \eta)$ on $[0,t]$. □

Proof of Theorem 4.44, (i) ⇔ (ii): This is again clear from Theorem 4.42.

(iii) ⇒ (i): As before, we may take the pairs (α_n, β_n) and (α, ν) to be non-random. Define $\tilde{\alpha}_n = \alpha_n / t_n$ and $\nu_n = \beta_n / t_n$, and write

$$\xi_n = \tilde{\alpha}_n \otimes \lambda + \sum\nolimits_j \beta_{nj} \otimes \delta_{\tau_{nj}}, \quad n \in \mathsf{N},$$

where $\beta_n = \sum_j \delta_{\beta_{nj}}$, and the τ_{nj} are i.i.d. $U(0, t_n)$ for each n. For fixed $t > 0$ and $n \in \mathsf{N}$ with $t_n \ge t$, Lemma 4.45 yields some independent Poisson processes η_{nj} on R_+ with constant rates t_n^{-1}, such that

$$E\|\delta_{\tau_{nj}} - \eta_{nj}\|_t^2 \le 2(t/t_n)^2, \quad n, j \in \mathsf{N}.$$

The η_{nj} may be combined into some Poisson processes η_n on $\mathcal{M}'_S \times \mathsf{R}_+$, given by

$$\eta_n = \sum\nolimits_j \delta_{\beta_{nj}} \otimes \eta_{nj}, \quad n \in \mathsf{N},$$

and since $E\eta_n = \nu_n$, we may form some λ-symmetric random measures $\tilde{\xi}_n$ on $S \times \mathsf{R}_+$, with directing pairs $(\tilde{\alpha}_n, \nu_n)$, by putting

$$\tilde{\xi}_n = \tilde{\alpha}_n \otimes \lambda + \iint (\mu \otimes \delta_t)\, \eta_n(d\mu\, dt), \quad n \in \mathsf{N}.$$

Now let $f : S \times \mathsf{R}_+ \to [0,1]$ be measurable with support in $S \times [0,t]$. Using the subadditivity of $x \wedge 1$, the inequality $cn \wedge 1 \le (c \wedge 1)n^2$, Fubini's theorem, Lemma 4.45, and (iii), we get

$$\begin{aligned}
E\Big(\big|(\xi_n - \tilde{\xi}_n)f\big| \wedge 1\Big) &= E\Big(\Big|\sum\nolimits_j \big\{\beta_{nj} \otimes (\delta_{\tau_{nj}} - \eta_{nj})\big\} f\Big| \wedge 1\Big) \\
&\le E\Big(\sum\nolimits_j \beta_{nj} S \,\|\, \delta_{\tau_{nj}} - \eta_{nj} \,\|_t \wedge 1\Big) \\
&\le \sum\nolimits_j E\Big(\beta_{nj} S \,\|\, \delta_{\tau_{nj}} - \eta_{nj} \,\|_t \wedge 1\Big) \\
&\le \sum\nolimits_j (\beta_{nj} S \wedge 1)\, E\|\, \delta_{\tau_{nj}} - \eta_{nj} \,\|_t^2 \\
&\lesssim (t/t_n)^2 \sum\nolimits_j (\beta_{nj} S \wedge 1) \\
&= (t/t_n)^2 \|\hat{\beta}'_n\| \to 0,
\end{aligned}$$

which shows that $\xi_n f - \tilde{\xi}_n f \xrightarrow{P} 0$. Since t and f were arbitrary and $\tilde{\xi}_n \xrightarrow{vd} \xi$ by Theorem 4.43, (i) follows by Theorem 4.11.

(i) $\Rightarrow$ (iii): For $t_n \ge 1$, let the restriction of ξ_n to $S \times [0,1]$ be directed by $(\tilde{\alpha}_n, \gamma_n)$, so that γ_n is a t_n^{-1}-thinning of β_n, and define $\hat{\gamma}'_n$ as in (24). Defining γ and $\hat{\gamma}'$ correspondingly for the random measure ξ, we see from (i) and Theorem 4.42 that $\hat{\gamma}'_n \xrightarrow{wd} \hat{\gamma}'$, and so by Theorem 4.10,

$$\lim_{t\to 0} \inf_n E \exp(-t\|\hat{\gamma}'_n\|) = \sup_{K \in \mathcal{K}'} \inf_n E \exp(-t' \hat{\gamma}'_n K^c) = 1, \tag{28}$$

where $\mathcal{K}' = \mathcal{K}_{S \times \mathcal{M}_S}$.

Now let $A \in \mathcal{S} \otimes \mathcal{M}_S$ be arbitrary, and define A_0, h, and t' as in (26). Using Lemma 3.1 and (20), we get as before

$$\begin{aligned}
E e^{-t\hat{\gamma}'_n A} &= E \exp(-t\tilde{\alpha}_n A_0 - t\gamma_n h) \\
&= E \exp\Big(-t\tilde{\alpha}_n A_0 + \beta_n \log\big\{1 - t_n^{-1}(1 - e^{-th})\big\}\Big) \\
&\le E \exp\big\{-t\tilde{\alpha}_n A_0 - \nu_n (1 - e^{-th})\big\} \\
&\le E \exp(-t'\tilde{\alpha}_n A_0 - t'\nu_n h) = E e^{-t'\hat{\nu}_n A}.
\end{aligned}$$

Hence, (28) extends to the sequence $(\hat{\nu}_n)$, and (iii) follows as before, by a standard compactness argument. □

Chapter 5

Stationarity in Euclidean Spaces

For any measure μ on R^d and constant $r \in \mathsf{R}^d$, we may introduce[1] the *shifted measure* $\theta_r\mu = \mu \circ \theta_r^{-1}$, where $\theta_r x = x + r$, so that $(\theta_r\mu)B = \mu(B - r)$ and $(\theta_r\mu)f = \mu(f \circ \theta_r)$. This amounts to shifting μ by an amount $|r|$ in the direction of the vector r. The shift operators θ_r clearly satisfy the semi-group property $\theta_{r+s} = \theta_r\theta_s$. They extend immediately to measures on any measurable product space $\mathsf{R}^d \times T$, where we may take T to be Borel.

A random measure ξ on $\mathsf{R}^d \times T$ is said to be *stationary*, if $\theta_r\xi \overset{d}{=} \xi$ for all $r \in \mathsf{R}^d$. More generally, we may restrict r to a closed additive subgroup $G \subset \mathsf{R}^d$, and say that ξ is *G-stationary* if $\theta_r\xi \overset{d}{=} \xi$ for every $r \in G$. Since stationarity is often confused with *invariance*, we emphasize that stationarity of ξ means invariance of its distribution, in the sense that $\mathcal{L}(\theta_r\xi) = \mathcal{L}(\xi)$ for all r. Invariance of ξ itself is much stronger and implies that $\xi = \lambda^d \otimes \nu$ a.s. for some random measure ν on T, where λ denotes Lebesgue measure on R.

Any stationary random measure ξ on R or Z satisfies the $0 - \infty$ law $\xi\mathsf{R}_+ = \xi\mathsf{R}_- = \infty$ a.s. on $\{\xi \neq 0\}$. Thus, any stationary simple point process $\xi \neq 0$ on R has infinitely many points τ_k, which may be enumerated as

$$\cdots < \tau_{-1} < \tau_0 \leq 0 < \tau_1 < \tau_2 < \cdots,$$

so that τ_1 is the first point of ξ in $(0, \infty)$. The mentioned law clearly implies the corresponding property for stationary processes on $\mathsf{R}^d \times T$.

For any stationary random measure ξ on R^d, we define the associated *Palm measure* Q_ξ on $\mathcal{M}_{\mathsf{R}^d}$ by

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

where I_1 denotes the unit cube in R^d. Here and below, f denotes an arbitrary measurable function on the appropriate space. The Palm measure is bounded only when $E\xi I_1 < \infty$, in which case it can be normalized to a *Palm distribution*. However, the stated version is usually preferable, since it often leads to simpler formulas.

It is suggestive to think of Q_ξ as the distribution of a random measure $\tilde{\xi}$ on S, defined on a measurable space $\tilde{\Omega}$ with a σ-finite *pseudo-probability*

[1] Recall that $\mu f = \int f d\mu$, whereas $(\mu \circ f^{-1})B = \mu(f^{-1}B)$ or $(\mu \circ f^{-1})g = \mu(g \circ f)$, where $(g \circ f)_s = g(f_s)$.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_5

measure $\tilde{P}$. This allows us to write $Q_\xi f = \tilde{E} f(\tilde{\xi})$, where $\tilde{E}$ denotes $\tilde{P}$-integration. In particular, we get for $d = 1$ the defining formula

$$\tilde{E} f(\tilde{\xi}) = E \int_0^1 f(\theta_{-r}\xi)\, \xi(dr), \quad f \geq 0. \tag{1}$$

In general, the Palm measure determines the underlying distribution $\mathcal{L}(\xi)$, which can be recovered through suitable inversion formulas.

If ξ is a simple point process on R, then $\tilde{\xi}$ is again simple with infinitely many points $\cdots < \tilde{\tau}_{-1} < \tilde{\tau}_0 < \tilde{\tau}_1 < \cdots$, where $\tilde{\tau}_0 = 0$ a.s. Though $\tilde{\xi}$ is not stationary in general, it has the striking property of *cycle stationarity*, defined by $\theta_{-\tilde{\tau}_n}\tilde{\xi} \overset{d}{=} \tilde{\xi}$ for all n. In other words, the spacing variables $\tilde{\tau}_n - \tilde{\tau}_{n-1}$ form a stationary sequence. By Theorem 5.4, the Palm transformation $\mathcal{L}(\xi) \mapsto \tilde{\mathcal{L}}(\tilde{\xi})$ provides a $1-1$ correspondence between the pseudo-distributions of all stationary point processes $\xi \neq 0$ and all cycle-stationary ones $\tilde{\xi}$. Here we also have the simple inversion formula

$$E f(\xi) = \tilde{E} \int_0^{\tilde{\tau}_1} f(\theta_{-r}\tilde{\xi})\, dr, \quad f \geq 0. \tag{2}$$

For simple point processes ξ on R^d, the original idea behind the Palm measures was to form the conditional distribution, given that ξ has a point at the origin. Since $\xi\{0\} = 0$ a.s., the conditioning can only be understood in an approximate sense. Assuming $I_1 \subset \mathsf{R}^d$ with $E\xi I_1 < \infty$ and considering neighborhoods B_n of 0 with $B_n \downarrow \{0\}$, we show in Theorem 5.5 that

$$\begin{aligned} &P\{\xi B_n = 1\} \sim P\{\xi B_n > 0\} \sim E\xi B_n, \\ &\left\| \mathcal{L}(\,\theta_{-\sigma_n}\xi \,|\, \xi B_n = 1) - \mathcal{L}(\tilde{\xi}) \right\| \to 0, \end{aligned} \tag{3}$$

where σ_n is defined for $\xi B_n = 1$ as the unique point of ξ in B_n, and $\mathcal{L}(\tilde{\xi})$ denotes the normalized Palm distribution of ξ.

For stationary point processes on R, formulas (1) and (2) exhibit a striking similarity, which suggests that we replace the Palm measure by the associated *spacing measure*, defined as follows. First, we introduce for any $\mu \in \mathcal{M}_{\mathsf{R}}$ the *distribution function* $\mu_t = \mu(0,t]$, interpreted as $-\mu(t,0]$ when $t < 0$. Next, we choose $\tilde{\mu}$ to be the image of Lebesgue measure λ under the mapping $t \mapsto \mu_t$, so that

$$\tilde{\mu} f = \int f(\mu_t)\, dt, \quad f \geq 0.$$

The distribution functions of μ and $\tilde{\mu}$ are essentially mutual inverses, which suggests that we refer to the map $\mu \mapsto \tilde{\mu}$ as *measure inversion*.

Given a stationary random measure ξ on R, we may now define the associated *spacing measure* P_ξ^0 by

$$P_\xi^0 f = E \int_0^{\xi_1} f(\theta_{-r}\tilde{\xi})\, dr, \quad f \geq 0.$$

In Theorem 5.6, we show that the Palm and spacing measures are essentially inversions of one another, that the spacing map preserves stationarity, and that a repeated application of the spacing transformation essentially leads back to the original distribution.

When ξ is a stationary random measure on R^d, the multivariate ergodic theorem yields $\xi B_n/\lambda^d B_n \to \bar{\xi}$ a.s., for any convex sets $B_n \uparrow \mathsf{R}^d$ with inner radii $r_n \to \infty$. Here the limit $\bar{\xi}$, known as the *sample intensity* of ξ, is given by $\bar{\xi} = E(\xi I_1 \,|\, \mathcal{I}_\xi)$, where $\mathcal{I}_\xi$ denotes the invariant σ-field for ξ, consisting of all ξ-measurable events invariant under shifts of ξ. Note that $\bar{\xi}$ can take any values in $[0,\infty]$. The convergence remains true in L^p, whenever $p \ge 1$ is such that $\|\xi I_1\|_p < \infty$.

The corresponding convergence in probability holds under much more general conditions. To make this precise, we say that the probability measures $\nu_1, \nu_2, \ldots$ are *asymptotically invariant*, if

$$\|\nu_n - \nu_n * \delta_x\| \to 0, \quad x \in \mathsf{R}^d.$$

They are also said to be *weakly asymptotically invariant*, if the convolutions $\nu_n * \mu$ are asymptotically invariant in the previous sense, for any probability measure $\mu \ll \lambda^d$. In Theorem 5.25, we show that $\xi * \nu_n \xrightarrow{vP} \bar{\xi}\lambda^d$ when the ν_n are weakly asymptotically invariant, whereas $\xi * \nu_n \xrightarrow{ulP} \bar{\xi}\lambda^d$ when they are asymptotically invariant in the strict sense. Here the latter, probabilistic modes of convergence, $\xrightarrow{vP}$ and $\xrightarrow{ulP}$, may be defined by the appropriate subsequence conditions, as in FMP 4.2, which avoids any reference to a specific metrization of the measure space $\mathcal{M}_{\mathsf{R}^d}$.

The previous results lead in particular to versions of some classical limit theorems, for ν_n-transforms of a stationary point process ξ with sample intensity $\bar{\xi} < \infty$ a.s., where the *ν_n-transforms* ξ_n are formed by independently shifting the points of ξ according to the distributions ν_n. Under the same conditions of weak or strict asymptotic invariance, we prove in Corollary 5.27 that $\xi_n \xrightarrow{vd} \eta$ or $\xi_n \xrightarrow{uld} \eta$, respectively, where η is a mixture of stationary Poisson processes with random rate $\bar{\xi}$. The results apply in particular to the case where the points of ξ perform independent random walks, based on a distribution μ satisfying some mild regularity conditions, as specified by Theorem 5.18.

We have seen how the Palm measure of a stationary random measure ξ on R^d, originally defined by an explicit formula, can also be formed, in the point process case, by local conditioning as in (3). A third approach is via averaging, where the appropriate ergodic theorems lead back and forth between the original distributions and their Palm versions. Among the numerous statements of this type given below, we quote only some results highlighting the striking relationship between the two distributions.

This requires some notation. Given a measurable function $f \ge 0$ on $\mathcal{M}_{\mathsf{R}}$

and a measure μ on R, we define the *average* $\bar{f}(\mu)$ by

$$t^{-1}\int_0^t f(\theta_{r+s}\mu)\,dr \to \bar{f}(\mu) \text{ as } t \to \pm\infty,$$

whenever the limit exists and is independent of s. Given a stationary random measure ξ on R with $0 < \bar{\xi} < \infty$ a.s., we next introduce a random measure η, whose $\tilde{P}$-distribution is given by the spacing measure P_ξ^0. Since η is again stationary, it has an associated sample intensity $\bar{\eta}$. We may also introduce the inverted measures $\tilde{\xi}$ and $\tilde{\eta}$. Then for any measurable function $f \geq 0$, we may derive the remarkable formulas

$$\begin{aligned} \tilde{E}\bar{f}(\tilde{\eta}) &= E\,\bar{\xi}f(\xi), &\qquad \tilde{E}f(\eta) &= E\,\bar{\xi}\bar{f}(\tilde{\xi}),\\ E\bar{f}(\tilde{\xi}) &= \tilde{E}\,\bar{\eta}f(\eta), &\qquad Ef(\xi) &= \tilde{E}\,\bar{\eta}\bar{f}(\tilde{\eta}). \end{aligned}$$

Introducing a random measure ζ with distribution $\tilde{E}(\,\bar{\eta};\, \eta \in \cdot\,)$, we arrive at the simplified versions

$$Ef(\zeta) = E\bar{f}(\tilde{\xi}), \qquad Ef(\xi) = E\bar{f}(\tilde{\zeta}).$$

We conclude with an extension of the classical *ballot theorem*, one of the earliest rigorous results in probability theory. The original claim was the fact that, if two candidates in an election, A and B, are getting the proportions p and $1-p$ of the votes, then A will lead throughout the counting of ballots with probability $(2p-1) \vee 0$. The most general version of the theorem, known to date, is based on some ergodic properties of a stationary, a.s. singular random measure ξ on R_+ or $[0,1)$, where *singularity* means that the absolutely continuous component vanishes, and stationarity on $[0,1)$ is defined cyclically, in the sense of addition modulo 1. Under the stated conditions, Theorem 5.51 asserts the existence of a $U(0,1)$ random variable $\sigma \perp\!\!\!\perp \mathcal{I}_\xi$ satisfying

$$\sigma \sup_{t>0} t^{-1}\xi[0,t] = \bar{\xi} \text{ a.s.}$$

5.1 Palm Measures and Cycle Stationarity

Let $(G, \mathcal{G})$ be a Borel measurable group with Haar measure λ, acting measurably on a space T, as explained in Section 7.1 below. Given a random measure ξ on G and a random element η in T, we say that the pair (ξ, η) is *G-stationary* if $\theta_r(\xi,\eta) \overset{d}{=} (\xi,\eta)$ for all $r \in G$, where $\theta_r\eta = r\eta$ and $\theta_r\xi = \xi \circ \theta_r^{-1}$, so that $(\theta_r\xi)B = \xi(\theta_r^{-1}B)$ and $(\theta_r\xi)f = \xi(f \circ \theta_r)$. We assume that $\xi B < \infty$ a.s. whenever $\lambda B < \infty$. Note that λ is an invariant supporting measure for ξ. The associated *Palm measure* $Q_{\xi,\eta}$ is given by

$$Q_{\xi,\eta}f = E\int f(r^{-1}\eta)\,g(r)\,\xi(dr), \quad f \in \mathcal{T}_+, \tag{4}$$

for any fixed $g \in \mathcal{G}_+$ with $\lambda g = 1$. In particular, we may choose $\eta = \xi$, or we may replace η by the pair (ξ, η).

We always assume (ξ, η) to be such that $Q_{\xi,\eta}$ is σ-finite. However, it is bounded only when $E\xi$ is σ-finite, in which case it can be normalized into a *Palm distribution* of η with respect to ξ. Even in general, it is suggestive and convenient to regard $Q_{\xi,\eta}$ as the pseudo-distribution of a random element $\tilde{\eta}$ in the same space T, so that the left-hand side of (4) becomes $\tilde{E}f(\tilde{\eta})$, where $\tilde{E}$ denotes expectation with respect to the underlying pseudo-probability measure $\tilde{P}$.

We show that $Q_{\xi,\eta}$ is independent of the choice of normalizing function g. The idea is that the pair (ξ, η) can be identified with a stationary, marked random measure on $G \times T$.

Lemma 5.1 *(coding) For a group G acting measurably on T, consider a stationary pair of a random measure ξ on G and a random element η in T, and define a random measure Ξ on $G \times T$ by*

$$\Xi f = \int f(r, r^{-1}\eta)\, \xi(dr), \quad f \geq 0.$$

Then Ξ is stationary under shifts in G only. It is further ergodic, whenever that is true for (ξ, η).

Proof: We may write the definition as $\Xi = \Phi(\xi, \eta)$. Introducing the partial shifts $\vartheta_r(p, t) = (rp, t)$, for $r, p \in G$ and $t \in T$, we get for any $r \in G$

$$\begin{aligned}(\Xi \circ \vartheta_r{}^{-1})f &= \Xi(f \circ \vartheta_r) = \int f(rp, p^{-1}\eta)\, \xi(dp) \\ &= \int f\big\{p, (r^{-1}p)^{-1}\eta\big\}\, (\xi \circ \theta_r^{-1})(dp) \\ &= \int f(p, p^{-1}r\eta)\, (\theta_r\xi)(dp),\end{aligned}$$

which shows that $\Xi \circ \vartheta_r^{-1} = \Phi \circ \theta_r(\xi, \eta)$. The asserted properties now follow from those of (ξ, η). $\square$

We now specialize to the case of $G = \mathsf{R}^d$, beginning with the following $0 - \infty$ laws for stationary random sets and measures. The version for sets will be needed in a later section.

Lemma 5.2 *($0-\infty$ laws) Any stationary random measure ξ on R^d satisfies $\|\xi\| \in \{0, \infty\}$ a.s., and similarly for random sets. In fact, we have*

(i) *for any stationary random measure ξ on R or Z,*

$$\xi\mathsf{R}_\pm = \infty \ \text{a.s. on } \{\xi \neq 0\},$$

(ii) *for any stationary random closed set φ in R or Z,*

$$\sup(\pm\varphi) = \infty \ \text{a.s. on } \{\varphi \neq \emptyset\}.$$

Proof: (i) First let ξ be a random measure on R. Using the stationarity of ξ and Fatou's lemma, we get for any $t \in \mathsf{R}$ and $h, \varepsilon > 0$

$$\begin{aligned} P\{\,\xi[t, t+h) > \varepsilon\} &= \limsup_{n\to\infty} P\Big\{\,\xi[(n-1)h, nh) > \varepsilon\Big\} \\ &\le P\Big\{\,\xi[(n-1)h, nh) > \varepsilon \ \text{ i.o.}\Big\} \\ &\le P\{\,\xi[0,\infty) = \infty\}. \end{aligned}$$

Letting $\varepsilon \to 0$, $h \to \infty$, and $t \to -\infty$, in this order, we obtain $P\{\xi \neq 0\} \le P\{\xi\mathsf{R}_+ = \infty\}$. Since $\xi\mathsf{R}_+ = \infty$ implies $\xi \neq 0$, the two events agree a.s. Applying this result to the reflected random measure $\tilde{\xi}B = \xi(-B)$, we see that also $\xi\mathsf{R}_- = \infty$ a.s. on $\{\xi \neq 0\}$. For random measures on Z, the same argument applies with t and h restricted to Z.

(ii) Here we write instead

$$\begin{aligned} P\Big\{\varphi \cap [t, t+h) \neq \emptyset\Big\} &= \limsup_{n\to\infty} P\Big\{\varphi \cap [(n-1)h, nh) \neq \emptyset\Big\} \\ &\le P\Big\{\varphi \cap [(n-1)h, nh) \neq \emptyset \ \text{ i.o.}\Big\} \\ &\le P\{\,\sup \varphi = \infty\}. \end{aligned}$$

The assertion follows as before, as we let $h \to \infty$ and then $t \to -\infty$. $\square$

The Palm definition (4), applied to the pairs (ξ, η), is essentially a $1-1$ correspondence, and we proceed to derive some useful inversion formulas. When ξ is a simple point process, then so is $\tilde{\xi}$, and we note that $\tilde{\xi}\{0\} = 1$. The result may then be stated in terms of the *Voronoi cell* around 0, given by

$$V_0(\mu) = \{x \in \mathsf{R}^d;\ \mu B_x^{|x|} = 0\}, \quad \mu \in \mathcal{N}(\mathsf{R}^d),$$

where B_x^r denotes the open ball around x of radius $r > 0$. When $d = 1$, the supporting points of μ may be enumerated in increasing order as $t_n(\mu)$, subject to the convention $t_0(\mu) \le 0 < t_1(\mu)$.

Theorem 5.3 *(uniqueness and inversion) For R^d acting measurably on T, consider a stationary pair of a random measure ξ on R^d and a random element η in T. Then $\mathcal{L}\{(\xi,\eta);\ \xi \neq 0\}$ is determined by the Palm measure $\mathcal{L}(\tilde{\xi}, \tilde{\eta})$, and when $f \in \mathcal{T}_+$, we have*

(i) *for any measurable function $g > 0$ on R^d with $\xi g < \infty$ a.s.,*

$$E\{f(\eta);\ \xi \neq 0\} = \tilde{E}\int \frac{f(\theta_x\tilde{\eta})}{\tilde{\xi}(g \circ \theta_x)}\, g(x)\, dx,$$

(ii) *for any simple point process ξ on R^d,*

$$E\{f(\eta);\ \xi \neq 0\} = \tilde{E}\int_{V_0(\tilde{\xi})} f(\theta_{-x}\tilde{\eta})\, dx,$$

(iii) *for any simple point process ξ on R,*

$$E\{f(\eta);\ \xi \neq 0\} = \tilde{E}\int_0^{t_1(\tilde{\xi})} f(\theta_{-r}\tilde{\eta})\, dr.$$

Replacing η by (ξ,η) and $\tilde{\eta}$ by $(\tilde{\xi},\tilde{\eta})$ yields inversion formulas for the associated Palm measures. A similar remark applies to subsequent theorems.

Proof: We may prove the stated formulas with η and $\tilde{\eta}$ replaced by $\zeta=(\xi,\eta)$ and $\tilde{\zeta}=(\tilde{\xi},\tilde{\eta})$, respectively.

(i) Multiplying (4) by $\lambda^d B$ and extending by a monotone-class argument, we get for any measurable function $f\geq 0$

$$\tilde{E}\int f(\tilde{\zeta},x)\,dx = E\int f(\theta_{-x}\zeta,x)\,\xi(dx),$$

and so, by a simple substitution,

$$\tilde{E}\int f(\theta_x\tilde{\zeta},x)\,dx = E\int f(\zeta,x)\,\xi(dx). \tag{5}$$

In particular, we have for any measurable functions $f,g\geq 0$

$$\tilde{E}\int f(\theta_x\tilde{\zeta})\,g(x)\,dx = Ef(\zeta)\,\xi g.$$

Choose $g>0$ with $\xi g<\infty$ a.s., so that $\xi g>0$ iff $\xi\neq 0$. The asserted formula now follows, as we replace f by the function

$$h(\mu,t)=\frac{f(\mu,t)}{\mu g}\,1\{\mu g>0\},\quad \mu\in\mathcal{M}(\mathsf{R}^d),\;\; t\in T.$$

(ii) Apply (5) to the function

$$h(\mu,t,x)=f(\mu,t)\,1\{\mu B_0^{|x|}=0\}.$$

Since $(\theta_x\mu)B_0^{|x|}=0$ iff $-x\in V_0(\mu)$, and since a stationary point process $\xi\neq 0$ has an a.s. unique point closest to 0, we get

$$\begin{aligned}\tilde{E}\int_{V_0(\tilde{\xi})} f(\theta_{-x}\tilde{\zeta})\,dx &= \tilde{E}\int f(\theta_x\tilde{\zeta})\,1\Big\{(\theta_x\tilde{\xi})B_0^{|x|}=0\Big\}\,dx\\ &= E\,f(\zeta)\int 1\{\xi B_0^{|x|}=0\}\,\xi(dx)\\ &= E\{f(\zeta);\,\xi\neq 0\}.\end{aligned}$$

(iii) Apply (5) to the function

$$h(\mu,t,r)=f(\mu,t)\,1\{t_0(\mu)=r\}.$$

Since $t_0(\theta_r\tilde{\xi})=r$ iff $-r\in[0,t_1(\tilde{\xi}))$, and since $|t_0(\xi)|<\infty$ a.s. when $\xi\neq 0$, by Lemma 5.2, we get

$$\begin{aligned}\tilde{E}\int_0^{t_1(\tilde{\xi})} f(\theta_{-r}\tilde{\zeta})\,dr &= \tilde{E}\int f(\theta_r\tilde{\zeta})\,1\Big\{t_0(\theta_r\tilde{\xi})=r\Big\}\,dr\\ &= E\,f(\zeta)\int 1\{t_0(\xi)=r\}\,\xi(dr)\\ &= E\{f(\zeta);\,\xi\neq 0\,\}.\end{aligned}$$

□

For suitable pairs of a simple point process ξ on R and a random element η in a space T, we give a striking relationship between stationarity in discrete and continuous time. Assuming $\xi\mathsf{R}_\pm = \infty$ a.s., we say that the pair (ξ, η) is *cycle-stationary*, if $\theta_{-\tau_n}(\xi, \eta) \overset{d}{=} (\xi, \eta)$ for every $n \in \mathsf{Z}$, where $\tau_n = t_n(\xi)$. Informally, this means that the excursions of (ξ, η) between the points of ξ form a stationary sequence. To keep the distinction clear, we may refer to ordinary stationarity on R as *time stationarity.*

Theorem 5.4 *(time and cycle stationarity, Kaplan) With R acting measurably on T, consider some simple point processes ξ, $\tilde{\xi}$ on R and random elements η, $\tilde{\eta}$ in T, and put $\zeta = (\xi, \eta)$ and $\tilde{\zeta} = (\tilde{\xi}, \tilde{\eta})$. Then the relations*

(i) $\quad \tilde{E} f(\tilde{\zeta}) = E \int_0^1 f(\theta_{-r}\zeta)\, \xi(dr),$

(ii) $\quad E f(\zeta) = \tilde{E} \int_0^{t_1(\tilde{\xi})} f(\theta_{-r}\tilde{\zeta})\, dr,$

provide a $1-1$ correspondence between the [pseudo-]distributions of all time-stationary pairs ζ with $\xi \neq 0$ a.s. and all cycle-stationary ones $\tilde{\zeta}$ with $\tilde{\xi}\{0\} = 1$ a.s. and $\tilde{E} t_1(\tilde{\xi}) = 1$.

Proof: First let ζ be stationary with $\xi \neq 0$ a.s., put $\sigma_k = t_k(\xi)$, and define $\mathcal{L}(\tilde{\zeta})$ by (i). Then for any $n \in \mathsf{N}$ and bounded, measurable $f \geq 0$, we have

$$\begin{aligned} n\, \tilde{E} f(\tilde{\zeta}) &= E \int_0^n f(\theta_{-r}\zeta)\, \xi(dr) \\ &= E \sum_{\sigma_k \in (0,n)} f(\theta_{-\sigma_k}\zeta). \end{aligned}$$

Writing $\tau_k = t_k(\tilde{\xi})$, we get by a suitable substitution

$$n\, \tilde{E} f(\theta_{-\tau_1}\tilde{\zeta}) = E \sum_{\sigma_k \in (0,n)} f(\theta_{-\sigma_{k+1}}\zeta),$$

and so, by subtraction,

$$\left| \tilde{E} f(\theta_{-\tau_1}\tilde{\zeta}) - \tilde{E} f(\tilde{\zeta}) \right| \leq 2n^{-1}\|f\|.$$

As $n \to \infty$, we get $\tilde{E} f(\theta_{-\tau_1}\tilde{\zeta}) = \tilde{E} f(\tilde{\zeta})$, and so $\theta_{-\tau_1}\tilde{\zeta} \overset{d}{=} \tilde{\zeta}$, which means that $\tilde{\zeta}$ is cycle-stationary. Note that (ii) holds in this case by Theorem 5.3 (iii). Taking $f \equiv 1$ gives $\tilde{E} t_1(\tilde{\xi}) = 1$.

Conversely, let $\tilde{\zeta}$ be cycle-stationary with $\tilde{E} t_1(\tilde{\xi}) = 1$, and define $\mathcal{L}(\zeta)$ by (ii). Then for n and f as above,

$$n\, E f(\zeta) = E \int_0^{\tau_n} f(\theta_{-r}\zeta)\, dr,$$

and so for any $h \in \mathsf{R}$,

$$\begin{aligned} n\, E f(\theta_{-h}\zeta) &= E \int_0^{\tau_n} f(\theta_{-r-h}\tilde{\zeta})\, dr \\ &= E \int_h^{\tau_n + h} f(\theta_{-r}\tilde{\zeta})\, dr, \end{aligned}$$

whence, by subtraction,

$$\left|Ef(\theta_{-h}\zeta) - Ef(\zeta)\right| \le 2n^{-1}|h|\,\|f\|.$$

As $n \to \infty$, we get $Ef(\theta_{-h}\zeta) = Ef(\zeta)$, and so $\theta_{-h}\zeta \overset{d}{=} \zeta$, which means that ζ is stationary.

Now introduce as in (i) a pair $\tilde\zeta' = (\tilde\xi', \tilde\eta')$ with associated expectation operator $\tilde E'$, satisfying

$$\tilde E' f(\tilde\zeta') = E\int_0^1 f(\theta_{-r}\zeta)\,\xi(dr). \tag{6}$$

Using Theorem 5.3 (iii) and comparing with (ii), we obtain

$$\tilde E' \int_0^{t_1(\tilde\xi')} f(\theta_{-r}\tilde\zeta')\,dr = \tilde E \int_0^{t_1(\tilde\xi)} f(\theta_{-r}\tilde\zeta)\,dr.$$

Replacing $f(\mu,t)$ by $f(\theta_{-t_0(\mu)}(\mu,t))$, and noting that $-t_0(\theta_{-r}\mu) = r$ when $\mu\{0\} = 1$ and $r \in [0, t_1(\mu))$, we get

$$\tilde E' t_1(\tilde\xi')\, f(\tilde\zeta') = \tilde E\, t_1(\tilde\xi)\, f(\tilde\zeta).$$

A further substitution yields $\tilde E' f(\tilde\zeta') = \tilde E\, f(\tilde\zeta)$, and so (i) holds by (6). □

When $E\xi$ is locally finite, the Palm distribution for the pair (ξ, η) may be thought of as the conditional distribution of η, given that $\xi\{0\} = 1$. The interpretation is justified by the following result, which also provides an asymptotic formula for the probabilities of hitting small Borel sets. Write $B_n \to \{0\}$ for the convergence $\sup\{|x|;\, x \in B_n\} \to 0$, and let $\|\cdot\|$ denote the total variation norm.

Theorem 5.5 *(local hitting and conditioning, Ryll-Nardzewski, König, Matthes) For R^d acting measurably on T, consider a stationary pair of a simple point process ξ on R^d and a random element η in T. Let $B_1, B_2, \ldots \in \mathcal{B}^d$ with $\lambda^d B_n > 0$ and $B_n \to \{0\}$, and let f be bounded, measurable, and shift-continuous. Assume $E\xi B_1 < \infty$, and let $\mathcal{L}(\tilde\eta)$ denote the normalized Palm distribution of η. When $\xi B_n = 1$, let σ_n be the unique supporting point of ξ in B_n. Then*

(i) $P\{\xi B_n = 1\} \sim P\{\xi B_n > 0\} \sim E\xi B_n$,

(ii) $\left\|\mathcal{L}(\,\theta_{-\sigma_n}\eta \,|\, \xi B_n = 1) - \mathcal{L}(\tilde\eta)\right\| \to 0$,

(iii) $E\{f(\eta) \,|\, \xi B_n > 0\} \to Ef(\tilde\eta)$.

Proof: (i) Since $\tilde{\xi}\{0\} = 1$ a.s., we have $(\theta_x \tilde{\xi})B_n > 0$ iff $x \in B_n$, and so by Theorem 5.3 (ii),

$$\begin{aligned}\frac{P\{\xi B_n > 0\}}{E\xi[0,1]^d} &= E\int_{V_0(\tilde{\xi})} 1\Big\{(\theta_{-x}\tilde{\xi})B_n > 0\Big\}\, dx \\ &\geq E\lambda^d\Big\{V_0(\tilde{\xi}) \cap (-B_n)\Big\}.\end{aligned}$$

Dividing by $\lambda^d B_n$ and using Fatou's lemma, we get

$$\begin{aligned}\liminf_{n\to\infty} \frac{P\{\xi B_n > 0\}}{E\xi B_n} &\geq \liminf_{n\to\infty} \frac{E\lambda^d\Big\{V_0(\tilde{\xi}) \cap (-B_n)\Big\}}{\lambda^d B_n} \\ &\geq E \liminf_{n\to\infty} \frac{\lambda^d\Big\{V_0(\tilde{\xi}) \cap (-B_n)\Big\}}{\lambda^d B_n} = 1,\end{aligned}$$

The assertions now follow from the elementary relations

$$\begin{aligned}2 \cdot 1\{k > 0\} - k &\leq 1\{k = 1\} \\ &\leq 1\{k > 0\} \leq k, \quad k \in \mathsf{Z}_+.\end{aligned}$$

(ii) Introduce on T the measures

$$\begin{aligned}\mu_n &= E\int_{B_n} 1\{\theta_{-x}\eta \in \cdot\}\, \xi(dx), \\ \nu_n &= \mathcal{L}(\theta_{-\sigma_n}\eta;\, \xi B_n = 1),\end{aligned}$$

and put $m_n = E\xi B_n$ and $p_n = P\{\xi B_n = 1\}$. By (4) the stated total variation equals

$$\begin{aligned}\left\|\frac{\nu_n}{p_n} - \frac{\mu_n}{m_n}\right\| &\leq \left\|\frac{\nu_n}{p_n} - \frac{\nu_n}{m_n}\right\| + \left\|\frac{\nu_n}{m_n} - \frac{\mu_n}{m_n}\right\| \\ &\leq p_n\left|\frac{1}{p_n} - \frac{1}{m_n}\right| + \frac{1}{m_n}|p_n - m_n| = 2\left|1 - \frac{p_n}{m_n}\right|,\end{aligned}$$

which tends to 0 in view of (i).

(iii) Here we write

$$\begin{aligned}&\Big|\, E\{f(\eta)\,|\,\xi B_n > 0\} - Ef(\tilde{\eta})\Big| \\ &\quad \leq \Big|\, E\{f(\eta)\,|\,\xi B_n > 0\} - E\{f(\eta)\,|\,\xi B_n = 1\}\,\Big| \\ &\quad\quad + \Big|\, E\Big\{f(\eta) - f(\theta_{-\sigma_n}\eta)\,\Big|\,\xi B_n = 1\Big\}\,\Big| \\ &\quad\quad + \Big|\, E\,\Big\{f(\theta_{-\sigma_n}\eta)\,|\,\xi B_n = 1\Big\} - Ef(\tilde{\eta})\,\Big|.\end{aligned}$$

By (i) and (ii), the first and last terms on the right tend to 0 as $n \to \infty$. To estimate the middle term, we introduce on T the bounded, measurable functions

$$g_\varepsilon(t) = \sup_{|x|<\varepsilon} |\, f(\theta_{-x}t) - f(t)\,|, \quad \varepsilon > 0,$$

and conclude from (ii) that, for large enough n,

$$\Big| E\Big\{ f(\eta) - f(\theta_{-\sigma_n}\eta) \,\Big|\, \xi B_n = 1 \Big\} \Big| \\ \le E\Big\{ g_\varepsilon(\theta_{-\sigma_n}\eta) \,\Big|\, \xi B_n = 1 \Big\} \to E g_\varepsilon(\tilde\eta).$$

Since $Eg_\varepsilon(\tilde\eta) \to 0$ as $\varepsilon \to 0$, by continuity and dominated convergence, the asserted convergence follows. □

5.2 Inversion and Spacing Measures

The similarity between the formulas in Theorem 5.4 suggests that we look for a symmetric version of the Palm correspondence. Our construction of spacing measures is based on an elementary notion of *measure inversion* $\mu \mapsto \tilde\mu$, defined as follows. First we introduce, for any $\mu \in \mathcal{M} = \mathcal{M}_{\mathsf{R}}$, the associated right-continuous *distribution function*

$$\mu_t = \begin{cases} \mu(0,t], & t \ge 0, \\ -\mu(t,0], & t < 0. \end{cases}$$

Then $\tilde\mu$ is defined as the image of Lebesgue measure λ under the mapping $t \mapsto \mu_t$, so that[2]

$$\tilde\mu f = \int f(\mu_t)\, dt, \quad f \ge 0. \tag{7}$$

Informally, $\tilde\mu$ is the unique measure on R, whose distribution function is the inverse of the distribution function of μ.

For some simple examples, note that the measure $c\lambda$ with $c > 0$ has inverse $c^{-1}\lambda$, and also that, if $\mu = \sum_n \delta_{t_n}$ for some strictly increasing sequence t_n, $n \in \mathsf{Z}$, with $t_{-1} \le 0 < t_0$, then $\tilde\mu = \sum_n (t_n - t_{n-1})\, \delta_n$. (In the latter case, the point masses of $\tilde\mu$ are given by the spacing sequence of μ, which motivates our subsequent terminology.) In general, $\tilde\mu$ is again locally finite iff $\mu\mathsf{R}_\pm = \infty$. Some further properties of the measure inversion will be given below.

Given a stationary random measure ξ on R, we define the associated *spacing measure* P^0_ξ by

$$P^0_\xi f = E\int_0^{\xi_1} f(\theta_{-r}\tilde\xi)\, dr, \quad f \ge 0. \tag{8}$$

Note that P^0_ξ is again restricted to $\mathcal{M}$, since $\tilde\xi \in \mathcal{M}$ a.s. on $\{\xi_1 > 0\}$ by

[2]Here and below, the functions $f, g, \ldots$ are understood to be non-negative and measurable with respect to the appropriate σ-fields.

Lemma 5.2. To see that it is also σ-finite, put $A_n = \{\mu;\, \mu_n \wedge |\mu_{-n}| > 1\}$. Then $P^0_\xi A_n < \infty$, since $\theta_{-r}\tilde{\xi} \in A_n$ for some $r \in (0, \xi_1)$ implies $\xi_1 \le 2n$; moreover, $P^0_\xi \bigcap_n A^c_n = 0$, since $\tilde{\xi}_n \wedge |\tilde{\xi}_{-n}| \to \infty$ a.s.

More generally, for any measure ν on $\mathcal{M}$, we define the associated spacing measure ν^0 by

$$\nu^0 f = \int \nu(d\mu) \int_0^{\mu_1} f(\theta_{-r}\tilde{\mu})\, dr, \quad f \ge 0. \tag{9}$$

For measures ν of interest, ν^0 is again σ-finite and restricted to $\mathcal{M}$. Iterating the spacing map leads to the repeated spacing measure $\nu^{00} = (\nu^0)^0$. The inversion $\mu \mapsto \tilde{\mu}$ on $\mathcal{M}$ induces an inversion $\tilde{\nu} = \nu\{\mu;\, \tilde{\mu} \in \cdot\}$ on $\mathcal{M}_{\mathcal{M}}$, and we write $\tilde{P}^0_\xi$ for the inversion of P^0_ξ. Note that the construction of $\tilde{P}^0_\xi$ from P_ξ actually involves two inversions, where the first one is implicit already in the definition of P^0_ξ.

We may now state our main result, which gives some basic properties of the spacing measure P^0_ξ, along with its relations to the Palm measure Q_ξ. Recall that f denotes an arbitrary measurable function ≥ 0.

Theorem 5.6 *(spacing measure) Let ξ be a stationary random measure on* R *with distribution P_ξ, Palm measure Q_ξ, and spacing measure P^0_ξ. Then*

(i) $Q_\xi = \tilde{P}^0_\xi$,

(ii) $P^0_\xi f = \int Q_\xi(d\mu) \int_0^1 f(\theta_{r\mu\{0\}}\tilde{\mu})\, dr$,

(iii) $P^0_\xi \circ \theta_h^{-1} = P^0_\xi, \quad h > 0$,

(iv) $P^{00}_\xi = P_\xi(\cdot \setminus \{0\})$.

Here (i) and (ii) show that the Palm and spacing measures are essentially inversions of one another, except that an extra randomization is needed in (ii), to resolve the possible discontinuity at 0. Equation (iii) shows that the spacing map preserves stationarity. Finally, (iv) shows how a repeated application of the spacing transformation essentially leads back to the original distribution.

If ξ is locally integer-valued, we may replace P^0_ξ by the measure on Z,

$$p^0_\xi f = E \sum_{k=1}^{\xi_1} f(\theta_{-k}\tilde{\xi}), \quad f \ge 0,$$

which is invariant under discrete shifts. Note that $p^0_\xi = P^0_\xi \circ \kappa^{-1}$, where $\kappa(\mu) = \mu \circ n^{-1}$ with $n(t) = [t]$. Conversely, P^0_ξ may be recovered from p^0_ξ through a uniform randomization, as in

$$P^0_\xi f = \int p^0_\xi(d\mu) \int_0^1 f(\theta_{-r}\mu)\, dr, \quad f \ge 0.$$

Parts (i) and (iv) of Theorem 5.6 remain valid with P_ξ^0 replaced by p_ξ^0, whereas (ii) needs to be replaced by

$$p_\xi^0 f = \int \frac{Q_\xi(d\mu)}{\mu\{0\}} \sum_{k=1}^{\mu\{0\}} f(\theta_k \tilde{\mu}), \quad f \geq 0.$$

For stationary random measures ξ on Z, we need to replace (iv) by the inversion formula

$$E\{f(\xi);\, \xi \neq 0\} = \int P_\xi^0(d\mu) \sum_{k=1}^{\mu_1} f(\theta_{-k}\tilde{\mu}), \quad f \geq 0.$$

To see this, we may apply the previous formulas to the stationary random measure $\theta_\vartheta \xi$ on R, where ϑ is $U(0,1)$ and independent of ξ. Finally, when ξ is a stationary and integer-valued random measure on Z, we may replace all integrals with respect to continuous shifts by the corresponding sums. In all three cases, counterparts of formulas (i)$-$(iv) can also be proved directly by similar arguments.

Several lemmas are needed for the proof of Theorem 5.6. We begin with some elementary relations involving inversions and inverses. Recall that $\mu \mapsto \mu_t$ denotes the right-continuous distribution function of μ, subject to the centering condition $\mu_0 = 0$, and write $r \mapsto \mu_r^{-1}$ for the right-continuous inverse of (μ_t). First, we list some basic properties of the measure inversion. Here a basic role is played by the quantity

$$\gamma = \mu_0^{-1} = \inf\{t > 0;\, \mu(0,t] > 0\}.$$

Lemma 5.7 *(inversions and inverses) Let $\mu \in \mathcal{M}$ with $\tilde{\mu} \in \mathcal{M}$, and put $\gamma = \mu_0^{-1}$. Then for any $t \in \mathsf{R}$,*

(i) $(\theta_{-t}\mu)^\sim = \theta_{-\mu_t}\, \tilde{\mu}$,

(ii) $\tilde{\mu}_t = \mu_t^{-1} - \gamma$,

(iii) $\tilde{\tilde{\mu}} = \theta_{-\gamma}\, \mu$,

(iv) $(\theta_{-t}\tilde{\mu})^\sim = \theta_{-\mu_t^{-1}}\mu$.

Proof: (i) For measurable $f \geq 0$ and any $t \in \mathsf{R}$, we have

$$\begin{aligned}
(\theta_{-t}\mu)^\sim f &= \int f((\theta_{-t}\mu)_s)\, ds \\
&= \int f(\mu_{s+t} - \mu_t)\, ds \\
&= \int (f \circ \theta_{-\mu_t})(\mu_{s+t})\, ds \\
&= \int (f \circ \theta_{-\mu_t})(\mu_s)\, ds \\
&= \tilde{\mu}(f \circ \theta_{-\mu_t}) = (\theta_{-\mu_t}\tilde{\mu}) f.
\end{aligned}$$

(ii) For any $t \geq 0$,

$$\tilde{\mu}_t = \tilde{\mu}(0,t] = \int 1_{(0,t]}(\mu_s)\, ds = \mu_t^{-1} - \gamma,$$
$$\tilde{\mu}_{-t} = -\tilde{\mu}(-t,0] = -\int 1_{(-t,0]}(\mu_s)\, ds = \mu_{-t}^{-1} - \gamma.$$

(iii) Using (ii) and the fact that the function μ^{-1} maps λ into μ, we get for measurable $f \geq 0$

$$\begin{aligned} \tilde{\tilde{\mu}} f &= \int f(\tilde{\mu}_t)\, dt = \int f(\mu_t^{-1} - \gamma)\, dt \\ &= \int (f \circ \theta_{-\gamma})(\mu_t^{-1})\, dt \\ &= \mu(f \circ \theta_{-\gamma}) = (\theta_{-\gamma}\mu) f. \end{aligned}$$

(iv) Combining (i), (ii), and (iii) gives

$$(\theta_{-t}\tilde{\mu})^{\sim} = \theta_{-\tilde{\mu}_t}\, \tilde{\tilde{\mu}} = \theta_{-\mu_t^{-1}+\gamma}\, \tilde{\tilde{\mu}} = \theta_{-\mu_t^{-1}}\mu. \qquad \square$$

Turning to inversions within integrals, we write $\tilde{f}(\mu) = f(\tilde{\mu})$, for typographical convenience.

Lemma 5.8 *(inversion in integrals) Let $\mu \in \mathcal{M}$ with $\tilde{\mu} \in \mathcal{M}$. Then for any constants $a, b > 0$ and measurable functions $f \geq 0$,*

(i) $\displaystyle\int_0^{\mu_1} \tilde{f}(\theta_{-r}\tilde{\mu})\, dr = \int_{(0,1]} f(\theta_{-s}\,\mu)\, \mu(ds),$

(ii) $\displaystyle\int_0^a (\theta_{-r}\tilde{\mu})_b\, dr = \int_0^b (\theta_{-r}\tilde{\mu})_a\, dr.$

Proof: (i) Using Lemma 5.7 (iv), and noting that the function μ^{-1} on R maps λ on $(0, \mu_1]$ into μ on $(0,1]$, we obtain

$$\begin{aligned} \int_0^{\mu_1} \tilde{f}(\theta_{-r}\tilde{\mu})\, dr &= \int_0^{\mu_1} f(\theta_{-\mu_r^{-1}}\mu)\, dr \\ &= \int_{(0,1]} f(\theta_{-s}\mu)\, \mu(ds). \end{aligned}$$

(ii) By (7) and Fubini's theorem, the assertion is equivalent to

$$\int dt \int_0^a 1_{[\mu_t - b,\, \mu_t)}(r)\, dr = \int dt \int_0^b 1_{[\mu_t - a,\, \mu_t)}(r)\, dr,$$

which holds since

$$\int_0^a 1_{[\mu_t - b,\, \mu_t)}(r)\, dr = \Big\{ \mu_t - (\mu_t - a)_+ - (\mu_t - b)_+ \Big\}_+,$$

and similarly with a and b interchanged. $\square$

When φ is a stationary random closed set in R, the process $X_t = f(\varphi - t)$ is clearly stationary for every measurable function, and so $E(X_t - X_0) = 0$, whenever X is integrable. To appreciate the next result, note that the variables γ_t below need not be integrable.

Lemma 5.9 *(right entry) Let φ be a stationary random closed set in* R *with $\varphi \neq \emptyset$ a.s., and define*

$$\gamma_t = \inf\Big\{h > 0;\ \varphi \cap (t, t+h) \neq \emptyset\Big\}, \quad t \in \mathsf{R}.$$

Then the increments $\gamma_t - \gamma_s$ are integrable with mean 0.

Proof: Since $\varphi \neq \emptyset$ a.s., we have a.s. $\gamma_t < \infty$ for all t, by Lemma 5.2 (ii). By scaling we may take $t = 1$. Since $\gamma_0 \leq \gamma_1 + 1$, we note that $E(\gamma_1 - \gamma_0)$ exists in $[-1, \infty]$. For fixed $n \in \mathsf{N}$, define $\varphi_n = \{k/n;\ \varphi((k-1, k]/n) \neq \emptyset\}$, and let γ_t^n denote the associated variables γ_t. Since φ_n is again stationary under shifts in Z/n and $|(\gamma_1^n - \gamma_0^n) - (\gamma_1 - \gamma_0)| \leq n^{-1}$, it is enough to prove that $E(\gamma_1^n - \gamma_0^n) = 0$ for each n. By scaling and stationarity, we may take $n = 1$, which reduces the assertion to the corresponding statement in discrete time.

Here we note that

$$\begin{aligned} E(\gamma_1 - \gamma_0) &= E(\gamma_1 - \gamma_0;\ \gamma_0 = 1) - P\{\gamma_0 > 1\} \\ &= E(\gamma_1;\ \gamma_0 = 1) - 1. \end{aligned}$$

Writing $\xi_k = 1\{k \in \varphi\}$ for $k \in \mathsf{Z}$, we get by stationarity and Lemma 5.2

$$\begin{aligned} E(\gamma_1;\ \gamma_0 = 1) &= E(\gamma_0;\ \xi_0 = 1) \\ &= \sum_{k \geq 0} P\{\xi_0 = 1,\ \gamma_0 > k\} \\ &= \sum_{k \geq 0} P\{\xi_0 = 1,\ \xi_1 = \cdots = \xi_k = 0\} \\ &= \sum_{k \geq 0} P\{\xi_{-k-1} = 1,\ \xi_{-k} = \cdots = \xi_{-1} = 0\} \\ &= P \bigcup_{k<0} \{\xi_k = 1\} = 1, \end{aligned}$$

and so $E(\gamma_1 - \gamma_0) = 1 - 1 = 0$. $\square$

We proceed with an elementary randomization property.

Lemma 5.10 *(uniform randomization) Fix any closed set $B \subset [0,1]$ with $\{0,1\} \subset B$, let σ and τ be i.i.d. $U(0,1)$, and define*

$$\sigma_- = \sup\,(B \cap [0, \sigma]), \qquad \sigma_+ = \inf\,(B \cap (\sigma, 1]).$$

Then $\sigma' \equiv \sigma_+ - (\sigma_+ - \sigma_-)\,\tau \overset{d}{=} \sigma$.

Proof: Let B^c have connected components $I_k = (a_k, b_k)$, $k \geq 1$. Since a.s. $\sigma_+ = b_k$ iff $\sigma \in [a_k, b_k]$ for each k, we have

$$\begin{aligned} P(\,\sigma \leq t \,|\, \sigma_+ = b_k) &= \frac{t - a_k}{b_k - a_k} \\ &= P(\,\sigma' \leq t \,|\, \sigma_+ = b_k), \quad t \in I_k,\ \ k \geq 1. \end{aligned}$$

On the other hand, we have a.s. $\sigma_+ \notin \{b_1, b_2, \ldots\}$ iff $\sigma \notin \bigcup_k [a_k, b_k]$, which implies $\sigma = \sigma' = \sigma_+$. Hence, by combination,

$$P(\sigma \in \cdot \,|\, \sigma_+) = P(\sigma' \in \cdot \,|\, \sigma_+) \ \text{a.s.} \qquad \square$$

The last result yields an interesting connection between inversion and randomization.

Lemma 5.11 *(inversions and shifts) Let $\mu \in \mathcal{M}$ with $\tilde{\mu} \in \mathcal{M}$, and let $c > 0$ with $\{0, c\} \subset \operatorname{supp} \mu$. Let σ/c and τ be i.i.d. $U(0,1)$, and define $\eta = \theta_{-\sigma}\mu$ and $\delta = \tilde{\eta}\{0\}$. Then $\eta \overset{d}{=} \theta_{\delta\tau}\tilde{\tilde{\eta}}$.*

Proof: Define $\sigma_\pm$ as in Lemma 5.10 with $B = \operatorname{supp} \mu$. Then Lemma 5.7 (iii) yields $\tilde{\tilde{\eta}} = \theta_{-\gamma}\eta$ with $\gamma = \eta_0^{-1} = \sigma_+ - \sigma$, and so

$$\begin{aligned} \tilde{\tilde{\eta}} &= \theta_{-\gamma}\eta = \theta_{-\gamma}\theta_{-\sigma}\mu \\ &= \theta_{-\sigma-\gamma}\,\mu = \theta_{-\sigma_+}\mu. \end{aligned}$$

Furthermore, we get from Lemma 5.10

$$\begin{aligned} \sigma &\overset{d}{=} \sigma_+ - (\sigma_+ - \sigma_-)\,\tau \\ &= \sigma_+ - \delta\,\tau. \end{aligned}$$

Hence, by combination,

$$\begin{aligned} \eta = \theta_{-\sigma}\mu &\overset{d}{=} \theta_{-\sigma_+ + \delta\tau}\,\mu \\ &= \theta_{\delta\tau}\,\theta_{-\sigma_+}\mu = \theta_{\delta\tau}\,\tilde{\tilde{\eta}}\,. \end{aligned} \qquad \square$$

The following scaling property is needed to prove the last two assertions of the main theorem.

Lemma 5.12 *(scaling) Let ξ be a stationary random measure on R, and let $f \geq 0$ be a measurable function on $\mathcal{M}$. Then*

$$E\int_0^{\xi_t} f(\theta_{-r}\,\tilde{\xi})\,dr = t\,E\int_0^{\xi_1} f(\theta_{-r}\,\tilde{\xi})\,dr, \quad t > 0.$$

Proof: Using an elementary substitution, Lemma 5.7 (i), and the stationarity of ξ, we get for any $t, h > 0$

$$\begin{aligned} \int_{\xi_t}^{\xi_{t+h}} f(\theta_{-r}\,\tilde{\xi})\,dr &= \int_0^{\xi_{t+h}-\xi_t} f(\theta_{-r-\xi_t}\,\tilde{\xi})\,dr \\ &= \int_0^{(\theta_{-t}\,\xi)_h} f\Big\{\theta_{-r}(\theta_{-t}\,\xi)^{\tilde{}}\Big\}\,dr \\ &\overset{d}{=} \int_0^{\xi_h} f(\theta_{-r}\,\tilde{\xi})\,dr. \end{aligned}$$

Taking expected values and summing over $t = 0, h, 2h, \dots$, we obtain the asserted formula for rational t. The general relation now follows by the monotonicity on both sides. □

Proof of Theorem 5.6: (i) By Lemma 5.8, we get for measurable $f \geq 0$

$$\begin{aligned} \tilde{P}^0_\xi f &= E\int_0^{\xi_1} \tilde{f}(\theta_{-r}\tilde{\xi})\,dr \\ &= E\int_0^1 f(\theta_{-s}\,\xi)\,\xi(ds) = Q_\xi f. \end{aligned}$$

(ii) Let σ and τ be i.i.d. $U(0,1)$ and independent of ξ. Since $\tilde{\xi} \in \mathcal{M}$ a.s. on $\{\xi_1 > 0\}$ by Lemma 5.2 (i), we may define, on the latter set, a random measure $\eta = \theta_{-\sigma\,\xi_1}\tilde{\xi}$. By an elementary substitution and Fubini's theorem, we get for measurable $f \geq 0$

$$\begin{aligned} P^0_\xi f &= E\int_0^{\xi_1} f(\theta_{-r}\,\tilde{\xi})\,dr \\ &= E\,\xi_1\int_0^1 f(\theta_{-r\xi_1}\tilde{\xi})\,dr \\ &= E\,\xi_1 f(\theta_{-\sigma\xi_1}\tilde{\xi}) = E\,\xi_1 f(\eta). \end{aligned} \tag{10}$$

Now $\{0, \xi_1\} \subset \operatorname{supp}\tilde{\xi}$, by the right-continuity of the function $\xi_t = \xi(0,t]$. Using (10) (twice), Lemma 5.11, Fubini's theorem, and (i), we obtain

$$\begin{aligned} P^0_\xi f &= E\,\xi_1 f(\eta) = E\,\xi_1 E\{f(\eta)\,|\,\xi\} \\ &= E\,\xi_1 E\Big\{ f(\theta_{\tau\tilde{\eta}\{0\}}\,\tilde{\tilde{\eta}})\,\Big|\,\xi\Big\} \\ &= E\,\xi_1 f(\theta_{\tau\tilde{\eta}\{0\}}\,\tilde{\tilde{\eta}}) \\ &= E\,\xi_1\int_0^1 f(\theta_{s\tilde{\eta}\{0\}}\,\tilde{\tilde{\eta}})\,ds \\ &= \int P^0_\xi(d\mu)\int_0^1 f(\theta_{s\tilde{\mu}\{0\}}\tilde{\tilde{\mu}})\,ds \\ &= \int Q_\xi(d\mu)\int_0^1 f(\theta_{s\mu\{0\}}\tilde{\mu})\,ds. \end{aligned}$$

(iii) Let $f \geq 0$ be a bounded, measurable function on $\mathcal{M}$ with $P^0_\xi f < \infty$. Using Lemma 5.12 and an elementary substitution, we get for any $t, h > 0$

$$\begin{aligned} P^0_\xi(f \circ \theta_h) &= E\int_0^{\xi_1} f(\theta_{h-r}\,\tilde{\xi})\,dr \\ &= t^{-1}E\int_0^{\xi_t} f(\theta_{h-r}\,\tilde{\xi})\,dr \\ &= t^{-1}E\int_{-h}^{\xi_t - h} f(\theta_{-r}\,\tilde{\xi})\,dr, \end{aligned}$$

and so

$$\Big| P^0_\xi(f \circ \theta_h) - P^0_\xi f \Big| \leq 2h\,t^{-1}\,\|f\|, \quad t, h > 0.$$

As $t \to \infty$, we get $P^0_\xi(f \circ \theta_h) = P^0_\xi f$ for $h > 0$ and bounded f, and the general relation follows by monotone convergence, since P^0_ξ is σ-finite.

(iv) First let ξ be ergodic with $0 < E\xi_1 < \infty$. Fix any bounded, measurable function $f \ge 0$ on $\mathcal{M}$. Using (i) and the continuous-time ergodic theorem, we get as $t \to \infty$

$$\begin{aligned}
P^0_\xi\Big\{\mu;\, t^{-1}\int_0^t f(\theta_{-r}\tilde\mu)\,dr \to Ef(\xi)\Big\}
&= E\int_0^1 1\Big\{t^{-1}\int_0^t f(\theta_{-r-s}\,\xi)\,dr \to Ef(\xi)\Big\}\,\xi(ds)\\
&= E\Big\{\xi_1;\, t^{-1}\int_0^t f(\theta_{-r}\,\xi)\,dr \to Ef(\xi)\Big\} = E\xi_1.
\end{aligned}$$

Since $\|P^0_\xi\| = E\xi_1$, we conclude that

$$t^{-1}\int_0^t f(\theta_{-r}\tilde\mu)\,dr \to Ef(\xi), \quad \mu \in \mathcal{M} \text{ a.e. } P^0_\xi.$$

Applying the same argument to the function $f(\mu) = h^{-1}\mu_h$, and letting $h \to 0$, we get $\tilde\mu_t/t \to E\xi_1$ a.e. P^0_ξ. Hence, Lemma 5.7 (ii) yields $t/\mu_t \to E\xi_1$ a.e. P^0_ξ, and so by combination

$$t^{-1}\int_0^{\mu_t} f(\theta_{-r}\tilde\mu)\,dr \to \frac{Ef(\xi)}{E\xi_1}, \quad \mu \in \mathcal{M} \text{ a.e. } P^0_\xi. \tag{11}$$

By (iii), the ergodic theorem also applies to P^0_ξ, and so the a.e. convergence $\mu_t/t \to 1/E\xi_1$ remains valid in $L^1(P^0_\xi)$. Hence, for large enough $t > 0$, the ratios μ_t/t are uniformly integrable with respect to P^0_ξ. Applying Lemma 5.12 to P^0_ξ and using (11), along with the mentioned uniform integrability, we get as $t \to \infty$

$$\begin{aligned}
\int P^0_\xi(d\mu)\int_0^{\mu_1} f(\theta_{-r}\tilde\mu)\,dr &= \int P^0_\xi(d\mu)\,t^{-1}\int_0^{\mu_t} f(\theta_{-r}\tilde\mu)\,dr\\
&\to \int P^0_\xi(d\mu)\,\frac{Ef(\xi)}{E\xi_1} = Ef(\xi).
\end{aligned}$$

Since the left-hand side is independent of t, the extreme members agree, which proves (iv) when ξ is ergodic with $0 < E\xi_1 < \infty$. The statement is also trivially true when $E\xi = 0$.

In the non-ergodic case, define

$$P^{0,\mathcal{I}_\xi}_\xi f = E^{\mathcal{I}_\xi}\int_0^{\xi_1} f(\theta_{-r}\,\tilde\xi)\,dr, \quad f \ge 0,$$

where $\mathcal{I}_\xi$ is the ξ-invariant σ-field and $E^{\mathcal{I}_\xi} = E(\,\cdot\,|\mathcal{I}_\xi\,)$. Arguing as before, with E and P^0_ξ replaced by $E^{\mathcal{I}_\xi}$ and $P^{0,\mathcal{I}_\xi}_\xi$, respectively, and using the disintegration theorem, we obtain for measurable $f \ge 0$

$$\int P^{0,\mathcal{I}_\xi}_\xi(d\mu)\int_0^{\mu_1} f(\theta_{-r}\,\tilde\mu)\,dr = E^{\mathcal{I}_\xi} f(\xi) \text{ a.s. on } \{\bar\xi < \infty\}.$$

Taking expected values, and noting that $P^0_\xi = EP^{0,\mathcal{I}_\xi}_\xi$ and $E\xi_1 = E\bar{\xi}$, we obtain (iv) again, under the same condition $E\xi_1 < \infty$.

It remains to eliminate the moment condition $E\xi_1 < \infty$. By elementary conditioning, we may then assume that $\xi \neq 0$ a.s., in which case (iv) can be written more explicitly as

$$E\int_0^{\xi_1} dt \int_0^{(\theta_{-t}\tilde{\xi})_1} f\Big(\theta_{-s}(\theta_{-t}\tilde{\xi})\tilde{}\,\Big)\, ds = Ef(\xi). \tag{12}$$

Now introduce the product-measurable processes

$$X^n_t = 1\Big\{\,\xi[t-1,t+1] \le n\Big\}, \quad t \in \mathsf{R},\ n \in \mathsf{N},$$

and form the stationary random measures $\xi^n = X^n \cdot \xi$ with intensities $E\xi^n_1 = E\xi^n(0,1] \le n$. Further note that a.s.

$$\lim_{n\to\infty} 1\Big\{\,\xi^n = \xi,\ \tilde{\xi}^n = \tilde{\xi} \text{ on } [-c,c]\Big\} = 1, \quad c > 0, \tag{13}$$

where the statement for $\tilde{\xi}$ and $\tilde{\xi}^n$ follows from that for ξ and ξ^n by Lemmas 5.2 and 5.7 (ii).

Now suppose that f is *local*, in the sense that that there exists a constant $c > 0$, such that $f(\mu) = f(\nu)$ whenever $\mu = \nu$ on $[-c,c]$. Writing (12) as $EAf(\xi) = Ef(\xi)$, we see from (13) that a.s. $Af(\xi^n) = Af(\xi)$ and $f(\xi^n) = f(\xi)$, for large enough n. Since (12) holds for each ξ^n, we get by Fatou's lemma and dominated convergence

$$\begin{aligned} EAf(\xi) &= E\lim_{n\to\infty} Af(\xi^n) \\ &\le \liminf_{n\to\infty} EAf(\xi^n) \\ &= \lim_{n\to\infty} Ef(\xi^n) = Ef(\xi). \end{aligned}$$

Letting $0 \le f \le 1$, and applying the same inequality to the local function $1-f$, we obtain

$$\begin{aligned} EA1(\xi) - EAf(\xi) &= EA(1-f)(\xi) \\ &\le E(1-f)(\xi) = 1 - Ef(\xi). \end{aligned}$$

If we can show that $EA1(\xi) \ge 1$, then $EAf(\xi) \ge Ef(\xi)$ holds by subtraction, and the desired equality follows.

To see this, we may use Lemmas 5.7 (ii) and 5.8 (ii), along with the right-continuity of $\tilde{\xi}_t$ and the definition of γ_t, to get

$$\begin{aligned} h^{-1}\int_0^{\xi_1} (\theta_{-r}\tilde{\xi})_h\, dr &= h^{-1}\int_0^h (\theta_{-r}\tilde{\xi})_{\xi_1}\, dr \\ &\to \tilde{\xi}_{\xi_1} - \tilde{\xi}_0 = \xi^{-1}_{\xi_1} - \xi^{-1}_0 \\ &= 1 + \gamma_1 - \gamma_0. \end{aligned}$$

Using (iii), Lemma 5.9, and Fatou's lemma, we get as $h \to 0$

$$\begin{aligned} E\int_0^{\xi_1} (\theta_{-r}\tilde{\xi})_1\, dr &= \int \mu_1\, P_\xi^0(d\mu) = h^{-1}\int \mu_h\, P_\xi^0(d\mu) \\ &= h^{-1}E\int_0^{\xi_1} (\theta_{-r}\tilde{\xi})_h\, dr \\ &\geq E(1+\gamma_1-\gamma_0) = 1, \end{aligned}$$

which means that $EA1(\xi) = 1$. This completes the proof of (iv), for bounded, local, measurable functions $f \geq 0$. The general result follows by a routine extension, based on the monotone-class theorem. □

We conclude this section with a couple of elementary lemmas that will be needed below. For any measure μ on R, we define the *intensity* $\bar{\mu}$ by

$$\bar{\mu} = \lim_{|t|\to\infty} t^{-1}(\theta_s\mu)_t, \quad s \in \mathsf{R},$$

whenever the limit exists and is independent of s. It is easy to check that the mapping $\mu \mapsto \bar{\mu}$ is measurable, along with its domain.

Lemma 5.13 *(intensity of inversion) For any $\mu \in \mathcal{M}$ with $\tilde{\mu} \in \mathcal{M}$, the intensities $\bar{\mu}$ and $\bar{\tilde{\mu}}$ exist simultaneously in $[0,\infty]$ and satisfy $\bar{\tilde{\mu}} = (\bar{\mu})^{-1}$.*

Proof: Since μ and $\tilde{\mu}$ are locally finite, the functions μ_t and $\tilde{\mu}_t$ are both unbounded as $|t| \to \infty$. Furthermore, $\tilde{\mu}_t = \mu_t^{-1} - \mu_0^{-1}$ by Lemma 5.7 (ii), and so

$$(\theta_s\tilde{\mu})_t = \mu_{s+t}^{-1} - \mu_s^{-1}, \quad s, t \in \mathsf{R}.$$

Since $|\mu_t| \to \infty$ as $|t| \to \infty$, we have $|\mu_s^{-1}| < \infty$ for all $s \in \mathsf{R}$, and so it suffices to prove the assertion for the functions μ_t and μ_t^{-1}. The statement is then elementary. □

Lemma 5.14 *(invariance) For any invariant function f on $\mathcal{M}$, the function $\tilde{f}(\mu) = f(\tilde{\mu})$ is invariant on the set $\{\mu \in \mathcal{M};\ \tilde{\mu} \in \mathcal{M}\}$.*

Proof: Using Lemma 5.7 (i) and the invariance of f, we get for any $t \in \mathsf{R}$ and $\mu \in \mathcal{M}$ with $\tilde{\mu} \in \mathcal{M}$

$$\begin{aligned} \tilde{f}(\theta_{-t}\mu) &= f \circ (\theta_{-t}\mu)^{\sim} = f(\theta_{-\mu_t}\tilde{\mu}) \\ &= f(\tilde{\mu}) = \tilde{f}(\mu). \end{aligned}$$

□

5.3 Asymptotic Invariance

The uniformly bounded measures μ_n on R^d, indexed by N or R_+, are said to be *asymptotically invariant* if[3]

$$\|\mu_n - \mu_n * \delta_x\| \to 0, \quad x \in \mathsf{R}^d, \tag{14}$$

where the norm denotes total variation. By a *weight function* on R^d, we mean a measurable function $p \geq 0$ with $\lambda^d p = 1$. Any such functions p_n are said to be *asymptotically invariant*, if the corresponding property holds for the measures[4] $\mu_n = p_n \cdot \lambda^d$. For any linear subspace $u \subset \mathsf{R}^d$, we may also consider the notions of *asymptotic u-invariance*, defined as above, but with x restricted to u. We usually consider only the case where $u = \mathsf{R}^d$, the general case being similar.

Asymptotic invariance of some measures μ_n implies the following seemingly stronger properties:

Lemma 5.15 *(asymptotic invariance) Let the measures μ_n on R^d be uniformly bounded and asymptotically invariant. Then* (14) *holds uniformly for bounded x, and for any probability measure ν on R^d, we have $\|\mu_n - \mu_n * \nu\| \to 0$. Moreover, the singular components μ_n'' of μ_n satisfy $\|\mu_n''\| \to 0$.*

Proof: For any probability measure ν, we get by dominated convergence

$$\begin{aligned}\|\mu_n - \mu_n * \nu\| &= \left\|\int (\mu_n - \mu_n * \delta_x)\, \nu(dx)\right\| \\ &\leq \int \|\mu_n - \mu_n * \delta_x\|\, \nu(dx) \to 0.\end{aligned}$$

If ν is absolutely continuous, then so is $\mu_n * \nu$ for all n, and we get

$$\|\mu_n''\| \leq \|\mu_n - \mu_n * \nu\| \to 0.$$

Letting ν_h denote the uniform distribution on $[0, h]^d$, and noting that $\|\mu * \nu\| \leq \|\mu\|\, \|\nu\|$, we get for any $x \in \mathsf{R}^d$ with $|x| \leq r$

$$\begin{aligned}\|\mu_n - \mu_n * \delta_x\| &\leq \|\mu_n - \mu_n * \nu_h\| + \|\mu_n * \nu_h - \mu_n * \nu_h * \delta_x\| \\ &\quad + \|\mu_n * \nu_h * \delta_x - \mu_n * \delta_x\| \\ &\leq 2\, \|\mu_n - \mu_n * \nu_h\| + \|\nu_h - \nu_h * \delta_x\| \\ &\leq 2\, \|\mu_n - \mu_n * \nu_h\| + 2\, r\, h^{-1} d^{1/2},\end{aligned}$$

which tends to 0 for fixed r, as we let $n \to \infty$ and then $h \to \infty$. This proves the asserted uniform convergence. □

[3] Convolutions are given by $(\mu * \nu) f = \iint \mu(dx)\nu(dy) f(x + y)$, so that $\mu * \delta_r = \theta_r \mu$.

[4] Recall that $f \cdot \mu$ denotes the measure $\nu \ll \mu$ with μ-density f, so that $(f \cdot \mu) g = \mu(fg)$.

The uniformly bounded measures μ_n are further said to be *weakly asymptotically invariant*, if the convolutions $\mu_n * \nu$ are asymptotically invariant for every bounded measure $\nu \ll \lambda^d$ on R^d, so that

$$\|\mu_n * \nu - \mu_n * \nu * \delta_x\| \to 0, \quad x \in \mathsf{R}^d. \tag{15}$$

Some criteria are given below. Write $I_h = [0,h]^d$, and let $e_1, \ldots, e_d$ denote the unit coordinate vectors in R^d.

Lemma 5.16 *(weak asymptotic invariance) For any uniformly bounded measures $\mu_1, \mu_2, \ldots$ on R^d, the weak asymptotic invariance is equivalent to each of the conditions:*

(i) $\|\mu_n * \nu - \mu_n * \nu'\| \to 0$ *for any probability measures* $\nu, \nu' \ll \lambda^d$,

(ii) $\displaystyle\int \Big| \mu_n(I_h + x) - \mu_n(I_h + x + h\,e_i) \Big|\, dx \to 0, \quad h > 0,\ i \le d.$

In (ii), *we may restrict h to the set $\{2^{-k};\, k \in \mathsf{N}\}$, and replace the integration by summation over $(h\mathsf{Z})^d$.*

Proof: (ii) Assuming $\nu = f \cdot \lambda^d$ and $\nu' = f' \cdot \lambda^d$, we note that

$$\|\mu * (\nu - \nu')\| \le \|\mu\|\, \|\nu - \nu'\| = \|\mu\|\, \|f - f'\|_1.$$

By FMP 1.35, it is then enough in (15) to consider measures $\nu = f \cdot \lambda^d$, where f is continuous with bounded support. It is also enough to take $r = he_i$ for some $h > 0$ and $i \in \{1, \ldots, d\}$, and by uniform continuity, we may restrict h to the values 2^{-n}. By uniform continuity, we may next approximate f in L^1 by simple functions f_n over the cubic grids $\mathcal{I}_n$ in R^d of mesh size 2^{-n}, which implies the equivalence with (ii). The last assertion is yet another consequence of the uniform continuity.

(i) The defining condition (15) is clearly a special case of (i). Conversely, to show that (15) implies (i), we may approximate as in (ii), to reduce to the case of finite sums $\nu = \sum_j a_j \lambda_{mj}$ and $\nu' = \sum_j b_j \lambda_{mj}$, where $\sum_j a_j = \sum_j b_j = 1$, and λ_{nj} denotes the uniform distribution over the cube $I_{mj} = 2^{-m}[j-1, j]$ in R^d. Assuming the μ_n to be weakly asymptotically invariant, and writing $h = 2^{-m}$, we get

$$\begin{aligned}
\|\mu_n * \nu - \mu_n * \nu'\| &\le \|\mu_n * \nu - \mu_n * \lambda_{m0}\| + \|\mu_n * \nu' - \mu_n * \lambda_{m0}\| \\
&\le \textstyle\sum_j (a_j + b_j) \Big\| \mu_n * \lambda_{mj} - \mu_n * \lambda_{m0} \Big\| \\
&= \textstyle\sum_j (a_j + b_j) \Big\| \mu_n * \lambda_{m0} * \delta_{jh} - \mu_n * \lambda_{m0} \Big\| \to 0,
\end{aligned}$$

by the criterion in (ii). □

We also say that the locally finite measures μ_n on R^d are *vaguely asymptotically invariant*, if $\sup_n \mu_n B < \infty$ for bounded $B \in \mathcal{B}^d$, and (15) holds on every such set B.

Lemma 5.17 *(vague asymptotic invariance) For any measures $\mu_n \in \mathcal{M}_d$, the vague asymptotic invariance is equivalent to each of the conditions:*

(i) $\sup_n \mu_n f < \infty$ *and* $\mu_n f - \mu_n g \to 0$, *for any continuous functions* $f, g \geq 0$ *on* R^d *with bounded supports and* $\lambda^d f = \lambda^d g$,

(ii) (μ_n) *is vaguely relatively compact, and every limiting measure equals* $c\lambda^d$, *for some* $c \geq 0$.

Proof: Assuming (i), we see from Theorem 4.2 that the sequence (μ_n) is vaguely relatively compact. If $\mu_n \xrightarrow{v} \mu$ along a sub-sequence, we obtain $\mu f = \mu g$ for any functions f and g as stated. In particular, μ is translation invariant, which implies $\mu = c\lambda^d$ for some $c \geq 0$. This shows that (i) implies (ii), and the converse implication is obvious.

Now assume the defining properties of vague asymptotic invariance. Then (μ_n) is vaguely relatively compact, and the local version of (15) shows that any limiting measure μ is translation invariant, which proves (ii). Conversely, we claim that (ii) implies

$$\Big| \mu_n(I_h + x) - \mu_n(I_h + x + h\, e_i) \Big| \to 0, \quad h > 0,\ x \in \mathsf{R},\ i \leq d.$$

Indeed, if the left-hand side exceeds $\varepsilon > 0$ along a sub-sequence $N' \subset \mathsf{N}$, for some h, x, and i, we have convergence $\mu_n \xrightarrow{v} c\lambda^d$ along a further sub-sequence $N'' \subset N'$ for some $c \geq 0$, which yields a contradiction. For any bounded set $B \in \mathcal{B}^d$, we conclude that the corresponding integral over B tends to 0 by dominated convergence, which shows that the μ_n are vaguely asymptotically invariant. $\square$

A measure ν on R^d is said to be *non-lattice*, if no translate of ν is supported by a proper closed subgroup of R^d.

Theorem 5.18 *(convolution powers, Maruyama) For any distribution μ on R^d, the convolution powers μ^{*n} are*

(i) *weakly asymptotically invariant iff μ is non-lattice,*

(ii) *asymptotically invariant iff μ^{*k} is non-singular for some $k \in \mathsf{N}$.*

Our proof depends on the following local version of the central limit theorem. For every $h > 0$, we introduce on R^d the probability density

$$p_h(x) = (\pi h)^{-d} \prod_{i \leq d} (1 - \cos hx_i)/x_i^2, \quad x = (x_1, \ldots, x_d) \in \mathsf{R}^d,$$

which has characteristic function

$$\hat{p}_h(t) = \prod_{i \leq d} (1 - |t_i|/h)_+, \quad t = (t_1, \ldots, t_d) \in \mathsf{R}^d.$$

Lemma 5.19 *(local central limit theorem) Let μ be a distribution on R^d with means* 0 *and covariances δ_{ij}, and let φ denote the standard normal density on R^d. Then*

$$\lim_{n\to\infty} n^{d/2}\,\|\mu^{*n} * p_h - \varphi^{*n}\| = 0, \quad h > 0.$$

Proof: Since $|\partial_i \varphi^{*n}| \lesssim n^{-1-d/2}$ for all $i \le d$, the assertion is trivially fulfilled for the standard normal distribution, and so it suffices to prove that, for any measures μ and ν as stated,

$$\lim_{n\to\infty} n^{d/2} \left\|(\mu^{*n} - \nu^{*n}) * p_h\right\| = 0, \quad h > 0.$$

By Fourier inversion,

$$(\mu^{*n} - \nu^{*n}) * p_h(x) = (2\pi)^{-d}\int e^{-ixt}\,\hat{p}_h(t)\left\{\hat{\mu}^n(t) - \hat{\nu}^n(t)\right\} dt,$$

and writing $I_h = [-h,h]^d$, we get

$$\left\|(\mu^{*n} - \nu^{*n}) * p_h\right\| \le (2\pi)^{-d}\int_{I_h}\left|\hat{\mu}^n(t) - \hat{\nu}^n(t)\right| dt.$$

It is then enough to show that the integral on the right declines faster than $n^{-d/2}$. For notational convenience we assume that $d = 1$, the general case being similar.

A standard Taylor expansion (FMP 5.10) yields

$$\begin{aligned}\hat{\mu}_n(t) &= \left\{1 - \tfrac{1}{2}t^2 + o(t^2)\right\}^n \\ &= \exp\left\{-\tfrac{1}{2}nt^2(1 + o(1))\right\} \\ &= e^{-nt^2/2}\{1 + nt^2 o(1)\},\end{aligned}$$

and similarly for $\hat{\nu}^n(t)$, and so

$$\left|\,\hat{\mu}^n(t) - \hat{\nu}^n(t)\right| = e^{-nt^2/2} n t^2 o(1).$$

Here clearly

$$n^{1/2}\int e^{-nt^2/2} n t^2\, dt = \int t^2 e^{-t^2/2} < \infty,$$

and since μ and ν are non-lattice, we have $|\hat{\mu}|\vee|\hat{\nu}| < 1$, uniformly on compacts in $\mathsf{R}\setminus\{0\}$. Splitting the interval I_h into I_ε and $I_h \setminus I_\varepsilon$, we get for any $\varepsilon > 0$

$$n^{1/2}\int_{I_h}\left|\,\hat{\mu}^n(t) - \hat{\nu}^n(t)\right| dt \lesssim r_\varepsilon + h\, n^{1/2}\, m_\varepsilon^n,$$

where $r_\varepsilon \to 0$ as $\varepsilon \to 0$, and $m_\varepsilon < 1$ for all $\varepsilon > 0$. The desired convergence now follows, as we let $n \to \infty$ and then $\varepsilon \to 0$. □

Proof of Theorem 5.18: (i) We may write $\mu = \frac{1}{2}(\mu_1 + \mu_2)$, where μ_1 is non-lattice with bounded support. Then

$$\begin{aligned}\mu^{*n} &= 2^{-n} \sum_{k=0}^{n} \binom{n}{k} \mu_1^{*k} * \mu_2^{*(n-k)} \\ &= \mu_1^{*m} * \alpha_{n,m} + \beta_{n,m},\end{aligned}$$

where $1 - \|\alpha_{n,m}\| = \|\beta_{n,m}\| \to 0$, as $n \to \infty$ for fixed m. Since $\|\mu * \nu\| \leq \|\mu\|\,\|\nu\|$ for any signed measures μ and ν, any weak asymptotic invariance of μ_1^{*n} would extend to μ^{*n}. This reduces the proof to the case of measures μ with bounded support. By a linear transformation, along with a shift, we may further assume that μ has means 0 and covariances δ_{ij}.

By Lemma 5.19 along with the estimate $|\partial_i \varphi^{*n}| \lesssim n^{-1-d/2}$, we have

$$n^{d/2} \left\| \mu^{*n} * \delta_x * p_h - \varphi^{*n} \right\| \to 0, \quad x \in \mathsf{R}^d, \;\; h > 0,$$

and so, for any finite convex combination p of functions $\delta_x * p_h$, we get

$$n^{d/2} \left\| (\mu^{*n} * \delta_x - \mu^{*n}) * p \right\| \to 0, \quad x \in \mathsf{R}^d.$$

Writing $\rho = p \cdot \lambda^d$, we have for any $r > 0$, $x \in \mathsf{R}^d$, and $n \in \mathsf{N}$

$$\begin{aligned}\left\| (\mu^{*n} * \delta_x - \mu^{*n}) * \rho \right\| \;\lesssim\;& r^d n^{d/2} \left\| (\mu^{*n} * \delta_x - \mu^{*n}) * p \right\| \\ &+ \left\{ \mu^{*n} * (\rho * \delta_x + \rho) \right\} I^c_{r\sqrt{n}}\,.\end{aligned}$$

By the central limit theorem, we get with $\nu = \varphi \cdot \lambda^d$

$$\limsup_{n\to\infty} \left\| (\mu^{*n} * \delta_x - \mu^{*n}) * \rho \right\| \lesssim 2\,\nu I^c_r,$$

and $r > 0$ being arbitrary, we conclude that $\|(\mu^{*n} * \delta_x - \mu^{*n}) * \rho\| \to 0$.

Now fix any $h > 0$, and let γ_h denote the uniform distribution on $[0, h]^d$. By an elementary approximation, we may choose measures ρ_m as above with $\|\rho_m - \gamma_h\| \to 0$, and we get

$$\left\| (\mu^{*n} * \delta_x - \mu^{*n}) * \gamma_h \right\| \leq \left\| (\mu^{*n} * \delta_x - \mu^{*n}) * \rho_m \right\| + 2\,\|\rho_m - \gamma_h\|,$$

which tends to 0 as $n \to \infty$ and then $m \to \infty$. The weak asymptotic invariance of μ^{*n} now follows by Lemma 5.16.

(ii) Write $\mu^{*n} = \mu'_n + \mu''_n$, where $\mu'_n \ll \lambda^d \perp \mu''_n$, and note that $\|\mu'_n\|$ is non-decreasing since $\mu'_n * \mu \ll \lambda^d$. If $\|\mu'_k\| > 0$ for some k, then $\|\mu''_{nk}\| \leq \|\mu''_k\|^n \to 0$, and so $\|\mu''_n\| \to 0$. In particular, the μ^{*n} are non-lattice. For any $x \in \mathsf{R}^d$, we have

$$\begin{aligned}\left\| \mu^{*n} - \delta_x * \mu^{*n} \right\| &= \left\| \mu^{*k} * \left(\mu^{*(n-k)} - \delta_x * \mu^{*(n-k)} \right) \right\| \\ &\leq \left\| \mu'_k * \left(\mu^{*(n-k)} - \delta_x * \mu^{*(n-k)} \right) \right\| + 2\,\|\mu''_k\|,\end{aligned}$$

which tends to 0 by part (i), as we let $n \to \infty$ and then $k \to \infty$. This proves the sufficiency of the stated condition, while the necessity holds by Lemma 5.15. □

We turn to a stochastic version of the asymptotic invariance. A sequence of random measures $\xi_1, \xi_2, \ldots$ on R^d is said to be *asymptotically invariant in probability*, if it is vaguely tight and satisfies

$$\xi_n f - (\delta_r * \xi_n) f \xrightarrow{P} 0, \quad r \in \mathsf{R}^d, \tag{16}$$

for any continuous functions with bounded supports $f \geq 0$. For any $p \geq 1$, we further say that the ξ_n are *asymptotically L^p-invariant*, if for any $r \in \mathsf{R}^d$ and functions $f \geq 0$ as above,

$$\sup_n \| \xi_n f \|_p < \infty, \qquad \lim_{n \to \infty} \left\| \xi_t f - (\delta_r * \xi_t) f \right\|_p = 0. \tag{17}$$

Here is the probabilistic version of Lemma 5.17:

Lemma 5.20 *(asymptotic invariance in probability) For any vaguely tight sequence of random measures $\xi_1, \xi_2, \ldots$ on R^d, their asymptotic invariance in probability is equivalent to each of the conditions:*

(i) *$\xi_n f - \xi_n g \xrightarrow{P} 0$, for any continuous functions $f, g \geq 0$ with bounded supports satisfying $\lambda^d f = \lambda^d g$,*

(ii) *if $\xi_n \xrightarrow{vd} \xi$ along a sub-sequence, then $\xi = \alpha \lambda^d$ a.s. for some random variable $\alpha \geq 0$.*

Proof: Consider any vaguely tight sequence of random measures ξ_n satisfying (i), and suppose that $\xi_n \xrightarrow{vd} \xi$ along a sub-sequence. Then for any functions $f, g \geq 0$ as in (i) and constants $a, b \geq 0$,

$$\begin{aligned} a\,\xi_n f + b\,\xi_n g &= \xi_n(af + bg) \\ &\xrightarrow{d} \xi(af + bg) \\ &= a\,\xi f + b\,\xi g, \end{aligned}$$

and so $(\xi_n f, \xi_n g) \xrightarrow{d} (\xi f, \xi g)$ by the Cramér–Wold theorem, which implies $\xi_n f - \xi_n g \xrightarrow{d} \xi f - \xi g$. By (i) it follows that $\xi f = \xi g$ a.s. Applying this to a measure-determining class of functions f_k and a dense set of shift parameters $r \in \mathsf{R}^d$, we conclude that ξ is a.s. translation invariant, which implies $\xi = \alpha \lambda^d$ a.s. for some random variable $\alpha \geq 0$, proving (ii).

Conversely, assume (ii), and suppose that (i) fails for some functions f and g as stated, so that

$$E(|\xi_n f - \xi_n g| \wedge 1) \geq \varepsilon > 0, \tag{18}$$

along a sub-sequence. By the tightness and condition (ii), we have $\xi_n \xrightarrow{vd} \alpha\lambda^d$ along a further sub-sequence, for some random variable $\alpha \geq 0$. As before, we get

$$\xi_n f - \xi_n g \xrightarrow{d} \alpha\lambda^d f - \alpha\lambda^d g = 0,$$

contradicting (18). This shows that indeed (ii) implies (i).

Since (16) is a special case of (i), it remains to show that it also implies (ii). This may be seen as in the proof of the implication (i) ⇒ (ii). □

We add some further criteria, which will be useful in Chapter 11.

Lemma 5.21 *(random shift) Consider some random measures ξ_t on R^d, $t > 0$, along with an independent, bounded random element ϑ in R^d. Then for any $p \geq 1$, the random measures ξ_t and $\xi_t \circ \theta_\vartheta^{-1}$ are simultaneously asymptotically L^p-invariant.*

Proof: We need to prove that (17) is equivalent to the relations

$$\begin{aligned} \limsup_{t\to\infty} \left\| \xi_t(f \circ \theta_\vartheta) \right\|_p &< \infty, \\ \lim_{t\to\infty} \left\| \xi_t(f \circ \theta_\vartheta) - \xi_t(f \circ \theta_{r+\vartheta}) \right\|_p &= 0, \end{aligned} \tag{19}$$

for arbitrary f and r. Since the first conditions in (17) and (19) are clearly equivalent, it is enough to consider the second ones. Then note that (17) implies (19), by Fubini's theorem and dominated convergence. To prove the converse, we may assume that 0 lies in the support of $\nu = \mathcal{L}(\vartheta)$. Fixing a sequence T of numbers $t \to \infty$, and using (19), Fubini's theorem, and the sub-sequence criterion for convergence in probability, we get

$$\left\| \xi_t(f \circ \theta_s) - \xi_t(f \circ \theta_{r+s}) \right\|_p \to 0, \quad s \in \mathsf{R}^d \text{ a.e. } \nu, \tag{20}$$

as $t \to \infty$ along a sub-sequence $T' \subset T$. Let S denote the set of non-exceptional s-values, and write

$$\begin{aligned} \left\| \xi_t f - \xi_t(f \circ \theta_r) \right\|_p \leq & \left\| \xi_t f - \xi_t(f \circ \theta_s) \right\|_p \\ & + \left\| \xi_t(f \circ \theta_s) - \xi_t(f \circ \theta_{r+s}) \right\|_p \\ & + \left\| \xi_t(f \circ \theta_{r+s}) - \xi_t(f \circ \theta_r) \right\|_p . \end{aligned}$$

Using (20), the first condition in (17), and the uniform continuity of f, we see that the right-hand side tends to 0, as $t \to \infty$ along T', and then $s \to 0$ along S. Since T was arbitrary, the convergence extends to R_+, which completes the proof of (17). □

We also note the following simple equivalence:

Lemma 5.22 *(convolutions) Let $\nu_1, \nu_2, \ldots$ be distributions on R^d with uniformly bounded supports satisfying $\nu_n \xrightarrow{w} \delta_0$, and fix any $p \geq 1$. Then the random measures ξ_t are asymptotically L^p-invariant, iff the convolutions $\nu_n * \xi_t$ have this property for every n.*

Proof: Noting that

$$\left\{\delta_r * (\nu_n * \xi_t)\right\} f = (\delta_r * \xi_t)(\nu_n * f), \quad n \in \mathsf{N},$$

where the functions $\nu_n * f$ are again continuous with bounded support, we see that the asymptotic L^p-invariance of the random measures ξ_t implies the same property for the measures $\nu_n * \xi_t$, for every fixed n.

Conversely, suppose that $\nu_n * \xi_t$ is asymptotically L^p-invariant for every n, and write

$$\begin{aligned}\left\| \xi_t f - (\delta_r * \xi_t) f \right\|_p \leq & \left\| \xi_t |f - \nu_n * f| \right\|_p \\ & + \left\| (\nu_n * \xi_t) f - (\delta_r * \nu_n * \xi_t) f \right\|_p \\ & + \left\| (\delta_r * \xi_t) |f - \nu_n * f| \right\|_p .\end{aligned}$$

Noting that the functions $\nu_n * f$ have uniformly bounded supports, and arguing by dominated convergence in two steps, we see that the right-hand side tends to 0, as $t \to \infty$ and then $n \to \infty$. This shows that the ξ_t are again asymptotically L^p-invariant. □

5.4 Averaging and Smoothing Limits

Here we begin with some basic ergodic theorems for stationary random measures ξ on R^d. Let $\mathcal{I}_\xi$ denote the invariant σ-field induced by ξ, and introduce the *sample intensity* $\bar{\xi} = E(\xi I_1 \,|\, \mathcal{I}_\xi)$, where $I_1 = [0,1]^d$. Note that $\bar{\xi}$ may take any value in $[0, \infty]$. Given a bounded, convex set $B \subset \mathsf{R}^d$, we define the *inner radius* r_B as the maximum radius of all open balls contained in B.

Theorem 5.23 *(sample intensity, Nguyen & Zessin) For any stationary random measure ξ on R^d and bounded, increasing, convex sets $B_n \in \mathcal{B}^d$ with inner radii $r_{B_n} \to \infty$, we have $\xi B_n / \lambda^d B_n \to \bar{\xi}$ a.s. This remains true in L^p with $p \geq 1$, whenever $\xi I_1 \in L^p$.*

Proof: By Fubini's theorem, we have for any $A, B \in \mathcal{B}^d$

$$\begin{aligned}\int_B (\theta_{-s}\xi) A \, ds &= \int_B ds \int 1_A(t-s)\, \xi(dt) \\ &= \int \xi(dt) \int_B 1_A(t-s)\, ds \\ &= \xi(1_A * 1_{\bar{B}}).\end{aligned}$$

Assuming $\lambda^d A = 1$ and $A \subset S_a = \{s;\, |s| < a\}$, and putting $B^+ = B + S_a$ and $B^- = (B^c + S_a)^c$, we note that also $1_A * 1_{B^-} \le 1_B \le 1_A * 1_{B^+}$. Applying this to the sets $B = B_n$ gives

$$\frac{\lambda^d B_n^-}{\lambda^d B_n} \cdot \frac{\xi(1_A * 1_{B_n^-})}{\lambda^d B_n^-} \le \frac{\xi B_n}{\lambda^d B_n} \le \frac{\lambda^d B_n^+}{\lambda^d B_n} \cdot \frac{\xi(1_A * 1_{B_n^+})}{\lambda^d B_n^+}.$$

Since $r_{B_n} \to \infty$, we have $\lambda^d B_n^\pm \sim \lambda^d B_n$ by Lemma A4.3. Applying Theorem A2.4 to the function $f(\mu) = \mu A$ and convex sets $B_n^\pm$, we further obtain

$$\frac{\xi(1_A * 1_{B_n^\pm})}{\lambda^d B_n^\pm} \to E(\,\xi A \,|\, \mathcal{I}_\xi\,) = \bar\xi,$$

in the appropriate sense. □

The L^p-version of the last theorem remains valid under much weaker conditions.

Theorem 5.24 *(averaging)* *For any stationary random measure ξ and asymptotically invariant weight functions $f_1, f_2, \ldots$ on R^d, we have*

(i) $\xi f_n \xrightarrow{P} \bar\xi$,

(ii) $\xi f_n \to \bar\xi$ *in L^1 iff $E\xi I_1 < \infty$,*

(iii) $\xi f_n \to \bar\xi$ *in L^p with $p > 1$, iff the variables $(\xi f_n)^p$ are uniformly integrable.*

In particular, (iii) *holds when $\xi I_1 \in L^p$ and $f_n \lesssim 1_{B_n}/\lambda^d B_n$, for some sets B_n as in Theorem 5.23.*

Proof: (ii) By Theorem 5.23 we may choose some weight functions g_m on R^d, such that $\xi g_m \to \bar\xi$ in L^1. Using Minkowski's inequality, the stationarity of ξ, the invariance of $\bar\xi$, and dominated convergence, we get as $n \to \infty$ and then $m \to \infty$

$$\begin{aligned}\|\xi f_n - \bar\xi\|_1 &\le \Big\|\,\xi f_n - \xi(f_n * g_m)\Big\|_1 + \Big\|\,\xi(f_n * g_m) - \bar\xi\,\Big\|_1 \\ &\le E\xi\Big|f_n - f_n * g_m\Big| + \int \Big\|\,\xi(g_m \circ \theta_s) - \bar\xi\,\Big\|_1 f_n(s)\,ds \\ &\le E\bar\xi \int \lambda^d\Big|f_n - f_n \circ \theta_t\Big|\, g_m(t)\,dt + \|\,\xi g_m - \bar\xi\,\|_1 \to 0.\end{aligned}$$

(i) Applying (ii) to the point processes $\xi^r = 1\{\bar\xi \le r\}\,\xi$ with $r > 0$, which are clearly stationary with

$$E\,\xi^r I_1 = E\,\bar\xi^r = E(\,\bar\xi;\, \bar\xi \le r) \le r, \quad r > 0,$$

we get $\xi^r f_n \xrightarrow{P} \bar\xi^r$, which implies $\xi f_n \xrightarrow{P} \bar\xi$ on $\{\bar\xi < \infty\}$. Next, we introduce the product-measurable processes

$$X_k(s) = 1\{\xi B_s^1 \le k\}, \quad s \in \mathsf{R}^d,\ k \in \mathsf{N},$$

and note that the random measures $\xi_k = X_k \cdot \xi$ are again stationary with $\bar{\xi} \precsim k$ a.s. Then, as before,

$$\xi f_n \geq \xi_k f_n \xrightarrow{P} \bar{\xi}_k, \quad k \in \mathsf{N}. \tag{21}$$

Since $X_k \uparrow 1$, we have $\xi_k \uparrow \xi$ and even $\bar{\xi}_k \uparrow \bar{\xi}$ a.s., by successive monotone convergence. For any sub-sequence $N' \subset \mathsf{N}$, the convergence in (21) holds a.s. along a further sub-sequence N'', and so $\liminf_n \xi f_n \geq \bar{\xi}$ a.s. along N''. In particular, $\xi f_n \xrightarrow{P} \bar{\xi}$ remains true on $\{\bar{\xi} = \infty\}$.

(iii) This follows from (i) by FMP 4.12. □

Theorem 5.24 yields the following limit theorem for convolutions $\xi * \nu_n$. Here the convergence $\eta_n \xrightarrow{vP} \eta$ means that, for every sub-sequence $N' \subset \mathsf{N}$, we have $\xi_n \xrightarrow{v} \xi$ a.s. along a further sub-sequence $N'' \subset N'$. By FMP 4.2 it is equivalent that $E\rho(\xi, \xi_n) \to 0$, for any bounded metrization ρ of $\mathcal{M}_S$.

Theorem 5.25 *(smoothing) For any stationary random measure ξ with $\bar{\xi} < \infty$ a.s. and distributions ν_n on R^d, we have*

(i) $\xi * \nu_n \xrightarrow{vP} \bar{\xi}\lambda^d$ *when the ν_n are weakly asymptotically invariant,*

(ii) $\xi * \nu_n \xrightarrow{ulP} \bar{\xi}\lambda^d$ *when the ν_n are asymptotically invariant.*

Proof: (i) By a simple approximation, we may choose a measure-determining sequence of weight functions $f_k \in \hat{\mathcal{C}}_+$ of the form $\mu_h * g_k$, where μ_h denotes the uniform distribution on I_h. By a compactness argument based on Theorem 4.2, we may assume that the f_k are even convergence determining for the limit λ^d. Since the functions $\nu_n * \mu_h * g_k$ are asymptotically invariant in n for fixed h and k, we get by Theorem 5.24 (i)

$$\begin{aligned}(\xi * \nu_n) f_k &= (\xi * \nu_n * \mu_h) g_k \\ &= \xi(\nu_n * \mu_h * g_k) \\ &\xrightarrow{P} \bar{\xi}\lambda^d g_k = \bar{\xi},\end{aligned}$$

which implies $\xi * \nu_n \xrightarrow{vP} \bar{\xi}\lambda^d$, by a sub-sequence argument.

(ii) Let the ν_n be asymptotically invariant. Then they are also weakly asymptotically invariant, and so by (i) we have $\xi * \nu_n \xrightarrow{vP} \bar{\xi}\lambda^d$. To see that even $\xi * \nu_n \xrightarrow{ulP} \bar{\xi}\lambda^d$, it suffices to prove, for any sub-sequence $N' \subset \mathsf{N}$, that $\xi * \nu_n \xrightarrow{ul} \bar{\xi}\lambda^d$ a.s. along a further sub-sequence $N'' \subset N'$. It is then enough to show that $\xi * \nu_n \xrightarrow{v} \bar{\xi}\lambda^d$ a.s. implies $\xi * \nu_n \xrightarrow{ul} \bar{\xi}\lambda^d$ a.s. along a sub-sequence. Replacing ξ by $\xi/\bar{\xi}$, we may further assume that $\bar{\xi} = 1$.

Now assume that $\xi * \nu_n \xrightarrow{v} \lambda^d$ a.s. Since the ν_n are asymptotically invariant, Lemma 5.15 yields $\|\nu_n - \nu_n * \rho\| \to 0$, for every probability measure ρ on

R^d. Choosing $\rho = f \cdot \lambda^d$ for some continuous function $f \geq 0$ with bounded support C, and writing $\gamma_n = |\nu_n - \nu_n * \rho|$, we get for any $B \in \mathcal{B}^d$

$$\begin{aligned} E\Big\| \xi * \nu_n - \xi * \nu_n * \rho \Big\|_B &= E(\xi * \gamma_n)B = \|\gamma_n\| \, \lambda^d B \\ &= \|\nu_n - \nu_n * \rho\| \, \lambda^d B \to 0, \end{aligned}$$

and so $\|\xi * \nu_n - \xi * \nu_n * \rho\|_B \to 0$ a.s., as $n \to \infty$ along a sub-sequence.

By the assumptions on f, the a.s. convergence $\xi * \nu_n \xrightarrow{v} \lambda^d$ yields a.s.

$$(\xi * \nu_n * f)_x \to (\lambda^d * f)_x = 1, \quad x \in \mathsf{R}^d.$$

Furthermore, we have a.s. for any compact set $B \subset \mathsf{R}^d$

$$\begin{aligned} \limsup_{n\to\infty} \sup_{x\in B} (\xi * \nu_n * f)_x &\leq \|f\| \limsup_{n\to\infty} (\xi * \nu_n)(B - C) \\ &\leq \|f\| \, \lambda^d(B - C) < \infty. \end{aligned}$$

Hence, by dominated convergence,

$$\Big\| \xi * \nu_n * \rho - \lambda^d \Big\|_B = \int_B \Big| (\xi * \nu_n * f)_x - 1 \Big| \, dx \to 0.$$

Combining the two estimates, we get a.s. for any $B \in \mathcal{B}^d$

$$\Big\| \xi * \nu_n - \lambda^d \Big\|_B \leq \Big\| \xi * \nu_n - \xi * \nu_n * \rho \Big\|_B + \Big\| \xi * \nu_n * \rho - \lambda^d \Big\|_B \to 0,$$

which means that $\xi * \nu_n \xrightarrow{ul} \lambda^d$ a.s. □

The associated invariance criterion holds without restrictions on $\bar{\xi}$.

Theorem 5.26 *(smoothing invariance) Let ξ be a stationary random measure on R^d, and fix any non-lattice distribution ν on R^d. Then $\xi * \nu \stackrel{d}{=} \xi$, iff $\xi = \bar{\xi}\lambda^d$ with $\bar{\xi} < \infty$ a.s.*

Proof: If a.s. $\xi = \bar{\xi}\lambda^d$ with $\bar{\xi} < \infty$, then the invariance of λ^d yields

$$\xi * \nu = \bar{\xi}(\lambda^d * \nu) = \bar{\xi}\lambda^d = \xi \text{ a.s.},$$

which implies $\xi * \nu \stackrel{d}{=} \xi$. Conversely, assuming $\xi * \nu \stackrel{d}{=} \xi$ and $\bar{\xi} < \infty$ a.s., we get by Theorems 5.18 (i) and 5.25 (i)

$$\xi \stackrel{d}{=} \xi * \nu^{*n} \xrightarrow{vP} \bar{\xi}\lambda^d,$$

which implies $\xi \stackrel{d}{=} \bar{\xi}\lambda^d$. Thus, ξ is a.s. invariant, and so $\xi = \bar{\xi}\lambda^d$ a.s.

For general $\bar{\xi}$, we may introduce the stationary and measurable process $X_s \equiv \xi B_s^1$ on R^d, and define $\xi_k = 1\{X \leq k\} \cdot \xi$ for all $k \in \mathsf{N}$. Then the ξ_k are again stationary with $\bar{\xi}_k \leq k$ a.s., and so as before

$$\xi \stackrel{d}{=} \xi * \nu^{*n} \geq \xi_k * \nu^{*n} \xrightarrow{vP} \bar{\xi}_k\lambda^d,$$

as $n \to \infty$ for fixed k. Hence, for any $f \in \hat{C}_+ \setminus \{0\}$ and $r > 0$,

$$\begin{aligned} P\{\xi f > r\} &\geq \liminf_{n\to\infty} P\Big\{(\xi_k * \nu^{*n})f > r\Big\} \\ &\geq P\{\bar{\xi}_k \lambda^d f > r\}, \end{aligned}$$

and so, as $k \to \infty$ and then $r \to \infty$,

$$\begin{aligned} 0 &= P\{\xi f = \infty\} \\ &\geq P\{\bar{\xi}\lambda^d f = \infty\} = P\{\bar{\xi} = \infty\}, \end{aligned}$$

which shows that $\bar{\xi} < \infty$ a.s. □

Given a point process ξ and a distribution ν on R^d, we may form a ν-*transform* of ξ by independently shifting the points of ξ, according to the common distribution ν.

Corollary 5.27 *(Cox convergence, Dobrushin) Given some distributions ν_n on R^d, let the ξ_n be ν_n-transforms of a stationary point process ξ on R^d with $\bar{\xi} < \infty$ a.s., and let η be a Cox process directed by $\bar{\xi}\lambda^d$. Then*

(i) $\xi_n \xrightarrow{vd} \eta$ *when the ν_n are weakly asymptotically invariant,*

(ii) $\xi_n \xrightarrow{uld} \eta$ *when the ν_n are asymptotically invariant.*

Proof: (i) Combine Theorems 4.40 and 5.25 (i).

(ii) Combine Theorems 4.36 (iii) and 5.25 (ii). □

We also note the following related invariance criterion.

Corollary 5.28 *(shift invariance) Let ξ be a stationary point process on R^d with ν-transform ξ_ν, where ν is a non-lattice distribution on R^d. Then $\xi_\nu \overset{d}{=} \xi$ iff ξ is mixed Poisson.*

Proof: Since ν is non-lattice, the convolution powers ν^{*n} are dissipative by Theorem 5.18 (i), and so by Corollary 4.37 we have $\xi_\nu \overset{d}{=} \xi$, iff ξ is Cox with a directing random measure η satisfying $\eta * \nu \overset{d}{=} \eta$. Since η is again stationary by Theorem 3.3 (i), Theorem 5.26 shows that $\eta * \nu \overset{d}{=} \eta$, iff a.s. $\eta = \bar{\eta}\lambda^d$ with $\bar{\eta} < \infty$, which means that ξ is mixed Poisson. □

Next we examine the possible preservation of independence, for a point process subject to i.i.d. random shifts. Some interesting implications for line processes and associated particle systems will be discussed in Chapter 11. Define

$$\varphi(s,r) = (s+r,r), \quad e_r(s) = e^{rs}, \qquad r,s \in \mathsf{R}^d.$$

Theorem 5.29 *(independence preserved) Let ξ_μ be a μ-randomization of a point process ξ on R^d, for some non-lattice distribution μ on R^d. Then $\xi_\mu \circ \varphi^{-1}$ is a ν-randomization of a point process on R^d, for some distribution ν on R^d, iff ξ is Cox and directed by $\rho\, e_r \cdot \lambda^d$, for some random variable $\rho \geq 0$ and constant $r \in \mathsf{R}^d$. In that case, $\nu = (e_{-r} \cdot \mu)/\mu e_{-r}$ with $\mu e_{-r} < \infty$.*

Proof: Assume the stated randomization property. Letting $\hat{\xi}_\mu$ denote a μ-transform of ξ, and writing $\tilde{\nu}$ for the reflection of ν, we note that ξ is distributed as a $\tilde{\nu}$-transform of $\hat{\xi}_\mu$, and so $\hat{\xi}_{\mu * \tilde{\nu}} \overset{d}{=} \xi$ by Theorem 3.2 (i). Since $\mu * \tilde{\nu}$ is again non-lattice, its convolution powers $(\mu * \tilde{\nu})^{*n}$ are dissipative by Theorem 5.18, and so ξ is Cox by Corollary 4.37, say with directing random measure η. We may exclude the trivial case where $\eta = 0$ a.s.

By hypothesis and Theorems 3.2 (i) and 3.3 (i), we have

$$(\eta \otimes \mu) \circ \varphi^{-1} \overset{d}{=} (\eta * \mu) \otimes \nu. \tag{22}$$

Here the transfer theorem gives $(\eta \otimes \mu) \circ \varphi^{-1} = \tilde{\eta} \otimes \nu$ a.s., for some random measure $\tilde{\eta}$ on R^d, and so a.s.

$$\eta * \mu = \left\{ (\eta \otimes \mu) \circ \varphi^{-1} \right\} (\cdot \times \mathsf{R}^d) = \tilde{\eta},$$

which implies $\tilde{\eta} = \eta * \mu$ a.s. Thus, (22) can be strengthened to

$$(\eta \otimes \mu) \circ \varphi^{-1} = (\eta * \mu) \otimes \nu \;\text{ a.s.} \tag{23}$$

(This would follow directly from Theorem 3.2, if we could only verify the required conditional independence.)

Writing $\hat{\mu}_B = 1_B \mu / \mu B$, for any $B \in \mathcal{B}^d$ with $\mu B > 0$, we get from (23)

$$\eta * \hat{\mu}_B = (\eta * \mu) \frac{\nu B}{\mu B} \;\text{ a.s.}, \quad B \in \mathcal{B}^d \text{ with } \mu B > 0.$$

For any $s \in \operatorname{supp} \mu$, we may choose some sets $B_n \in \mathcal{B}^d$ with $\mu B_n > 0$ and $B_n \downarrow \{s\}$, and conclude by compactness that

$$\eta * \delta_s = a_s (\eta * \mu) \;\text{ a.s.}, \quad s \in \operatorname{supp} \mu, \tag{24}$$

for some constants $a_s \in [0, \infty]$. Since $P\{\eta \neq 0\} > 0$, we have in fact $a_s \in (0, \infty)$ for all s. Comparing (24) for any $s, t \in \operatorname{supp} \mu$, we obtain $\eta * \delta_{s-t} = (a_s / a_t) \eta$ a.s., and since μ is non-lattice, we get by iteration and continuity

$$\eta * \delta_s = c_s\, \eta \;\text{ a.s.}, \quad s \in \mathsf{R}^d,$$

for some constants $c_s > 0$. A further iteration yields the Cauchy equation $c_{s+t} = c_s\, c_t$, and since c is clearly continuous, we conclude that $c_s = e^{-rs}$ for some $r \in \mathsf{R}^d$. Thus, a.s.

$$\eta * \delta_s = e^{-rs} \eta, \quad s \in \mathsf{R}^d, \tag{25}$$

where the exceptional P-null set can be chosen, by the continuity of both sides, to be independent of s.

Now define $\zeta = e_{-r} \cdot \eta$. Noting that, for measurable functions $f \geq 0$,

$$\begin{aligned}\Big\{(e_{-r} \cdot \eta) * \delta_s\Big\} f &= \int f(x+s)\, e^{-rs}\, \eta(dx) \\ &= e^{rs} \int f(x)\, e^{-rs}\, (\eta * \delta_s)(dx) \\ &= e^{rs} \Big\{e_{-r} \cdot (\eta * \delta_s)\Big\} f,\end{aligned}$$

and using (25), we get a.s.

$$\begin{aligned}\zeta * \delta_s &= e^{rs} e_{-r} \cdot (\eta * \delta_s) \\ &= e_{-r} \cdot \eta = \zeta, \qquad s \in \mathsf{R}^d,\end{aligned}$$

which shows that ζ is a.s. invariant. Thus, $\zeta = \rho\, \lambda^d$ a.s. for some random variable $\rho \geq 0$, which means that $\eta = \rho\, e_r \cdot \lambda^d$ a.s.

Using Fubini's theorem and the definition and invariance of ζ, we get a.s., for any measurable function $f \geq 0$,

$$\begin{aligned}\Big\{(\eta \otimes \mu) \circ \varphi^{-1}\Big\} f &= \int \eta(ds) \int \mu(dt)\, f(s+t, t) \\ &= \int \mu(dt) \int \zeta(ds)\, e^{rs} f(s+t, t) \\ &= \int \mu(dt) \int \zeta(ds)\, e^{rs-rt} f(s, t) \\ &= \Big\{\eta \otimes (e_{-r} \cdot \mu)\Big\} f. \end{aligned} \tag{26}$$

Comparing with (23) yields

$$(\eta * \mu) \otimes \nu = \eta \otimes (e_{-r} \cdot \mu) \ \text{ a.s.}$$

Since the measures on the left are locally finite, so are those on the right, and so the four measures are pairwise proportional. In particular, $\mu e^{-r} < \infty$ and $\nu = (e_{-r} \cdot \mu)/\mu e^{-r}$.

This completes the proof of the necessity. The sufficiency holds by the same calculation (26), combined with Theorem 3.2 (i). □

We conclude with a special property of Cox processes.

Proposition 5.30 *(stationary Cox processes) Let ξ be a Cox process on $\mathsf{R}^d \times S$, directed by some stationary random measure η. Then $\mathcal{I}_\xi \subset \sigma(\eta)$ a.s., with equality when η is a.s. invariant. Furthermore, $\bar{\xi} = \bar{\eta}$ a.s.*

Proof: If $I \in \mathcal{I}_\xi$, there exist some sets $B_n \in \hat{\mathcal{B}}^d \otimes \mathcal{S}$ and $A_n \in \sigma(1_{B_n}\xi)$, $n \in \mathsf{N}$, such that $P(I \Delta A_n) \leq 2^{-n}$. For any $B \in \mathcal{B}^d \otimes \mathcal{S}$, we may then use the stationarity of ξ to construct some events $\tilde{A}_n \in \sigma(1_{B^c}\xi)$ with $P(I \Delta \tilde{A}_n) =$

$P(I\Delta A_n) \leq 2^{-n}$. Hence, the Borel-Cantelli lemma yields $P\{I\Delta\tilde{A}_n \text{ i.o.}\} = 0$, which implies $I = \limsup_n \tilde{A}_n$ a.s. Since B was arbitrary, we conclude that I is a.s. $\mathcal{T}_\xi$-measurable, where $\mathcal{T}_\xi$ denotes the tail σ-field of ξ. By the conditional form of Kolmogorov's 0−1 law, we obtain $\mathcal{I}_\xi \subset \mathcal{T}_\xi \subset \sigma(\eta)$ a.s.

When η is a.s. invariant, the ergodic theorem yields $E(\xi\,|\mathcal{I}_\xi) = \eta$ a.s., which shows that η is a.s. $\mathcal{I}_\xi$-measurable. Hence, in this case, $\sigma(\eta) \subset \mathcal{I}_\xi$ a.s., and so the two σ-fields agree a.s. □

5.5 Palm and Spacing Averages

We begin with some ergodic theorems for Palm distributions, showing how the distributions of η and $\tilde{\eta}$ arise as limits of suitable averages. Given a function $f \in \mathcal{T}_+$ and a bounded measure $\nu \neq 0$ on R^d, we introduce the average

$$\bar{f}_\nu(t) = \|\nu\|^{-1} \int f(\theta_x t)\, \nu(dx), \quad t \in T.$$

Here and below, we will use Theorem A2.6 to obtain short and transparent proofs. Since the latter result applies only when the sample intensity $\bar{\xi}$ is a constant, we also need to introduce the *modified Palm distribution* $Q'_{\xi,\eta}$, defined for measurable functions $f \geq 0$ by

$$Q'_{\xi,\eta} f = E\,(\bar{\xi})^{-1} \int_{I_1} f(\theta_{-x}\eta)\, \xi(dx),$$

whenever $0 < \bar{\xi} < \infty$ a.s. For ergodic ξ we have $\bar{\xi} = E\bar{\xi}$ a.s., and $Q'_{\xi,\eta}$ agrees with $Q_{\xi,\eta}$. We may think of $Q'_{\xi,\eta}$ as the distribution of η when the pair (ξ, η) is shifted to a "typical" point of ξ. This interpretation clearly fails for $Q_{\xi,\eta}$, in general.

Theorem 5.31 *(pointwise averages) With R^d acting measurably on T, consider a stationary pair (ξ, η) with modified Palm version $(\tilde{\xi}, \tilde{\eta})$, where ξ is a random measure on R^d with $0 < \bar{\xi} < \infty$ a.s., and η is a random element in T. Then for any asymptotically invariant distributions μ_n or weight functions p_n on R^d, and any bounded, measurable functions f, we have*

(i) $\bar{f}_{\mu_n}(\tilde{\eta}) \xrightarrow{P} Ef(\eta)$,

(ii) $\bar{f}_{p_n \cdot \xi}(\eta) \xrightarrow{P} Ef(\tilde{\eta})$.

The convergence holds a.s. when $\mu_n = 1_{B_n}\lambda^d$ or $p_n = 1_{B_n}$, respectively, for some bounded, convex, increasing Borel sets B_n in R^d with $r_{B_n} \to \infty$.

Our proofs, here and below, are based on a powerful coupling property of independent interest.

Theorem 5.32 *(shift coupling, Thorisson) With R^d acting measurably on T, consider a stationary pair (ξ, η) with modified Palm version $(\tilde{\xi}, \tilde{\eta})$, where ξ*

is a random measure on R^d *with* $0 < \bar{\xi} < \infty$ *a.s., and* η *is a random element in* T. *Then there exist some random vectors* σ *and* τ *in* R^d, *such that*

$$\eta \stackrel{d}{=} \theta_\sigma \tilde{\eta}, \qquad \tilde{\eta} \stackrel{d}{=} \theta_\tau \eta.$$

Proof: Let $\mathcal{I}$ be the invariant σ-field in $\mathcal{M}_{\mathsf{R}^d} \times T$, put $I_1 = [0,1]^d$, and note that $E(\,\xi I_1\,|\,\mathcal{I}_{\xi,\eta}\,) = \bar{\xi}$ a.s. Then for any $I \in \mathcal{I}$,

$$\begin{aligned} P\{\tilde{\eta} \in I\} &= E\int_{I_1} 1\{\theta_{-x}\eta \in I\}\, \xi(dx)/\bar{\xi} \\ &= E\left\{\xi I_1/\bar{\xi};\, \eta \in I\right\} = P\{\eta \in I\}, \end{aligned}$$

which shows that $\eta \stackrel{d}{=} \tilde{\eta}$ on $\mathcal{I}$. The assertions now follow by Theorem A2.6. $\square$

Proof of Theorem 5.31: (i) By Theorem 5.32, we may assume that $\tilde{\eta} = \theta_\tau \eta$ for some random vector τ in R^d. Using Theorem A2.5 and the asymptotic invariance of the μ_n, we get

$$\left|\bar{f}_{\mu_n}(\tilde{\eta}) - Ef(\eta)\right| \le \|\,\mu_n - \theta_\tau \mu_n\|\, \|f\| + \left|\bar{f}_{\mu_n}(\eta) - Ef(\eta)\right| \xrightarrow{P} 0.$$

The a.s. version follows in the same way from Theorem A2.4.

(ii) For fixed f, we define a random measure ξ_f on R^d by

$$\xi_f B = \int_B f(\theta_{-x}\eta)\, \xi(dx), \quad B \in \mathcal{B}^d, \tag{27}$$

which is again stationary by Lemma 5.1. Applying Theorem 5.24 to both ξ and ξ_f, and using (4), we get with $I_1 = [0,1]^d$

$$\begin{aligned} \bar{f}_{g_n \cdot \xi}(\eta) &= \frac{\xi_f g_n}{\lambda^d g_n} \frac{\lambda^d g_n}{\xi g_n} \xrightarrow{P} \frac{\bar{\xi}_f}{\bar{\xi}} \\ &= \frac{E\xi_f I_1}{E\xi I_1} = Ef(\tilde{\eta}). \end{aligned}$$

The a.s. version follows by a similar argument, based on Theorem 5.23. $\square$

Taking expected values in Theorem 5.31, we get for bounded f the formulas

$$E\bar{f}_{\mu_n}(\tilde{\eta}) \to Ef(\eta), \qquad E\bar{f}_{p_n \cdot \xi}(\eta) \to Ef(\tilde{\eta}),$$

which may be regarded as limit theorems for suitable space averages of the distributions of η and $\tilde{\eta}$. We proceed to show that both statements hold uniformly for bounded f. For a striking formulation, we introduce the possibly defective distributions $\overline{\mathcal{L}}_\mu(\tilde{\eta})$ and $\overline{\mathcal{L}}_{p\cdot\xi}(\tilde{\eta})$, given for measurable functions $f \ge 0$ by

$$\overline{\mathcal{L}}_\mu(\tilde{\eta}) f = E\bar{f}_\mu(\tilde{\eta}), \qquad \overline{\mathcal{L}}_{p\cdot\xi}(\eta) f = E\bar{f}_{p\cdot\xi}(\eta).$$

Theorem 5.33 *(distributional averages, Slivnyak, Zähle) With* R^d *acting measurably on* T*, consider a stationary pair* (ξ, η) *with modified Palm version* $(\tilde{\xi}, \tilde{\eta})$*, where* ξ *is a random measure on* R^d *with* $0 < \bar{\xi} < \infty$ *a.s., and* η *is a random element in* T*. Then for any asymptotically invariant distributions* μ_n *or weight functions* p_n *on* R^d*, we have*

(i) $\left\| \overline{\mathcal{L}}_{\mu_n}(\tilde{\eta}) - \mathcal{L}(\eta) \right\| \to 0,$

(ii) $\left\| \overline{\mathcal{L}}_{p_n \cdot \xi}(\eta) - \mathcal{L}(\tilde{\eta}) \right\| \to 0.$

Proof: (i) By Theorem 5.32, we have $\tilde{\eta} = \theta_\tau \eta$ for some random vector τ. Using Fubini's theorem and the stationarity of η, we get for any measurable function $f \geq 0$

$$\begin{aligned} \overline{\mathcal{L}}_{\mu_n}(\eta) f &= \int E f(\theta_x \eta)\, \mu_n(ds) \\ &= E f(\eta) = \mathcal{L}(\eta) f. \end{aligned}$$

Hence, by Fubini's theorem and dominated convergence,

$$\begin{aligned} \left\| \overline{\mathcal{L}}_{\mu_n}(\tilde{\eta}) - \mathcal{L}(\eta) \right\| &= \left\| \overline{\mathcal{L}}_{\mu_n}(\theta_\tau \eta) - \overline{\mathcal{L}}_{\mu_n}(\eta) \right\| \\ &\leq E \left\| \int 1\{\theta_x \eta \in \cdot\}\, (\mu_n - \theta_\tau \mu_n)(dx) \right\| \\ &\leq E \|\mu_n - \theta_\tau \mu_n\| \to 0. \end{aligned}$$

(ii) Defining ξ_f by (27) with $0 \leq f \leq 1$, we get

$$\xi_f p_n = \int f(\theta_{-x} \eta)\, p_n(x)\, \xi(dx) \leq \xi p_n,$$

and so

$$\begin{aligned} \left| \overline{\mathcal{L}}_{p_n \cdot \xi}(\eta) f - \mathcal{L}(\tilde{\eta}) f \right| &= \left| E \bar{f}_{p_n \cdot \xi}(\eta) - E f(\tilde{\eta}) \right| \\ &\leq E \left| \frac{\xi_f p_n}{\xi p_n} - \frac{\xi_f p_n}{\bar{\xi}} \right| \leq E \left| 1 - \frac{\xi p_n}{\bar{\xi}} \right|, \end{aligned}$$

where $0/0 = 0$. Since $\xi p_n / \bar{\xi} \xrightarrow{P} 1$ by Theorem 5.24, and

$$\begin{aligned} E(\, \xi p_n / \bar{\xi}\,) &= E\left\{ E(\xi p_n | \mathcal{I}_{\xi,\eta}) \,/\, \bar{\xi} \right\} \\ &= E(\, \bar{\xi} / \bar{\xi}\,) = 1, \end{aligned}$$

we get $\xi p_n / \bar{\xi} \to 1$ in L^1 by FMP 4.12, and the assertion follows. $\Box$

We turn to some basic ergodic theorems for spacing measures. Given a stationary random measure ξ on R with spacing measure P^0_ξ, we introduce a random measure η with pseudo-distribution P^0_ξ, defined on an abstract space $\tilde{\Omega}$ with pseudo-probability $\tilde{P}$ and associated integration $\tilde{E}$. For ease of reference, we first restate some of the basic formulas of previous sections in terms of η.

Corollary 5.34 *(spacing measures and inversions) For any stationary random measure ξ on R with spacing measure P^0_ξ, let η be a random measure with pseudo-distribution P^0_ξ. Then*

(i) $\tilde{E}f(\eta) = E\int_0^{\xi_1} f(\theta_{-s}\tilde{\xi})\,ds,$

(ii) $\tilde{E}f(\tilde{\eta}) = E\int_0^1 f(\theta_{-s}\xi)\,\xi(ds),$

(iii) $E\{f(\xi);\, \xi \neq 0\} = \tilde{E}\int_0^{\eta_1} f(\theta_{-s}\tilde{\eta})\,ds,$

(iv) $E\{f(\tilde{\xi});\, \xi \neq 0\} = \tilde{E}\int_0^1 f(\theta_{-s}\eta)\,\eta(ds).$

Proof: (i) This agrees with (8).

(ii) Use (4) and Theorem 5.6 (i).

(iii)$-$(iv): By Theorem 5.6 (iii)$-$(iv), η is again stationary with spacing measure $\mathcal{L}(\xi;\, \xi \neq 0)$, and so (i)$-$(ii) apply to η. □

Before stating our main results, we need some basic lemmas, beginning with the local finiteness of the inversions $\tilde{\xi}$ and $\tilde{\eta}$.

Lemma 5.35 *(local finiteness)*

(i) $\tilde{\eta} \in \mathcal{M}$ *a.e.* $\tilde{P}$,

(ii) $\tilde{\xi} \in \mathcal{M}$ *a.s. on* $\{\xi \neq 0\}$.

Proof: (i) By Corollary 5.34 (ii), we have for any $B \in \hat{\mathcal{B}}$

$$\tilde{P}\{\tilde{\eta}B = \infty\} = E\int_0^1 1\Big\{\xi(B - s) = \infty\Big\}\,\xi(ds) = 0.$$

(ii) By Lemma 5.2 (i), we have $|\xi_t| \to \infty$ as $|t| \to \infty$, a.s. on $\{\xi \neq 0\}$, and so for any $B \in \hat{\mathcal{B}}$,

$$\tilde{\xi}B = \lambda\Big\{t \in \mathsf{R};\, \xi_t \in B\Big\} < \infty \ \text{ a.s. on } \ \Big\{\xi \neq 0\Big\}. \qquad \square$$

Next, we consider the existence, finiteness, and positivity of the sample intensity $\bar{\eta}$ of η, defined as before as the limit of $t^{-1}(\theta_s\eta)_t$ as $|t| \to \infty$. Recall that $\mathcal{I}_\eta = \mathcal{I} \circ \eta^{-1}$, where $\mathcal{I}$ denotes the invariant σ-field in $\mathcal{M}$.

Lemma 5.36 *(sample intensity of η)*

(i) $\bar{\eta}$ *exists and is finite a.e.* $\tilde{P}$,

(ii) *if* $\bar{\xi} < \infty$ *a.s., then* $\bar{\eta} > 0$ *a.e.* $\tilde{P}$,

(iii) $\tilde{P}$ *is* σ*-finite on* $\mathcal{I}_\eta \cap \{\bar{\eta} > 0\}$.

Proof: (i) Let $A = \{\mu \in \mathcal{M};\ \bar\mu \text{ exists}\}$. By Corollary 5.34 (i), Lemmas 5.13 and 5.35 (ii), and Theorem 5.23, we have

$$\begin{aligned}\tilde P\{\eta \notin A\} &= E\int_0^{\xi_1} 1\{\theta_{-s}\tilde\xi \notin A\}\,ds\\ &= E(\,\xi_1;\ \tilde\xi \notin A\,)\\ &= E(\,\xi_1;\ \xi \notin A\,) = 0.\end{aligned}$$

Hence, by the same results,

$$\begin{aligned}\tilde P\{\bar\eta = \infty\} &= E\int_0^{\xi_1} 1\Big\{\,\overline{\theta_{-s}\tilde\xi} = \infty\Big\}\,ds\\ &= E(\,\xi_1;\ \bar{\tilde\xi} = \infty\,) = E(\,\xi_1;\ \bar\xi = 0\,)\\ &= E\Big\{E(\xi_1|\,\mathcal{I}_\xi);\ \bar\xi = 0\Big\}\\ &= E(\,\bar\xi;\ \bar\xi = 0\,) = 0.\end{aligned}$$

(ii) If $\bar\xi < \infty$ a.s., then as before,

$$\begin{aligned}\tilde P\{\bar\eta = 0\} &= E(\,\xi_1;\ \bar{\tilde\xi} = 0\,)\\ &= E(\,\xi_1;\ \bar\xi = \infty\,) = 0.\end{aligned}$$

(iii) For every $n \in \mathsf{N}$, we get as before

$$\begin{aligned}\tilde P\{\bar\eta > n^{-1}\} &= E\int_0^{\xi_1} 1\Big\{\,\overline{\theta_{-s}\tilde\xi} > n^{-1}\Big\}\,ds\\ &= E(\,\xi_1;\ \bar\xi < n\,)\\ &= E(\,\bar\xi;\ \bar\xi < n\,) < n.\end{aligned}$$

Since $\bar\eta$ is $\mathcal{I}_\eta$-measurable and $\{\bar\eta > 0\} = \bigcup_n\{\bar\eta > n^{-1}\}$, this completes the proof. □

For any measurable function $f \geq 0$ on $\mathcal{M}$, we define the *average* $\bar f(\mu)$ by

$$\bar f(\mu) = \lim_{|t|\to\infty} t^{-1}\int_0^t f(\theta_{r+s}\,\mu)\,dr, \quad s \in \mathsf{R},$$

whenever the limit exists and is independent of s.

Lemma 5.37 *(functional averages) For any measurable function $f \geq 0$ on $\mathcal{M}$, we have*

(i) *$\bar f(\tilde\eta)$ exists a.e. $\tilde P$,*

(ii) *$\bar f(\tilde\xi)$ exists a.s. on $\{0 < \bar\xi < \infty\}$.*

Proof: (i) For fixed f, we introduce the measurable set $A = \{\mu \in \mathcal{M};\ \bar{f}(\mu) \text{ exists}\}$. By Corollary 5.34 (ii) and the pointwise ergodic theorem, we have

$$\begin{aligned}\tilde{P}\{\tilde{\eta} \notin A\} &= E\int_0^1 1\{\theta_{-s}\xi \notin A\}\,\xi(ds)\\ &= E(\,\xi_1;\, \xi \notin A\,) = 0.\end{aligned}$$

(ii) By Lemma 5.36 (iii) and the pointwise ergodic theorem, $\bar{f}(\eta)$ exists a.e. $\tilde{P}$ on $\{\bar{\eta} > 0\}$. Further note that $\{\bar{\xi} = 0\} = \{\xi = 0\}$ a.s., since

$$\begin{aligned}E(\,\bar{\xi};\, \xi = 0\,) &= E(\,\xi_1;\, \xi = 0\,) = 0,\\ E(\,\xi_t;\, \bar{\xi} = 0\,) &= t\,E(\,\bar{\xi};\, \bar{\xi} = 0\,) = 0, \quad t > 0.\end{aligned}$$

Using Corollary 5.34 (iv) and Lemma 5.13, we get as before

$$\begin{aligned}P\Big\{\tilde{\xi} \notin A,\, \bar{\xi} \in (0,\infty)\Big\} &= P\Big\{\tilde{\xi} \notin A,\, \xi \neq 0,\ \bar{\tilde{\xi}} > 0\Big\}\\ &= \tilde{E}\int_0^1 1\Big\{\theta_{-s}\eta \notin A,\, \overline{\theta_{-s}\eta} > 0\Big\}\,\eta(ds)\\ &= \tilde{E}\Big\{\,\eta_1;\, \eta \notin A,\, \bar{\eta} > 0\Big\} = 0,\end{aligned}$$

and the assertion follows. □

We may now state the main result of this section, which exhibits a striking symmetry between ξ, $\tilde{\xi}$, η, and $\tilde{\eta}$, apart from the conditions of validity.

Theorem 5.38 *(inversions and averages) For any stationary random measure ξ on R, let η be a random measure with $\tilde{P}$-distribution P_ξ^0. Then for any measurable function $f \geq 0$ on $\mathcal{M}$, we have*

(i) $\tilde{E}\bar{f}(\tilde{\eta}) = E\,\bar{\xi}f(\xi)$,

(ii) $\tilde{E}f(\eta) = E\,\bar{\xi}\bar{f}(\tilde{\xi})$ *when* $\bar{\xi} < \infty$ *a.s.,*

(iii) $E\bar{f}(\tilde{\xi}) = \tilde{E}\,\bar{\eta}f(\eta)$ *when* $0 < \bar{\xi} < \infty$ *a.s.,*

(iv) $Ef(\xi) = \tilde{E}\,\bar{\eta}\bar{f}(\tilde{\eta})$ *when* $0 < \bar{\xi} < \infty$ *a.s.*

A simple substitution reveals an even more remarkable symmetry.

Corollary 5.39 *(modified spacing measure) For ξ and η as above, with $0 < \bar{\xi} < \infty$ a.s., let ζ be a random measure with $\mathcal{L}(\zeta) = \tilde{E}(\bar{\eta};\, \eta \in \cdot)$. Then for any measurable function $f \geq 0$ on $\mathcal{M}$,*

(i) $Ef(\zeta) = E\bar{f}(\tilde{\xi})$,

(ii) $Ef(\xi) = E\bar{f}(\tilde{\zeta})$.

Proof: (i) Define $g(\mu) = \bar\mu f(\mu)$ on the domain of $\bar\mu$, and note that $\bar g(\mu) = \bar\mu \bar f(\mu)$ by the invariance of $\bar\mu$. By Theorem 5.38 (ii) and Lemma 5.13,

$$\begin{aligned} Ef(\zeta) &= \tilde E\, \bar\eta\, f(\eta) = \tilde E g(\eta) \\ &= E\,\bar\xi\, \bar g(\tilde\xi) = E\,\bar\xi\, \bar{\tilde\xi}\, \bar f(\tilde\xi) \\ &= E\bar f(\tilde\xi). \end{aligned}$$

(ii) Use Theorem 5.38 (iv). □

Proof of Theorem 5.38: (i) Using Corollary 5.34 (ii) and the invariance of $\bar f$, we get

$$\begin{aligned} \tilde E \bar f(\tilde\eta) &= E \int_0^1 \bar f(\theta_{-s}\xi)\, \xi(ds) \\ &= E\,\xi_1 \bar f(\xi) = E\,\xi_1 E\{f(\xi)\,|\,\mathcal I_\xi\} \\ &= E\,E(\xi_1|\mathcal I_\xi)\, f(\xi) = E\,\bar\xi f(\xi). \end{aligned}$$

(ii) When $\bar\xi < \infty$ a.s., the measure $\tilde P$ is σ-finite on $\mathcal I_\eta$, by Lemma 5.36 (iii), and so $\bar f(\eta)$ exists, by the pointwise ergodic theorem, and satisfies

$$\bar f(\eta) = \tilde E\{f(\eta)\,|\,\mathcal I_\eta\} \ \text{ a.e. } \tilde P.$$

The function $\bar f(\mu)$ is clearly invariant whenever it exists, and the invariance extends to $\bar f(\tilde\mu)$ by Lemma 5.14, which shows that $\bar f(\tilde\xi)$ is a.s. $\mathcal I_\xi$-measurable. Using Corollary 5.34 (i), and proceeding as before, we obtain

$$\begin{aligned} \tilde E f(\eta) &= \tilde E \tilde E\{f(\eta)\,|\,\mathcal I_\eta\} = \tilde E \bar f(\eta) \\ &= E \int_0^{\xi_1} \bar f(\theta_{-s}\tilde\xi)\, ds \\ &= E\,\xi_1 \bar f(\tilde\xi) = E\,\bar\xi \bar f(\tilde\xi). \end{aligned}$$

(iii) Let $\bar\xi \in (0,\infty)$ a.s., and note that $\bar\eta = \tilde E(\eta_1\,|\,\mathcal I_\eta)$ and $\bar f(\eta) = \tilde E\{f(\eta)\,|\,\mathcal I_\eta\}$ a.e. $\tilde P$, by Lemma 5.36. Using Corollary 5.34 (iv), and noting that $\xi \neq 0$ a.s., we get as in case of (i)

$$\begin{aligned} E\bar f(\tilde\xi) &= \tilde E \int_0^1 \bar f(\theta_{-s}\eta)\, \eta(ds) \\ &= \tilde E\,\eta_1 \bar f(\eta) = \tilde E\,\bar\eta f(\eta). \end{aligned}$$

(iv) Let $\bar\xi \in (0,\infty)$ a.s., and note as before that $\bar f(\xi) = E\{f(\xi)\,|\,\mathcal I_\xi\}$ a.s. and $\bar\eta = \tilde E(\eta_1\,|\,\mathcal I_\eta)$ a.e. $\tilde P$. Further note that $\bar f(\tilde\eta)$ is a.e. invariant and hence $\mathcal I_\eta$-measurable. Using Corollary 5.34 (iii), we get as in part (ii)

$$\begin{aligned} Ef(\xi) &= E\bar f(\xi) = \tilde E \int_0^{\eta_1} \bar f(\theta_{-s}\tilde\eta)\, ds \\ &= \tilde E\,\eta_1 \bar f(\tilde\eta) = \tilde E\,\bar\eta \bar f(\tilde\eta). \end{aligned}$$ □

The next result shows that ξ and η are essentially simultaneously ergodic. For η, this means that $P^0_\xi I \wedge P^0_\xi I^c = 0$ for every $I \in \mathcal I$.

Theorem 5.40 *(ergodicity) For any stationary random measure ξ on* R*, the measures $\mathcal{L}(\xi;\, \xi \neq 0)$ and P^0_ξ are simultaneously ergodic.*

Proof: First, let ξ be ergodic on $\{\xi \neq 0\}$. Fix any $I \in \mathcal{I}$, and note that $J = \{\mu;\, \tilde{\mu} \in I\}$ is invariant on $\{\mu;\, \tilde{\mu} \in \mathcal{M}\}$ by Lemma 5.14, so that either $\xi \in J^c \setminus \{0\}$ a.s. or $\xi \in J \setminus \{0\}$ a.s., by Lemma 5.35 (ii). In the former case, we get by Corollary 5.34 (i) and the invariance of I

$$\begin{aligned} P^0_\xi I &= E\int_0^{\xi_1} 1\Big\{\theta_{-s}\tilde{\xi} \in I\Big\}\, ds \\ &= E(\xi_1;\, \tilde{\xi} \in I) \\ &= E(\xi_1;\, \xi \in J) = 0. \end{aligned}$$

Similarly, $P^0_\xi I^c = 0$ when $\xi \in J \setminus \{0\}$ a.s. This shows that P^0_ξ is ergodic.

Conversely, suppose that P^0_ξ is ergodic, and fix any $I \in \mathcal{I}$. Put $J = \{\mu;\, \tilde{\mu} \in I\}$ as before, and note that either $P^0_\xi J = 0$ or $P^0_\xi J^c = 0$. Writing $\gamma = \xi_0^{-1}$, and using Corollary 5.34 (i), Lemmas 5.7 (iii), 5.14, and 5.35 (ii), and the invariance of I, we get

$$\begin{aligned} P^0_\xi J &= E\int_0^{\xi_1} 1\Big\{\theta_{-s}\tilde{\xi} \in J\Big\}\, ds \\ &= E(\xi_1;\, \tilde{\xi} \in J) = E\Big\{\xi_1;\, \tilde{\tilde{\xi}} \in I\Big\} \\ &= E\Big\{\xi_1;\, \theta_{-\gamma}\xi \in I\Big\} = E(\xi_1;\, \xi \in I) \\ &= E\Big\{E(\xi_1 \,|\, \mathcal{I}_\xi);\, \xi \in I\Big\} = E(\bar{\xi};\, \xi \in I), \end{aligned}$$

and similarly $P^0_\xi J^c = E(\bar{\xi};\, \xi \in I^c)$. Hence,

$$E(\bar{\xi};\, \xi \in I) \wedge E(\bar{\xi};\, \xi \in I^c) = 0, \quad I \in \mathcal{I},$$

and so either $\xi \in I^c$ a.s. or $\xi \in I$ a.s., on the set $\{\bar{\xi} > 0\}$. Since $\{\bar{\xi} = 0\} = \{\xi = 0\}$ a.s., this shows that ξ is ergodic on $\{\xi \neq 0\}$. □

In the ergodic case, we can strengthen Corollary 5.39 to a pointwise ergodic theorem.

Corollary 5.41 *(functional averages) Let ξ be a stationary and ergodic random measure on* R*, with $0 < \bar{\xi} < \infty$ a.s. Then for any measurable function $f \geq 0$ on $\mathcal{M}$, we have*

(i) $\bar{f}(\tilde{\xi}) = Ef(\zeta)$ *a.s.,*

(ii) $\bar{f}(\tilde{\zeta}) = Ef(\xi)$ *a.s.*

Proof: Since $\bar{\xi} \in (0, \infty)$ a.s., $\bar{f}(\tilde{\xi})$ and $\bar{f}(\tilde{\zeta})$ exist a.s. by Lemma 5.37. Noting that $\bar{f}$ is invariant whenever it exists, we see from Lemma 5.14 that the two random variables are even a.s. invariant under shifts of ξ and ζ, respectively. Since ξ and ζ are both ergodic, by Theorem 5.40, we conclude that $\bar{f}(\tilde{\xi})$ and $\bar{f}(\tilde{\zeta})$ are a.s. constant, and the assertions follow from Corollary 5.39. □

We proceed with a symmetric version of the distributional limit Theorem 5.33, where we use the same notation as before.

Corollary 5.42 *(distributional averages) Let ξ be a stationary random measure on $\mathbb{R}$ with $0 < \bar{\xi} < \infty$ a.s., and consider any asymptotically invariant distributions μ_n on $\mathbb{R}$. Then as $n \to \infty$,*

(i) $\left\| \overline{\mathcal{L}}_{\mu_n}(\tilde{\zeta}) - \mathcal{L}(\xi) \right\| \to 0,$

(ii) $\left\| \overline{\mathcal{L}}_{\mu_n}(\tilde{\xi}) - \mathcal{L}(\zeta) \right\| \to 0.$

Proof: For any invariant function $f \geq 0$ on $\mathcal{M}$, we have $\bar{f} = f$, and so by Corollary 5.39

$$Ef(\xi) = Ef(\tilde{\zeta}), \qquad Ef(\zeta) = Ef(\tilde{\xi}).$$

Hence, $\xi \overset{d}{=} \tilde{\zeta}$ and $\zeta \overset{d}{=} \tilde{\xi}$ on $\mathcal{I}$, and so by Theorem 5.32 there exist some random variables σ and τ satisfying

$$\tilde{\zeta} \overset{d}{=} \theta_\sigma \xi, \qquad \tilde{\xi} \overset{d}{=} \theta_\tau \zeta.$$

Now choose some random variables γ_n with distributions μ_n, independent of ξ, ζ, σ, and τ. Using the stationarity of ξ, Fubini's theorem, and dominated convergence, we get

$$\begin{aligned}
\left\| \overline{\mathcal{L}}_{\mu_n}(\tilde{\zeta}) - \mathcal{L}(\xi) \right\| &= \left\| \mathcal{L}(\theta_{\gamma_n+\sigma}\xi) - \mathcal{L}(\theta_{\gamma_n}\xi) \right\| \\
&\leq \left\| \mathcal{L}(\gamma_n + \sigma) - \mathcal{L}(\gamma_n) \right\| \\
&\leq E\left\| \mu_n - \mu_n * \delta_\sigma \right\| \to 0,
\end{aligned}$$

which proves part (i). The proof of (ii) is similar. □

We conclude with an a.s. limit theorem for the empirical distributions

$$\hat{\mathcal{L}}_t(\tilde{\xi}) = t^{-1}\int_0^t 1\{\theta_s\tilde{\xi} \in \cdot\}\, ds, \qquad \hat{\mathcal{L}}_t(\tilde{\zeta}) = t^{-1}\int_0^t 1\{\theta_s\tilde{\zeta} \in \cdot\}\, ds,$$

suggested by Corollary 5.41.

Corollary 5.43 *(empirical distributions) Let ξ be a stationary and ergodic random measure on $\mathbb{R}$ with $0 < \bar{\xi} < \infty$ a.s. Then as $t \to \infty$,*

(i) $\hat{\mathcal{L}}_t(\tilde{\zeta}) \xrightarrow{vw} \mathcal{L}(\xi)$ *a.s.,*

(ii) $\hat{\mathcal{L}}_t(\tilde{\xi}) \xrightarrow{vw} \mathcal{L}(\zeta)$ *a.s.*

Proof: Let $\hat{C}$ be the class of bounded, continuous functions on $\mathcal{M}$, and define

$$\bar{f}_t(\mu) = t^{-1} \int_0^t f(\theta_s \mu)\, ds, \quad t > 0.$$

We need to prove that, as $t \to \infty$,

$$\bar{f}_t(\tilde{\zeta}) \to Ef(\xi), \quad \bar{f}_t(\tilde{\xi}) \to Ef(\zeta), \qquad f \in \hat{C},$$

outside a fixed P-null set. Since the convergence holds for every fixed $f \in \hat{C}$, by Corollary 5.41, it remains to show that $\hat{C}$ contains a countable, convergence-determining subclass for the vague topology on $\mathcal{M}$, which is clear by an elementary construction. □

5.6 Local Invariance

Introducing the scaling function $S_t(x) = tx$, we say that a measure $\mu \in \mathcal{M}_{\mathsf{R}^d}$ is *locally invariant*, if the measures $\mu_t = (1_B \mu) \circ S_t^{-1}$ are weakly asymptotically invariant, for every bounded rectangle $B \subset \mathsf{R}^d$. By Lemma 5.16, the stated condition is equivalent to

$$\lim_{h \to 0} h^{-d} \int_B \Big| \mu(I_h + x) - \mu(I_h + x + h\, e_i) \Big|\, dx = 0, \quad i \le d,\ B \in \hat{\mathcal{B}}^d, \qquad (28)$$

where we may restrict h to the numbers 2^{-n} and replace the integration by summation over $\mathcal{I}_n$. We may also consider the more general property of *local u-invariance*, where the stated invariance is only assumed under shifts x in the subspace u. For simplicity of notation, we consider only the case of $u = \mathsf{R}^d$, the general statements and proofs being similar.

It might also seem natural to consider the stronger version of *strict local invariance*, defined by the requirement that the scaled measures μ_t be asymptotically invariant in the strict sense. However, this notion gives nothing new:

Lemma 5.44 *(strict local invariance) For any probability measure μ on R^d, these conditions are equivalent:*

(i) *μ is strictly locally invariant,*

(ii) *μ is absolutely continuous.*

Proof: Assume (i), and let μ'' denote the singular component of μ. Then Lemma 5.15 yields $\|\mu''\| = \|\mu'' \circ S_t^{-1}\| \to 0$ as $t \to \infty$, which implies $\mu'' = 0$ and hence $\mu \ll \lambda^d$, proving (ii).

Now assume (ii), so that $\mu = f \cdot \lambda^d$ for some measurable function $f \ge 0$. By FMP 1.35, we may choose some continuous functions $f_n \ge 0$ with bounded supports on R^d, such that $\|f - f_n\|_1 \to 0$. Writing $\mu_t = \mu \circ S_t^{-1}$ and $h = 1/t$, we get for any $r \in \mathsf{R}^d$

$$\begin{aligned}\|\mu_t - \mu_t * \delta_r\| &= \int \big| f(x) - f(x+hr) \big|\, dx \\ &\le \int \big| f_n(x) - f_n(x+hr) \big|\, dx + 2\|f - f_n\|_1,\end{aligned}$$

which tends to 0 as $h \to 0$ and then $n \to \infty$. Since r was arbitrary, this proves (i). □

We proceed with a basic closure property, for the class of locally invariant measures on R^d.

Theorem 5.45 *(absolute continuity) If the measure μ on R^d is locally invariant, then so is any locally finite measure $\mu' \ll \mu$.*

Proof: We may assume that μ is bounded, so that its local invariance becomes equivalent to weak asymptotic invariance of the measures $\mu_t = \mu \circ S_t^{-1}$. Since $\mu' \ll \mu$ is locally finite, we may further assume that $\mu' = f \cdot \mu$, for some μ-integrable function $f \ge 0$ on R^d. By FMP 1.35, there exist some continuous functions f_n with bounded supports, such that $f_n \to f$ in $L^1(\mu)$. Writing $\mu'_n = f_n \cdot \mu$ and $\mu'_{n,t} = \mu'_n \circ S_t^{-1}$, we get

$$\begin{aligned}\Big\| \mu'_{n,t} * \nu - \mu'_t * \nu \Big\| &\le \|\mu'_{n,t} - \mu'_t\| = \|\mu'_n - \mu'\| \\ &= \|(f_n - f) \cdot \mu\| \\ &= \mu|f_n - f| \to 0,\end{aligned}$$

and so it suffices to verify (15) for the measures $\mu'_{n,t}$. Thus, we may henceforth assume that $\mu' = f \cdot \mu$, for some continuous function f with bounded support.

To verify (15), let r and ν be such that the measures ν and $\nu * \delta_a$ are both supported by the ball B_0^a. Writing

$$\begin{aligned}\mu'_t &= (f \cdot \mu)_t = (f \cdot \mu) \circ S_t^{-1} \\ &= (f \circ S_t) \cdot (\mu \circ S_t^{-1}) = f_t \cdot \mu_t, \\ \nu - \nu * \delta_r &= g \cdot \lambda^d - (g \circ \theta_{-r}) \cdot \lambda^d \\ &= \Delta g \cdot \lambda^d,\end{aligned}$$

we get for any bounded, measurable function h on R^d

$$\begin{aligned}\Big(\mu'_t * \nu - \mu'_t * \nu * \delta_r\Big) h &= \Big\{ (f_t \cdot \mu_t) * (\Delta g \cdot \lambda^d) \Big\} h \\ &= \int f_t(x)\, \mu_t(dx) \int \Delta g(y)\, h(x+y)\, dy \\ &= \int f_t(x)\, \mu_t(dx) \int \Delta g(y-x)\, h(y)\, dy \\ &= \int h(y)\, dy \int f_t(x)\, \Delta g(y-x)\, \mu_t(dx) \\ &\approx \int h(y)\, f_t(y)\, dy \int \Delta g(y-x)\, \mu_t(dx) \\ &= \Big\{ \mu_t * (\Delta g \cdot \lambda^d) \Big\} f_t h.\end{aligned}$$

Letting m denote the modulus of continuity of f, and writing $m_t = m(r/t)$, we can estimate the approximation error in the fifth step by

$$\begin{aligned}\int |h(y)|\,dy \int |f_t(x) - f_t(y)|\,|\Delta g(y-x)|\,\mu_t(dx)&\\ \le m_t \int \mu_t(dx) \int \Big| h(y)\,\Delta g(y-x)\Big|\,dy&\\ \le m_t\,\|h\|\,\|\mu_t\| \int |\Delta g(y)|\,dy&\\ \le 2\,m_t\,\|h\|\,\|\mu\|\,\|\nu\| \to 0,&\end{aligned}$$

by the uniform continuity of f. It remains to note that

$$\Big\| f_t \cdot \big\{ \mu_t * (\Delta g \cdot \lambda^d) \big\} \Big\| \le \|f\|\, \Big\| \mu_t * \nu - \mu_t * \nu * \delta_r \Big\| \to 0,$$

by the local invariance of μ. □

The various regularity properties of measures on R^d are related as follows:

Corollary 5.46 *(regularity classes) The classes of absolutely continuous, locally invariant, and diffuse measures on* R^d *are related by*

$$\mathcal{M}_a \subset \mathcal{M}_l \subset \mathcal{M}_d,$$

where both inclusions are strict.

Proof: Clearly $\lambda^d \in \mathcal{M}_l$ by the shift invariance of λ, and so $\mathcal{M}_a \subset \mathcal{M}_l$ by Theorem 5.45. The second relation holds since

$$\liminf_{h\to 0} h^{-d} \int \Big| \mu(I_h + x) - \mu(I_h + x + h\,e_i)\Big|\,dx \ge 2\,\textstyle\sum_x \mu\{x\}.$$

To see that both inclusions are strict, we may first take $d = 1$. Letting $\sigma_1, \sigma_2, \ldots$ form a Bernoulli sequence with $E\sigma_n = \frac{1}{2}$, and putting

$$\alpha_1 = \sum_{n\ge 1} n^{-1} 2^{-n} \sigma_n, \qquad \alpha_2 = \sum_{n\ge 1} 3^{-n} \sigma_n, \tag{29}$$

we note that the distributions $\mu_1 = \mathcal{L}(\alpha_1)$ and $\mu_2 = \mathcal{L}(\alpha_2)$ satisfy $\mu_1 \in \mathcal{M}_l \setminus \mathcal{M}_a$ and $\mu_2 \in \mathcal{M}_d \setminus \mathcal{M}_l$. Indeed, both measures are singular, since their supports have Lebesgue measure 0, and they are also diffuse, since the maps in (29) are injective from $\{\pm 1\}^\infty$ to R. Next, we note that $\mu_2 \notin \mathcal{M}_l$, since the summation version of (28) fails along the sequence $h = 3^{-n}$. On the other hand, $\mu_1 \in \mathcal{M}_l$ may be seen from the local approximation

$$\lim_{m\to\infty} \left| \sum_{n\ge m} \frac{m}{n}\, 2^{m-n}\,\sigma_n - \sum_{n\ge m} 2^{m-n}\,\sigma_n \right| \to 0,$$

where the distribution of the second term is $U(0,1)$.

To get examples for general d, we may form products of μ_1 or μ_2 above with Lebesgue measure λ^{d-1}. □

The definition of local invariance extends to random measures ξ on R^d, as follows. Writing $\|\cdot\|_p$ for the norm in $L^p(P)$, we say that ξ is *locally* L^p*-invariant*, if for all bounded sets $B \in \mathcal{B}^d$,

$$\limsup_{h\to 0} h^{-d}\int_B \|\,\xi(I_h+x)\|_p\,dx < \infty, \tag{30}$$

$$\lim_{h\to 0} h^{-d}\int_B \Big\|\,\xi(I_h+x)-\xi(I_h+x+h\,e_i)\Big\|_p\,dx = 0, \quad i\le d. \tag{31}$$

The definition of local L^p-invariance under shifts in u is similar. By Jensen's inequality, the versions for $p>1$ are more restrictive than those for $p=1$.

Theorem 5.47 *(local L^p-invariance) Let ξ be a random measure on R^d.*

(i) *For fixed $p>1$, ξ is locally L^p-invariant, iff $\|\xi\|_p \in \mathcal{M}$ and $E\xi \in \mathcal{M}_l$.*

(ii) *ξ is locally L^1-invariant, iff $E\xi \in \mathcal{M}$ and ξ satisfies* (28) *in L^1. When $\xi \ll \mu$ a.s. for some $\mu\in\mathcal{M}$, it is equivalent that $E\xi\in\mathcal{M}_l$.*

Combining (i) with Theorem 2.13, we see that ξ is locally L^p-invariant for some $p>1$, iff $\xi \ll E\xi \in \mathcal{M}_l$ a.s. and $\|\xi\|_p \in \mathcal{M}$. However, by (ii) and Corollary 5.46, the latter conditions are strictly stronger for $p=1$. They will therefore be referred to, for any $p\ge 1$, as *strong local L^p-invariance.*

Proof: (i) We may assume that ξ has bounded support, which allows us to take $B=\mathsf{R}^d$ in (30) and (31). In R^d we introduce the standard dissecting system $\mathcal{I}=(I_{nj})$, where the I_{nj} are cubes of side length 2^{-n}. By an easy approximation, (30) is equivalent to

$$\lim_{n\to\infty} \sum\nolimits_j \|\,\xi I_{nj}\,\|_p < \infty,$$

where the limit exists by Minkowski's inequality. By Theorem 2.13, this is equivalent to the existence of the L^p-intensity $\|\xi\|_p$, which also implies $E\xi\in\mathcal{M}$. Using Jensen's and Minkowski's inequalities, we get

$$\begin{aligned} h^{-d}\int \Big\|\,\xi(I_h+x)-\xi(I_h+x+h\,e_i)\Big\|_p\,dx \\ &\ge h^{-d}\int E\Big|\,\xi(I_h+x)-\xi(I_h+x+h\,e_i)\Big|\,dx \\ &\ge h^{-d}\int \Big|E\,\xi(I_h+x)-E\,\xi(I_h+x+h\,e_i)\Big|\,dx, \end{aligned}$$

which shows that (31) implies $E\xi\in\mathcal{M}_l$. This proves the necessity in (i).

Conversely, suppose that $\|\xi\|_p\in\mathcal{M}$ and $E\xi\in\mathcal{M}_l$ for some $p>1$. Then Theorem 2.13 yields an L^p-valued, measurable process $X\ge 0$ on R^d, such that

$$\xi = X\cdot E\xi \;\text{ a.s.,} \qquad \|\xi\|_p = \|X\|_p\cdot E\xi.$$

Using dominated convergence twice, we obtain

$$E\,\xi\|X - X\wedge n\|_p \to 0. \tag{32}$$

Since $E\,\xi\|X\wedge n\|_1 < \infty$, there exist as in FMP 1.35 some continuous processes Y_{nk} in $[0,n]$ with uniformly bounded supports, satisfying

$$\lim_{k\to\infty} E\,\xi\|X\wedge n - Y_{nk}\|_1 = 0, \quad n \in \mathsf{N}, \tag{33}$$

which implies $\|X\wedge n - Y_{nk}\|_1 \to 0$ in $E\xi$-measure. This extends by boundedness to $\|X\wedge n - Y_{nk}\|_p \to 0$, which yields the L^p-version of (33), by the boundedness of both $E\xi$ and the norm. By Minkowski's inequality and (32), we may then construct some uniformly bounded, continuous processes $X_n \geq 0$ with uniformly bounded supports, satisfying $E\,\xi\|X - X_n\|_p \to 0$.

Now define $\xi_n = X_n \cdot E\xi$. Writing $a_n \approx b_n$ for the approximation $a_n - b_n \to 0$, we obtain

$$\begin{aligned}
h^{-d}\!\int \Big\| \xi(I_h+x) - \xi(I_h+x+h\,e_i)\Big\|_p\,dx \\
&\approx h^{-d}\!\int \Big\| \xi_n(I_h+x) - \xi_n(I_h+x+h\,e_i)\Big\|_p\,dx \\
&\approx h^{-d}\!\int \|X_n(x)\|_p \Big| E\,\xi(I_h+x) - E\,\xi(I_h+x+h\,e_i)\Big|\,dx \\
&\approx h^{-d}\!\int \Big| \|\xi_n\|_p(I_h+x) - \|\xi_n\|_p(I_h+x+h\,e_i)\Big|\,dx,
\end{aligned}$$

where all three estimates will be justified below. Here the right-hand side tends to 0, by the local invariance of $E\xi$ and Theorem 5.45.

To justify the first approximation, we may use the extended Minkowski inequality (FMP 1.30), Fubini's theorem, and the invariance of λ, to see that the corresponding difference is bounded by twice the amount

$$\begin{aligned}
h^{-d}\!\int \Big\| \xi(I_h+x) - \xi_n(I_h+x)\Big\|_p\,dx \\
&= h^{-d}\!\int \Big\|\Big\{(X - X_n)\cdot E\xi\Big\}(I_h+x)\,\Big\|_p\,dx \\
&\leq h^{-d}\!\int \Big(\|X - X_n\|_p \cdot E\xi\Big)(I_h+x)\,dx \\
&= h^{-d}\!\int \Big(\|X - X_n\|_p \cdot E\xi\Big)(dy)\int 1_{I_h}(y-x)\,dx \\
&= E\,\xi\|X - X_n\|_p \to 0,
\end{aligned}$$

uniformly h. It remains to prove that the other two approximation errors tend to 0, as $h \to 0$ for fixed n.

For convenience, we may write $Y = X_n$, $\mu = E\xi$, and $\Delta I_x^h = (I_h + x) - (I_h + x + he_i)$, where the latter difference is understood in the sense of indicator functions. Let ρ denote the modulus of continuity of Y. Then the difference in the second step becomes

$$\begin{aligned}
h^{-d}\Big|\int \Big\|(Y\cdot\mu)\Delta I_x^h\Big\|_p\,dx - \int \|Y_x\|_p\,|\mu(\Delta I_x^h)|\,dx\Big| \\
&\leq h^{-d}\!\int \Big\|\Big\{(Y - Y_x)\cdot\mu\Big\}\Delta I_x^h\Big\|_p\,dx
\end{aligned}$$

$$\begin{aligned}
&\leq h^{-d} \int \Big(\|Y - Y_x\|_p \cdot \mu \Big) |\Delta I_x^h| \, dx \\
&\leq 2\, h^{-d} \, \|\rho_h\|_p \int \mu(dy) \int |\Delta I_0^h(y-x)| \, dx \\
&= 4\, \|\rho_h\|_p \, \|\mu\| \to 0,
\end{aligned}$$

in view of the regular and extended versions of Minkowski's inequality, Fubini's theorem, the invariance of λ, the a.s. uniform continuity of Y, the boundedness of Y and μ, and dominated convergence. Similarly, the difference in the third step equals

$$\begin{aligned}
h^{-d} \bigg| \int \Big| (\|Y\|_p \cdot \mu) \Delta I_x^h \Big| \, dx - \int \|Y_x\|_p \, |\mu(\Delta I_x^h)| \, dx \bigg| \\
\leq h^{-d} \int \Big| (\|Y\|_p - \|Y_x\|_p) \cdot \mu) \Delta I_x^h \Big| \, dx \\
\leq h^{-d} \int (\|Y - Y_x\|_p \cdot \mu) |\Delta I_x^h| \, dx,
\end{aligned}$$

which tends to 0 as before. This completes the proof for $p > 1$.

(ii) For $p = 1$, condition (30) means that $E\xi \in \mathcal{M}$, whereas (31) means that (28) holds in L^1 for the random measure ξ. Since the sum in (28) is uniformly integrable by (30), this is equivalent to convergence in L^1. We also see as before that $E\xi \in \mathcal{M}_l$.

Conversely, suppose that $\xi \ll \mu$ a.s. and $E\xi \in \mathcal{M}_l$. Then $\xi \ll E\xi$ a.s. by Theorem 2.12, and so Theorem 5.45 yields $\xi \in \mathcal{M}_l$ a.s. The local L^1-invariance in (31) now follows by dominated convergence. □

The following stochastic version of Theorem 5.45 will be needed in a later chapter.

Corollary 5.48 *(absolute continuity) For any $p \geq 1$, consider a locally L^p-invariant random measure ξ on R^d, and let $\eta \ll \xi$ a.s. with $\|\eta\|_p \in \mathcal{M}$. Then η is again locally L^p-invariant.*

Proof, ($p = 1$): By Theorem 5.47 (i), we need to show that if ξ satisfies (28) in probability, then so does η. By the sub-sequence criterion in FMP 4.2, it suffices to prove the corresponding implication for a.s. convergence along a sub-sequence, which holds by an obvious extension of Theorem 5.45.

($p > 1$): By Theorem 5.47 (ii), we need to show that $E\xi \in \mathcal{M}_l$ implies $E\eta \in \mathcal{M}_l$, which holds by Theorem 5.45, since $E\eta \in \mathcal{M}$ with $E\eta \ll E\xi$. □

We have also a stochastic version of Corollary 5.46:

Corollary 5.49 *(classes of L^p-regularity) Corollary 5.46 remains valid for the corresponding classes of random measures with locally finite L^p-intensities, where the local invariance is understood in the L^p-sense.*

Proof: First let $\xi \ll \lambda^d$ a.s. with $\|\xi\|_p \in \mathcal{M}$. For $p = 1$, we get $E\xi \in \mathcal{M}$, and ξ is a.s. locally invariant by Theorem 5.45. Hence, ξ is locally L^1-invariant by Theorem 5.47 (i). If instead $p > 1$, then Theorem 2.13 yields $E\xi \in \mathcal{M}$ with $E\xi \ll \lambda^d$, and so ξ is locally L^p-invariant, by Theorem 5.47 (ii).

To prove the second relation, we may assume that ξ has uniformly bounded support. If ξ is locally L^p-invariant, then by Fatou's lemma and the extended Minkowski inequality,

$$\begin{aligned}\Big\|\textstyle\sum_x \xi\{x\}\Big\|_p &\le \Big\|\liminf_{h\to 0} h^{-d}\int \Big|\xi(I_h + x) - \xi(I_h + x + he_i)\Big|\,dx\Big\|_p \\ &\le \liminf_{h\to 0}\Big\|h^{-d}\int \Big|\xi(I_h + x) - \xi(I_h + x + he_i)\Big|\,dx\Big\|_p \\ &\le \liminf_{h\to 0} h^{-d}\int \Big\|\xi(I_h + x) - \xi(I_h + x + he_i)\Big\|_p\,dx,\end{aligned}$$

which shows that ξ is a.s. diffuse. □

Finally, we show how the local L^p-invariance of a random measure is preserved, under smooth, invertible transformations of the space. Say that a mapping $f\colon \mathbb{R}^d \to \mathbb{R}^d$ is *locally affine*, if it is continuously differentiable with a Jacobian $J \neq 0$. Then f is known to be locally invertible, with an inverse function f^{-1} of the same kind.

As before, we consider only the case of local invariance in all directions. For local u-invariance, we need f to preserve the set of lines parallel to u, in which case the same proof applies.

Theorem 5.50 *(invertible transformations) Fix any $p \ge 1$ and a locally affine bijection f on $\mathbb{R}^d$, and let ξ be a random measure on $\mathbb{R}^d$. Then ξ and $\xi \circ f^{-1}$ are simultaneously locally L^p-invariant.*

Proof: Let ξ be locally L^p-invariant. For any compact set C and vector $v = e_1, \dots, e_d$, we need to verify that, as $h \to 0$,

$$h^{-d}\int_C \Big\|\xi \circ f^{-1}(I_h + x) - \xi \circ f^{-1}(I_h + hv + x)\Big\|_p\,dx \to 0.$$

Since the Jacobian of f is locally bounded, it suffices to prove that

$$h^{-d}\int_{C'} \Big\|\xi \circ f^{-1}\{I_h + f(x)\} - \xi \circ f^{-1}\{I_h + hv + f(x)\}\Big\|_p\,dx \to 0,$$

where $C' = f^{-1}C$. For any $x_0 \in C$, let f_0 denote the affine map on $\mathbb{R}^d$, such that the functions f and f_0 agree at x_0, along with their first order partial derivatives. Define

$$\begin{aligned} I'_h &= f_0^{-1}\{I_h + f(x_0)\} - x_0\,, \\ v' &= h^{-1}\Big(f_0^{-1}\{hv + f(x_0)\} - x_0\Big).\end{aligned}$$

Fixing any $\varepsilon > 0$, let $(\partial I_h')^{\varepsilon h}$ denote the εh-neighborhood of the boundary $\partial I_h'$. Since f is locally affine, we have

$$f^{-1}\{I_h + f(x)\} \,\triangle\, (I_h' + x) \subset (\partial I_h')^{\varepsilon h} + x,$$
$$f^{-1}\{I_h + he_i + f(x)\} \,\triangle\, (I_h' + hv' + x) \subset (\partial I_h')^{\varepsilon h} + hv' + x,$$

for x in some neighborhood B of x_0, and for all sufficiently small $h > 0$.

Now choose a finite cover of C' by disjoint sets B_i as above, also satisfying $\xi\partial B_i = 0$ a.s. Writing ν_h^ε for the uniform distribution on the set $-(\partial I_h')^{\varepsilon h}$, we get as $h \to 0$

$$\begin{aligned}
&\Big| h^{-d}\int_{C'} \Big\| \xi\circ f^{-1}\{I_h + f(x)\} - \xi\circ f^{-1}\{I_h + hv + f(x)\}\Big\|_p\, dx \\
&\qquad - h^{-d}\sum\nolimits_i \int_{B_i} \Big\| \xi(I_h' + x) - \xi(I_h' + hv' + x)\Big\|_p\, dx\Big| \\
&\le\ h^{-d}\sum\nolimits_i \int_{B_i}\Big(\Big\|\xi\Big\{(\partial I_h')^{\varepsilon h} + x\Big\}\Big\|_p + \Big\|\xi\Big\{(\partial I_h')^{\varepsilon h} + hv' + x\Big\}\Big\|_p\Big)\, dx \\
&\le\ h^{-d}\sum\nolimits_i \int_{B_i}\Big(\|\xi\|_p\Big\{(\partial I_h')^{\varepsilon h} + x\Big\} + \|\xi\|_p\Big\{(\partial I_h')^{\varepsilon h} + hv' + x\Big\}\Big)\, dx \\
&=\ h^{-d}\sum\nolimits_i \lambda^d(\partial I_h')^{\varepsilon h}\,\Big(\|\xi\|_p * \nu_h^\varepsilon + \|\xi\|_p * \nu_h^\varepsilon * \delta_{-hv'}\Big) B_i \\
&\lesssim\ \varepsilon \sum\nolimits_i \|\xi\|_p\, B_i = \varepsilon\|\xi\|_p\, C',
\end{aligned}$$

by Minkowski's inequality and the fact that $\|\xi\|_p * \nu_h^\varepsilon \xrightarrow{v} \|\xi\|_p$. Since ξ is locally L^p-invariant, the second term on the left tends to 0 as $h \to 0$, and so the first term is ultimately bounded by a constant times $\varepsilon\|\xi\|_p\, C'$. Since $\varepsilon > 0$ was arbitrary, this completes the proof. □

5.7 Ballot Theorems and Sojourn Laws

The following extended version of the classical ballot theorem is a striking maximum identity for stationary, a.s. singular random measures ξ on $I = \mathsf{R}_+$ or $[0,1)$. When $I = [0,1)$, we define stationarity in terms of addition modulo 1, and put $\bar\xi = \xi[0,1)$. Recall that ξ is said to be *singular*, if its absolutely continuous component vanishes. This holds in particular for any purely atomic random measure.

Theorem 5.51 *(continuous-time ballot theorem) For any stationary, a.s. singular random measure ξ on R_+ or $[0,1)$, there exists a $U(0,1)$ random variable $\sigma \perp\!\!\!\perp \mathcal{I}_\xi$, such that*

$$\sigma \sup_{t>0} t^{-1}\xi[0,t] = \bar\xi \ \ a.s.$$

Proof: When ξ is stationary on $[0,1)$, the periodic continuation $\eta = \sum_{n\ge 0}\theta_n\xi$ is stationary on R_+, and moreover $\mathcal{I}_\eta = \mathcal{I}_\xi$ and $\bar\eta = \bar\xi$. Using the elementary inequality

$$\frac{x_1 + \cdots + x_n}{t_1 + \cdots + t_n} \le \max_{k \le n} \frac{x_k}{t_k}, \quad n \in \mathsf{N},$$

valid for any $x_1, x_2, \ldots \ge 0$ and $t_1, t_2, \ldots > 0$, we get $\sup_t t^{-1}\eta[0,t] = \sup_t t^{-1}\xi[0,t]$. It is then enough to consider random measures on R_+.

In that case, let $X_t = \xi(0,t]$, and define for $t \ge 0$

$$A_t = \inf_{s \ge t}(s - X_s), \qquad \alpha_t = 1\{A_t = t - X_t\}. \tag{34}$$

Since $A_t \le t - X_t$ and X is non-decreasing, we get for any $s < t$

$$\begin{aligned} A_s &= \inf_{r \in [s,t)} (r - X_r) \wedge A_t \\ &\ge (s - X_t) \wedge A_t \\ &\ge (s - t + A_t) \wedge A_t \\ &= s - t + A_t. \end{aligned}$$

If A_0 is finite, then so is A_t for every t, and so by subtraction

$$0 \le A_t - A_s \le t - s \;\text{ on }\; \{A_0 > -\infty\}, \quad s < t. \tag{35}$$

Thus, A is non-decreasing and absolutely continuous on $\{A_0 > -\infty\}$.

Now fix a singular path of X such that A_0 is finite, and let $t \ge 0$ with $A_t < t - X_t$. Then $A_t + X_{t\pm} < t$ by monotonicity. Hence, by the left and right continuity of A and X, we may choose some $\varepsilon > 0$ with

$$A_s + X_s < s - 2\varepsilon, \quad |s - t| < \varepsilon.$$

Then (35) yields

$$s - X_s > A_s + 2\varepsilon > A_t + \varepsilon, \quad |s - t| < \varepsilon,$$

and so by (34) we get $A_s = A_t$ whenever $|s - t| < \varepsilon$. In particular, A has derivative $A'_t = 0 = \alpha_t$ at t.

We turn to the complementary set $D = \{t \ge 0;\, A_t = t - X_t\}$. By Lebesgue's differentiation theorem, both A and X are differentiable a.s., the latter with derivative 0, and we may form a set D' by excluding the corresponding null sets, as well as the countably many isolated points of D. For any $t \in D'$, we may choose some $t_n \to t$ in $D \setminus \{t\}$. By the definition of D,

$$\frac{A_{t_n} - A_t}{t_n - t} = 1 - \frac{X_{t_n} - X_t}{t_n - t}, \quad n \in \mathsf{N},$$

and as $n \to \infty$, we get $A'_t = 1 = \alpha_t$. Combining with the result in the previous case gives $A' = \alpha$ a.e., and since A is absolutely continuous, the differentiation theorem yields

$$A_t - A_0 = \int_0^t \alpha_s\, ds \;\text{ on }\; \{A_0 > -\infty\}, \quad t \ge 0. \tag{36}$$

The ergodic theorem gives $X_t/t \to \bar\xi$ a.s. as $t \to \infty$. When $\bar\xi < 1$, we see from (34) that $-\infty < A_t/t \to 1 - \bar\xi$ a.s. Also

$$\begin{aligned} A_t + X_t - t &= \inf_{s \geq t} \Big\{ (s-t) - (X_s - X_t) \Big\} \\ &= \inf_{s \geq 0} \Big\{ s - \theta_t \xi(0,s] \Big\}, \end{aligned}$$

and hence

$$\alpha_t = 1\Big\{ \inf_{s \geq 0} (s - \theta_t \xi(0,s]) = 0 \Big\}, \quad t \geq 0.$$

Dividing (36) by t and using the ergodic theorem again, we obtain a.s. on $\{\bar\xi < 1\}$

$$\begin{aligned} P\Big\{ \sup_{t>0} (X_t/t) \leq 1 \Big| \mathcal{I}_\xi \Big\} &= P\Big\{ \sup_{t>0} (X_t - t) = 0 \Big| \mathcal{I}_\xi \Big\} \\ &= P(\, A_0 = 0 \,|\, \mathcal{I}_\xi \,) \\ &= E(\alpha_0 \,|\, \mathcal{I}_\xi) = 1 - \bar\xi. \end{aligned}$$

Replacing ξ by $r\xi$ and taking complements, we get more generally

$$P\Big\{ r \sup_{t>0} (X_t/t) > 1 \Big| \mathcal{I}_\xi \Big\} = r\bar\xi \wedge 1 \;\text{ a.s.}, \quad r \geq 0.$$

Since the same relation holds trivially when $\bar\xi = 0$ or ∞, we see that σ is conditionally $U(0,1)$ given $\mathcal{I}_\xi$, which means that σ is $U(0,1)$ and independent of $\mathcal{I}_\xi$. □

The last theorem implies a similar discrete-time result. Here the basic relation holds only with inequality, unless the random variables are Z_+-valued. For an infinite, stationary sequence $\xi = (\xi_1, \xi_2, \ldots)$ in R_+ with invariant σ-field $\mathcal{I}_\xi$, we define $\bar\xi = E(\xi_1|\mathcal{I}_\xi)$ a.s. For finite sequences $\xi = (\xi_1, \ldots, \xi_n)$, stationarity is defined in terms of addition modulo n, and we put $\bar\xi = n^{-1} \sum_k \xi_k$.

Corollary 5.52 *(discrete-time ballot theorem) For any finite or infinite, stationary sequence of random variables $\xi_1, \xi_2, \ldots \geq 0$ with partial sums $S_k = \sum_{j \leq k} \xi_j$, there exists a $U(0,1)$ random variable $\sigma \perp\!\!\!\perp \mathcal{I}_\xi$, such that*

$$\sigma \sup_{k>0} (S_k/k) \leq \bar\xi \;\; \textit{a.s.}$$

If the ξ_k are Z_+-valued, then also

$$P\Big\{ \sup_{k>0} (S_k - k) \geq 0 \Big| \mathcal{I}_\xi \Big\} = \bar\xi \wedge 1 \;\; \textit{a.s.}$$

Proof: Arguing by periodic continuation, as before, we may reduce to the case of infinite sequences ξ. Now let $\vartheta \perp\!\!\!\perp \xi$ be $U(0,1)$, and define $X_t = S_{[t+\vartheta]}$ for $t \geq 0$. Then X has stationary increments, and we note that $\mathcal{I}_X = \mathcal{I}_\xi$ and $\bar X = \bar\xi$. By Theorem 5.51, there exists a $U(0,1)$ random variable $\sigma \perp\!\!\!\perp \mathcal{I}_X$, such that a.s.

$$\begin{aligned}\sup_{k>0}(S_k/k) &= \sup_{t>0}(S_{[t]}/t)\\ &\le \sup_{t>0}(X_t/t) = \bar\xi/\sigma.\end{aligned}$$

When the ξ_k are Z_+-valued, the same result yields a.s.

$$\begin{aligned}P\Big\{\sup_{k>0}(S_k-k)\ge 0\Big|\,\mathcal{I}_\xi\Big\} &= P\Big\{\sup_{t\ge 0}(X_t-t)>0\Big|\,\mathcal{I}_\xi\Big\}\\ &= P\Big\{\sup_{t>0}(X_t/t)>1\Big|\,\mathcal{I}_\xi\Big\}\\ &= P(\,\bar\xi>\sigma\,|\,\mathcal{I}_\xi) = \bar\xi\wedge 1.\end{aligned}$$ □

We conclude with some general sojourn laws for stationary processes X on R_+ or $[0,1)$. The subject is related to random measures, via the notions of occupation time and occupation measure. Assuming X to be product-measurable, we define the *sojourn times* of X, above or below the initial value X_0, by

$$\Lambda_t^\pm = \lambda\{s\in[0,t);\ \pm(X_s-X_0)>0\},\quad t\in\mathsf{R}_+ \text{ or } [0,1].$$

In discrete time, we use the same definition, but with Lebesgue measure λ replaced by counting measure on Z_+ or $\mathbb{Z}_n$.

Theorem 5.53 *(sojourn times) Let X be a real-valued, stationary, and measurable process on R_+ or Z_+. Then*

(i) *there exist some random variables $\rho_\pm$ in $[0,1]$, such that*

$$\lim_{t\to\infty} t^{-1}\Lambda_t^\pm = \rho_\pm \quad a.s.,$$

(ii) *if $\vartheta\perp\!\!\!\perp X$ is $U(0,1)$, then so is the random variable*

$$\sigma = \rho_- + \vartheta(1-\rho_+-\rho_-).$$

Here (ii) *remains true for any cyclically stationary process on $[0,1)$ or Z_n, provided we define $\rho_\pm = \Lambda_1^\pm$ or $\Lambda_n^\pm/n$, respectively.*

Our proof is based on the following lemma.

Lemma 5.54 *(uniform sampling) Let the random variable τ and probability measure η on R be such that $\mathcal{L}(\tau|\,\eta)=\eta$ a.s., and let $\vartheta\perp\!\!\!\perp(\tau,\eta)$ be $U(0,1)$. Then the random variable $\sigma=\eta(-\infty,\tau)+\vartheta\,\eta\{\tau\}$ is again $U(0,1)$.*

Proof: First reduce by conditioning to the case where η is non-random with $\mathcal{L}(\tau)=\eta$. Introduce the distribution function $F(t)=\eta(\infty,t]$ with right-continuous inverse F^{-1}, and let $\gamma\perp\!\!\!\perp\vartheta$ be $U(0,1)$. Then $\tau'=F^{-1}(\gamma)$ is independent of ϑ with distribution η, and so we may assume for convenience that $\tau'=\tau$. Next define $\gamma_\pm=F(\tau\pm)$, and note that $\sigma=\gamma_-+\vartheta(\gamma_+-\gamma_-)$. Write A for the closed range of F, and let $I_k=(s_k,t_k)$ denote the connected components of the open set $(0,1)\setminus A$. Then for any k we have a.s. $\gamma_-=s_k$ iff $\gamma\in I_k$, in which case also $\gamma_+=t_k$. Thus,

$$\begin{aligned}P(\gamma\le t\,|\,\gamma_-=s_k)&=\frac{t-s_k}{t_k-s_k}\\&=P(\sigma\le t\,|\,\gamma_-=s_k),\quad t\in I_k.\end{aligned}$$

If instead $\gamma_-\notin\{s_1,s_2,\dots\}$, then a.s. $\gamma\notin\bigcup_k[s_k,t_k]$, which implies $\sigma=\gamma$. Hence, by combination, $\mathcal{L}(\sigma\,|\,\gamma_-)=\mathcal{L}(\gamma\,|\,\gamma_-)$ a.s., and so $\sigma\overset{d}{=}\gamma$. □

Proof of Theorem 5.53: It is enough to consider processes on R_+ or Z_+, since the results for $[0,1)$ or Z_n will then follow by periodic continuation. Define the mean occupation measure η and empirical distributions η_t, for any $B\in\mathcal{B}$, by

$$\begin{aligned}\eta B&=P(X_0\in B\,|\,\mathcal{I}_X),\\\eta_t B&=t^{-1}\lambda\{s<t;\,X_s\in B\},\quad t>0.\end{aligned}$$

Then the Glivenko–Cantelli theorem (FMP 4.24) yields a.s.

$$\begin{aligned}\lim_{t\to\infty}\sup\nolimits_x\Big|\eta_t(-\infty,x]-\eta(-\infty,x]\Big|&=0,\\\lim_{t\to\infty}\sup\nolimits_x\Big|\eta_t(-\infty,x)-\eta(-\infty,x)\Big|&=0,\end{aligned}$$

and so as $t\to\infty$,

$$\begin{aligned}t^{-1}\Lambda_t^+&=\eta_t(X_0,\infty)\to\eta(X_0,\infty),\\t^{-1}\Lambda_t^-&=\eta_t(-\infty,X_0)\to\eta(-\infty,X_0),\end{aligned}$$

which proves (i) with

$$\rho_+=\eta(X_0,\infty),\qquad\rho_-=\eta(-\infty,X_0).$$

(ii) The definition of η gives, a.s. for fixed $B\in\mathcal{B}$,

$$\begin{aligned}P(X_0\in B\,|\,\eta)&=E\Big(P\{X_0\in B\,|\,\mathcal{I}_X\}\Big|\,\eta\Big)\\&=E(\eta B\,|\,\eta)=\eta B,\end{aligned}$$

which shows that η is a regular version of $\mathcal{L}(X_0\,|\,\eta)$. Since $\eta\{X_0\}=1-\rho_+-\rho_-$, the variable σ is again $U(0,1)$ by Lemma 5.54. □

Since $\rho_-\le\sigma\le1-\rho_+$ in Theorem 5.53, we conclude that ρ_+ and ρ_- are both $U(0,1)$, whenever

$$\lim_{t\to\infty}t^{-1}\lambda\Big\{s\in[0,t);\,X_s=X_0\Big\}=0\ \text{ a.s.},$$

or, for processes on $[0,1)$, when $\lambda\{s;\, X_s = X_0\} = 0$ a.s. We state some less obvious criteria for the variables $\rho_\pm$ to be $U(0,1)$. The result translates immediately into similar criteria for processes on $[0,1]$ with cyclically stationary *increments*. Here the *mean occupation measure* η of X is defined by

$$\begin{aligned}\eta B &= \lim_{t\to\infty} t^{-1}\lambda\{s \le t;\, X_s \in B\}\\ &= P(X_0 \in B \,|\, \mathcal{I}_X), \qquad\qquad B \in \mathcal{B},\end{aligned}$$

where the a.s. limit exists by the ergodic theorem.

Corollary 5.55 *(uniform law) Let X be a stationary and measurable process on R_+, Z_+, $[0,1)$, or Z_n with mean occupation measure η, and define $\rho_\pm$ as in Theorem 5.53. Then these statements are equivalent:*
(i) *ρ_+ (or ρ_-) is $U(0,1)$,*
(ii) *$\rho_+ + \rho_- = 1$ a.s.,*
(iii) *η is a.s. diffuse.*

Proof: Let ϑ and σ be such as in Theorem 5.53. Condition (ii) implies $\rho_- = \sigma$ a.s., and (i) follows. Conversely, (i) yields

$$\begin{aligned}\tfrac{1}{2} = E\sigma &= E\rho_- + E\vartheta(1 - \rho_+ - \rho_-)\\ &= \tfrac{1}{2} + \tfrac{1}{2}\, E(1 - \rho_+ - \rho_-),\end{aligned}$$

which implies (ii). Next, recall that $1 - \rho_+ - \rho_- = \eta\{X_0\}$ and $\mathcal{L}(X_0 \,|\, \eta) = \eta$ a.s., and conclude from the disintegration theorem that

$$\begin{aligned}E(1 - \rho_+ - \rho_-) &= E\eta\{X_0\} = E\,E(\,\eta\{X_0\} \,|\, \eta\,)\\ &= E\int \eta(dx)\,\eta\{x\}\\ &= E\textstyle\sum_x (\eta\{x\})^2,\end{aligned}$$

which shows that (ii) and (iii) are equivalent. □

The last theorem leads easily to a uniform law for certain processes on $[0,1]$ with stationary *increments*.

Corollary 5.56 *(stationary increments) Let X be a measurable process on $[0,1)$, with cyclically stationary increments and $X_0 = X_1 = 0$ a.s., and assume that $\lambda\{s \in [0,1);\, X_s = 0\} = 0$ a.s. Then the variables $\lambda\{s \in [0,1);\, \pm X_s > 0\}$ are $U(0,1)$.*

Proof: Apply Theorem 5.53 to the stationary process $Y_t = X_{\tau+t}$, where $\tau \perp\!\!\!\perp X$ is $U(0,1)$, and note that $\rho_\pm \overset{d}{=} \lambda\{s;\, \pm X_s > 0\}$. □

We finally note the close relationship between the uniform laws for processes on $[0,1]$ and the deepest arcsine law for Brownian motion. Here the *arcsine distribution* may be defined as the distribution of $\sin^2\alpha$ when α is $U(0,2\pi)$. By an easy calculation, its Lebesgue density equals

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}, \quad x \in (0,1). \tag{37}$$

Corollary 5.57 *(third arcsine law, Lévy) For a Brownian motion B in R, the following random variables are arcsine distributed:*

$$\rho_\pm = \lambda\Big\{t \in [0,1];\ \pm B_t > 0\Big\}.$$

We show how this follows from the uniform law in Corollary 5.56, along with the more elementary arcsine law for the last zero on $[0,1]$.

Proof: Put $\tau = \sup\{t \le 1;\ B_t = 0\}$, and note that the restrictions of B to $[0,\tau]$ and $[\tau,1]$ are conditionally independent, given τ. Furthermore, B has conditionally stationary increments on $[0,\tau]$, and $\lambda\{t < \tau;\ B_t = 0\} = 0$ a.s. Writing $A = \{B_1 > 0\}$, $\gamma_\pm = \lambda\{t < \tau;\ \pm B_t > 0\}$, and $\sigma_\pm = \gamma_\pm/\tau$, we get

$$\begin{aligned} \rho_+ &= \gamma_+ + 1_A\,(1-\tau) \\ &= 1_{A^c}\,\tau\,\sigma_+ + 1_A\,(1-\tau\,\sigma_-). \end{aligned}$$

Now τ is arcsine distributed (FMP 13.16), and so its probability density is given by (37). Since A, τ, and the $\sigma_\pm$ are independent, and the latter are $U(0,1)$ by Corollary 5.56, the $\rho_\pm$ have density

$$g(x) = \frac{1}{2\pi}\int_x^1 \frac{du}{u\sqrt{u(1-u)}} + \frac{1}{2\pi}\int_{1-x}^1 \frac{du}{u\sqrt{u(1-u)}}, \quad x \in (0,1).$$

An easy calculation gives

$$g'(x) = \frac{2x-1}{\pi\sqrt{x(1-x)}} = f'(x), \quad x \in (0,1),$$

which shows that f and g differ by a constant. Since both functions are probability densities, they must then agree, which shows that even the $\rho_\pm$ are arcsine distributed. $\square$

Chapter 6

Palm and Related Kernels

Palm measures play a fundamental role for general random measures, regardless of any notion of stationarity or invariance. In general, we have a whole family of Palm measures, one for each point in the underlying state space S. We may also consider the more general case of a random pair (ξ, η), where ξ is a random measure on S, and η is a random element in a possibly different Borel space T. As an important special case, we may take $T = \mathcal{M}_S$ and choose $\eta = \xi$.

The Palm measures of η with respect to ξ, here denoted by $\mathcal{L}(\eta \,\|\, \xi)_s$, extend the notion of regular conditional distributions, and the two notions agree when $\xi = \delta_\sigma$ for some random element σ in S. Just as the conditional distributions $\mathcal{L}(\eta \,|\, \sigma)_s$ are obtained by disintegration of the joint distribution $\mathcal{L}(\sigma, \eta)$ on the product space $S \times T$, we may form the general Palm measures by disintegration[1] of the *Campbell measure* $C_{\xi,\eta}$ on $S \times T$, defined by

$$C_{\xi,\eta} f = E \int \xi(ds)\, f(s, \eta), \quad f \geq 0, \tag{1}$$

where the function f on $S \times T$ is understood to be measurable. Note in particular that $C_{\xi,\eta}(\cdot \times T) = E\xi$.

When $E\xi$ is σ-finite, so is $C_{\xi,\eta}$, and there exists a probability kernel $\mathcal{L}(\eta \,\|\, \xi)$ from S to T satisfying the disintegration $C_{\xi,\eta} = E\xi \otimes \mathcal{L}(\eta \,\|\, \xi)$, or in explicit notation

$$C_{\xi,\eta} f = \int E\xi(ds)\, E\{ f(s, \eta) \| \, \xi \,\}_s, \quad f \geq 0.$$

Here $\mathcal{L}(\eta \,\|\, \xi)$ is known as the *Palm kernel* of η with respect to ξ, and the measures $\mathcal{L}(\eta \,\|\, \xi)_s$ are called *Palm distributions* of η with respect to ξ. If the intensity measure $E\xi$ fails to be σ-finite, it may be replaced by a σ-finite *supporting measure* $\nu \sim E\xi$. Assuming $C_{\xi,\eta}$ to remain σ-finite, which holds under weak regularity conditions, we now have a disintegration $C_{\xi,\eta} = \nu \otimes \mathcal{L}(\xi \,\|\, \eta)$, where the *Palm measures* $\mathcal{L}(\eta \,\|\, \xi)_s$ are σ-finite on T. The latter are then unique a.e. $E\xi$, up to normalizations depending on the choice of ν.

[1] From this point on, we will rely heavily on the theory of kernels with associated compositions and disintegrations, as treated in Sections 1.3–4. In particular, we will use the notation $(\mu \otimes \nu) f = \int \mu(ds) \int \nu_s(dt) f(s,t)$ without further comment.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_6

However, they are bounded only when $E\xi$ is σ-finite, in which case they can be normalized to Palm *distributions*.

For simple point processes ξ 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 a point at s. The interpretation is justified by some local approximations in Theorem 6.32 and subsequent results, which show that if $B \downarrow \{s\}$ in a suitable sense, then for $s \in S$ a.e. $E\xi$,

$$P\{\xi B > 0\} \sim P\{\xi B = 1\} \sim E\xi B,$$

$$\mathcal{L}(\eta \,|\, \xi B > 0) \approx \mathcal{L}(\eta \,|\, \xi B = 1) \approx \mathcal{L}(\eta \,\|\, \xi)_s.$$

Here we write $a \sim b$ for $a/b \to 1$, and $a \approx b$ for $a - b \to 0$. Note the similarity with the approximations in Theorem 5.5, for stationary point processes on R^d.

For a more direct conditioning interpretation of Palm distributions, we may introduce, on a suitable pseudo-probability space $(\tilde{\Omega}, \tilde{P})$, a random element σ in S with $\tilde{\mathcal{L}}(\sigma |\, \xi, \eta) = \xi$ a.s. In other words, we choose σ by sampling from S, according to the random pseudo-distribution ξ. Assuming $E\xi$ to be σ-finite, we have

$$\mathcal{L}(\eta \,\|\, \xi)_s = \tilde{\mathcal{L}}(\eta \,|\, \sigma)_s \;\text{ a.e. } E\xi,$$

which extends the elementary result for $\xi = \delta_\sigma$, mentioned earlier.

If ξ is a random measure on S, then $\xi^{\otimes n}$ is a random measure on S^n, and we may form the *n-th order Palm measures* $\mathcal{L}(\eta \,\|\, \xi^{\otimes n})_s$ of η with respect to ξ, for any $s = (s_1, \ldots, s_n) \in S^n$. For simple point processes ξ, they may be thought of as conditional distributions of η, given that ξ has atoms at $s_1, \ldots, s_n$. Specializing to the case where $\eta = \xi$, we see that, under the distribution $\mathcal{L}(\xi \,\|\, \xi^{\otimes n})_s$ for $s \in S^n$ a.e. $E\xi^{\otimes n}$, all the points $s_1, \ldots, s_n$ belong a.s. to the support of ξ. This suggests that we consider instead the *reduced Palm measures*

$$\mathcal{L}\Big\{\xi - \textstyle\sum_{k\le n}\delta_{\sigma_k}\Big\|\, \xi^{(n)}\Big\}_s, \quad s = (s_1, \ldots, s_n) \in S^{(n)},$$

where $\xi^{(n)}$ denotes the restriction of $\xi^{\otimes n}$ to the non-diagonal part $S^{(n)}$ of S^n. Those measures also arise directly, by disintegration of the *reduced Campbell measure* $C_\xi^{(n)}$ on $S^{(n)} \times \mathcal{N}_S$, given for measurable $f \ge 0$ by

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

Since $C_\xi^{(n)}$ is clearly symmetric in the coordinates $s_1, \ldots, s_n$, the same thing holds a.e. for the associated Palm measures, which may then be regarded as functions of the associated counting measures $\mu = \sum_{k\le n} \delta_{s_k}$. Combining the versions for different n, we may define the *compound Campbell measure* C_ξ by

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

where $C_\xi^{(0)} = \mathcal{L}(\xi)$, and the last sum extends over all measures $\mu \le \xi$ in $\mathcal{N}_S$. Here C_ξ is automatically σ-finite, and yields the reduced Palm measures of arbitrary order through suitable disintegrations. It also plays an important role for the theory of external conditioning in Chapter 8.

For Poisson processes ξ on S, the reduced Palm distributions agree a.e. with the original distributions. This property characterizes the class of Poisson processes, so that ξ is Poisson iff

$$\mathcal{L}(\xi \,\|\, \xi)_s = \mathcal{L}(\xi + \delta_s), \quad s \in S \text{ a.e. } E\xi.$$

Writing this as $\mathcal{L}(\xi \,\|\, \xi)_s = \mathcal{L}(\xi) * \rho_s$, where ρ_s denotes a unit mass at δ_s, we see that $\mathcal{L}(\xi)$ is a.e. a convolution factor of $\mathcal{L}(\xi \,\|\, \xi)_s$, here expressed as $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)_s$. By Theorem 6.17, the latter property characterizes the class of infinitely divisible random measures. In fact, under a weak regularity condition, it is enough to assume that $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi f)$, for a single measurable function $f > 0$ with $\xi f < \infty$ a.s.

The elementary conditional distributions $P_B = P(\cdot\,|B)$ satisfy $P_{B \cap C} = (P_B)_C = (P_C)_B$, which suggests that the general conditional distributions $P_\mathcal{F} = P(\cdot\,|\mathcal{F})$ would obey the *chain rules*

$$P_{\mathcal{F} \vee \mathcal{G}} = (P_\mathcal{F})_\mathcal{G} = (P_\mathcal{G})_\mathcal{F}, \qquad P_\mathcal{F} = E_\mathcal{F}(P_\mathcal{F})_\mathcal{G}.$$

In Theorem 6.21, we prove that this holds a.s. on $\mathcal{H}$, under suitable regularity conditions on the σ-fields $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{H}$. Those rather subtle formulas should not be confused with the elementary relations $E_\mathcal{F} E_\mathcal{G} = E_\mathcal{G} E_\mathcal{F} = E_{\mathcal{F} \cap \mathcal{G}}$, valid when $\mathcal{F} \perp\!\!\!\perp_{\mathcal{F} \cap \mathcal{G}} \mathcal{G}$, and $P_\mathcal{F} = E_\mathcal{F} P_{\mathcal{F} \vee \mathcal{G}}$ for general $\mathcal{F}$ and $\mathcal{G}$. The conditioning interpretation of Palm measures suggests the more general relations

$$P_\mathcal{F}(\cdot \,\|\, \xi) = \{P(\cdot \,\|\, \xi)\}_\mathcal{F}, \qquad P(\cdot \,\|\, \xi)_s = E\{P_\mathcal{F}(\cdot \,\|\, \xi)_s \| \, \xi\}_s,$$
$$P(\cdot \,\|\, \xi \,\|\, \eta)_{s,t} = P(\cdot \,\|\, \eta \,\|\, \xi)_{t,s} = P(\cdot \,\|\, \xi \otimes \eta)_{t,s},$$

established in Theorems 6.22 and 6.23, under suitable regularity conditions on ξ, η, and $\mathcal{F}$.

Using the indicated iteration approach, we may easily derive the Palm distributions of randomizations and Cox processes. For a typical example, let ξ be a Cox process directed by η, and choose ξ_n and η_n with distributions $\mathcal{L}(\xi' \,\|\, \xi^{(n)})$ and $\mathcal{L}(\eta \,\|\, \eta^{\otimes \eta})$, respectively, where the former denotes the n-th order reduced Palm distribution of ξ. Then ξ_n is distributed as a Cox process directed by η_n. Here the Palm distributions are indexed by vectors $s \in S^n$, and the mentioned property holds a.e. with respect to $E\eta^{\otimes n}$. Results of this type play important roles in the context of super-processes, as will be seen in Chapter 13.

We already indicated how the Palm distributions of a simple point process can be approximated by elementary conditional distributions. In fact, we can say much more. Under suitable regularity conditions, we prove in Theorem 6.36 that, as $B \downarrow \{s\}$ with $s \in S^{(n)}$ a.e. $E\xi^{(n)}$, the restrictions of ξ to B

and B^c are asymptotically conditionally independent, given $\xi B > 0$. Indeed, assuming $E\xi B \in (0, \infty)$, and letting τ_B be a random element in B with distribution $\mathcal{L}(\tau_B) = E(1_B \xi)/E\xi B$, we show that a.e.

$$\mathcal{L}(1_B \xi, 1_{B^c} \xi \,|\, \xi B > 0) \approx \mathcal{L}(\delta_{\tau_B}) \otimes \mathcal{L}(\xi - \delta_s \,\|\, \xi)_s,$$

and similarly for the hitting of several disjoint targets $B_1, \ldots, B_n$ around the points $s_1, \ldots, s_n \in S$. This illustrates the phenomenon of *decoupling*, recurring frequently in subsequent chapters.

Any σ-finite measure ρ on a product space $S \times T$ admits the *dual* disintegrations $\rho = \nu \otimes \mu \cong \nu' \otimes \mu'$, or explicitly

$$\begin{aligned} \rho f &= \int \nu(ds) \int f(s,t)\, \mu_s(dt) \\ &= \int \nu'(dt) \int f(s,t)\, \mu'_t(ds), \end{aligned}$$

for some σ-finite measures ν on S and ν' on T, along with associated kernels $\mu \colon S \to T$ and $\mu' \colon T \to S$. We can often use one of those disintegrations to draw conclusions about the other. This applies in particular to the Campbell measure $C_{\xi,\eta}$ of a random element η, with respect to a random measure ξ, where the dual disintegrations become

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

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 can choose a common density function $f \colon S \times T \to \mathsf{R}_+$, in the sense that

$$E(\xi \,|\, \eta) = f(\cdot, \eta) \cdot \nu \;\text{ a.s.} \quad \Leftrightarrow \quad \mathcal{L}(\eta \,\|\, \xi)_s = f(s, \cdot) \cdot \mathcal{L}(\eta) \;\text{ a.e. } \nu.$$

We now specialize to the case where η is the identity map on Ω, with associated filtration $\mathcal{F} = (\mathcal{F}_t)$ on R_+. Here the corresponding densities form an adapted process $M_t^s \geq 0$ on $S \times \mathsf{R}_+$, chosen to be $(\mathcal{S} \otimes \mathcal{F}_t)$-measurable for every $t \geq 0$. We further assume $E\xi$ to be σ-finite. Then for every $t > 0$, the previous equivalence becomes

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

Fixing the process M, we can use the latter relation to define a version of the kernel $P(\mathcal{F}_t \| \, \xi)_s$ on $S \times \mathsf{R}_+$. Since the $\mathcal{F}_t$ are non-decreasing in t, we may hope for *consistency*, in the sense that $P(\mathcal{F}_t \| \, \xi) = P(\mathcal{F}_u \| \, \xi)$ on $\mathcal{F}_t$, for all $t \leq u$. Assuming S to be Polish, and imposing a weak regularity condition on the filtration $\mathcal{F}$, we prove in Theorem 6.42 that

(a) for fixed $t \geq 0$, the process $P(\mathcal{F}_t \| \, \xi)_s$ is continuous in total variation in $s \in S$, iff M_t^s is L^1-continuous in s,

(b) for fixed $s \in S$, the process $P(\mathcal{F}_t \| \, \xi)_s$ is consistent in $t \geq 0$, iff M_t^s is a martingale in t,

(c) continuity as in (a) for all $t \ge 0$ implies consistency as in (b) for all $s \in \operatorname{supp} E\xi$.

Those results will be especially useful in Chapters 12 and 13, where they are needed to establish some deep hitting and conditioning properties of regenerative sets and super-processes. The duality theory will also play important roles in Chapters 7 and 8.

6.1 Campbell Measures and Palm Kernels

For any random measure ξ on S, we may introduce[2] the associated *Campbell measure* $C_\xi = P \otimes \xi$ on $\Omega \times S$. Since $C_\xi(\Omega \times \cdot) = E\xi$, we may regard C_ξ as a refinement of the intensity measure $E\xi$. Note that C_ξ is automatically σ-finite, and that $\xi Y < \infty$ a.s., for some product-measurable process $Y > 0$ on S. Unfortunately, the associated Palm kernel may not exist in this case, unless special conditions are imposed on Ω. However, for most purposes we may choose Ω to be the sample space $\mathcal{M}_S$ of ξ with probability measure $\mathcal{L}(\xi)$, in which case the desired disintegration does exist, since $\mathcal{M}_S$ is again Borel by Theorem 1.5.

More generally, we may consider any random element η in a measurable space T, and define the Campbell measure $C_{\xi,\eta}$ of the pair (ξ, η) as in (1). When T is Borel and $C_{\xi,\eta}$ is σ-finite, Lemma 1.23 yields a disintegration $C_{\xi,\eta} = \nu \otimes \mu$, in terms of a σ-finite *supporting measure* $\nu \sim E\xi$ of ξ, and a σ-finite kernel μ of *Palm measures* μ_s on T. When $E\xi$ is σ-finite, and only then, can we take $\nu = E\xi$, and choose the μ_s to be probability measures on T, called the *Palm distributions* of η with respect to ξ. For general $E\xi$, the Palm measures are always σ-finite, but they may be unbounded. We shall often use the suggestive notation

$$\mu_s = \mathcal{L}(\eta \,\|\, \xi)_s, \qquad \mu_s B = P(\eta \in B \,\|\, \xi)_s, \qquad \mu_s f = E(f(\eta) \,\|\, \xi)_s.$$

For a simple example, suppose[3] that $\xi = X \cdot \lambda$ for some σ-finite measure λ on S, and a measurable process $X \ge 0$ on S with $EX_s < \infty$ for all s, so that $\xi f = \lambda(fX) = \int fX\, d\lambda$. Then

$$\begin{aligned} C_{\xi,\eta}(B \times A) &= E(\xi B;\, \eta \in A) \\ &= \int_B \lambda(ds)\, E(X_s;\, \eta \in A) \\ &= \int_B E\xi(ds)\, \frac{E(X_s;\, \eta \in A)}{EX_s}, \end{aligned}$$

which shows that $\mathcal{L}(\eta \,\|\, \xi)_s = E(X_s;\, \eta \in \cdot)/EX_s$ a.e.

[2] For kernel composition and disintegration, see the previous remarks and Sections 1.3-4.

[3] Recall that $\mu f = \int f d\mu$, whereas $f \cdot \mu$ denotes the measure with μ-density f.

It is often sufficient to consider the Palm distributions of the random measure ξ with respect to itself, since the Palm distributions of the pair (ξ, η) will then follow by a simple conditioning. In the following formula, we assume the Palm kernels $\mathcal{L}(\xi, \eta \,\|\, \xi)$ and $\mathcal{L}(\xi \,\|\, \xi)$ to be defined with respect to a common supporting measure for ξ.

Lemma 6.1 *(conditioning) For any random measure ξ on S and random element η in T, we have*

$$\mathcal{L}(\xi, \eta \,\|\, \xi) = \mathcal{L}(\xi \,\|\, \xi) \otimes \mathcal{L}(\eta \,|\, \xi) \quad a.e.\ E\xi.$$

Proof: Write C_ξ and $C_{\xi,\eta}$ for the Campbell measures of ξ and (ξ, η), respectively, with respect to a common supporting measure ν on S, let Q be an associated Palm kernel for ξ, and introduce some regular conditional distributions $\mu(\xi, \cdot) = \mathcal{L}(\eta \,|\, \xi)$. Then for any measurable function $f \geq 0$ on $S \times \mathcal{M}_S \times T$, we get by the disintegration theorem (FMP 6.4)

$$\begin{aligned} C_{\xi,\eta} f &= E \int \xi(ds)\, f(s, \xi, \eta) \\ &= E \int \xi(ds)\, E\{f(s, \xi, \eta) \,|\, \xi\} \\ &= E \int \xi(ds) \int \mu(\xi, dt)\, f(s, \xi, t) \\ &= E \iint C_\xi(ds\, dm) \int \mu(m, dt)\, f(s, m, t) \\ &= (C_\xi \otimes \mu) f = (\nu \otimes Q \otimes \mu) f, \end{aligned}$$

which shows that the pair (ξ, η) has Palm kernel $Q \otimes \mu$. □

Multi-variate Palm measures with respect to ξ may be defined as Palm measures $\mathcal{L}(\eta \,\|\, \xi^{\otimes n})_s$ with respect to the product measures $\xi^{\otimes n}$, provided that the associated n-th order Campbell measures $C^n_{\xi,\eta}$ on $S^n \times T$ are σ-finite. We state some elementary properties of Palm measures.

Lemma 6.2 *(a.s. properties)*

(i) *Consider a random measure ξ on S and a random element η in T, such that $C_{\xi,\eta}$ is σ-finite. Then for any $A \in \mathcal{T}$ with $\eta \in A$ a.s.,*

$$P(\eta \notin A \,\|\, \xi)_s = 0, \quad s \in S \ \ a.e.\ E\xi.$$

(ii) *For any point process ξ on S and $n \in \mathsf{N}$,*

$$P\Big(\textstyle\prod_{k \leq n} \xi\{s_k\} = 0 \,\Big\|\, \xi^{\otimes n}\Big)_s = 0, \quad s \in S^n \ \ a.e.\ E\xi^{\otimes n}.$$

In particular, (i) shows that if ξ is a [simple] point process or a diffuse random measure, then $\mathcal{L}(\xi \,\|\, \xi)_s$ is a.e. restricted to the set of [simple] counting or diffuse measures, respectively. In the simple case, the inner equality in part (ii) can then be strengthened to $\neq 1$.

Proof: (i) For any supporting measure ν of ξ, we have

$$\int \nu(ds)\, P(\eta \notin A \,\|\, \xi)_s = E(\xi S;\, \eta \notin A) = 0,$$

and the assertion follows.

(ii) In $S^n \times \mathcal{N}_S$, we consider the set

$$A = \Big\{(s,\mu) \in S^n \times \mathcal{N}_S;\, \textstyle\prod_{i\le n} \mu\{s_i\} = 0\Big\}.$$

Since the mapping $(s,\mu) \mapsto \mu\{s\}$ is measurable by Lemma 1.6, we see that A is measurable in $S^n \times \mathcal{N}_S$. Assuming $\xi = \sum_{i\in I} \delta_{\sigma_i}$, and fixing a supporting measure ν of $\xi^{\otimes n}$, we get

$$\begin{aligned}\int \nu(ds)\, P\Big\{(s,\xi) \in A \,\Big\|\, \xi^{\otimes n}\Big\}_s &= E\int \xi^{\otimes n}(ds)\, 1_A(s,\xi)\\ &= E\sum_{i\in I^n} \delta_{\sigma_{i_1},\dots,\sigma_{i_n}}(ds)\, 1_A(s,\xi)\\ &= E\sum_{i\in I^n} 1_A(\sigma_{i_1},\dots,\sigma_{i_n};\, \xi)\\ &= E\sum_{i\in I^n} 1\Big\{\textstyle\prod_{k\le n} \xi(\sigma_{i_k}) = 0\Big\} = 0,\end{aligned}$$

which implies $P\Big\{(s,\xi) \in A \,\Big\|\, \xi^{\otimes n}\Big\} = 0$ a.e. $E\xi^{\otimes n}$. □

To ensure existence of the various Palm measures, we need to check that the corresponding Campbell measures are σ-finite:

Lemma 6.3 *(σ-finiteness) The Campbell measure $C^n_{\xi,\eta}$ is σ-finite, whenever $E\xi^{\otimes n}$ is σ-finite or ξ is η-measurable.*

The measurability condition on ξ is rather harmless, since if $C^n_{\xi,\eta}$ fails to be σ-finite, we can replace η by the pair $\tilde\eta = (\xi,\eta)$. Then the Palm measures with respect to ξ always exist, though their T-marginals may fail to be σ-finite.

Proof: It is enough to show that $C_{\xi,\eta}$ is σ-finite when ξ is η-measurable. Then fix any partition $B_1, B_2, \ldots \in \hat{\mathcal{S}}$ of S, and let $A_{nk} = \{k-1 \le \xi B_n < k\}$ for all $n,k \in \mathbb{N}$. Since ξ is η-measurable, we have $A_{nk} = \{\eta \in C_{nk}\}$ for some $C_{nk} \in \mathcal{T}$ by FMP 1.13, where the latter sets may be modified to form a partition of T, for every fixed n. The function

$$f = \sum_{n\ge1} 2^{-n} \sum_{k\ge1} k^{-1} 1_{B_n\times C_{nk}}$$

is product-measurable with $f > 0$, and we note that

$$\begin{aligned}C_{\xi,\eta} f &= \sum_{n\ge1} 2^{-n} \sum_{k\ge1} k^{-1} E(\xi B_n;\, \eta \in C_{nk})\\ &\le \sum_{n\ge1} 2^{-n} \sum_{k\ge1} P\{\eta \in C_{nk}\} = \sum_{n\ge1} 2^{-n} = 1,\end{aligned}$$

which proves the required σ-finiteness. □

We turn to some simple cases, where the Palm measures are given by elementary formulas. First suppose that $\xi \ll \mu$ a.s., for some fixed measure $\mu \in \mathcal{M}_S$ with a measurable density X.

Lemma 6.4 *(absolute continuity) Let $\xi = X \cdot \mu$ a.s., where $\mu \in \mathcal{M}_S$, and $X \geq 0$ is a measurable process on S, and let η be a random element in T.*

(i) *If $EX^n < \infty$ a.e. μ, then the Palm distributions $\mathcal{L}(\eta \,\|\, \xi^n)$ exist and are given, a.e. $E\xi^n$, by*

$$\mathcal{L}(\eta \,\|\, \xi^n)_{s_1,\dots,s_n} = \frac{E(X_{s_1} \cdots X_{s_n};\, \eta \in \cdot)}{E(X_{s_1} \cdots X_{s_n})}, \quad s_1, \dots, s_n \in S.$$

(ii) *If $\xi = 1_\Xi \cdot \mu$ for some random set Ξ in S, then a.e.*

$$\mathcal{L}(\eta \,\|\, \xi^n)_{s_1,\dots,s_n} = \mathcal{L}(\eta \mid s_1, \dots, s_n \in \Xi), \quad s_1, \dots, s_n \in S.$$

Proof: (i) Write $X_s^{\otimes n} = X_{s_1} \cdots X_{s_n}$ and $Q_s = \mathcal{L}(\eta \in \cdot \,\|\, \xi^n)_s$ for $s = (s_1, \dots, s_n) \in S^n$. By the definition of Q_s and Fubini's theorem, we have for any $A \in \mathcal{T}$ and $B \in \mathcal{S}^n$

$$\begin{aligned}
\int_B Q_s(A)\, EX_s^{\otimes n}\, \mu^n(ds) &= \int_B Q_s(A)\, E\xi^n(ds) \\
&= E(\,\xi^n B;\, \eta \in A) \\
&= E\Big\{\int_B X_s^{\otimes n}\, \mu^n(ds);\, \eta \in A\Big\} \\
&= \int_B E\Big(X_s^{\otimes n};\, \eta \in A\Big)\, \mu^n(ds).
\end{aligned}$$

Since B was arbitrary, we obtain

$$E(X_s^{\otimes n};\, \eta \in A) = Q_s(A)\, EX_s^{\otimes n}, \quad s \in S^n \text{ a.e. } \mu^n,$$

and the asserted formula follows.

(ii) This follows from (i) with $X = 1_\Xi$. □

Next, we consider Palm measures with respect to the projection $\xi(\cdot \times T)$, where ξ is a random measure on a product space $S \times T$.

Lemma 6.5 *(projection) Let ξ be a random measure on $S \times T$, with projection $\hat{\xi} = \xi(\cdot \times T)$ onto S, such that $E\hat{\xi}$ is σ-finite with $E\xi = E\hat{\xi} \otimes \nu$, for some probability kernel ν from S to T. Then for any random element η in a Borel space U, the Palm kernels $Q = \mathcal{L}(\eta \,\|\, \xi)$ and $\hat{Q} = \mathcal{L}(\eta \,\|\, \hat{\xi})$ are related by*

$$\hat{Q}_s = \int \nu_s(dt)\, Q_{s,t}, \quad s \in S \text{ a.e. } E\hat{\xi}.$$

Proof: For any $B \in \mathcal{S}$ and $C \in \mathcal{U}$, we have

$$\begin{aligned}\int_B \hat{Q}_s(C)\, E\hat{\xi}(ds) &= E(\hat{\xi}B;\, \eta \in C)\\ &= E\Big\{\xi(B \times T);\, \eta \in C\Big\}\\ &= \iint_{B\times T} Q_{s,t}(C)\, E\xi(ds\, dt)\\ &= \int_B E\hat{\xi}(ds) \int_T Q_{s,t}(C)\, \nu_s(dt).\end{aligned}$$

Since B was arbitrary, it follows that

$$\hat{Q}_s(C) = \int_T Q_{s,t}(C)\, \nu_s(dt), \quad s \in \mathcal{S} \text{ a.e. } E\hat{\xi},$$

and since U is Borel, the exceptional null set can be chosen to be independent of C. □

We turn to another connection between Palm measures and regular conditional distributions.

Lemma 6.6 *(Palm measures along paths) Let X be a measurable process from S to T, fix a σ-finite measure μ on S, and define $\xi = \mu \circ \tilde{X}^{-1}$, where $\tilde{X}_s = (s, X_s)$. Then for any random element η in U, these conditions are equivalent:*

(i) $Q_{s,t} = \mathcal{L}(\eta \,\|\, \xi)_{s,t}$, $(s,t) \in S \times T$ *a.e.* $E\xi$,

(ii) $Q_{s,X_s} = \mathcal{L}(\eta \,|\, X_s)$, $s \in S$ *a.e.* μ.

Proof, (ii) ⇒ (i): Assuming (ii) and using Fubini's theorem, conditioning on X_s, and the definition of ξ, we get for any $A \in \mathcal{S}$, $B \in \mathcal{T}$, and $C \in \mathcal{U}$

$$\begin{aligned}E\{\,\xi(A \times B);\, \eta \in C\,\} &= E\Big\{\int_A 1_B(X_s)\, \mu(ds);\, \eta \in C\Big\}\\ &= \int_A P\{X_s \in B,\, \eta \in C\}\, \mu(ds)\\ &= \int_A E\Big\{P(\eta \in C \,|\, X_s);\, X_s \in B\Big\}\, \mu(ds)\\ &= \int_A E\Big\{Q_{s,X_s}(C);\, X_s \in B\Big\}\, \mu(ds)\\ &= E\int_A Q_{s,X_s}(C)\, 1_B(X_s)\, \mu(ds)\\ &= E\iint_{A\times B} Q_{s,X_s}(C)\, \xi(ds\, dt).\end{aligned}$$

This extends by a monotone-class argument to any set in $\mathcal{S} \otimes \mathcal{T}$, and (i) follows.

(i) ⇒ (ii): Assuming (i) and letting $A \in \mathcal{S}$, $B \in \mathcal{T}$, and $C \in \mathcal{U}$, we get by the same calculations

$$\int_A E\Big\{P(\eta \in C|\, X_s); X_s \in B\Big\}\, \mu(ds) = \int_A E\Big\{Q_{s,X_s}(C); X_s \in B\Big\}\, \mu(ds).$$

Since A was arbitrary, it follows that

$$E\Big\{P(\eta \in C|\, X_s); X_s \in B\Big\} = E\Big\{Q_{s,X_s}(C); X_s \in B\Big\}, \quad s \in S \text{ a.e. } \mu,$$

and since T and U are Borel, we can choose the exceptional μ-null set to be independent of B and C. Hence, for any non-exceptional $s \in S$,

$$P(\eta \in C\,|X_s) = Q_{s,X_s}(C) \text{ a.s.}, \quad C \in \mathcal{U}.$$

Here the exceptional P-null set can be chosen to be independent of C, and (ii) follows. □

The following closure property is needed in Chapter 12.

Lemma 6.7 *(closure under conditioning) Given a probability kernel $\nu\colon S \to T$, let A be the set of distributions $\mathcal{L}(\xi,\eta)$ on $S\times T$, such that $\mathcal{L}(\eta\,|\,\xi) = \nu(\xi,\cdot)$ a.s. Assuming S, T, and U to be Borel, we have*

- (i) *A is measurable in $\mathcal{M}_{S\times T}$,*
- (ii) *if $\mathcal{L}(\xi,\eta) \in A$, and ζ is a ξ-measurable random element in U, then $\mathcal{L}(\xi,\eta\,|\,\zeta) \in A$ a.s.,*
- (iii) *if $\mathcal{L}(\xi,\eta) \in A$, and ζ is a bounded, ξ-measurable random measure on U, then $\mathcal{L}(\xi,\eta \,\|\, \zeta) \in A$ a.e.*

Proof: (i) The condition $\mathcal{L}(\xi,\eta) \in A$ is equivalent to

$$Ef(\xi,\eta) = E\int f(\xi,t)\,\nu(\xi,dt), \quad f \geq 0. \tag{2}$$

Letting $\mu = \mathcal{L}(\xi,\eta)$, we may write (2) in the form $\mu = \bar\mu \otimes \nu$, where $\bar\mu = \mu(\cdot \times T)$. Here $\bar\mu$ is clearly a measurable function of μ, and Lemma 1.17 (i) shows that $\bar\mu \otimes \nu$ depends measurably on $\bar\mu$. Hence, $\bar\mu \otimes \nu$ is a measurable function of μ, which ensures the measurability of the sets

$$A_B = \Big\{\mu \in \hat{\mathcal{M}}_{S\times T};\ (\bar\mu \otimes \nu)B = \mu B\Big\}, \quad B \in \mathcal{S} \otimes \mathcal{T}.$$

Since $S \times T$ is again Borel, the set $A = \{\mu;\ \bar\mu \otimes \nu = \mu\}$ is the intersection of countably many sets A_B, and the assertion follows.

(ii) Replacing f in (2) by the function $f(s,t)\,1_B(s)$, for arbitrary $B \in \mathcal{S}$, we obtain

$$E\{f(\xi,\eta);\, \xi \in B\,\} = E\Big\{\int f(\xi,t)\,\nu(\xi,dt);\, \xi \in B\Big\}, \quad B \in \mathcal{S}.$$

Since ζ is ξ-measurable, we have in particular

$$E\{f(\xi,\eta)\,|\,\zeta\} = E\Big\{\int f(\xi,t)\,\nu(\xi,dt)\Big|\,\zeta\Big\} \ \text{ a.s.,}$$

which agrees with (2) for the conditional distributions.

(iii) Since ζ is ξ-measurable, we can write ζB in the form $g(\xi)$, for any $B \in \mathcal{U}$. Applying (2) to the function $f(x,y)\,g(x)$, and using repeatedly the definition of Q_s, we obtain

$$\begin{aligned}\int_B E\zeta(ds)\,Q_s f &= E\,\zeta(B) f(\xi,\eta)\\ &= E\,\zeta B \int f(\xi,t)\,\nu(\xi,dt)\\ &= \int_B E\zeta(ds)\int Q_s(dx\,dy)\int f(x,t)\,\nu(x,dt).\end{aligned}$$

Since B was arbitrary, it follows that

$$Q_s f = \int Q_s(dx\,dy)\int f(x,t)\,\nu(x,dt), \quad s \in S \ \text{ a.e. } E\zeta,$$

and since S and T are Borel, the exceptional null set can be chosen to be independent of f. Thus, (2) holds a.e. for the univariate Palm measures Q_s. To obtain the corresponding multi-variate result, we need only replace ζ by the product measure ζ^n for arbitrary n. □

The following criterion for uniform approximation is needed in Chapter 13. Given a random measure ξ on R^d, we introduce the associated *centered Palm distributions* $P^s = \mathcal{L}(\theta_{-s}\xi \,\|\, \xi)_s$, $s \in \mathsf{R}^d$.

Lemma 6.8 *(uniform approximation) Let ξ_n and η_n be random measures on R^d, with centered Palm distributions P_n^s and Q_n^s, respectively, and fix any $B \in \mathcal{B}^d$. Then $\sup_{s\in B}\|P_n^s - Q_n^s\| \to 0$, under these conditions:*

(i) $E\xi_n B \asymp 1$,

(ii) $\Big\| E(\xi_n B;\, \xi_n \in \cdot\,) - E(\eta_n B;\, \eta_n \in \cdot\,)\Big\| \to 0$,

(iii) $\sup_{r,s\in B}\|P_n^r - P_n^s\| + \sup_{r,s\in B}\|Q_n^r - Q_n^s\| \to 0.$

Proof: Writing

$$f_A(\mu) = (\mu B)^{-1}\int_B \mu(ds)\,1_A(\theta_{-s}\mu), \quad A \subset \mathcal{M}_d,$$

we get

$$\begin{aligned}\int_B E\xi_n(ds)\,P_n^s A &= E\xi_n(B)\,f_A(\xi_n)\\ &= \int E(\xi_n B;\, \xi_n \in d\mu)\,f_A(\mu), \qquad (3)\end{aligned}$$

and similarly for η_n. For any $s \in B$, we have by (i)

$$\begin{aligned}\|P_n^s - Q_n^s\| &\lesssim E\xi_n(B)\,\|P_n^s - Q_n^s\| \\ &\le \Big\|E\xi_n(B)\,P_n^s - E\eta_n(B)\,Q_n^s\Big\| + \Big|E\xi_n B - E\eta_n B\Big|.\end{aligned}$$

Here (ii) shows that the second term on the right tends to 0, and from (3) we see that the first term is bounded by

$$\begin{aligned}&\left\|E\xi_n(B)\,P_n^s - \int_B E\xi_n(dr)\,P_n^r\right\| + \left\|E\eta_n(B)\,Q_n^s - \int_B E\eta_n(dr)\,Q_n^r\right\| \\ &\qquad + \left\|\int_B E\xi_n(dr)\,P_n^r - \int_B E\eta_n(dr)\,Q_n^r\right\| \\ &\quad\le E\xi_n(B)\sup_{r,s\in B}\|P_n^r - P_n^s\| + E\eta_n(B)\sup_{r,s\in B}\|Q_n^r - Q_n^s\| \\ &\qquad + \Big\|E(\xi_n B;\, \xi_n \in \cdot\,) - E(\eta_n B;\, \eta_n \in \cdot\,)\Big\|,\end{aligned}$$

which tends to 0 by (i)−(iii). □

We conclude with an extension of Lemma 6.1, needed in Chapter 13.

Lemma 6.9 *(conditional independence) Consider a random measure ξ on a Polish space S, and some random elements α and β in Borel spaces. Choose some kernels μ and ν with $\mu_{\alpha,\xi} = \mathcal{L}(\beta \,|\, \alpha, \xi)$ and $\nu_\alpha = \mathcal{L}(\beta \,|\, \alpha)$ a.s. Then*

(i) $\mathcal{L}(\xi, \alpha, \beta \,\|\, \xi)_s = \mathcal{L}(\xi, \alpha \,\|\, \xi)_s \otimes \mu$, $\quad s \in S$ *a.e.* $E\xi$.

Assuming $\beta \perp\!\!\!\perp_\alpha \xi$, we have also

(ii) $\mathcal{L}(\alpha, \beta \,\|\, \xi)_s = \mathcal{L}(\alpha \,\|\, \xi)_s \otimes \nu$, $\quad s \in S$ *a.e.* $E\xi$,

(iii) *if $\mathcal{L}(\alpha \,\|\, \xi)$ has a version that is continuous in total variation, then so has $\mathcal{L}(\beta \,\|\, \xi)$.*

Proof: First we prove (ii) when ξ is α-measurable and $\nu_\alpha = \mathcal{L}(\beta \,|\, \alpha)$ a.s. Assuming $E\xi$ to be σ-finite, we get

$$\begin{aligned}\int E\xi(ds)\,E\{f(s,\alpha,\beta)\|\,\xi\}_s &= E\int \xi(ds)\,f(s,\alpha,\beta) \\ &= E\int \xi(ds)\int \nu_\alpha(dt)\,f(s,\alpha,t) \\ &= \int E\xi(ds)\,E\left\{\int \nu_\alpha(dt)\,f(s,\alpha,t)\,\middle\|\,\xi\right\}_s,\end{aligned}$$

and so for $s \in S$ a.e. $E\xi$,

$$\begin{aligned}\mathcal{L}(\alpha,\beta \,\|\, \xi)_s f &= E\{f(\alpha,\beta)\|\,\xi\}_s \\ &= E\left\{\int \nu_\alpha(dt)\,f(\alpha,t)\,\middle\|\,\xi\right\}_s \\ &= \int P(\alpha \in dr \,\|\, \xi)_s \int \nu_r(dt)\,f(r,t) \\ &= \{\mathcal{L}(\alpha \,\|\, \xi)_s \otimes \nu\} f,\end{aligned}$$

as required. To obtain (i), it suffices to replace α in (ii) by the pair (ξ, α). If $\beta \perp\!\!\!\perp_\alpha \xi$, then $\mu_{\xi,\alpha} = \nu_\alpha$ a.s., and (ii) follows from (i). In particular,

$$\mathcal{L}(\beta \,\|\, \xi)_s = \mathcal{L}(\nu_\alpha \| \, \xi)_s, \quad s \in S \text{ a.e. } E\xi. \tag{4}$$

Assuming $\mathcal{L}(\alpha \,\|\, \xi)$ to be continuous in total variation, and defining $\mathcal{L}(\beta \,\|\, \xi)$ by (4), we get for any $s, s' \in S$

$$\begin{aligned}\left\|\mathcal{L}(\beta \,\|\, \xi)_s - \mathcal{L}(\beta \,\|\, \xi)_{s'}\right\| &= \left\|\mathcal{L}(\nu_\alpha \,\|\, \xi)_s - \mathcal{L}(\nu_\alpha \,\|\, \xi)_{s'}\right\| \\ &\le \left\|\mathcal{L}(\alpha \,\|\, \xi)_s - \mathcal{L}(\alpha \,\|\, \xi)_{s'}\right\|,\end{aligned}$$

and the assertion in (iii) follows. □

6.2 Reduced Palm Measures and Conditioning

For point processes $\xi = \sum_{i \in I} \delta_{\sigma_i}$ on S, Lemma 6.2 (ii) suggests that we "reduce" the Palm measure at $s \in S^n$ by omitting the "trivial" atoms at $s_1, s_2, \ldots$. To this aim, we introduce for every $n \in \mathsf{N}$ the *factorial moment measure* $E\xi^{(n)}$ on S^n, along with the associated *reduced Campbell measure* $C_\xi^{(n)}$ on $S^n \times \mathcal{N}_S$, given for measurable functions $f \ge 0$ on $S^n \times \mathcal{N}_S$ by

$$\begin{aligned}C_\xi^{(n)} f &= E \int \xi^{(n)}(ds)\, f\Big(s;\, \xi - \textstyle\sum_{k \le n} \delta_{s_k}\Big) \\ &= E \sum_{i \in I^{(n)}} f\Big(\sigma_{i_1}, \ldots, \sigma_{i_n};\, \textstyle\sum_{j \in I_i^{(n)}} \delta_{\sigma_j}\Big),\end{aligned}$$

where $I_i = I \setminus \{i_1, \ldots, i_n\}$, and the measurability on the right follows from Lemma 1.6. The reduced and compound Campbell measures $C_\xi^{(n)}$ and C_ξ are always σ-finite. For consistency, we also define $C_\xi^{(0)} = \mathcal{L}(\xi)$.

Disintegrating $C_\xi^{(n)}$ with respect to a σ-finite supporting measure $\nu \sim E\xi^{(n)}$ yields a kernel of *reduced Palm measures* $\mathcal{L}(\xi - \sum_i \delta_{s_i} \| \, \xi^{(n)})_s$, $s \in S^n$. We list some basic properties of the latter. Write $\mu_A = 1_A \mu$ for the restriction of the measure μ to the set A.

Lemma 6.10 *(reduced Palm measures) For any point process ξ on S,*

(i) $E\xi^{(n)} = E\xi^{\otimes n}$ *on* $S^{(n)}$, $\quad n \in \mathsf{N}$.

Letting $\nu \sim E\xi^{\otimes n}$ and $\nu' \sim E\xi^{(n)}$ be σ-finite with $\nu = \nu'$ on $S^{(n)}$, writing $\mathcal{L}_s$ and $\mathcal{L}'_s$ for the associated Palm and reduced Palm distributions, and putting $A_s = \{s_1, \ldots, s_n\}$, we have

(ii) $\mathcal{L}_s(\xi_{A_s^c}) = \mathcal{L}'_s(\xi_{A_s^c})$, $\quad s \in S^{(n)}$ *a.e.* $E\xi^{(n)}$,

(iii) *when ξ is a.s. simple,*

$$\mathcal{L}'_s\{\xi A_s > 0\} = 0, \quad s \in S^{(n)} \text{ a.e. } E\xi^{(n)}.$$

Proof: (i) Assuming $\xi = \sum_{i\in I} \delta_{\sigma_i}$, we have

$$\xi^{\otimes n} = \sum_{i\in I^n} \delta_{\sigma_{i_1},\dots,\sigma_{i_n}} \geq \sum_{i\in I^{(n)}} \delta_{\sigma_{i_1},\dots,\sigma_{i_n}} = \xi^{(n)}.$$

Since $(\sigma_{i_1},\dots,\sigma_{i_n}) \in S^{(n)}$ implies $(i_1,\dots,i_n) \in I^{(n)}$, we get $\xi^{\otimes n} = \xi^{(n)}$ on $S^{(n)}$, and the assertion follows.

(ii) Since the mapping $(s,\mu) \mapsto 1_{A_s^c}\mu$ is product-measurable on $S^n \times \mathcal{N}_S$ by Lemma 1.6, we get for any measurable sets $B \subset S^{(n)}$ and $M \subset \mathcal{N}_S$

$$\begin{aligned}\int_B \nu'(ds)\, \mathcal{L}'_s\{\xi_{A_s^c} \in M\} &= E\int_B \xi^{(n)}(ds)\, 1_M \circ \Big\{\xi - \textstyle\sum_{i\leq n}\delta_{s_i}\Big\}_{A_s^c} \\ &= E\int_B \xi^{\otimes n}(ds)\, 1_M(\xi_{A_s^c}) \\ &= \int_B \nu(ds)\, \mathcal{L}_s\{\xi_{A_s^c} \in M\}.\end{aligned}$$

Since $\nu = \nu'$ on $S^{(n)}$, and B is arbitrary, the assertion follows.

(iii) Letting $\xi = \sum_{i\in I} \delta_{\sigma_i}$, and noting that $\sigma_i \neq \sigma_j$ whenever $i \neq j$, we get

$$\begin{aligned}&\int \nu'(ds)\, \mathcal{L}'_s\{\xi A_s > 0\} \\ &\quad = E\int \xi^{(n)}(ds)\, 1\Big\{\Big(\xi - \textstyle\sum_{k\leq n}\delta_{s_k}\Big)A_s > 0\Big\} \\ &\quad = E\sum_{i\in I^{(n)}} \delta_{\sigma_{i_1},\dots,\sigma_{i_n}} 1\Big\{\textstyle\sum_{h\in I_i}\sum_{k\leq n}\delta_{\sigma_h}\{\sigma_{i_k}\} > 0\Big\} = 0,\end{aligned}$$

and the assertion follows. □

The measures $C^n_{\xi,\eta}$ and $C^{(n)}_\xi$ are clearly invariant under permutations of coordinates in S^n, which allows us to choose symmetric versions of the associated Palm measures $\mathcal{L}(\eta \,\|\, \xi^{\otimes n})_s$ and $\mathcal{L}(\xi' \,\|\, \xi^{(n)})_s$. Identifying the n-tuples $s = (s_1,\dots,s_n) \in S^n$ with the corresponding counting measures $\mu = \sum_i \delta_{s_i}$, we may then regard $C^n_{\xi,\eta}$ and $C^{(n)}_\xi$ as defined on $\hat{\mathcal{N}}_S \times T$ and $\hat{\mathcal{N}}_S \times \mathcal{N}_S$, and consider the associated Palm measures

$$\mathcal{L}(\eta \,\|\, \xi^{\mu S})_\mu, \quad \mathcal{L}(\xi - \mu \,\|\, \xi^{(\mu S)})_\mu, \qquad \mu \in \hat{\mathcal{N}}_S,$$

as kernels from $\hat{\mathcal{N}}_S$ to T and $\mathcal{N}_S$, respectively.

Lemma 6.11 *(compound Campbell measure) For any point process ξ on S, the reduced Palm kernel $\mathcal{L}(\xi - \mu \,\|\, \xi^{(\mu S)})_\mu$, $\mu \in \hat{\mathcal{N}}_S$, may be obtained by disintegration of the measure C_ξ on $\hat{\mathcal{N}}_S \times \mathcal{N}_S$, given by*

$$C_\xi f = \sum_{n\geq 0} \frac{C^{(n)}_\xi f}{n!} = E\sum_{\mu\leq\xi} f(\mu, \xi - \mu).$$

Proof: The last equality is immediate from Theorem 1.13 (i), which may also be consulted for the precise interpretation of the right-hand side. □

It is often useful to think of Palm measures as generalized conditional distributions. The connection is obvious when $\xi = \delta_\sigma$ for some random element σ in S, since in that case $E\xi = \mathcal{L}(\sigma)$, and so by definitions

$$\mathcal{L}(\,\eta\,\|\,\xi)_s = \mathcal{L}(\,\eta\,|\,\sigma)_s \ \text{ a.e. } E\xi, \tag{5}$$

where the right-hand side is a version of the regular conditional distribution of η given σ, regarded as a kernel from S to T. More generally, if ξ is a random measure on S with non-random total mass $\|\xi\| \in (0,\infty)$, and σ is a randomly chosen point from ξ, in the sense of

$$\mathcal{L}(\sigma\,|\,\xi,\eta) = \frac{\xi}{\|\xi\|} = \hat{\xi} \ \text{ a.s.}, \tag{6}$$

then (5) will be shown to remain valid for the supporting measure $\nu = E\hat{\xi}$.

For general ξ, the implication (6) ⇒ (5) fails and needs to be modified, as follows. Instead of assuming (6), we choose σ to satisfy $\mathcal{L}(\sigma\,|\,\xi,\eta) = \xi$ a.s. To make this precise, we may introduce a pseudo-probability space[4] $(\tilde{\Omega}, \tilde{P})$, with associated pseudo-expectation ($\tilde{P}$-integration) $\tilde{E}$, carrying a random pair $(\sigma, \tilde{\eta})$ satisfying

$$\tilde{E}f(\sigma,\tilde{\eta}) = E\int \xi(ds)\, f(s,\eta) = C_{\xi,\eta}f, \quad f \in (\mathcal{S}\otimes\mathcal{T})_+. \tag{7}$$

The easiest choice is to take $(\sigma, \tilde{\eta})$ to be the identity map on $\tilde{\Omega} = S \times T$, endowed with the pseudo-probability measure $\tilde{P} = C_{\xi,\eta}$. We may also construct $(\tilde{\Omega}, \tilde{P})$ by a suitable extension of the original probability space (Ω, P). When there is no risk for confusion, we may identify $\tilde{\eta}$ with η, and write $\tilde{P} = P$ and $\tilde{E} = E$.

In the general case, we also need to make sense of the right-hand side of (5). Assuming $E\xi = \mathcal{L}(\sigma)$ to be σ-finite, for the sake of simplicity, we may define $\mathcal{L}(\eta\,|\,\sigma)$ in the usual way, as the a.e. unique probability kernel from S to T satisfying

$$E\Big(\tilde{E}\{f(\eta)\,|\,\sigma\}_\sigma;\, \sigma \in B\Big) = E\{f(\eta);\, \sigma \in B\}, \quad f \in \mathcal{T}_+,\ B \in \mathcal{S}. \tag{8}$$

With the stated definitions and assumptions, (5) remains true with essentially the same proof.

Since the multi-variate Palm measures $\mathcal{L}(\eta\,\|\,\xi^{\otimes n})$ are univariate Palm measures with respect to the product measures $\xi^{\otimes n}$, they too may be obtained by a similar conditioning approach. However, it is often more natural to

[4]The prefix "pseudo" signifies that the measure $\tilde{P}$ on $\tilde{\Omega}$ is not normalized and may even be unbounded, though it is always taken to be σ-finite.

regard the measures $\mathcal{L}(\eta \,\|\, \xi^{\otimes n})$ as indexed by $\hat{\mathcal{N}}_S$. Writing $\delta_n(s) = \sum_i \delta_{s_i}$ for $s = (s_1, \dots, s_n) \in S^n$, we may then introduce a pseudo-point process ζ_n on S with $\mathcal{L}(\zeta_n \,|\, \xi, \eta) = \xi^{\otimes n} \circ \delta_n^{-1}$, and define $\mathcal{L}(\eta \,\|\, \xi^{\otimes n})_\mu = \mathcal{L}(\eta \,|\, \zeta_n)_\mu$. More precisely, we consider a pseudo-probability space carrying a random pair $(\tilde{\eta}, \zeta_n)$, such that for any measurable function $f \geq 0$ on $T \times \hat{\mathcal{N}}_S$,

$$\tilde{E} f(\tilde{\eta}, \zeta_n) = E \int f\Big(\eta, \, \textstyle\sum_i \delta_{s_i}\Big)\, \xi^{\otimes n}(ds_1 \cdots ds_n).$$

The conditioning $\mathcal{L}(\eta \,|\, \zeta_n)$ can then be interpreted as before.

The conditioning approach becomes even more natural for the reduced Palm distributions of a point process ξ. Here we may introduce a single pseudo-point process ζ on S with $\mathcal{L}(\zeta \,|\, \xi) = \sum_n (\xi^{(n)}/n!) \circ \delta_n^{-1}$, where the last conditioning can be made precise, on a suitable pseudo-probability space, by the formula

$$\begin{aligned}\tilde{E} f(\tilde{\xi}, \zeta) &= \sum_{n \geq 0} \frac{1}{n!}\, E \int f\Big(\xi, \textstyle\sum_i \delta_{s_i}\Big)\, \xi^{(n)}(ds) \\ &= E \sum_{\mu \leq \xi} f(\xi, \mu),\end{aligned} \tag{9}$$

for any measurable functions $f \geq 0$ on $\mathcal{N}_S \times \hat{\mathcal{N}}_S$. Omitting the “tilde” $\sim$ as before, and assuming for simplicity that all moment measures $E\xi^n$ are σ-finite, we claim that

$$\mathcal{L}\{\, \xi - \mu \,\|\, \xi^{(n)}\}_\mu = \mathcal{L}(\, \xi - \zeta \,|\, \zeta \,)_\mu, \quad \mu \in \hat{\mathcal{N}}_S \ \text{ a.e. } E(\zeta;\, \|\zeta\| = n). \tag{10}$$

To summarize, we have the following constructions of Palm measures, in terms of elementary conditioning. To avoid distracting technicalities, we may again assume the appropriate moment measures to be σ-finite.

Theorem 6.12 *(conditioning approach) Consider a random measure ξ on S, and a random element η in T.*

(i) *When $\|\xi\|$ is non-random in $(0, \infty)$, choose a random element σ in S with $\mathcal{L}(\sigma \,|\, \xi, \eta) = \hat{\xi}$ a.s. Then $\mathcal{L}(\eta \,\|\, \xi) = \mathcal{L}(\eta \,|\, \sigma)$ a.e. $E\xi$.*

(ii) *For general ξ with σ-finite intensity $E\xi$, choose a pseudo-random element σ in S with $\tilde{\mathcal{L}}(\sigma \,|\, \xi, \eta) = \xi$ a.s., in the sense of* (7)*. Then $\mathcal{L}(\eta \,\|\, \xi) = \mathcal{L}(\eta \,|\, \sigma)$ a.e. $E\xi$, in the sense of* (8)*.*

(iii) *Given a point process ξ on S with σ-finite moment measures $E\xi^{\otimes n}$, choose a finite pseudo-point process ζ on S satisfying* (9)*. Then the reduced Palm distributions $\mathcal{L}(\, \xi - \mu \,\|\, \xi^{(n)})_\mu$ of ξ are given by* (10)*, a.e. $E\xi^{(n)} \circ \delta_n^{-1}$.*

Proof: (i) For any $f \in \mathcal{T}_+$ and $B \in \mathcal{S}$, we have

$$\begin{aligned} E\Big\{E(f(\eta) \,\|\, \xi)_\sigma;\, \sigma \in B\Big\} &= E \int_B E\{f(\eta) \,\|\, \xi\}_s \, \hat{\xi}(ds) \\ &= \int_B E\{f(\eta) \,\|\, \xi\}_s \, E\hat{\xi}(ds) \\ &= E\, f(\eta)\, \hat{\xi} B = E\{f(\eta);\, \sigma \in B\}, \end{aligned}$$

where the first and last steps hold by (6), the second step holds by the definition of $E\xi$, and the third step holds by Palm disintegration. The desired equation now follows, by the definition of conditional distributions.

(ii) Proceed as in (i), though with $\hat{\xi}$ replaced by ξ, where the pairs (σ, η) are now defined on the pseudo-probability space $(\tilde{\Omega}, \tilde{P})$.

(iii) Proceeding as in (i), we get for any measurable function $f \geq 0$ on $\mathcal{N}_S$, and for $B \subset \hat{\mathcal{N}}_S$ restricted to measures μ with $\|\mu\| = n$,

$$\begin{aligned} E\Big\{f(\xi-\zeta);\, \zeta \in B\Big\} &= E\sum_{\mu\leq\xi} f(\xi-\mu)\, 1_B(\mu) = C_\xi(f\otimes 1_B) \\ &= E\sum_{\mu\leq\xi} E\Big\{f(\xi-\mu)\,\Big\|\,\xi^{(n)}\Big\}_\mu 1_B(\mu) \\ &= E\Big(E\{f(\xi-\zeta)\,\|\,\xi^{(n)}\}_\zeta;\, \zeta\in B\Big). \end{aligned}$$

Here the first and last equalities hold by (9), the second equality holds by the definition of C_ξ, and the third equality holds by Palm disintegration. This shows that $E\{f(\xi-\zeta)\,\|\,\xi^{(n)}\}_\zeta$ satisfies the defining relation for $E\{f(\xi-\zeta)\,|\,\zeta\}$. □

The reduced Palm kernel plays a fundamental role in Chapter 8, as dual to the equally important Gibbs kernel. The following result, related to Lemma 6.10 above, will be needed in Chapter 13.

Lemma 6.13 *(singleton mass) Given a random measure ξ on S and a random element τ in a Borel space T, we have*

$$P\{\tau \neq t\,\|\,\xi\otimes\delta_\tau\}_{s,t} = 0, \quad (s,t)\in S\times T \;\; a.e.\;\; E(\xi\otimes\delta_\tau).$$

Proof: Since T is Borel, the diagonal $\{(t,t);\, t\in T\}$ in T^2 is measurable. Letting ν be the associated supporting measure for $\xi\otimes\delta_\tau$, we get by Palm disintegration

$$\begin{aligned} \iint \nu(ds\,dt)\, P\{\tau\neq t\,\|\,\xi\otimes\delta_\tau\}_{s,t} &= E\iint (\xi\otimes\delta_\tau)(ds\,dt)\, 1\{\tau\neq t\} \\ &= E\int \xi(ds)\, 1\{\tau\neq\tau\} = 0, \end{aligned}$$

and the assertion follows. □

6.3 Slivnyak Properties and Factorization

The Palm distributions of a random measure can often be obtained, most conveniently, by a simple computation based on Laplace transforms. To illustrate the method, we consider three basic propositions, involving Poisson and related processes. They will all be extended in various ways in subsequent sections.

First we fix a σ-finite measure λ on S. If ξ is a mixed Poisson process on S directed by $\rho\lambda$, for some random variable $\rho \geq 0$, then clearly

$$P\{\xi B = 0\} = \varphi(\lambda B), \quad B \in \mathcal{S},$$

where $\varphi(t) = Ee^{-t\rho}$ denotes the Laplace transform of ρ. When $\lambda S \in (0, \infty)$, the same formula holds for a mixed binomial process ξ, based on the probability measure $\lambda/\lambda S$ and a Z_+-valued random variable κ, provided that we choose $\varphi(t) = E(1 - t/\lambda S)^\kappa$ for $t \in [0, \lambda S]$. In either case, ξ has Laplace functional

$$Ee^{-\xi f} = \varphi\{\lambda(1 - e^{-f})\}, \quad f \in \mathcal{S}_+, \tag{11}$$

and so $\mathcal{L}(\xi)$ is uniquely determined by the pair (λ, φ), which justifies the notation $\mathcal{L}(\xi) = M(\lambda, \varphi)$. For convenience, we may also allow P to be σ-finite, so that $M(\lambda, -\varphi')$ makes sense as the pseudo-distribution of a point process on S.

Lemma 6.14 *(mixed Poisson and binomial processes) Let $\mathcal{L}(\xi) = M(\lambda, \varphi)$ for some σ-finite measure λ on S. Then the reduced Palm measures of ξ equal*

$$\mathcal{L}(\xi - \delta_s \,\|\, \xi)_s = M(\lambda, -\varphi'), \quad s \in S \ \text{a.e. } \lambda,$$

which agrees with $\mathcal{L}(\xi)$ iff ξ is Poisson with $E\xi = \lambda$.

Proof: First assume that $E\xi$ is σ-finite. For any $f \in \mathcal{S}_+$ and $B \in \mathcal{S}$, with $\lambda f < \infty$ and $\lambda B < \infty$, we get by (11)

$$Ee^{-\xi f - t\xi B} = \varphi\Big\{\lambda(1 - e^{-f - t1_B})\Big\}, \quad t \geq 0.$$

Taking right derivatives at $t = 0$ gives

$$E(\xi B\, e^{-\xi f}) = -\varphi'\{\lambda(1 - e^{-f})\}\, \lambda(1_B\, e^{-f}),$$

where the differentiation on both sides is justified by dominated convergence. Hence, by disintegration,

$$\int_B \lambda(ds)\, E(e^{-\xi f} \,\|\, \xi)_s = -\varphi'\{\lambda(1 - e^{-f})\} \int_B e^{-f(s)} \lambda(ds).$$

Since B was arbitrary, we get

$$E(e^{-\xi f} \,\|\, \xi)_s = -\varphi'\{\lambda(1 - e^{-f})\}\, e^{-f(s)}, \quad s \in S \ \text{a.e. } \lambda,$$

which implies

$$\begin{aligned} E\Big\{e^{-(\xi - \delta_s)f} \Big\| \,\xi\Big\}_s &= E(e^{-\xi f} \,\|\, \xi)_s\, e^{f(s)} \\ &= -\varphi'\{\lambda(1 - e^{-f})\}. \end{aligned}$$

The first assertion now follows by (11). For general $E\xi$, we may take $f \geq \varepsilon 1_B$ for fixed $\varepsilon > 0$, to ensure the validity of the differentiation. The resulting formulas extend by monotone convergence to arbitrary $f \geq 0$. To prove the last assertion, we note that the equation $\varphi = -\varphi'$, with initial condition $\varphi(0) = 1$, has the unique solution $\varphi(t) = e^{-t}$. □

We continue with a classical Poisson characterization.

Lemma 6.15 *(Poisson criterion, Slivnyak, Kerstan & Matthes, Mecke) Let ξ be a random measure on S with σ-finite intensity $E\xi$. Then ξ is Poisson iff*

$$\mathcal{L}(\xi \,\|\, \xi)_s = \mathcal{L}(\xi + \delta_s), \quad s \in S \;\; a.e.\;\, E\xi.$$

Proof: The necessity being part of Lemma 6.14, it remains to prove the sufficiency. Then assume the stated condition, and put $E\xi = \lambda$. Fixing an $f \geq 0$ with $\lambda f < \infty$, and writing $\varphi(t) = Ee^{-t\xi f}$, we get by dominated convergence

$$\begin{aligned} -\varphi'(t) &= E\,\xi f\, e^{-t\xi f} \\ &= \int \lambda(ds)\, f(s)\, E(e^{-t\xi f} \,\|\, \xi)_s \\ &= \varphi(t)\, \lambda(f e^{-tf}). \end{aligned}$$

Noting that $\varphi(0) = 1$ and $\varphi(1) = Ee^{-\xi f}$, we obtain

$$\begin{aligned} -\log Ee^{-\xi f} &= \int_0^1 \frac{\varphi'(t)}{\varphi(t)}\, dt = \int_0^1 \lambda(f e^{-tf})\, dt \\ &= \lambda \int_0^1 f e^{-tf(s)}\, dt \\ &= \lambda(1 - e^{-f}), \end{aligned}$$

and so ξ is Poisson λ by Lemma 3.1. □

We turn to the Palm distributions of an infinitely divisible random measure ξ, directed as in Theorem 3.20 by some measures α on S and λ on $\mathcal{M}_S \setminus \{0\}$. The associated Campbell measure Λ on $S \times (\mathcal{M}_S \setminus \{0\})$ is defined by

$$\Lambda f = \int f(s,0)\, \alpha(ds) + \int \mu(ds) \int f(s,\mu)\, \lambda(d\mu). \tag{12}$$

Since $E\xi = \Lambda(\cdot \times \mathcal{M}_S)$, any supporting measure ν of ξ is also a supporting measure for Λ.

Lemma 6.16 *(infinitely divisible random measures) Let ξ be an infinitely divisible random measure on S directed by (α, λ), and define Λ by (12). Fix a supporting measure $\nu \sim E\xi$ and a disintegration $\Lambda = \nu \otimes \rho$. Then*

$$\mathcal{L}(\xi \,\|\, \xi)_s = \mathcal{L}(\xi) * \rho_s, \quad s \in S \;\; a.e.\;\, \nu. \tag{13}$$

Proof: By Theorem 3.20, we have

$$-\log Ee^{-\xi f} = \alpha f + \int (1 - e^{-\mu f})\, \lambda(d\mu), \quad f \in \mathcal{S}_+,$$

Replacing f by $f + t1_B$ and differentiating at $t = 0$, we get

$$E(\xi B\, e^{-\xi f}) = Ee^{-\xi f} \Big\{ \alpha B + \int \mu B\, e^{-\mu f}\, \lambda(d\mu) \Big\},$$

where the formal differentiation may be justified as before. Hence, by disintegration,

$$\int_B \nu(ds)\, E(e^{-\xi f} \,\|\, \xi)_s = Ee^{-\xi f} \int_B \nu(ds) \int \rho_s(d\mu)\, e^{-\mu f},$$

and so

$$\int_B \nu(ds)\, \mathcal{L}(\xi \,\|\, \xi)_s = \mathcal{L}(\xi) * \int_B \nu(ds)\, \rho_s,$$

by the uniqueness theorem for Laplace transforms. Since B was arbitrary, the asserted factorization follows by the uniqueness in Theorem 1.23. □

For any σ-finite measures μ and ν on an additive semigroup H, we define the *convolution* $\mu * \nu$ as the image of the product measure $\mu \otimes \nu$, under the mapping $(a, b) \mapsto a + b$. Say that μ *divides* ν and write $\mu \prec \nu$, if $\nu = \mu * \rho$ for some measure ρ on H. The last result shows that, if ξ is an infinitely divisible random measure on S, then $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)_s$ for $s \in S$ a.e. $E\xi$. We show that the two conditions are in fact equivalent.

It is enough to assume $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi f)$ for a single function $f > 0$, provided we add a suitable condition on the constant part of ξ. Then define the set function $\operatorname{ess\,inf} \xi$ on $\mathcal{S}$ by

$$\begin{aligned} (\operatorname{ess\,inf} \xi) B &= \operatorname{ess\,inf}\, (\xi B) \\ &= \sup\Big\{ r \geq 0;\ \xi B \geq r \ \text{ a.s.} \Big\}. \end{aligned}$$

Theorem 6.17 *(infinite divisibility via factorization) Let S be Borel with a generating semiring $\mathcal{I} \subset \hat{\mathcal{S}}$. Then a point process ξ on S is infinitely divisible as a random measure, iff*

(i) $\mathcal{L}(\xi f) \prec \mathcal{L}(\xi f \,\|\, \xi f)$ *for all* $f \in \mathcal{I}_+$.

Furthermore, for any random measure ξ on S, each of the following conditions is equivalent to the infinite divisibility of ξ:

(ii) $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi I)$ *for all* $I \in \mathcal{I}$,

(iii) $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)_s$ *for* $s \in S$ *a.e.* $E\xi$,

(iv) $\operatorname{ess\,inf}\, (f \cdot \xi)$ *is finitely sub-additive and* $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi f)$, *for some measurable function* $f > 0$ *on* S *with* $\xi f < \infty$ *a.s.*

Our proof will be based on two lemmas.

Lemma 6.18 *(infinitely divisible random variables) A random variable $\xi \geq 0$ is infinitely divisible iff $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)$.*

Though the necessity is a special case of Lemma 6.16, we give a complete proof, for the sake of clarity.

Proof: Put $\varphi(t) = Ee^{-t\xi}$. If ξ is infinitely divisible, then Theorem 3.20 yields

$$-\log \varphi(t) = t\alpha + \int_0^\infty (1 - e^{-tx})\, \lambda(dx), \quad t \geq 0,$$

for some constant $\alpha \geq 0$ and measure λ on $(0, \infty)$ with $\int (x \wedge 1)\, \lambda(dx) < \infty$. Hence, for $t > 0$, we get by dominated convergence

$$\begin{aligned} -\frac{\varphi'(t)}{\varphi(t)} &= \alpha + \int_0^\infty x\, e^{-tx}\, \lambda(dx) \\ &= \int_0^\infty e^{-tx}\, \rho(dx), \end{aligned}$$

where $\rho\{0\} = \alpha$ and $\rho(dx) = x\, \lambda(dx)$ for $x > 0$. Noting that $E(e^{-t\xi} \| \,\xi) = E(\xi\, e^{-t\xi}) = -\varphi'(t)$, and using Lemma 2.2, we obtain $\mathcal{L}(\xi \,\|\, \xi) = \mathcal{L}(\xi) * \rho$, and so $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)$.

Conversely, let $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)$, so that $\mathcal{L}(\xi \,\|\, \xi) = \mathcal{L}(\xi) * \rho$ for some σ-finite measure ρ on R_+. Then for any $t > 0$,

$$\begin{aligned} -\varphi'(t) &= E(\xi\, e^{-t\xi}) = E(e^{-t\xi} \| \,\xi) \\ &= \varphi(t) \int e^{-tx}\, \rho(dx). \end{aligned}$$

Dividing by $\varphi(t)$, integrating from 0 to u, and using Fubini's theorem, we get

$$\begin{aligned} -\log \varphi(u) &= \int_0^u dt \int_0^\infty e^{-tx}\, dx \\ &= \int_0^\infty dx \int_0^u e^{-tx}\, dt \\ &= u\alpha + \int_0^\infty (1 - e^{-ux})\, \lambda(dx), \end{aligned}$$

where $\alpha = \rho\{0\}$ and $\lambda(dx) = x^{-1}\, \rho(dx)$ for $x > 0$. Since $\varphi(u) > 0$ implies $\alpha < \infty$ and $\int (x \wedge 1)\, \lambda(dx) < \infty$, we conclude that ξ is infinitely divisible. □

Lemma 6.19 *(Cox transform) Let ξ be a Cox process on S directed by η, and fix any $s \in S$ with $P\{\eta\{s\} > 0\} > 0$. Then $\mathcal{L}(\eta) \prec \mathcal{L}(\eta \,\|\, \eta)_s$ implies $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)_s$.*

Proof: For any σ-finite measure μ on $\mathcal{M}_S$, write $C\mu$ for the Cox transform of μ and $P_s\mu$ for the Palm measure of μ at s. We claim that

$$C(\mu * \nu) = C\mu * C\nu, \qquad CP_s\mu * \delta_{\delta_s} = P_sC\mu. \tag{14}$$

To prove the first relation, put $L\mu(f) = \int e^{-mf}\mu(dm)$ and $g = 1 - e^{-f}$, and conclude from Lemma 3.1 and Fubini's theorem that

$$\begin{aligned} LC(\mu * \nu)(f) &= L(\mu * \nu)(g) = L\mu(g)\, L\nu(g) \\ &= LC\mu(f)\, LC\nu(f) \\ &= L(C\mu * C\nu)(f). \end{aligned}$$

The assertion now follows by Lemma 2.2.

To prove the second relation, write the Laplace transform of μ as

$$L\mu(t, f) = \int e^{-m_s t - mf}\mu(dm), \quad t \in \mathsf{R}_+,\ f \in \mathcal{S}'_+,$$

with $S' = S \setminus \{s\}$, and let $(L\mu)'$ denote the partial t-derivative of $L\mu$. Then by dominated convergence and Lemma 3.1,

$$\begin{aligned} LP_s\mu(t, f) &= \int m_s\, e^{-m_s t - mf}\mu(dm) \\ &= -(L\mu)'(t, f), \\ LC\mu(t, f) &= L\mu\Big(1 - e^{-t}, 1 - e^{-f}\Big), \end{aligned}$$

and so, by combination,

$$\begin{aligned} LP_sC\mu(t, f) &= -(L\mu)'\Big(1 - e^{-t}, 1 - e^{-f}\Big)\, e^{-t} \\ &= LCP_s\mu(t, f)\, e^{-t}. \end{aligned}$$

Again it remains to use Lemma 2.2.

From (14) we get

$$\mu \prec \nu \;\Rightarrow\; C\mu \prec C\nu, \qquad\qquad CP_s\mu \prec P_sC\mu.$$

Writing $\mu = \mathcal{L}(\eta)$ and assuming $\mu \prec P_s\mu$, we get $C\mu \prec CP_s\mu \prec P_sC\mu$, which implies $C\mu \prec P_sC\mu$. □

Proof of Theorem 6.17: The necessity being clear from Lemma 6.16, it remains to prove the sufficiency.

(i) Use Lemma 6.18 and Theorem 3.24 (ii).

(ii) By Theorem 3.24 (i), we may take $S = \{1, \ldots, n\}$ and $\mathcal{I} = 2^S$. For any $c > 0$, let η_c be a Cox process directed by $c\xi$, and note that $\mathcal{L}(\eta_c) \prec \mathcal{L}(\eta_c \| \eta_c)_s$ for all s, by Lemma 6.19. Then $\mathcal{L}(t\eta_c) \prec \mathcal{L}(t\eta_c \| t\eta_c)$ for all $t \in \mathsf{R}^n_+$, and η_c is infinitely divisible by (i). Since $c > 0$ was arbitrary, even ξ is infinitely divisible by Corollary 4.31.

(iii) Assume (13) for some measures ρ_s. Taking Laplace transforms and using Lemma 1.18, we see that the ρ_s can be chosen to form a kernel from S to $\mathcal{M}_S$. Hence, (13) extends to (ii), and the infinite divisibility follows.

(iv) If ξ is infinitely divisible and directed by (α,λ), then clearly $\operatorname{ess\,inf}\xi=\alpha$, which is a measure, hence finitely subaddive. Conversely, suppose that ξ satisfies the stated conditions for some function $f>0$. Since ξ and $f\cdot\xi$ are simultaneously infinitely divisible, and the stated factorization implies $\mathcal{L}(f\cdot\xi)\prec\mathcal{L}(f\cdot\xi\,\|\,\xi f)$, we may assume that $f\equiv 1$. Furthermore, it suffices by Theorem 3.24 to take $S=\{1,\dots,n\}$ for some $n\in\mathsf{N}$, so that ξ may be regarded as a random vector $(\xi_1,\dots,\xi_n)$ in R_+^n. Now for any random variables $\eta_1,\eta_2\geq 0$, the a.s. relations $\eta_1\geq r_1$ and $\eta_2\geq r_2$ imply $\eta_1+\eta_2\geq r_1+r_2$ a.s., and so

$$\begin{aligned}\operatorname{ess\,inf}(\eta_1+\eta_2)&=\sup\Big\{r;\ \eta_1+\eta_2\geq r\ \text{ a.s.}\Big\}\\&\geq\sup\Big\{r;\ \eta_1\geq r\ \text{ a.s.}\Big\}+\sup\Big\{r;\ \eta_2\geq r\ \text{ a.s.}\Big\}\\&=\operatorname{ess\,inf}\eta_1+\operatorname{ess\,inf}\eta_2.\end{aligned}$$

Thus, $\alpha=\operatorname{ess\,inf}\xi$ is even finitely superadditive, hence a measure. Since ξ and $\xi-\alpha$ are simultaneously infinitely divisible, and the factorization properties of ξ and $\xi-\alpha$ are equivalent, we may henceforth assume that $\operatorname{ess\,inf}\xi=0$.

For $c>0$ and $1=\bar 1=(1,\dots,1)$, let (γ_c,η_c) be a Cox process on $\{0,1,\dots,n\}$ directed by $c(\|\xi\|,\xi)$, and note that $\mathcal{L}(\gamma_c,\eta_c)\prec\mathcal{L}(\gamma_c,\eta_c\|\,\gamma_c)$ by Lemma 6.19. Hence, the generating function $\psi_c(s,t)$ of (γ_c,η_c) and its partial s-derivative ψ_c' satisfy

$$s\,\psi_c'(s,t)=\psi_c(s,t)\int s^r\,t_1^{z_1}\cdots t_n^{z_n}\,\rho_c(dr\,dz),$$

for some measure ρ_c on Z_+^{n+1}. Since $P\{\gamma_c=0\}=Ee^{-c\bar\xi}>0$, we have $\psi_c>0$, and so s divides the integral on the right, which means that ρ_c is restricted to $\mathsf{N}\times\mathsf{Z}_+^n$. Dividing by $s\,\psi_c(s,t)$ and integrating over $s\in[0,1]$, we get

$$\log\hat\psi_c(t)=\log\psi_c(0,t)+\int t_1^{z_1}\cdots t_n^{z_n}\,\lambda_c(dz),$$

where $\hat\psi_c(t)=\psi_c(1,t)$ and $\lambda_c=\int r^{-1}\,\rho_c(dr\times\cdot)$. Subtracting the corresponding equation for $t=\bar 1$ gives

$$\begin{aligned}-\log\hat\psi_c(t)&=\log\psi_c(0,\bar 1)-\log\psi_c(0,t)+\int\{1-t_1^{z_1}\cdots t_n^{z_n}\}\,\lambda_c(dz)\\&=\log\psi_c(0,\bar 1)-\log\psi_c(0,t)-\log\vartheta_c(t),\end{aligned}$$

where ϑ_c is the generating function of an infinitely divisible distribution on Z_+^n by Theorem 3.20. Writing $p=1-q=c^{-1}$ and $\bar q=q\bar 1$, we see from Lemma 3.2 that η_1 is distributed as a p-thinning of η_c, and so by Lemma 3.1

$$\begin{aligned}-\log\hat\psi_1(t)&=-\log\hat\psi_c(\bar q+pt)\\&=\log\psi_c(0,\bar 1)-\log\psi_c(0,\bar q+pt)-\log\vartheta_c(\bar q+pt).\end{aligned}\tag{15}$$

Now $\|\xi\|$ is infinitely divisible by Lemma 6.18, and its constant part equals $\operatorname{ess\,inf}\|\xi\|=0$. Letting λ be the Lévy measure of $\|\xi\|$, and using Lemma 3.1, we get

$$\begin{aligned}0 &\le \log\psi_c(0,\bar{1}) - \log\psi_c(0,\bar{q}+pt)\\ &\le \log\psi_c(0,\bar{1}) - \log\psi_c(0,\bar{q})\\ &= \log P\{\gamma_c = 0\} - \log E(q^{\bar{\eta}_c};\, \gamma_c = 0)\\ &= \log Ee^{-c\|\xi\|} - \log Ee^{-(c+1)\|\xi\|}\\ &= \int\Big\{e^{-cr} - e^{-(c+1)r}\Big\}\lambda(dr)\\ &= \int e^{-cr}\,(1-e^{-r})\,\lambda(dr),\end{aligned}$$

which tends to 0 by dominated convergence as $c \to \infty$. Hence, (15) yields $\hat{\vartheta}_c(\bar{q}+pt) \to \hat{\psi}_1(t)$ for all $t \in [0,1]^n$. Since the generating functions $\hat{\vartheta}_c(\bar{q}+pt)$, corresponding to p-thinnings of $\hat{\vartheta}_c$, are again infinitely divisible, so is $\hat{\psi}_1$ by Lemma 4.24, which means that η_1 is infinitely divisible. Replacing ξ by $c\xi$, we see that η_c is infinitely divisible for every $c > 0$, and so ξ is again infinitely divisible by Corollary 4.31. □

We may also give a related characterization of infinite divisibility. For any $\mu \in \mathcal{M}_S$ and $\varepsilon > 0$, we define the measure μ^*_ε by $\mu^*_\varepsilon B = \sum_{s\in B} 1\{\mu\{s\} \ge \varepsilon\}$, and say that μ is *discrete*, if it is purely atomic with $\mu^*_{0+}B < \infty$ for all $B \in \hat{\mathcal{S}}$.

Theorem 6.20 *(factorization by conditioning) Let ξ be a purely atomic random measure on S with $E\xi\{s\} \equiv 0$, and fix a dissection system $\mathcal{I} \subset \hat{\mathcal{S}}$. Then ξ is infinitely divisible, iff $\mathcal{L}(\xi \,|\, \xi^*_\varepsilon B = 0) \prec \mathcal{L}(\xi)$ for every $B \in \mathcal{I}$ and $\varepsilon > 0$ with $P\{\xi^*_\varepsilon B = 0\} > 0$. When ξ is discrete, it suffices to take $\varepsilon = 0+$.*

Proof: Let ξ be infinitely divisible and directed by $(0,\lambda)$, so that $\xi = \int \mu\,\eta(d\mu)$ for some Poisson process on $\mathcal{M}_S \setminus \{0\}$ with $E\eta = \lambda$. Let $B \in \hat{\mathcal{S}}$ and $\varepsilon > 0$ with $P\{\xi^*_\varepsilon B = 0\} > 0$. Putting $M = \{\mu;\, \mu^*_\varepsilon B = 0\}$, and noting that $\int \mu\{s\}\,\lambda(d\mu) = E\xi(ds) = 0$ for all $s \in S$, we get

$$\{\xi^*_\varepsilon B = 0\} = \{\xi \in M\} = \{\eta M^c = 0\} \ \text{ a.s.}$$

Since $\mathcal{L}(\eta \,|\, \eta M^c = 0)$ is again Poisson with intensity $1_M\lambda$, the measure $\mathcal{L}(\xi \,|\, \xi \in M)$ is infinitely divisible and directed by $(0, 1_M\lambda)$, and so $\mathcal{L}(\xi \,|\, \xi \in M) \prec \mathcal{L}(\xi)$ by the independence property of Poisson processes.

To prove the converse assertion, we first take ξ to be a point process satisfying the stated condition with $\varepsilon = 0+$. Let $B \in \hat{\mathcal{I}}$ with $P\{\xi B = 0\} > 0$, so that $\mathcal{L}(\xi \,|\, \xi B = 0) \prec \mathcal{L}(\xi)$. Then $\xi \overset{d}{=} \eta_B + \zeta_B$, where $\mathcal{L}(\eta_B) = \mathcal{L}(\xi \,|\, \xi B = 0)$ and $\eta_B \perp\!\!\!\perp \zeta_B$. Noting that $\eta_B B = 0$ a.s., we get for any $I \in \mathcal{S}$ in B

$$\begin{aligned}E(\xi I;\, \xi \in \cdot) &= E\Big\{\zeta_B I;\, \eta_B + \zeta_B \in \cdot\Big\}\\ &= \mathcal{L}(\eta_B) * E(\zeta_B I;\, \zeta_B \in \cdot),\end{aligned}$$

and so for $f \in \mathcal{S}_+$ and $s \in B$ a.e. $E\xi$,

$$\mathcal{L}(\xi f \,\|\, \xi)_s = \mathcal{L}(\xi f \,|\, \xi B = 0) * \mathcal{L}(\zeta_B f \,\|\, \zeta_B)_s, \tag{16}$$

provided that the same supporting measure is used on both sides. Since $\mathcal{I}$ is countable, this holds simultaneously for all $B \in \mathcal{I}$, outside a fixed $E\xi$-null set. As $B \downarrow \{s\}$ along $\mathcal{I}$, for a non-exceptional $s \in S$, we have $\mathcal{L}(\xi f \,\|\, \xi B = 0) \xrightarrow{w} \mathcal{L}(\xi f)$, and so $\mathcal{L}(\zeta_B f \| \, \zeta_B)$ converges vaguely to some measure ν_s, by the continuity theorem for Laplace transforms. Then the same theorem shows that the convolution in (16) tends vaguely to $\mathcal{L}(\xi f) * \nu_s$, which implies $\mathcal{L}(\xi f) \prec \mathcal{L}(\xi f \,\|\, \xi)_s$. Since s and f were arbitrary, ξ is infinitely divisible by Theorem 6.17 (i).

For general ξ, define the point processes ξ_n by $\xi_n B = \sum_{s \in B} [n\xi\{s\}]$, where $[x]$ denotes the integral part of x. Since $E\xi$ is non-atomic, the stated condition implies $\mathcal{L}(\xi_n | \, \xi_n B = 0) \prec \mathcal{L}(\xi_n)$, for every $B \in \mathcal{I}$ with $P\{\xi_n B = 0\} > 0$, and so as before ξ_n is infinitely divisible. Since $n^{-1}\xi_n \uparrow \xi$, the infinite divisibility extends to ξ by Lemma 4.24. □

We conclude with a second proof of Lemma 6.14. It is enough to consider the restrictions to bounded sets, and so we may take ξ to be a mixed binomial process, based on $\hat{\lambda} = \lambda / \|\lambda\|$ and $\kappa = \|\xi\|$. Then write $\xi = \sum_{k \le \kappa} \delta_{\sigma_k}$, where $\sigma_1, \sigma_2, \ldots$ are i.i.d. $\hat{\lambda}$ and independent of κ. Choosing $\sigma = \sigma_1$, we introduce an independent random variable $\tilde{\kappa}$ with pseudo-distribution $\tilde{P}\{\tilde{\kappa} = k\} = kP\{\kappa = k\}$ on N. Putting $\tilde{\xi} = \sum_{k \le \tilde{\kappa}} \delta_{\sigma_k}$, we note that $(\sigma, \tilde{\xi} - \delta_\sigma)$ has pseudo-distribution $C_\xi^{(1)}$. Since $\tilde{\xi} - \delta_\sigma = \sum_{k=2}^{\tilde{\kappa}} \delta_{\sigma_k}$, we get $C_\xi^{(1)} = \hat{\lambda} \otimes M(\hat{\lambda}, \psi)$, where

$$\begin{aligned} \psi(t) &= \tilde{E}(1-t)^{\tilde{\kappa}-1} = E\kappa(1-t)^{\kappa-1} \\ &= -\frac{d}{dt} E(1-t)^{\kappa} \\ &= -\|\lambda\| \, \varphi'(t\|\lambda\|). \end{aligned}$$

This implies $M(\hat{\lambda}, \psi) = \|\lambda\| \, M(\lambda, -\varphi')$. □

6.4 Iterated Conditioning and Palm Recursion

Here we consider the basic iteration principles for Palm disintegration, beginning with the special case of ordinary conditioning. This corresponds to the case of point processes $\xi = \delta_\sigma$, but it also applies to random measures ξ with finite and non-random total mass $\|\xi\|$. The general case is similar but technically more sophisticated.

To explain the notation, let ξ, η, ζ be random elements in the Borel spaces S, T, U. Since $S \times U$ and $T \times U$ are again Borel, there exist some probability kernels (regular conditional distributions) $\mu\colon S \to T \times U$ and $\mu' \colon T \to S \times U$, such that a.s.

$$\mathcal{L}(\eta, \zeta \,|\, \xi) = \mu(\xi, \cdot), \qquad \mathcal{L}(\xi, \zeta \,|\, \eta) = \mu'(\eta, \cdot),$$

which amounts to the dual disintegrations

$$\mathcal{L}(\xi, \eta, \zeta) = \mathcal{L}(\xi) \otimes \mu \cong \mathcal{L}(\eta) \otimes \mu'.$$

For fixed s or t, we may regard μ_s and μ'_t as probability measures in their own right, and proceed with a second round of conditioning, leading to disintegrations of the form

$$\mu_s = \bar{\mu}_s \otimes \nu_s, \qquad \mu'_t = \bar{\mu}'_t \otimes \nu'_t,$$

with $\bar{\mu}_s = \mu_s(\cdot \times U)$ and $\bar{\mu}'_t = \mu'_t(\cdot \times U)$, for some kernels $\nu_s : T \to U$ and $\nu'_t : S \to U$. It is suggestive to write even the latter relations in the form of conditioning, as in

$$\mu_s(\tilde{\zeta} \in \cdot \,|\, \tilde{\eta}) = \nu_s(\tilde{\eta}, \cdot), \qquad \mu'_t(\tilde{\zeta} \in \cdot \,|\, \tilde{\xi}) = \nu'_t(\tilde{\xi}, \cdot),$$

where $\tilde{\xi}$, $\tilde{\eta}$, $\tilde{\zeta}$ denote the coordinate variables in S, T, U.

Since all three spaces are Borel, Theorem 1.25 allows us to choose ν_s and ν'_t to be product-measurable, hence as kernels $S \times T \to U$. With a slight abuse of notation, we may write the iterated conditioning in the form

$$\begin{aligned} P(\cdot \,|\, \eta \,|\, \xi)_{s,t} &= \{P(\cdot \,|\, \xi)_s\}(\cdot \,|\, \eta)_t, \\ P(\cdot \,|\, \xi \,|\, \eta)_{t,s} &= \{P(\cdot \,|\, \eta)_t\}(\cdot \,|\, \xi)_s. \end{aligned}$$

Substituting $s = \xi$ and $t = \eta$, and putting $\mathcal{F} = \sigma(\xi)$, $\mathcal{G} = \sigma(\eta)$, and $\mathcal{H} = \sigma(\zeta)$, we get some (ξ, η)-measurable random probability measures on $(\Omega, \mathcal{H})$, here written suggestively as

$$(P_\mathcal{F})_\mathcal{G} = P(\cdot \,|\, \eta \,|\, \xi\,)_{\xi,\eta}, \qquad (P_\mathcal{G})_\mathcal{F} = P(\cdot \,|\, \xi \,|\, \eta\,)_{\eta,\xi}.$$

Using this notation, we may state the basic commutativity properties of iterated conditioning in their most striking form. Say that a σ-field $\mathcal{F}$ is *Borel generated*, if $\mathcal{F} = \sigma(\xi)$ for some random element in a Borel space.

Theorem 6.21 *(iterated conditioning) Let $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ be Borel generated σ-fields in Ω. Then we have, a.s. on $\mathcal{H}$,*

(i) $(P_\mathcal{F})_\mathcal{G} = (P_\mathcal{G})_\mathcal{F} = P_{\mathcal{F} \vee \mathcal{G}}$,

(ii) $P_\mathcal{F} = E_\mathcal{F}(P_\mathcal{F})_\mathcal{G}$.

Here (i) should be distinguished from the elementary commutativity property

$$E_\mathcal{F} E_\mathcal{G} = E_\mathcal{G} E_\mathcal{F} = E_{\mathcal{F} \cap \mathcal{G}} \quad \text{a.s.},$$

valid iff $\mathcal{F} \perp\!\!\!\perp_{\mathcal{F} \cap \mathcal{G}} \mathcal{G}$. Likewise, (ii) should be distinguished from the elementary chain rule

$$P_\mathcal{F} = E_\mathcal{F} P_{\mathcal{F} \vee \mathcal{G}} \quad \text{a.s.}$$

Assertion (i) may seem plausible, as a kind of Fubini or Clairaut theorem for conditional distributions. Depending on viewpoint, it may also seem surprising, even contradictory. Indeed, after disintegrating twice, to form a kernel on the product space Ω^2, we are taking diagonal values (setting

$\omega_1 = \omega_2$) to form a kernel $(P_\mathcal{F})_\mathcal{G}$ on Ω. This is claimed to be equivalent to the single disintegration with respect to $\mathcal{F} \vee \mathcal{G}$. Since each kernel is only determined up to a null set, the diagonal values would seem to be arbitrary. We leave it to the reader to resolve the apparent paradox.

Proof: Writing

$$\mathcal{F} = \sigma(\xi), \quad \mathcal{G} = \sigma(\eta), \quad \mathcal{H} = \sigma(\zeta),$$

for some random elements ξ, η, ζ in the Borel spaces S, T, U, and putting

$$\mu_1 = \mathcal{L}(\xi), \qquad \mu_2 = \mathcal{L}(\eta), \qquad \mu_{12} = \mathcal{L}(\xi, \eta),$$
$$\mu_{13} = \mathcal{L}(\xi, \zeta), \qquad \mu_{123} = \mathcal{L}(\xi, \eta, \zeta),$$

we get by Theorem 1.27 the relations

$$\begin{aligned} \mu_{3|2|1} &= \mu_{3|12} \stackrel{\sim}{=} \mu_{3|1|2} \text{ a.e. } \mu_{12}, \\ \mu_{3|1} &= \mu_{2|1}\, \mu_{3|1|2} \text{ a.e. } \mu_1, \end{aligned}$$

which are equivalent to (i) and (ii). We can also obtain (ii) directly from (i), by noting that

$$E_\mathcal{F}(P_\mathcal{F})_\mathcal{G} = E_\mathcal{F} P_{\mathcal{F}\vee\mathcal{G}} = P_\mathcal{F},$$

by the chain rule for conditional expectations. □

Turning to the iteration of Palm disintegration and ordinary conditioning, we consider a random measure ξ on S and some random elements η in T and ζ in U, where S, T, U are Borel, and the Campbell measures $C_{\xi,\eta}$ and $C_{\xi,\zeta}$ are σ-finite. Regarding ξ as a random element in the Borel space $\mathcal{M}_S$, and fixing a supporting measure ρ of ξ, we may introduce the Palm measures and conditional distributions

$$\mathcal{L}(\eta, \zeta \,\|\, \xi\,)_s = \mu(s, \cdot), \qquad \mathcal{L}(\xi, \zeta \,|\, \eta\,)_t = \mu'(t, \cdot),$$

in terms of some kernels $\mu\colon S \to T \times U$ and $\mu'\colon T \to \mathcal{M}_S \times U$. For fixed s and t, we may next form the Palm and conditional measures

$$\mu_s(\tilde\zeta \in \cdot \,|\, \tilde\eta\,) = \nu_s(\tilde\eta, \cdot), \qquad \mu'_t(\tilde\zeta \in \cdot \,\|\, \tilde\xi\,)_s = \nu'_t(s, \cdot),$$

where the former conditioning makes sense, since the measures μ_s are a.e. σ-finite. By Theorem 1.25, we can choose ν and ν' to be product-measurable, hence as kernels on $S \times T$, in which case we may write suggestively

$$\begin{aligned} P(\cdot \,|\, \eta \,\|\, \xi\,)_{s,t} &= \{P(\cdot \,\|\, \xi)_s\}(\cdot \,|\, \eta\,)_t, \\ P(\cdot \,\|\, \xi \,|\, \eta\,)_{t,s} &= \{P(\cdot \,|\, \eta)_t\}(\cdot \,\|\, \xi\,)_s. \end{aligned}$$

Taking $t = \eta$ and putting $\mathcal{F} = \sigma(\eta)$ and $\mathcal{H} = \sigma(\zeta)$, we may finally introduce on S the product-measurable, measure-valued processes

$$\begin{aligned} P_{\mathcal{F}}(\cdot \,\|\, \xi\,)_s &= P(\cdot \,\|\, \xi \,|\, \eta\,)_{s,\eta}, \\ \{P(\cdot \,\|\, \xi\,)_s\}_{\mathcal{F}} &= P(\cdot \,|\, \eta \,\|\, \xi\,)_{\eta,s}. \end{aligned}$$

Under suitable integrability conditions, we show that Palm disintegration commutes with ordinary conditioning, so that the Palm kernel $P(\cdot \,\|\, \xi)$ can be obtained from its version for the conditional probability measure $P_{\mathcal{F}}$. When the random measure $E_{\mathcal{F}}\xi$ is a.s. σ-finite, it will be used as a supporting measure of ξ under $P_{\mathcal{F}}$.

Theorem 6.22 *(Palm kernels via conditioning) Let ξ be a random measure on a Borel space S, and let $\mathcal{F}$ and $\mathcal{H}$ be Borel-generated σ-fields in Ω, such that $C_{\xi,\mathcal{F}}$ and $C_{\xi,\mathcal{H}}$ are σ-finite. Fixing an unconditional supporting measure $\nu \sim E\xi$, we get on $\mathcal{H}$*

(i) $E_{\mathcal{F}}\xi$ *is a.s. σ-finite,*

(ii) $P_{\mathcal{F}}(\cdot \,\|\, \xi\,) = \{P(\cdot \,\|\, \xi\,)\}_{\mathcal{F}}$ *a.e. $C_{\xi,\mathcal{F}}$,*

(iii) $P(\cdot \,\|\, \xi\,)_s = E\{P_{\mathcal{F}}(\cdot \,\|\, \xi\,)_s \| \,\xi\,\}_s$ *a.e. $E\xi$.*

Here (iii) is the most useful, but also the most surprising of the three assertions. Indeed, after performing three disintegrations to create a kernel from S^2 to Ω, we take diagonal values (setting $s_1 = s_2$) to form a kernel on S, which is claimed to agree with the Palm kernel obtained in a single disintegration. The problem is that disintegration kernels are only determined up to a null set, so that their values along the diagonal may seem to be arbitrary. Once again, we invite the reader to resolve the apparent paradox.

It may also seem puzzling that, although the left-hand side of (ii) is independent of ν, its choice does affect the normalization of the inner disintegration kernel $P(\cdot \,\|\, \xi)$ on the right. This, however, is compensated by the normalization of the outer disintegration $\{P(\cdot \,\|\, \xi)\}_{\mathcal{F}}$, which depends reciprocally on the same measure. In (iii) there is no problem, since the same unconditional Palm disintegrations are formed on both sides.

Proof: Let $\mathcal{F} = \sigma(\eta)$ and $\mathcal{H} = \sigma(\zeta)$, for some random elements η and ζ in Borel spaces T and U, and introduce the measures

$$\begin{gathered} \mu_1 = \nu, \qquad \mu_2 = \mathcal{L}(\eta), \qquad \mu_3 = \mathcal{L}(\zeta), \\ \mu_{12} = C_{\xi,\eta}, \qquad \mu_{13} = C_{\xi,\zeta}, \qquad \mu_{123} = C_{\xi,\eta,\zeta}, \end{gathered}$$

which are all σ-finite. Here all support conditions are clearly fulfilled, and so by Theorem 1.27,

$$\mu_{3|2|1} \overset{\sim}{=} \mu_{3|1|2} \text{ a.e. } \mu_{12}, \qquad \mu_{3|1} = \mu_{2|1}\, \mu_{3|1|2} \text{ a.e. } \mu_1. \tag{17}$$

Further note that

$$\begin{aligned}\mu_{12} = C_{\xi,\eta} &\overset{\sim}{=} \mathcal{L}(\eta) \otimes E_{\mathcal{F}}\xi \\ &= \mu_2 \otimes E_{\mathcal{F}}\xi,\end{aligned}$$

which implies $\mu_{1|2} = E_{\mathcal{F}}\xi$ a.s., by the uniqueness in Theorem 1.23. In particular, this implies (i), and the two relations in (17) are equivalent to (ii) and (iii). □

We finally consider the iterated Palm disintegrations with respect to two random measures ξ and η, defined on some Borel spaces S and T. Here the construction is similar, and we write

$$P(\cdot \,\|\, \eta \,\|\, \xi\,)_{s,t} = \{P(\cdot \,\|\, \xi\,)_s\}(\cdot \,\|\, \eta\,)_t, \quad s \in S,\ t \in T.$$

Again, we show that the two operations commute.

Theorem 6.23 *(Palm iteration) Let ξ and η be random measures on some Borel spaces S and T, and let $\mathcal{H}$ be a Borel-generated σ-field in Ω, such that $C_{\xi\otimes\eta,\mathcal{H}}$ is σ-finite. Fix any supporting measures μ_1 of ξ, μ_2 of η, and μ_{12} of $\xi \otimes \eta$, with associated disintegrations $\mu_{12} = \mu_1 \otimes \mu_{2|1} \overset{\sim}{=} \mu_2 \otimes \mu_{1|2}$. Then the corresponding Palm kernels on $\mathcal{H}$ satisfy the a.e. relations*

(i) $\mu_{2|1} \sim E(\eta \,\|\, \xi\,)$, $\quad \mu_{1|2} \sim E(\xi \,\|\, \eta\,)$,

(ii) $P(\cdot \,\|\, \xi \,\|\, \eta\,)_{s,t} = P(\cdot \,\|\, \eta \,\|\, \xi\,)_{t,s} = P(\cdot \,\|\, \xi \otimes \eta\,)_{s,t}$.

If the intensity measures are σ-finite, we can choose $\mu_1 = E\xi$, $\mu_2 = E\eta$, and $\mu_{12} = E(\xi \otimes \eta)$, in which case the a.s. relations in (i) *become equalities.*

Here the main assertion is (ii), which shows how the Palm kernel with respect to a product random measure $\xi \otimes \eta$ can be formed by iterated disintegration, with respect to one random measure at a time.

Proof: Assume $\mathcal{H} = \sigma(\zeta)$ for some random element ζ in a Borel space U, and define $\mu_{123} = C_{\xi\otimes\eta,\zeta}$ on $S \times T \times U$. Then all support properties are satisfied for μ_1, μ_2, μ_{12}, μ_{123}, and so (i) and (ii) hold by Theorem 1.27. To prove the last assertion, let the intensities be σ-finite, and write for any $B \in \mathcal{S}$ and $C \in \mathcal{T}$

$$\begin{aligned}E(\xi \otimes \eta)(B \times C) &= \int_B E\xi(ds)\, E(\eta C \,\|\, \xi)_s \\ &= \Big\{E\xi \otimes E(\eta \,\|\, \xi)\Big\}\,(B \times C),\end{aligned}$$

which extends to $E(\xi \otimes \eta) = E\xi \otimes E(\eta \,\|\, \xi)$. This gives $\mu_{2|1} = E(\eta \,\|\, \xi)$ a.s., and the other relation follows by the symmetric argument. □

In particular, for any random measure ξ on S, the multi-variate Palm measures $\mathcal{L}(\,\xi \,\|\, \xi^{\otimes n})$ can be obtained recursively, by disintegration in n steps. More generally, we have the following result:

Corollary 6.24 *(Palm recursion for product random measures) Let ξ be a random measure on a Borel space S, and fix any $m, n \in \mathbb{N}$. Given a supporting measure ν_m of $\xi^{\otimes m}$, with associated Palm kernel $\mu_m = \mathcal{L}(\xi \,\|\, \xi^{\otimes m})$, choose a supporting kernel $\nu_{n|m} \sim E(\,\xi^{\otimes n} \| \,\xi^{\otimes m})$, with associated iterated Palm kernel $\mu_{n|m} = \mathcal{L}(\,\xi \,\|\, \xi^{\otimes n} \| \,\xi^{\otimes m})$. Then a supporting measure ν_{m+n} for $\xi^{\otimes(m+n)}$, with associated Palm kernel μ_{m+n}, are given by*

(i) $\nu_{m+n} = \nu_m \otimes \nu_{n|m} \sim E\xi^{\otimes(m+n)}$,

(ii) $\mu_{m+n} = \mathcal{L}(\,\xi \,\|\, \xi^{\otimes(m+n)}) = \mathcal{L}(\,\xi \,\|\, \xi^{\otimes n} \| \,\xi^{\otimes m})$.

When the moment measures are σ-finite, we may take

(iii) $\nu_m = E\xi^{\otimes m}, \quad \nu_{n|m} = E(\,\xi^{\otimes n} \| \,\xi^{\otimes m}), \quad \nu_{m+n} = E\xi^{\otimes(m+n)}$.

This is an immediate consequence of Theorem 6.23. We may also give a short direct proof, as follows:

Proof: Writing

$$\nu_{m,n} = \nu_m \otimes \nu_{n|m}, \quad Q_s^m = \mathcal{L}(\xi \,\|\, \xi^{\otimes m})_s, \quad Q_{s,t}^{m,n} = \mathcal{L}(\xi \,\|\, \xi^{\otimes n} \| \,\xi^{\otimes m})_{s,t},$$

we get for any measurable function $f \geq 0$ on $S^{m+n} \times \mathcal{M}_S$

$$\begin{aligned}
(\nu_{m,n} \otimes Q^{m,n})f &= \int \nu_m(ds) \int \nu_{n|m}(s, dt) \int Q_{s,t}^{m,n}(d\mu)\, f(s,t,\mu) \\
&= \int \nu_m(ds) \int Q_s^m(d\mu) \int \mu^{\otimes n}(dt)\, f(s,t,\mu) \\
&= E \int \xi^{\otimes m}(ds) \int \xi^{\otimes n}(dt)\, f(s,t,\xi) = C_\xi^{m+n} f,
\end{aligned}$$

where C_ξ^{m+n} denotes the Campbell measure for ξ of order $m+n$. This shows that $C_\xi^{m+n} = \nu_{m,n} \otimes Q^{m,n}$. In particular, $\xi^{\otimes m+n}$ has supporting measure $\nu_{m+n} = \nu_{m,n}$, as asserted in (i), and the associated Palm measures $\mathcal{L}(\xi \,\|\, \xi^{\otimes m+n})_{s,t}$ satisfy (ii). If the moment measures are σ-finite, then (iii) holds by the same calculation, with $f(s,t,\mu)$ independent of μ. □

In the point process case, a similar recursive construction applies to the reduced, multi-variate Palm kernels, which justifies our definitions. For simplicity, we may take ξ to be the identify map on $\Omega = \mathcal{N}_S$.

Theorem 6.25 *(reduced Palm recursion for point processes) Let ξ be a point process on a Borel space S, and fix any $m, n \in \mathbb{N}$. Given a supporting measure ν_m of $\xi^{(m)}$, with associated reduced Palm kernel μ_m of ξ, choose a supporting kernel $\nu_{n|m}$ of $\xi^{(n)}$ under μ_m, with associated reduced Palm kernel $\mu_{n|m}$ of ξ. Then $\nu_{m+n} = \nu_m \otimes \nu_{n|m}$ is a supporting measure of $\xi^{(m+n)}$, and $\mu_{n|m}$ agrees a.e. with the associated reduced Palm kernel of ξ. When the factorial moment measures are σ-finite, we may take $\nu_m = E\xi^{(m)}$, $\nu_{n|m} = \mu_m \xi^{(n)}$, and $\nu_{m+n} = E\xi^{(m+n)}$.*

Proof: For clarity, we first consider the supporting measure. Given any supporting measure ν_m and kernel $\nu_{n|m}$, as stated, we get by Lemma 1.12 (iii) for any $f \in (\mathcal{S}^{\otimes n})_+$

$$\begin{aligned} E\,\xi^{(m+n)} f &= E\int \xi^{(m)}(ds) \int \Big(\xi - \textstyle\sum_{i\le m}\delta_{s_i}\Big)^{(n)}(dt)\, f(s,t) \\ &= \int \nu_m(ds)\, E\Big\{\int \Big(\xi - \textstyle\sum_{i\le m}\delta_{s_i}\Big)^{(n)}(dt)\, f(s,t)\,\Big\|\, \xi^{(m)}\Big\}_s \\ &= \int \nu_m(ds) \int E\Big\{\Big(\xi - \textstyle\sum_{i\le m}\delta_{s_i}\Big)^{(n)}\Big\|\, \xi^{(m)}\Big\}_s(dt)\, f(s,t) \\ &\sim \int \nu_m(ds) \int \nu_{n|m}(s,dt)\, f(s,t) = (\nu_m \otimes \nu_{n|m}) f, \end{aligned}$$

which shows that $\xi^{(m+n)}$ has supporting measure $\nu_m \otimes \nu_{n|m}$. When the factorial moment measures are σ-finite, we may choose $\mu_m = E\xi^{(m)}$ and $\nu_{n|m} = \mu_m \xi^{(n)}$, in which case the equivalence relation $\sim$ becomes an equality, and the last assertion follows.

To identify the associated reduced Palm kernel, write $\nu_{m,n} = \nu_m \otimes \nu_{n|m}$, and put $Q^m = \mu_m$ and $Q^{m,n} = \mu_{n|m}$. Then for any measurable functions $f \ge 0$, we get by Lemma 1.12

$$\begin{aligned} &(\nu_{m,n} \otimes Q^{m,n}) f \\ &\quad = \int \nu_m(ds) \int \nu_{m,n}(s,dt) \int Q^{m,n}_{s,t}(d\mu)\, f(s,t,\mu) \\ &\quad = \int \nu_m(ds) \int Q^m_s(d\mu) \int \mu^{(n)}(dt)\, f\Big(s,t,\mu - \textstyle\sum_{i\le n}\delta_{t_i}\Big) \\ &\quad = E\int \xi^{(m)}(ds) \int \Big(\xi - \textstyle\sum_{i\le m}\delta_{s_i}\Big)^{(n)}(dt)\, f\Big(s,t,\xi - \textstyle\sum_{i\le m}\delta_{s_i} - \textstyle\sum_{k\le n}\delta_{t_k}\Big) \\ &\quad = E\iint \xi^{(m+n)}(ds\,dt)\, f\Big(s,\xi - \textstyle\sum_{i\le m}\delta_{s_i} - \textstyle\sum_{k\le n}\delta_{t_k}\Big) = C_\xi^{(m+n)} f, \end{aligned}$$

which shows that indeed $C_\xi^{(m+n)} = \nu_{m,n} \otimes Q^{m,n}$. □

6.5 Randomizations and Cox Processes

Here we apply the preceding conditioning and iteration methods to the reduced Palm measures of randomizations and Cox processes. For any point process ξ and random measure η, we write $\mathcal{L}(\eta \,\|\, \eta^{\otimes n})$ and $\mathcal{L}(\eta \,\|\, \xi^{(n)})$ for the ordinary Palm distributions of η, with respect to the random measures $\eta^{\otimes n}$ and $\xi^{(n)}$, respectively. For the reduced Palm distributions of a point process ξ, we use the shorthand notation

$$\mathcal{L}(\xi' \,\|\, \xi^{(n)})_s = \mathcal{L}\Big\{\xi - \textstyle\sum_{i\le n}\delta_{s_i}\Big\|\, \xi^{(n)}\Big\}_s, \quad s \in S^n \ \text{ a.e. } E\xi^{(n)},$$

with the obvious extensions to suitable pairs of point processes ξ and η. In this context, the "distributions" are normalized only when the moment measures $E\eta^{\otimes n}$ and $E\xi^{(n)}$ are σ-finite, and chosen as supporting measures of $\eta^{\otimes n}$ and $\xi^{(n)}$, respectively.

We begin with the ν-randomization ζ of a point process ξ on S, where ν is a probability kernel from S to T. Here the kernel $\nu^{\otimes n}\colon S^n \to T^n$ is defined by $\nu_s^{\otimes n} = \bigotimes_{i\le n} \nu_{s_i}$, for $s = (s_1, \dots, s_n) \in S^n$.

Theorem 6.26 *(randomization) Given a probability kernel ν between two Borel spaces S and T, let ζ be a ν-randomization of a point process ξ on S. Then for any $n \in \mathsf{N}$,*

(i) $E\zeta^{(n)} = E\xi^{(n)} \otimes \nu^{\otimes n}$,

(ii) $\mathcal{L}(\,\xi \,\|\, \zeta^{(n)}) = \mathcal{L}(\,\xi \,\|\, \xi^{(n)})$ *a.e.* $E\zeta^{(n)}$,

(iii) $\mathcal{L}(\,\zeta' \,\|\, \zeta^{(n)}) = \mathcal{L}(\,\zeta' \,\|\, \xi^{(n)})$ *a.e.* $E\zeta^{(n)}$,

(iv) $\mathcal{L}(\,\zeta' \,\|\, \zeta^{(n)})$ *is a ν-randomization of* $\mathcal{L}(\,\xi' \,\|\, \xi^{(n)})$, *a.e.* $E\zeta^{(n)}$,

provided the supporting measures in (ii)$-$(iv) *are related as in* (i).

By (iv) we mean that, for $(s,t) \in (S \times T)^n$ a.e. $E\zeta^{(n)}$, a point process on $S \times T$ with distribution $\mathcal{L}(\,\zeta' \,\|\, \zeta^{(n)})_{s,t}$ is a ν-randomization of some point process on S with distribution $\mathcal{L}(\,\xi' \,\|\, \xi^{(n)})_s$. This extends the result for mixed binomial processes in Theorem 6.14, corresponding to the case of a singleton S, where ν reduces to a probability measure on T.

Proof: (i) First reduce by conditioning to the case of a non-random measure $\xi = \sum_{i\in I} \delta_{s_i}$. Then $\zeta = \sum_{i\in I} \delta_{s_i,\tau_i}$, where the τ_i are independent random elements in T with distributions ν_{s_i}, and we get for measurable functions $f \ge 0$ on $S^n \times T^n$

$$
\begin{aligned}
E\zeta^{(n)} f &= E \sum_{i\in I^{(n)}} f(s_i, \tau_i) = \sum_{i\in I^{(n)}} Ef(s_i, \tau_i) \\
&= \sum_{i\in I^{(n)}} \int \nu_{s_i}^{\otimes n}(dt)\, f(s_i, t) \\
&= \int \xi^{(n)}(ds) \int \nu_s^{\otimes n}(dt)\, f(s,t) \\
&= (\xi^{(n)} \otimes \nu^{\otimes n}) f.
\end{aligned}
$$

(ii) Letting ρ be a supporting measure of $\xi^{(n)}$, and using (i), we get for any measurable function $f \ge 0$ on $S^n \times T^n \times \mathcal{N}_S$

$$
\begin{aligned}
E \iint \zeta^{(n)}(ds\, dt)\, f(s,t,\xi) &= EE\Big\{ \iint \zeta^{(n)}(ds\, dt)\, f(s,t,\xi) \Big| \,\xi \Big\} \\
&= E \iint E(\zeta^{(n)} |\, \xi)(ds\, dt)\, f(s,t,\xi) \\
&= E \iint (\xi^{(n)} \otimes \nu^{\otimes n})(ds\, dt)\, f(s,t,\xi) \\
&= E \int \xi^{(n)}(ds) \int \nu_s^{\otimes n}(dt)\, f(s,t,\xi)
\end{aligned}
$$

$$\begin{aligned}
&= \int \rho(ds)\, E\Big\{ \int \nu_s^{\otimes n}(dt)\, f(s,t,\xi) \,\Big\|\, \xi^{(n)} \Big\}_s \\
&= \int \rho(ds) \int \nu_s^{\otimes n}(dt)\, E\big\{ f(s,t,\xi) \,\big\|\, \xi^{(n)} \big\}_s \\
&= \iint (\rho \otimes \nu^{\otimes n})(ds\, dt)\, E\big\{ f(s,t,\xi) \,\big\|\, \xi^{(n)} \big\}_s .
\end{aligned}$$

Since f was arbitrary, we obtain $E(f(\xi) \,\|\, \zeta^{(n)})_{s,t} = E(f(\xi) \,\|\, \xi^{(n)})_s$ a.e. $E\zeta^{(n)}$, for any measurable function $f \ge 0$ on $\mathcal{N}_S$, and the assertion follows.

(iv) For non-random ξ and $n = 1$, the result follows easily from Theorem 6.12 (ii), or directly from Theorem 6.14, together with the independence of the increments. Proceeding by induction based on Theorem 6.25, using the intensities as supporting measures, we obtain the statement for non-random ξ and arbitrary $n \ge 1$. To extend the result to random ξ, we may apply Theorem 6.22 (iii), with conditioning on ξ. This amounts to averaging ξ with respect to the Palm distributions $\mathcal{L}(\,\xi \,\|\, \zeta^{(n)})$. By (ii), the latter measures agree with $\mathcal{L}(\,\xi \,\|\, \xi^{(n)})$, and the assertion follows.

(iii) By (iv), we need to show that $\mathcal{L}(\zeta' \,\|\, \xi^{(n)})$ is also a ν-randomization of $\mathcal{L}(\xi' \,\|\, \xi^{(n)})$. This is obvious for non-random ξ, and the general result follows by conditioning on ξ, using Theorem 6.22 (iii). □

We turn to the case of Cox processes.

Theorem 6.27 *(Cox processes) Let ξ be a Cox process on a Borel space S, directed by the random measure η. Then for any $n \in \mathbb{N}$,*

(i) $E\xi^{(n)} = E\eta^{\otimes n}$,

(ii) $\mathcal{L}(\,\eta \,\|\, \xi^{(n)}) = \mathcal{L}(\,\eta \,\|\, \eta^{\otimes n})$ *a.e.* $E\eta^{\otimes n}$,

(iii) $\mathcal{L}(\,\xi' \,\|\, \xi^{(n)}) = \mathcal{L}(\,\xi \,\|\, \eta^{\otimes n})$ *a.e.* $E\eta^{\otimes n}$,

(iv) $\mathcal{L}(\,\xi' \,\|\, \xi^{(n)})$ *is a Cox process directed by* $\mathcal{L}(\,\eta \,\|\, \eta^{\otimes n})$, *a.e.* $E\eta^{\otimes n}$,

provided the supporting measures in (ii)$-$(iv) *are related as in* (i).

By (iv) we mean that, for $s \in S^n$ a.e. $E\xi^{(n)}$, a point process on S with distribution $\mathcal{L}(\,\xi' \,\|\, \xi^{(n)})_s$ is a Cox process, directed by some random measure on S with distribution $\mathcal{L}(\,\eta \,\|\, \eta^{\otimes n})_s$. This extends the result for mixed Poisson processes in Theorem 6.14, corresponding to the case where $\eta = \rho\lambda$, for some σ-finite measure λ on S and random variable $\rho \ge 0$. If ξ is a Poisson process with intensity $\mu = E\xi$, then (i) reduces to $E\xi^{(n)} = \mu^{\otimes n}$, and we may apply (iv) to the function $g(s,m) = f(s, m + \sum_i \delta_{s_i})$, to get

$$E \int \xi^{(n)}(ds)\, f(s,\xi) = \int \mu^{\otimes n}(ds)\, Ef\Big(s,\, \xi + \textstyle\sum_{i \le n} \delta_{s_i}\Big), \tag{18}$$

which extends the Slivnyak property in Theorem 6.15.

Proof, (i) & (iv) for non-random η: For $n = 1$, (i) holds by hypothesis and (iv) holds by Theorem 6.14. To extend the results to general $n \geq 1$, we may proceed by induction based on Theorem 6.25, using the intensities as supporting measures.

(i) This follows from the previous case, by conditioning on η.

(ii) Fixing a supporting measure ρ for $\xi^{(n)}$, and using (i), we get for any measurable function $f \geq 0$ on $S \times \mathcal{N}_S$

$$\begin{aligned} E\int \xi^{(n)}(ds)\, f(s,\eta) &= EE\Big\{\int \xi^{(n)}(ds)\, f(s,\eta)\,\Big|\,\eta\Big\} \\ &= E\int E\{\xi^{(n)}|\,\eta\}(ds)\, f(s,\eta) \\ &= E\int \eta^{\otimes n}(ds)\, f(s,\eta) \\ &= \int \rho(ds)\, E\Big\{f(s,\eta)\,\Big\|\,\eta^{\otimes n}\Big\}_s. \end{aligned}$$

Since f was arbitrary, we obtain $E\{f(\eta)\,\|\,\xi^{(n)}\} = E\{f(\eta)\,\|\,\eta^{\otimes n}\}$ a.e. $E\xi^{(n)}$, and the assertion follows.

(iv) To extend the result to random η, we may apply Theorem 6.22 (iii), with conditioning on η. This amounts to averaging η with respect to the Palm distributions $\mathcal{L}(\eta\,\|\,\xi^{(n)})$. By (ii), the latter measures agree with $\mathcal{L}(\eta\,\|\,\eta^{\otimes n})$, and the assertion follows.

(iii) By (iv), we need to show that $\mathcal{L}(\xi\,\|\,\eta^{\otimes n})$ is also a Cox process directed by $\mathcal{L}(\eta\,\|\,\eta^{\otimes n})$. This is obvious for non-random η, and the general result follows by conditioning on η, using Theorem 6.22 (iii). □

The last two theorems lead to an interesting extension of Theorem 6.14. Here we refer to Lemma 3.36, for the characterizations of λ-symmetric, marked point processes in terms of randomizations and Cox processes.

Corollary 6.28 *(exchangeable point processes) For any Borel spaces S and T, and a diffuse measure λ on T with $\|\lambda\| \in \{1, \infty\}$, let ξ be a λ-symmetric, marked point process on $S \times T$, directed by β or ν, respectively.*

(i) *When $\|\lambda\| = 1$, we have $E\xi^{(n)} = E\beta^{(n)} \otimes \lambda^{\otimes n}$, and $\mathcal{L}(\xi'\,\|\,\xi^{(n)})$ is a.e. λ-symmetric and directed by $\mathcal{L}(\beta'\,\|\,\beta^{(n)})$.*

(ii) *When $\|\lambda\| = \infty$, we have $E\xi^{(n)} = E\nu^{\otimes n} \otimes \lambda^{\otimes n}$, and $\mathcal{L}(\xi'\,\|\,\xi^{(n)})$ is a.e. λ-symmetric and directed by $\mathcal{L}(\nu\,\|\,\nu^{\otimes n})$.*

In either case, we need the supporting measures to be related as the associated intensities. Our proof is based on a simple lemma, where the same convention applies.

Lemma 6.29 *(λ-invariance) Consider a random measure ξ on S, a σ-finite measure λ on T, and a random element η in U, where S, T, and U are Borel. Then*

$$\mathcal{L}(\eta \,\|\, \xi \otimes \lambda)_{s,t} = \mathcal{L}(\eta \,\|\, \xi)_s, \quad (s,t) \in S \times T \ \text{ a.e. } E\xi \otimes \lambda.$$

Proof: Let ρ be a supporting measure of ξ. Then for any measurable function $f \geq 0$ on $S \times T \times U$, we have

$$\begin{aligned} E \iint (\xi \otimes \lambda)(ds\, dt)\, f(s,t,\eta) &= E \int \xi(ds) \int \lambda(dt)\, f(s,t,\eta) \\ &= \int \rho(ds)\, E\Big\{ \int \lambda(dt)\, f(s,t,\eta) \Big\| \, \xi \Big\}_s \\ &= \int \rho(ds) \int \lambda(dt)\, E\{f(s,t,\eta) \,\|\, \xi\}_s \\ &= \iint (\rho \otimes \lambda)(ds\, dt)\, E\{f(s,t,\eta) \,\|\, \xi\}_s, \end{aligned}$$

and the assertion follows. □

Proof of Corollary 6.28: (i) Since ξ is a λ-randomization of β, Theorem 6.26 shows that $E\xi^{(n)} = E\beta^{(n)} \otimes \lambda^{\otimes n}$, and also that $\mathcal{L}(\xi' \,\|\, \xi^{(n)})$ is a.e. a λ-randomization of $\mathcal{L}(\beta' \,\|\, \beta^{(n)})$, which implies the stated property.

(ii) Since ξ is a Cox process directed by $\nu \otimes \lambda$, Theorem 6.27 shows that

$$E\xi^{(n)} = E(\nu \otimes \lambda)^{\otimes n} = E\nu^{\otimes n} \otimes \lambda^{\otimes n},$$

and that $\mathcal{L}(\xi' \,\|\, \xi^{(n)})_s$ is a.e. a Cox process directed by

$$\mathcal{L}\Big\{ \nu \otimes \lambda \,\|\, (\nu \otimes \lambda)^{\otimes n} \Big\}_s = \mathcal{L}(\nu \otimes \lambda \,\|\, \nu^{\otimes n})_s = \mathcal{L}(\nu_s \otimes \lambda),$$

where $\mathcal{L}(\nu_s) = \mathcal{L}(\nu \,\|\, \nu^{\otimes n})_s$. Here the equalities hold by Lemma 6.29, and the substitution rule for integrals (FMP 1.22). Thus, $\mathcal{L}(\xi' \,\|\, \xi^{(n)})$ is a.e. λ-symmetric and directed by $\mathcal{L}(\eta \,\|\, \eta^{\otimes n})$. □

Given a point process $\zeta = \sum_i \delta_{\tau_i}$ on T and a probability kernel $\nu \colon T \to \mathcal{M}_S$, we may form an associated *cluster process* $\xi = \sum_i \xi_i$, where the ξ_i are conditionally independent random measures on S with distributions ν_{τ_i}. We assume ζ and ν to be such that ξ is a.s. locally finite, hence a random measure on S. If ζ is Cox and directed by a random measure η, we call ξ a *Cox cluster process generated by ζ and ν*. In this case, $\chi = \sum_i \delta_{\xi_i}$ is again Cox and directed by $\eta\nu$. If $\eta = \mu$ is non-random, so that χ is Poisson, then ξ becomes infinitely divisible with Lévy measure $\lambda = \mu\nu$. Here we write $\tilde{\xi}$ for a random measure with pseudo-distribution λ.

The following result, extending the elementary Theorem 6.16, gives the multi-variate moment measures and Palm distributions of a Cox cluster process, needed in the context of super-processes. For any σ-finite measure $\mu = \sum_n \mu_n$, the functions $p_n = d\mu_n/d\mu$ are referred to as *relative densities* of

μ_n with respect to μ. For a cluster process generated by a Cox process with directing measure η, we write $\mathcal{L}_\mu$ and E_μ for the conditional distributions and expectations, given $\eta = \mu$. The convolution $*$ is defined with respect to addition in $\mathcal{M}_S$, and we write $(*)_J$ for the convolution of a set of measures indexed by J.

Theorem 6.30 *(Cox cluster processes) Let ξ be a cluster process on S, generated by a Cox process on T directed by η, and a probability kernel $\nu: T \to \mathcal{M}_S$, where S and T are Borel. For non-random $\eta = \mu$, we have*

(i) $$E_\mu\, \xi^{\otimes n} = \sum_{\pi \in \mathcal{P}_n} \bigotimes_{J \in \pi} E_\mu\, \tilde{\xi}^{\otimes J}, \quad n \in \mathsf{N},$$

and when $E_\mu \xi^{\otimes n}$ is σ-finite with relative densities p_μ^π given by (i), *we have*

(ii) $$\mathcal{L}_\mu(\xi \,\|\, \xi^{\otimes n})_s = \mathcal{L}_\mu(\xi) * \sum_{\pi \in \mathcal{P}_n} p_\mu^\pi(s) \underset{J \in \pi}{(*)} \mathcal{L}_\mu(\tilde{\xi} \,\|\, \tilde{\xi}^{\otimes J})_{s_J}.$$

The formulas for general η are obtained by averaging with respect to $\mathcal{L}(\eta)$ and $\mathcal{L}(\eta \,\|\, \xi^{\otimes n})_s$, respectively.

Here we assume $\mathcal{L}(\xi \,\|\, \xi^{\otimes n})$ and $\mathcal{L}(\eta \,\|\, \xi^{\otimes n})$ to be based on a common supporting measure for $\xi^{\otimes n}$. For a probabilistic interpretation of (ii), when $\eta = \mu$ and $s \in S^n$ are fixed, let ξ_s have distribution $\mathcal{L}(\xi \,\|\, \xi^{\otimes n})_s$, choose a random partition $\pi_s \perp\!\!\!\perp \xi$ in $\mathcal{P}_n$, distributed according to the relative densities $p_\mu^\pi(s)$ in (i), and consider some independent Palm versions $\tilde{\xi}_{s_J}^J$ of $\tilde{\xi}$, $J \in \pi_s$. Then (ii) is equivalent to

$$\xi_s \overset{d}{=} \xi + \sum_{J \in \pi_s} \tilde{\xi}_{s_J}^J, \quad s \in S^n \text{ a.e. } E_\mu \xi^{\otimes n}.$$

The result extends immediately to Palm measures of the form $\mathcal{L}(\xi' \,\|\, \xi^{\otimes n})_s$, where ξ and ξ' are random measures on S and S', respectively, such that the pair (ξ, ξ') is a Cox cluster process directed by η. Indeed, taking S and S' to be disjoint, we may apply the previous result to the cluster process $\xi'' = \xi + \xi'$ on $S'' = S \cup S'$.

Proof: Writing $\xi = \sum_{i \in I} \xi_i$ for the cluster representation of ξ, we note that $\chi = \sum_i \delta_{\xi_i}$ is a ν-randomization of the Cox process ζ with $\chi \perp\!\!\!\perp_\zeta \eta$. Since χ is again Cox and directed by $\eta\nu$, by Lemma 3.2 (i), we may assume that $T = \mathcal{M}_S$ and $\chi = \zeta$. We may further take $\eta = \mu$ to be non-random, so that ζ becomes Poisson with intensity μ, since the general case will then follow by Theorem 6.22 (iii).

Since $\xi = \sum_i \xi_i$ and $\zeta = \sum_i \delta_{\xi_i}$, we note that

$$\begin{aligned} \xi^{\otimes n} &= \left(\textstyle\sum_{i \in I} \xi_i\right)^{\otimes n} = \sum_{i \in I^n} \bigotimes_{k \le n} \xi_{i_k} \\ &= \sum_{\pi \in \mathcal{P}_n} \sum_{i \in I^{(\pi)}} \bigotimes_{J \in \pi} \xi_{i_J}^{\otimes J} \\ &= \sum_{\pi \in \mathcal{P}_n} \int \zeta^{(\pi)}(dm_\pi) \bigotimes_{J \in \pi} m_J^{\otimes J}, \end{aligned} \tag{19}$$

where $m_\pi \in \mathcal{M}_S^\pi$ with components m_J, $J \in \pi$. We further need the Palm disintegration

$$\begin{aligned} E_\mu\, g(\tilde{\xi})\, \tilde{\xi}^{\otimes n} f &= \int \mu(dm)\, g(m)\, m^{\otimes n} f \\ &= \int E_\mu\, \tilde{\xi}^{\otimes n}(ds)\, f(s)\, E_\mu g(\tilde{\xi}_s), \end{aligned} \tag{20}$$

where $f, g \geq 0$ are measurable functions on S^n and $\mathcal{M}_S$, respectively, and $\tilde{\xi}_s$ is such that $\mathcal{L}(\tilde{\xi}_s) = \mathcal{L}(\tilde{\xi} \,\|\, \tilde{\xi}^{\otimes n})_s$. Assuming $f = \bigotimes_{k \leq n} f_k$, and writing $f_J = \bigotimes_{j \in J} f_j$, we get

$$\begin{aligned} E_\mu\, g(\xi)\, \xi^{\otimes n} f &= E_\mu\, g(\xi) \sum_{\pi \in \mathcal{P}_n} \int \zeta^{(\pi)}(dm_\pi) \prod_{J \in \pi} m_J^{\otimes J} f_J \\ &= \sum_{\pi \in \mathcal{P}_n} E_\mu \int \cdots \int g\Big(\xi + \textstyle\sum_{J \in \pi} m_J\Big) \displaystyle\prod_{J \in \pi} \mu(dm_J)\, m_J^{\otimes J} f_J \\ &= \sum_{\pi \in \mathcal{P}_n} \int \cdots \int E_\mu\, g\Big(\xi + \textstyle\sum_{J \in \pi} \tilde{\xi}^J_{s_J}\Big) \displaystyle\prod_{J \in \pi} E_\mu \tilde{\xi}^{\otimes J}(ds_J)\, f_J(s_J) \\ &= \int E_\mu\, \xi^{\otimes n}(ds)\, f(s) \sum_{\pi \in \mathcal{P}_n} p_\mu^\pi(s)\, E_\mu\, g\Big(\xi + \textstyle\sum_{J \in \pi} \tilde{\xi}^J_{s_J}\Big), \end{aligned}$$

where the first equality holds by the diagonal decomposition in (19), the second equality holds by the Slivnyak formula in (18), and the third step holds by the Palm disintegration in (20). Taking $g \equiv 1$ in the resulting formula gives

$$E_\mu\, \xi^{\otimes n} f = \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} E_\mu\, \tilde{\xi}^{\otimes J} f_J,$$

which extends to (i) by a monotone-class argument. This justifies the density representation

$$\bigotimes_{J \in \pi} E_\mu\, \tilde{\xi}^J = p_\mu^\pi \cdot E_\mu\, \xi^{\otimes n}, \quad \pi \in \mathcal{P}_n,$$

which yields the fourth equality. Relation (ii) now follows by the uniqueness of the Palm disintegration. □

We conclude the section with an application to exchangeable random measures. Recall from Theorem 3.37 that, for any $\lambda \in \mathcal{M}_T^*$ with $\|\lambda\| = \infty$, a random measure ξ on $S \times T$ is λ-symmetric iff it has a representation

$$\xi = \alpha \otimes \lambda + \textstyle\sum_j (\beta_j \otimes \delta_{\tau_j}) \ \text{ a.s.},$$

where α is a random measure on S, and the pairs (β_j, τ_j) form a Cox process $\zeta \perp\!\!\!\perp_\nu \alpha$ on $\mathcal{M}_S \times T$ directed by $\nu \otimes \lambda$, for a suitable random measure ν on $\mathcal{M}_S \setminus \{0\}$. Then ξ is said to be directed by the pair (α, ν).

The following result extends the version for marked point processes in Corollary 6.28. For simplicity, we assume that the invariant component $\alpha \otimes \lambda$ vanishes. In this statement only, we write $\xi^{(n)}$ for the restriction of $\xi^{\otimes n}$ to $S^n \times T^{(n)}$, which is properly contained in $(S \times T)^{(n)}$.

Theorem 6.31 *(symmetric random measures) Let ξ be a λ-symmetric random measure on $S \times T$ directed by $(0, \nu)$, where λ is a diffuse, σ-finite measure on T with $\|\lambda\| = \infty$, and ν is a random measure on $\mathcal{M}_S \setminus \{0\}$. Let $\mathcal{L}(\beta \,|\, \nu) = \nu$, and assume $\bar\beta = E(\beta \,|\, \nu)$ to be σ-finite. Then for non-random ν, we have*

(i) $E_\nu\, \xi^{(n)} = (\bar\beta \otimes \lambda)^{\otimes n}$,

(ii) $\mathcal{L}_\nu(\, \xi \,\|\, \xi^{(n)})_{s,t} = \mathcal{L}_\nu(\xi) * \underset{j \le n}{(*)} \mathcal{L}_\nu(\, \beta \otimes \delta_{t_j} \,\|\, \beta\,)_{s_j}$.

The formulas for general ν are obtained by averaging with respect to $\mathcal{L}(\nu)$ and $\mathcal{L}(\, \nu \,\|\, \bar\beta^{\otimes n})_s$, respectively.

In the general version of (ii), we assume the underlying supporting measures of $\xi^{(n)}$ and $\bar\beta^{\otimes n}$ to be related as in (i). For non-random ν, we may write (ii) in probabilistic form, as

$$\xi_{s,t} \overset{d}{=} \xi + \sum_{j \le n} (\beta^j_{s_j} \otimes \delta_{t_j}), \quad (s,t) \in S^n \times T^{(n)} \text{ a.e. } E\xi^{(n)},$$

where the $\beta^j_{s_j}$ are mutually independent Palm versions of β, also assumed to be independent of ξ.

Proof: First let ν be non-random, so that the pairs (β_j, τ_j) form a Poisson process ζ on $\mathcal{M}_S \times T$ with intensity $\nu \otimes \lambda$. Noting that $\int m\, \nu(dm) = \bar\beta$, and using Theorem 6.27 (i), we get

$$\begin{aligned} E\xi^{(n)} &= E \sum_{i \in \mathsf{N}^{(n)}} \bigotimes_{k \le n} (\beta_{i_k} \otimes \delta_{\tau_{i_k}}) \\ &= \int \!\cdots\! \int \zeta^{(n)}(dm\, dt) \bigotimes_{k \le n} (m_k \otimes \delta_{t_k}) \\ &= \int \!\cdots\! \int (\nu \otimes \lambda)^{\otimes n}(dm\, dt) \bigotimes_{k \le n} (m_k \otimes \delta_{t_k}) \\ &= (\bar\beta \otimes \lambda)^{\otimes n}, \end{aligned}$$

proving (i). For general ν, we then obtain $E(\xi^{(n)} |\, \nu) = (\bar\beta \otimes \lambda)^{\otimes n}$ a.s., and so by Lemma 6.29,

$$\begin{aligned} \mathcal{L}(\, \nu \,\|\, \xi^{(n)})_{s,t} &= \mathcal{L}\Big\{ \nu \,\Big\|\, (\bar\beta \otimes \lambda)^{\otimes n} \Big\}_{s,t} \\ &= \mathcal{L}\Big\{ \nu \,\Big\|\, \bar\beta^{\otimes n} \Big\}_s, \end{aligned}$$

which proves the last assertion, given the Poisson version of (ii).

To prove the latter, we see from (i) that $E_\nu \xi^{(n)} = (E_\nu \tilde\xi)^{\otimes n}$, and so the formula in Theorem 6.30 (ii) reduces to

$$\mathcal{L}_\nu(\xi \,\|\, \xi^{(n)})_{s,t} = \mathcal{L}_\nu(\xi) * \underset{i \le n}{(*)}\, \mathcal{L}_\nu(\tilde\xi \,\|\, \tilde\xi)_{s_i,t_i}. \tag{21}$$

To identify the Palm measures on the right, consider a random pair (β, τ) in $\mathcal{M}_S \times T$ with pseudo-distribution $\nu \otimes \lambda$, and assume that $E\beta$ is σ-finite. Then for any measurable function $f \geq 0$ on $S \times \mathcal{M}_S \times T^2$,

$$\begin{aligned} E \int \cdots \int (\beta \otimes \delta_\tau)(ds\, dt)\, f(s, \beta, t, \tau) \\ &= \int \lambda(dt)\, E \int \beta(ds)\, f(s, \beta, t, t) \\ &= \int \lambda(dt) \int E\beta(ds)\, E\{f(s, \beta, t, t) \,\|\, \beta\}_s. \end{aligned}$$

Since f was arbitrary and $E(\beta \otimes \delta_\tau) = E\beta \otimes \lambda$, we obtain

$$\mathcal{L}(\beta, \tau \,\|\, \beta \otimes \delta_\tau)_{s,t} = \mathcal{L}(\,\beta, t \,\|\, \beta)_s, \quad (s,t) \in S \times T \text{ a.e. } E\beta \otimes \lambda,$$

and so we have, a.e. $E\beta \otimes \lambda$,

$$\begin{aligned} \mathcal{L}(\tilde{\xi} \,\|\, \tilde{\xi})_{s,t} &= \mathcal{L}(\beta \otimes \delta_\tau \,\|\, \beta \otimes \delta_\tau)_{s,t} \\ &= \mathcal{L}(\beta \otimes \delta_t \,\|\, \beta)_s. \end{aligned}$$

Inserting this into (21) yields (ii). □

6.6 Local Hitting and Conditioning

For a simple point process ξ on a space S, the associated Palm probabilities and expectations $P(A \,\|\, \xi)_s$ and $E(\eta \,\|\, \xi)_s$ may be thought of as conditional probabilities and expectations, given that ξ has an atom at s. Though the latter events have typically probability 0, the stated interpretation still holds approximately, in the sense of elementary conditioning on the events $\xi B > 0$ or $\xi B = 1$, as we let $B \downarrow \{s\}$ along some sets $B \subset S$ with $E\xi B > 0$. Our argument also yields some related approximations for the associated hitting probabilities $P\{\xi B > 0\}$ and $P\{\xi B = 1\}$.

The present discussion is based on the general differentiation theory in Section 1.4, where the notions of standard differentiation basis and dominating dissection system are defined. When $E\xi B \in (0, \infty)$, we write $1_B \tilde{\xi} = \delta_{\tau_B}$, where $\mathcal{L}(\tau_B) = E 1_B \xi / E\xi B$.

Theorem 6.32 *(local approximation) Let ξ be a simple point process on a Borel space S, such that $E\xi$ is σ-finite, and fix a standard differentiation basis $\mathcal{I}$ for $E\xi$. Then for $s \in S$ a.e. $E\xi$, we have as $B \downarrow \{s\}$ along $\mathcal{I}$*

(i) $P\{\xi B > 0\} \sim P\{\xi B = 1\} \sim E\xi B$,

(ii) $\mathcal{L}(1_B \xi \,|\, \xi B > 0) \overset{u}{\approx} \mathcal{L}(1_B \xi \,|\, \xi B = 1) \overset{u}{\approx} \mathcal{L}(1_B \tilde{\xi})$,

(iii) *for any random variable $\eta \geq 0$, such that even $E\eta\,\xi$ is σ-finite,*

$$E(\eta \,|\, \xi B > 0) \approx E(\eta \,|\, \xi B = 1) \to E(\,\eta \,\|\, \xi\,)_s.$$

Proof, (i) & (iii): Choose a dominating dissection system (I_{nj}) as in Lemma 1.30. Putting $I_n(s) = I_{nj}$ when $s \in I_{nj}$, we define some simple point processes $\xi_n \le \xi$ on S, for $B \in \mathcal{S}$, by

$$\begin{aligned}\xi_n B &= \int_B 1\{\xi I_n(s) = 1\}\, \xi(ds)\\ &= \sum_{s\in B} \delta_s\, 1\{\xi\{s\} = \xi I_n(s) = 1\}.\end{aligned}$$

Since ξ is simple and (I_{nj}) is separating, we have $\xi_n \uparrow \xi$, and so by monotone convergence $E\eta\,\xi_n \uparrow E\eta\,\xi$. Since $E\eta\,\xi_n \le E\eta\,\xi \ll E\xi$, where all three measures are σ-finite, the Radon–Nikodym theorem yields $E\eta\,\xi_n = f_n \cdot E\xi$, for some increasing sequence of measurable functions $f_n \ge 0$. Assuming $f_n \uparrow f$, we get by monotone convergence $E\eta\,\xi = f \cdot E\xi$, and so for any $B \in \mathcal{S}$,

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

which implies $f(s) = E(\eta \,\|\, \xi)_s$ a.e. $E\xi$.

By the choice of (I_{nj}), there exist for every $n \in \mathsf{N}$, and for $s \in S$ a.e. $E\xi$, some sets $B \in \mathcal{I}_s$ with $E\xi(B \setminus I_n(s)) = 0$. Since $\xi_n B \le 1$ a.s., we get

$$\begin{aligned}E\eta\,\xi_n B &= E(\eta;\, \xi_n B > 0)\\ &\le E(\eta;\, \xi B > 0) \le E\eta\,\xi B,\end{aligned}$$

and so

$$\frac{E\eta\,\xi_n B}{E\xi B} \le \frac{E(\eta;\, \xi B > 0)}{E\xi B} \le \frac{E\eta\,\xi B}{E\xi B}.$$

For fixed n and $s \in S$ a.e. $E\xi$, the differentiation property gives

$$\begin{aligned}f_n(s) &\le \liminf_{B\downarrow\{s\}} \frac{E(\eta;\, \xi B > 0)}{E\xi B}\\ &\le \limsup_{B\downarrow\{s\}} \frac{E(\eta;\, \xi B > 0)}{E\xi B} \le f(s).\end{aligned}$$

Since $f_n \uparrow f$, we get as $B \downarrow \{s\}$ along $\mathcal{I}$

$$\frac{E(\eta;\, \xi B > 0)}{E\xi B} \approx \frac{E\eta\,\xi B}{E\xi B} \to f(s), \quad s \in S \text{ a.e. } E\xi. \tag{22}$$

Noting that

$$\begin{aligned}0 &\le 1\{k > 0\} - 1\{k = 1\}\\ &= 1\{k > 1\} \le k - 1\{k > 0\}, \quad k \in \mathsf{Z}_+,\end{aligned}$$

we see from (22) that

$$\begin{aligned}0 &\le \frac{E(\eta;\, \xi B > 0)}{E\xi B} - \frac{E(\eta;\, \xi B = 1)}{E\xi B}\\ &\le \frac{E\eta\,\xi B}{E\xi B} - \frac{E(\eta;\, \xi B > 0)}{E\xi B} \to 0,\end{aligned}$$

and using (22) again gives

$$\frac{E(\eta;\,\xi B=1)}{E\xi B}\approx\frac{E(\eta;\,\xi B>0)}{E\xi B}\to f(s),\quad s\in S\ \text{ a.e. } E\xi.$$

Here (i) follows for $\eta\equiv 1$, and dividing the two versions yields

$$\frac{E(\eta;\,\xi B=1)}{P\{\xi B=1\}}\approx\frac{E(\eta;\,\xi B>0)}{P\{\xi B>0\}}\to f(s),\quad s\in S\ \text{ a.e. } E\xi,$$

which is equivalent to (iii).

(ii) Using (i), along with some basic definitions and elementary estimates, we get

$$\begin{aligned}
&\Big\|\mathcal{L}(1_B\xi\,|\,\xi B=1)-\mathcal{L}(1_B\tilde{\xi})\Big\|\\
&\quad=\|\mathcal{L}(\sigma_B|\,\xi B=1)-\mathcal{L}(\tau_B)\|\\
&\quad=\left\|\frac{E(1_B\xi;\,\xi B=1)}{P\{\xi B=1\}}-\frac{E1_B\xi}{E\xi B}\right\|\\
&\quad\le\frac{E(\,\xi B;\,\xi B>1)}{E\xi B}+E\xi B\left|\frac{1}{P\{\xi B=1\}}-\frac{1}{E\xi B}\right|\\
&\quad=2\left(\frac{E\xi B}{P\{\xi B=1\}}-1\right)\to 0,
\end{aligned}$$

$$\begin{aligned}
&\Big\|\mathcal{L}(1_B\xi\,|\,\xi B>0)-\mathcal{L}(1_B\xi\,|\,\xi B=1)\Big\|\\
&\quad\le\frac{P\{\xi B>1\}}{P\{\xi B=1\}}+P\{\xi B>0\}\left|\frac{1}{P\{\xi B=1\}}-\frac{1}{P\{\xi B>0\}}\right|\\
&\quad=2\left(\frac{P\{\xi B>0\}}{P\{\xi B=1\}}-1\right)\to 0,
\end{aligned}$$

and the assertion follows by combination. □

One might expect the previous approximations to extend to any measurable process $Y\ge 0$ on S, so that

$$E(Y_s|\,\xi B>0)\approx E(Y_s|\,\xi B=1)\to E(Y_s\|\,\xi)_s.$$

Though the latter statement fails in general, it becomes true if we replace Y_s in $E(Y_s|\,\xi B=1)$ by the composition $Y(\sigma_B)$, where σ_B denotes the unique atom of ξ in B. Writing $Y_B=1_BY$, we note that $\xi Y_B=Y(\sigma_B)$ on $\{\xi B=1\}$, which suggests that we replace Y_s by the integral ξY_B. This leads to the following useful approximation.

Theorem 6.33 *(integral approximation) Consider a simple point process ξ and a measurable process Y on a Borel space S, where $E\xi$ and $E(Y\cdot\xi)$ are σ-finite, and fix a standard differentiation basis $\mathcal{I}$ for $E\xi$. Then for $s\in S$ a.e. $E\xi$, we have as $B\downarrow\{s\}$ along $\mathcal{I}$*

$$E(\xi Y_B|\,\xi B>0)\approx E(\xi Y_B|\,\xi B=1)\to E(\,Y_s\,\|\,\xi\,)_s.$$

Proof: Noting that $E\xi Y_B = \int_B E\xi(ds)\, E(Y_s \,\|\, \xi)_s$, we get $E(Y\cdot\xi) = f\cdot E\xi$ with $f(s) = E(Y_s \,\|\, \xi)_s$. Using the differentiation property of $\mathcal{I}$ and Theorem 6.32 (i), we get for $s \in S$ a.e. $E\xi$, as $B \downarrow \{s\}$ along $\mathcal{I}$

$$\begin{aligned} E(\xi Y_B \,|\, \xi B > 0) &= \frac{E\xi Y_B}{P\{\xi B > 0\}} \\ &= \frac{E\xi Y_B}{E\xi B}\,\frac{E\xi B}{P\{\xi B > 0\}} \\ &\to E(Y_s \,\|\, \xi)_s, \end{aligned}$$

as required. Similarly,

$$\begin{aligned} E(\xi Y_B \,|\, \xi B = 1) &= \frac{E(\xi Y_B;\, \xi B = 1)}{P\{\xi B = 1\}} \le \frac{E\xi Y_B}{P\{\xi B = 1\}} \\ &= \frac{E\xi Y_B}{E\xi B}\,\frac{E\xi B}{P\{\xi B = 1\}} \\ &\to E(Y_s \,\|\, \xi)_s. \end{aligned} \tag{23}$$

To proceed in the opposite direction, we may choose a dominating dissection system (I_{nj}) of S, and form the corresponding approximating point processes ξ_n. Then $E(Y\cdot\xi_n) = f_n \cdot E\xi$, for some increasing sequence of measurable functions $f_n \ge 0$. Assuming $f_n \uparrow \tilde f$, we get $E(Y\cdot\xi) = \tilde f \cdot E\xi$ by monotone convergence, and so $\tilde f = f$ a.e. $E\xi$, which implies $f_n(s) \uparrow E(Y_s \,\|\, \xi)_s$ a.e. When $B \in \mathcal{I}_s$ with $E\xi\{B \setminus I_n(s)\} = 0$, we have a.s.

$$\{\xi_n Y_B > 0\} \subset \{\xi_n B > 0\} \subset \{\xi B = 1\},$$

and so by Theorem 6.32 (i), as $B \downarrow \{s\}$ for fixed n,

$$\begin{aligned} E(\xi Y_B \,|\, \xi B = 1) &\ge E\{\xi_n Y_B \,|\, \xi B = 1\} \\ &= \frac{E\{\xi_n Y_B;\, \xi B = 1\}}{P\{\xi B = 1\}} = \frac{E\xi_n Y_B}{P\{\xi B = 1\}} \\ &= \frac{E\xi_n Y_B}{E\xi B}\,\frac{E\xi B}{P\{\xi B = 1\}} \to f_n(s). \end{aligned}$$

Combining this with (23), we get for $s \in S$ a.e. $E\xi$

$$\begin{aligned} f_n(s) &\le \liminf_{B\downarrow\{s\}} E(\xi Y_B \,|\, \xi B = 1) \\ &\le \limsup_{B\downarrow\{s\}} E(\xi Y_B \,|\, \xi B = 1) \\ &\le E(Y_s \,\|\, \xi)_s, \end{aligned}$$

and letting $n \to \infty$ yields $E(\xi Y_B \,|\, \xi B = 1) \to E(Y_s \,\|\, \xi)_s$. □

The last theorem yields the following weak convergence result on R_+, for the integrals ξY_B.

Corollary 6.34 *(weak integral approximation) Consider a simple point process ξ and a measurable process Y on a Borel space S, where $E\xi$ is locally finite, and fix a standard differentiation basis $\mathcal{I}$ for $E\xi$. Then for $s \in S$ a.e. $E\xi$, we have as $B \downarrow \{s\}$ along $\mathcal{I}$*

$$\mathcal{L}(\xi Y_B | \, \xi B > 0) \approx \mathcal{L}(\xi Y_B | \, \xi B = 1) \xrightarrow{w} \mathcal{L}(Y_s \, \| \, \xi \,)_s.$$

Proof: For any bounded, measurable function f on $\mathbb{R}_+$, we note that $f(\xi Y_B) = \xi f(Y) 1_B$ on $\{\xi B = 1\}$. Hence, as $B \downarrow \{s\}$ for $s \in S$ a.e. $E\xi$, we get by Theorem 6.33

$$\begin{aligned} E\{f(\xi Y_B) \, | \, \xi B = 1\} \; &= \; E\{\xi f(Y) 1_B | \, \xi B = 1\} \\ &\to E\{f(Y_s) \, \| \, \xi\}_s. \end{aligned}$$

Applying this to a convergence-determining sequence of bounded, continuous functions f, we obtain the corresponding weak convergence. By elementary estimates and Theorem 6.32 (i),

$$\Big\| \mathcal{L}(\xi Y_B | \, \xi B > 0) - \mathcal{L}(\xi Y_B | \, \xi B = 1) \Big\| \le 2 \left(1 - \frac{P\{\xi B = 1\}}{P\{\xi B > 0\}} \right) \to 0, \qquad (24)$$

which yields the asserted approximation. □

We also have the following approximations of ordinary and reduced Palm distributions, with limits in the sense of weak convergence, with respect to the vague topology.

Corollary 6.35 *(weak Palm approximation) Let ξ be a simple point process on a Polish space S, such that $E\xi$ is σ-finite, and fix a standard differentiation basis $\mathcal{I}$ for $E\xi$. Then for $s \in S$ a.e. $E\xi$, we have as $B \downarrow \{s\}$ along $\mathcal{I}$*

(i) $\mathcal{L}(\xi \, | \, \xi B > 0) \approx \mathcal{L}(\xi \, | \, \xi B = 1) \xrightarrow{vw} \mathcal{L}(\, \xi \, \| \, \xi \,)_s$,

(ii) $\mathcal{L}(1_{B^c} \xi \, | \, \xi B > 0) \approx \mathcal{L}(1_{B^c} \xi \, | \, \xi B = 1) \xrightarrow{vw} \mathcal{L}(\, \xi - \delta_s \, \| \, \xi \,)_s$.

Proof: (i) For any bounded, measurable function $f \ge 0$ on $\mathcal{N}_S$, Theorem 6.32 (ii) yields for $s \in S$ a.e. $E\xi$, as $B \downarrow \{s\}$ along $\mathcal{I}$,

$$\begin{aligned} E\{f(\xi) \, | \, \xi B > 0\} \; &\approx \; E\{f(\xi) \, | \, \xi B = 1\} \\ &\to E\{f(\xi) \, \| \, \xi\}_s. \end{aligned}$$

Applying this to a convergence-determining sequence of bounded, continuous functions f, which exists since $\mathcal{N}_S$ is again Polish by Lemma 4.6, we obtain the corresponding weak convergence.

(ii) Fix any bounded, measurable function $f \ge 0$ on $\mathcal{N}_S$, and note that the process $Y_s = f(1_{\{s\}^c} \xi)$ is product-measurable by Lemma 1.6. Noting that

$$f(1_{B^c} \xi) = \int_B f(\xi - \delta_s) \, \xi(ds) = \xi Y_B \;\; \text{a.s. on } \{\xi B = 1\},$$

we get by Theorem 6.33, for $s \in S$ a.e. $E\xi$,

$$\begin{aligned} E\{f(1_{B^c}\xi)\,|\,\xi B = 1\} &= E\{\xi Y_B\,|\,\xi B = 1\} \\ &\to E(Y_s\,\|\,\xi)_s = E\{f(\xi - \delta_s)\,\|\,\xi\}_s, \end{aligned}$$

and the corresponding weak convergence follows as before. Proceeding as in (24), we obtain a similar statement for $\mathcal{L}(1_{B^c}\xi\,|\,\xi B > 0)$. □

Under a simple regularity condition, we may strengthen the last convergence to the sense of total variation. We can even establish asymptotic independence between the interior and exterior components $1_B\xi$ and $1_{B^c}\xi$, with respect to a set B, and show that $\mathcal{L}(1_B\xi\,|\,\xi B > 0) \approx \mathcal{L}(1_B\tilde{\xi})$.

Theorem 6.36 *(inner and outer conditioning) Let ξ be a simple point process on a Borel space S, such that $E\xi$ is σ-finite and $C_\xi^{(1)} \ll E\xi \otimes \rho$ for some σ-finite measure ρ. Fix a standard differentiation basis $\mathcal{I}$ for $E\xi$. Then for $s \in S$ a.e. $E\xi$, we have as $B \downarrow \{s\}$ along $\mathcal{I}$*

$$\begin{aligned} \mathcal{L}(1_B\xi, 1_{B^c}\xi\,|\,\xi B > 0) &\approx \mathcal{L}(1_B\tilde{\xi}) \otimes \mathcal{L}(\,\xi - \delta_s\,\|\,\xi\,)_s \\ &\approx \mathcal{L}(1_B\xi\,|\,\xi B > 0) \otimes \mathcal{L}(1_{B^c}\xi\,|\,\xi B > 0), \end{aligned}$$

in the sense of total variation, and similarly for conditioning on $\xi B = 1$.

Proof: For measures of the form $\mu = \delta_\sigma$, the correspondence $\mu \leftrightarrow \sigma$ is bimeasurable by Lemma 1.6, and so

$$\begin{aligned} &\Big\| \mathcal{L}(1_B\xi, 1_{B^c}\xi\,|\,\xi B = 1) - \mathcal{L}(1_B\tilde{\xi}) \otimes \mathcal{L}(\xi - \delta_s\|\,\xi)_s \Big\| \\ &\quad = \left\| E(1_B\xi;\ 1_{B^c}\xi \in \cdot\,|\,\xi B = 1) - \frac{E1_B\xi}{E\xi B} \otimes \mathcal{L}(\xi - \delta_s\|\,\xi)_s \right\| \\ &\quad \le \left\| E(1_B\xi;\ 1_{B^c}\xi \in \cdot\,|\,\xi B = 1) - \frac{E(1_B\xi;\ 1_{B^c}\xi \in \cdot)}{E\xi B} \right\| \\ &\qquad + \left\| \frac{E(1_B\xi;\ 1_{B^c}\xi \in \cdot)}{E\xi B} - \frac{E1_B\xi}{E\xi B} \otimes \mathcal{L}(1_{B^c}\xi\,\|\,\xi)_s \right\| \\ &\qquad + \Big\| \mathcal{L}(1_{B^c}\xi\,\|\,\xi)_s - \mathcal{L}(\xi - \delta_s\|\,\xi)_s \Big\|. \end{aligned}$$

Here the first term on the right is bounded by

$$\left\| E(1_B\xi;\ \xi \in \cdot\,|\,\xi B = 1) - \frac{E(1_B\xi;\ \xi \in \cdot)}{E\xi B} \right\| \le 2\left(\frac{E\xi B}{P\{\xi B = 1\}} - 1 \right),$$

which tends to 0 as $B \downarrow \{s\}$, for $s \in S$ a.e. $E\xi$, by Theorem 6.32 (i). Writing $\nu = E\xi$ and $\mu_s = \mathcal{L}(\xi - \delta_s\|\,\xi)_s$, we see that the second term on the right is bounded by $\|(1_B\nu/\nu B) \otimes (\mu - \mu_s)\|$, which tends a.e. to 0 by Lemma 1.32. Finally, the third term on the right is bounded by $P\{\xi(B \setminus \{s\}) > 0\,\|\,\xi\}_s$, which tends to 0 for every s, by dominated convergence. This proves the first approximation, with conditioning on $\xi B = 1$, and the corresponding assertion for $\xi B > 0$ follows as before by Theorem 6.32 (i).

In particular, the first relation gives as $B \downarrow \{s\}$, for $s \in S$ a.e. $E\xi$,

$$\mathcal{L}(1_B\xi \,|\, \xi B > 0) \;\approx\; \mathcal{L}(1_B\tilde{\xi}),$$
$$\mathcal{L}(1_{B^c}\xi \,|\, \xi B > 0) \;\to\; \mathcal{L}(\,\xi - \delta_s \| \,\xi)_s.$$

The second approximation now follows for $\xi B > 0$, since for any probability measures μ, μ' on S and ν, ν' on T,

$$\left\| \mu \otimes \nu - \mu' \otimes \nu' \right\| \leq \|\mu - \mu'\| + \|\nu - \nu'\|.$$

The proof for $\xi B = 1$ is similar. □

Under suitable regularity conditions, we may extend the previous results to the case of several targets $B_1, \ldots, B_n$. In particular, we have the following decoupling properties, where we write $B = B_1 \times \cdots \times B_n$:

Corollary 6.37 *(asymptotic independence) Let ξ be a simple point process on a Borel space S, such that $E\xi^{\otimes n}$ is σ-finite with $E\xi^{(n)} = p_n \cdot (E\xi)^{\otimes n}$, for some measurable functions p_n. Fix any differentiation basis $\mathcal{I}$ for $E\xi$. Then for $s \in S^{(n)}$ a.e. $E\xi^{(n)}$, we have as $B \downarrow \{s\}$ along $\mathcal{I}^{(n)}$*

(i) $\displaystyle P \bigcap_{k \leq n} \{\xi B_k > 0\} \sim p_n(s) \prod_{k \leq n} E\xi B_k,$

(ii) $\displaystyle \Big\| \mathcal{L}\Big(\xi_{B_1}, \ldots, \xi_{B_n} \Big| \textstyle\bigcap_{k \leq n} \{\xi B_k > 0\}\Big) - \textstyle\bigotimes_{k \leq n} \mathcal{L}(\tilde{\xi}_{B_k}) \Big\| \to 0,$

and similarly for the events $\bigcap_k \{\xi B_k = 1\}$.

Proof: (i) Applying Theorem 6.32 and the differentiation property of $\mathcal{I}^{(n)}$ to the sets $B = B_1 \times \cdots \times B_n$, we get

$$\begin{aligned} P \bigcap_{k \leq n} \{\xi B_k > 0\} &= P\{\xi^{(n)} B > 0\} \sim E\xi^{(n)} B \\ &\sim p_n(s) \prod_{k \leq n} E\xi B_k, \end{aligned}$$

and similarly for the events $\xi B_k = 1$.

(ii) Here Theorems 1.31 and 6.36 give

$$\begin{aligned} &\Big\| \mathcal{L}\Big(\xi_{B_1}, \ldots, \xi_{B_n} \Big| \textstyle\bigcap_{k \leq n} \{\xi B_k > 0\}\Big) - \textstyle\bigotimes_{k \leq n} \mathcal{L}(\tilde{\xi}_{B_k}) \Big\| \\ &\quad \leq \left\| \mathcal{L}\Big\{ 1_B \xi^{(n)} \,\Big|\, \xi^{(n)} B > 0 \Big\} - \frac{1_B E\xi^{(n)}}{E\xi^{(n)} B} \right\| \\ &\qquad + \left\| \frac{1_B E\xi^{(n)}}{E\xi^{(n)} B} - \frac{1_B (E\xi)^n}{(E\xi)^n B} \right\| \to 0, \end{aligned}$$

and similarly for conditioning on the events $\xi B_k = 1$. □

For general intensity measures $E\xi$, the Palm distributions may not exist, and the previous propositions may fail. Some statements can still be salvaged in the form of ratio limit theorems, often obtainable from the previous versions by a change to an equivalent probability measure $\tilde{P} \sim P$, chosen such that the associated intensity measure $\tilde{E}\xi$ becomes σ-finite. We begin with the required construction of $\tilde{P}$.

Lemma 6.38 *(change of measure) For any random measure ξ on S, there exists a random variable $\rho > 0$ with $E\rho = 1$, such that $\nu = E\rho\xi$ is σ-finite. Then ν is a supporting measure of ξ with $\tilde{E}\xi = \nu$, where $\tilde{E}$ denotes expectation with respect to the probability measure $\tilde{P} = \rho \cdot P$. Furthermore, for any measurable process $Y \geq 0$ on S, the associated Palm measures satisfy*

$$\tilde{E}(\,\rho^{-1}Y_s \,\|\, \xi\,)_s = E(\,Y_s \,\|\, \xi\,)_s, \quad s \in S \ \text{a.e. } \nu.$$

Proof: Since ξ is locally finite, a routine construction based on the Borel–Cantelli lemma yields a measurable function $f > 0$ on S with $\xi f < \infty$ a.s. Putting $\gamma = (\xi f \vee 1)^{-1} \in (0,1]$ a.s., we may define $\rho = \gamma/E\gamma$. Then $\|\tilde{P}\| = E\rho = 1$ and $\tilde{E}\,\xi = E\,\rho\,\xi = \nu$, and ν is σ-finite since

$$\nu f = E\rho\xi f = \frac{1}{E\gamma}\, E\left(\frac{\xi f}{\xi f \vee 1}\right) \leq \frac{1}{E\gamma} < \infty.$$

Finally, for Y as stated and any $B \in \mathcal{S}$,

$$\begin{aligned}
\int_B \nu(ds)\, \tilde{E}(\rho^{-1}Y_s \,\|\, \xi)_s &= \tilde{E}\int_B \xi(ds)\,(\rho^{-1}Y_s) \\
&= E\int_B \xi(ds)\, Y_s \\
&= \int_B \nu(ds)\, E(\,Y_s \,\|\, \xi)_s,
\end{aligned}$$

which yields the last assertion. □

Using the previous construction, we may establish some ratio versions of Theorems 6.32 and 6.33, as well as of Corollary 6.35 (ii).

Theorem 6.39 *(ratio convergence) Given a simple point process ξ on a Borel space S, choose a supporting measure $\nu = E\rho\xi$ of ξ, where ρ is such as in Lemma 6.38. Fix a standard differentiation basis $\mathcal{I}$ for ν. Then for $s \in S$ a.e. ν, we have as $B \downarrow \{s\}$ along $\mathcal{I}$*

(i) *for any random variable $\eta \geq 0$, such that $E\eta\,\xi$ is σ-finite,*

$$\frac{E(\eta \,|\, \xi B > 0)}{E(\rho \,|\, \xi B > 0)} \approx \frac{E(\eta \,|\, \xi B = 1)}{E(\rho \,|\, \xi B = 1)} \to E(\eta \,\|\, \xi\,)_s\,,$$

(ii) *for any measurable process $Y \geq 0$, such that $E(Y \cdot \xi)$ is σ-finite,*

$$\frac{E(\xi Y_B |\, \xi B > 0)}{E(\rho \,|\, \xi B > 0)} \approx \frac{E(\xi Y_B |\, \xi B = 1)}{E(\rho \,|\, \xi B = 1)} \to E(Y_s \,\|\, \xi\,)_s\,,$$

(iii) *when* $Y_s = f(\xi - \xi\{s\}\delta_s)$, *for a measurable function* $f \geq 0$ *on* $\mathcal{N}_S$ *such that* $E(Y \cdot \xi)$ *is* σ*-finite, we have*

$$\frac{E\{f(1_{B^c}\xi)\,|\,\xi B = 1\}}{E(\rho\,|\,\xi B = 1)} \to E\{f(\xi - \delta_s)\,\|\,\xi\,\}_s\,,$$

and similarly with conditioning on $\xi B > 0$, *whenever* $E\eta\,\xi$ *is* σ*-finite for some random variable* $\eta \geq \sup_{B\in\mathcal{I}} f(\xi_{B^c})$.

Proof: (i) Since $\tilde{E}\xi = \nu$ and $\tilde{E}(\rho^{-1}\eta)\,\xi = E\eta\,\xi$ are σ-finite, we may apply Theorem 6.32 to the probability measure $\tilde{P}$ and random variable η/ρ, and then use Lemma 6.38 to get

$$\begin{aligned}
\frac{E(\eta\,|\,\xi B > 0)}{E(\rho\,|\,\xi B > 0)} &= \frac{E(\,\eta;\,\xi B > 0)}{E(\,\rho;\,\xi B > 0)} \\
&= \frac{\tilde{E}(\rho^{-1}\eta;\,\xi B > 0)}{\tilde{P}\{\xi B > 0\}} \\
&= \tilde{E}(\rho^{-1}\eta\,|\,\xi B > 0) \\
&\to \tilde{E}(\rho^{-1}\eta\,\|\,\xi)_s = E(\eta\,\|\,\xi)_s.
\end{aligned}$$

The argument for conditioning on $\xi B = 1$ is similar.

(ii) Proceed as in (i), now using Theorem 6.33.

(iii) By (ii) and Lemma 6.2 (ii),

$$\begin{aligned}
\frac{E\{f(1_{B^c}\xi)\,|\,\xi B = 1\}}{E(\rho\,|\,\xi B = 1)} &= \frac{E(\,\xi Y_B\,|\,\xi B = 1)}{E(\rho\,|\,\xi B = 1)} \\
&\to E(\,Y_s\,\|\,\xi\,)_s \\
&= E\{f(\xi - \delta_s)\,\|\,\xi\,\}_s.
\end{aligned}$$

For η as stated, we may combine this result with Theorem 6.32 to get

$$\begin{aligned}
&\left|\frac{E\{f(1_{B^c}\xi)\,|\,\xi B > 0\}}{E(\rho\,|\,\xi B > 0)} - \frac{E\{f(1_{B^c}\xi)\,|\,\xi B = 1\}}{E(\rho\,|\,\xi B = 1)}\right| \\
&\qquad \approx \tilde{E}\Big\{\rho^{-1}f(1_{B^c}\xi);\,\xi B > 1\,\Big|\,\xi B > 0\Big\} \\
&\qquad \leq \tilde{E}\{\,\rho^{-1}\eta;\,\xi B > 1\,|\,\xi B > 0\,\} \\
&\qquad = \frac{\tilde{E}(\rho^{-1}\eta;\,\xi B > 0) - \tilde{E}(\rho^{-1}\eta;\,\xi B = 1)}{\tilde{P}\{\xi B > 0\}} \to 0,
\end{aligned}$$

which yields the version with conditioning on $\xi B > 0$. □

In Theorem 6.32 (i), we proved that $P\{\xi I > 0\} \sim P\{\xi I = 1\}$ a.e. when $E\xi$ is σ-finite. The following version, requiring no moment condition, will be needed in a subsequent chapter.

Lemma 6.40 *(local equivalence) Let ξ be a simple, bounded point process on S, fix any dissection system $\mathcal{I}$ on S, and consider a random variable $\eta \geq 0$ with $E(\eta;\, \xi S > 0) < \infty$. Then for $s \in S$ a.e. $E(\eta\,\xi;\, \xi S = 1)$, we have as $I \downarrow \{s\}$ along $\mathcal{I}$*

$$E\Big\{\eta;\, \xi I = \xi S > 0\Big\} \sim E\Big\{\eta;\, \xi I = \xi S = 1\Big\}.$$

Proof: Replacing P by the measure $Q = E(\eta;\, \xi S > 0,\, \cdot\,)$, and discarding a Q-null set, we may reduce to the case where $\eta \equiv 1$, and $\xi S > 0$ holds identically. The assertion then simplifies to

$$P\{\,\xi I = \xi S\} \sim P\{\,\xi I = \xi S = 1\} \;\text{ a.e. } E(\,\xi;\, \xi S = 1),$$

as $I \downarrow \{s\}$ along $\mathcal{I}$. For each $n \in \mathsf{N}$, let $\mathcal{F}_n$ denote the σ-field in Ω generated by the events $\{\xi I_{nj} = 0\}$, and note as in Lemma 1.7 that ξ is $\mathcal{F}_\infty$-measurable, where $\mathcal{F}_\infty = \bigvee_n \mathcal{F}_n$. Hence, by martingale convergence,

$$\begin{aligned} P(\,\xi S = 1 \,|\, \mathcal{F}_n) \to{}& P(\,\xi S = 1 \,|\, \mathcal{F}_\infty) \\ &= 1\{\,\xi S = 1\} = 1 \quad \text{a.s. on } \{\xi S = 1\}. \end{aligned}$$

On the other hand, since each set $A_{nj} = \{\xi I^c_{nj} = 0\}$ is an atom of $\mathcal{F}_n$, we have on those sets

$$\begin{aligned} P(\,\xi S = 1 \,|\, \mathcal{F}_n) &= P(\,\xi S = 1 \,|\, \xi I^c_{nj} = 0\,) \\ &= \frac{P\{\,\xi S = 1,\, \xi I^c_{nj} = 0\}}{P\{\,\xi I^c_{nj} = 0\}} \\ &= \frac{P\{\,\xi I_{nj} = \xi S = 1\}}{P\{\,\xi I_{nj} = \xi S\}}\,. \end{aligned}$$

Hence, by combination,

$$\sum\nolimits_j \frac{P\{\,\xi I_{nj} = \xi S = 1\}}{P\{\,\xi I_{nj} = \xi S\}}\, 1\{\,\xi I^c_{nj} = 0\} \to 1 \;\text{ a.s. on } \{\xi S = 1\},$$

and the assertion follows. □

6.7 Duality and Kernel Representation

Consider a random measure ξ on S and a random element η in T, where the spaces S and T are Borel. If the Campbell measure $C_{\xi,\eta}$ on $S \times T$ is σ-finite, it may be disintegrated with respect to either variable, as in

$$C_{\xi,\eta} = \nu \otimes \mu \cong \nu' \otimes \mu',$$

where $\cong$ means equality apart from the order of component spaces. Here ν and ν' are σ-finite measures on S and T, whereas the kernels $\mu\colon S \to T$ and $\mu'\colon T \to S$ are σ-finite.

The method of duality amounts to using either disintegration to gain information about the dual one. This leads to some striking connections, often useful in applications. Here we consider the case of absolute continuity. Further duality relations will be explored in a later chapter.

Theorem 6.41 *(duality and densities) Consider a random measure ξ on S and a random element η in T, where S and T are Borel, and fix a supporting measure ν of ξ. Then*

(i) *$E(\xi\,|\,\eta)$ and $\mathcal{L}(\eta\,\|\,\xi)$ exist simultaneously, as σ-finite kernels between S and T,*

(ii) *for $E(\xi\,|\,\eta)$ and $\mathcal{L}(\eta\,\|\,\xi)$ as in* (i),

$$E(\xi\,|\,\eta) \ll \nu \ \text{ a.s.} \quad \Leftrightarrow \quad \mathcal{L}(\eta\,\|\,\xi)_s \ll \mathcal{L}(\eta) \ \text{ a.e. } \nu,$$

in which case both kernels have product-measurable a.e. densities,

(iii) *for any measurable function $f \geq 0$ on $S \times T$,*

$$E(\xi\,|\,\eta) = f(\cdot, \eta) \cdot \nu \ \text{ a.s.} \quad \Leftrightarrow \quad \mathcal{L}(\eta\,\|\,\xi)_s = f(s, \cdot) \cdot \mathcal{L}(\eta) \ \text{ a.e. } \nu.$$

Proof: (i) If $E\xi$ is σ-finite, then for any $B \in \mathcal{S}$ and $C \in \mathcal{T}$,

$$\begin{aligned} E(\xi B;\, \eta \in C) &= E\Big\{E(\xi B\,|\,\eta);\, \eta \in C\Big\} \\ &= \int_C P\{\eta \in dt\}\, E(\xi B\,|\,\eta)_t, \end{aligned}$$

which extends immediately to the dual disintegration

$$C_{\xi,\eta} \stackrel{\sim}{=} \mathcal{L}(\eta) \otimes E(\xi|\,\eta). \tag{25}$$

In general, ξ is a countable sum of bounded random measures, and we may use Fubini's theorem to extend (25) to the general case, for a suitable version of $E(\xi|\,\eta)$. In particular, the latter kernel has a σ-finite version iff $C_{\xi,\eta}$ is σ-finite. Since this condition is also necessary and sufficient for the existence of a σ-finite Palm kernel, the assertion follows.

(iii) Assuming $E(\xi\,|\,\eta) = f(\cdot, \eta) \cdot \nu$ a.s., we get for any $B \in \mathcal{S}$ and $C \in \mathcal{T}$

$$\begin{aligned} E(\,\xi B;\, \eta \in C) &= E\Big\{E(\xi B\,|\,\eta);\, \eta \in C\Big\} \\ &= E \int_B f(s, \eta)\, \nu(ds)\, 1_C(\eta) \\ &= \int_B \nu(ds)\, E\Big\{f(s, \eta);\, \eta \in C\Big\} \\ &= \int_B \nu(ds) \int_C P\{\eta \in dt\}\, f(s, t), \end{aligned}$$

which implies $\mathcal{L}(\eta \,\|\, \xi)_s = f(s,\cdot)\cdot \mathcal{L}(\eta)$ a.e. ν. Conversely, the latter relation yields

$$\begin{aligned} E(\xi B;\, \eta \in C) &= \int_B \nu(ds)\, P(\eta \in C \,\|\, \xi)_s \\ &= \int_B \nu(ds) \int_C P\{\eta \in dt\}\, f(s,t) \\ &= \int_B \nu(ds)\, E\big\{f(s,\eta);\, \eta \in C\big\} \\ &= E\int_B f(s,\eta)\,\nu(ds)\, 1_C(\eta), \end{aligned}$$

which shows that

$$E(\xi B \,|\, \eta) = \int_B f(s,\eta)\,\nu(ds) = \{f(\cdot,\eta)\cdot\nu\}B \;\text{ a.s.,}$$

and therefore $E(\xi \,|\, \eta) = f(\cdot,\eta)\cdot \nu$ a.s.

(ii) Since S and T are Borel, Lemma 1.28 yields a product-measurable density in each case. The assertion now follows from (iii). □

The duality approach is often useful to construct versions of the Palm distributions with desirable regularity properties. Here we choose T to be an abstract space, partially ordered by a relation $\prec$. By a *filtration* on T we mean a family of sub-σ-fields $\mathcal{F} = \{\mathcal{F}_t;\, t \in T\}$ on Ω, such that $\mathcal{F}_s \subset \mathcal{F}_t$ for any $s \prec t$. An $\mathcal{F}$*-martingale* on T is a process $M = (M_t)$, such that M_t is $\mathcal{F}_t$-measurable for each t, and $E(M_t|\, \mathcal{F}_s) = M_s$ a.s. for all $s \prec t$.

Given a T-indexed filtration $\mathcal{F} = (\mathcal{F}_t)$ on the basic probability space Ω, we consider versions $M_{s,t}\cdot P$ of the Palm measures $P(\cdot \,\|\, \xi)_s$ on $\mathcal{F}_t$, where the process $M \geq 0$ on $S \times T$ is such that $M(s,t,\omega)$ is $(\mathcal{S}\otimes\mathcal{F}_t)$-measurable in (s,ω) for fixed t. Write $M_t = M(\cdot,t,\cdot)$ and $M_{s,t} = M(s,t,\cdot)$. The following theorem gives criteria for continuity and consistency of those versions, in terms of the continuity and martingale properties of M.

Theorem 6.42 *(continuity and consistency) Let ξ be a random measure on a Polish space S, such that $E\xi$ is σ-finite. Given a filtration $\mathcal{F} = (\mathcal{F}_t)$, indexed by a partially ordered set T, consider for every $t \in T$ an $(\mathcal{S}\otimes\mathcal{F}_t)$-measurable process $M_t \geq 0$ on S. Then for any $t \in T$,*

(i) $E(\xi \,|\, \mathcal{F}_t) = M_t \cdot E\xi \;\Leftrightarrow\; P(\cdot \,\|\, \xi)_s = M_{s,t}\cdot P$ *a.e. on* $\mathcal{F}_t$.

For M as in (i), *the Palm versions $P(\cdot \,\|\, \xi)_{s,t} = M_{s,t}\cdot P$ on $\mathcal{F}_t$ satisfy:*

(ii) *for fixed $t \in T$, the process $P(\mathcal{F}_t \,\|\, \xi)_{s,t}$ is continuous in total variation in $s \in S$, iff $M_{s,t}$ is L^1-continuous on S,*

(iii) *for fixed $s \in S$, the process $P(\mathcal{F}_t \,\|\, \xi)_{s,t}$ is consistent in $t \in T$, iff $M_{s,t}$ is a martingale on T,*

(iv) *when the $\mathcal{F}_t$ are countably generated, continuity as in* (ii) *for all $t \in T$ implies consistency as in* (iii) *for all $s \in \operatorname{supp} E\xi$.*

Proof: (i) This can be proved as in Theorem 6.41 (iii). Note that no regularity condition on $\mathcal{F}_t$ is needed for this proof.

(ii) This holds by the $L^\infty - L^1$ isometry

$$\begin{aligned}\Big\|P(\cdot\,\|\xi)_{s,t} - P(\cdot\,\|\xi)_{s',t}\Big\|_t &= \Big\|(M_{s,t} - M_{s',t})\cdot P\Big\|_t \\ &= E|M_{s,t} - M_{s',t}|,\end{aligned}$$

where $\|\cdot\|_t$ denotes total variation on $\mathcal{F}_t$.

(iii) Since $M_{s,t}$ is $\mathcal{F}_t$-measurable, we get for any $A \in \mathcal{F}_t$ and $t' \succ t$

$$\begin{aligned}P(A\,\|\,\xi)_{s,t} - P(A\,\|\,\xi)_{s,t'} &= E(\,M_{s,t} - M_{s,t'}; A) \\ &= E\Big\{M_{s,t} - E(M_{s,t'}|\mathcal{F}_t); A\Big\},\end{aligned}$$

which vanishes for every A, iff $M_{s,t} = E(M_{s,t'}|\mathcal{F}_t)$ a.s.

(iv) For any $t \prec t'$ and $A \in \mathcal{F}_t$, we have $P(A\,\|\,\xi)_{s,t} = P(A\,\|\,\xi)_{s,t'}$ a.e., by the uniqueness of the Palm disintegration. Since the $\mathcal{F}_t$ are countably generated, a monotone-class argument yields $P(\cdot\,\|\,\xi)_{s,t} = P(\cdot\,\|\,\xi)_{s,t'}$, a.e. on $\mathcal{F}_t$, and so by (ii) we have $M_{s,t} = E(M_{s,t'}\,|\,\mathcal{F}_t)$ a.s., for $s \in S$ a.e. $E\xi$. Using the L^1-continuity of $M_{s,t}$ and $M_{s,t'}$, along with the L^1-contractivity of conditional expectations, we may extend the latter relation to all $s \in \operatorname{supp} E\xi$, and so by (ii) the measures $P(\cdot\,\|\,\xi)_{s,t}$ are consistent for every $s \in \operatorname{supp} E\xi$. $\square$

Often, in applications, we have $E\xi = p \cdot \lambda$ for some σ-finite measure λ on S and continuous density $p > 0$ on S. When $E(\xi\,|\,\mathcal{F}_t) = X_t \cdot \lambda$ a.s., for some $(\mathcal{S} \otimes \mathcal{F}_t)$-measurable processes X_t on S, the last theorem applies with $M_t = X_t/p$. Furthermore, the L^1-continuity in (i) and martingale property in (ii) hold simultaneously for X and M.

We turn to the special case of Palm distributions $\mathcal{L}(\xi\,\|\,\xi)_s$, where ξ is a simple or diffuse random measure on S. Though the measure $E(\xi\,|\,\xi) = \xi$ is typically singular, the conditional intensity $E(\xi\,|\,1_{G^c}\xi)$ may still be absolutely continuous on G with a continuous density, for every open set G. We may then form consistent versions of the Palm distributions of the restrictions $1_{G^c}\xi$, for arbitrary neighborhoods G of s. A simple extension yields versions Q_s of the Palm distributions of ξ itself. We need the latter measures to be tight at s, in the sense that $Q_s\{\mu G < \infty\} = 1$ for every bounded neighborhood G of s. We may also want to extend the continuity of the restricted Palm distributions to the extended versions Q_s.

With such objectives in mind, we consider a random measure ξ with locally finite intensity $E\xi$, defined on a Polish space S with open sets $\mathcal{G}$ and topological base $\mathcal{I}$.

Theorem 6.43 *(extension, tightness, and continuity) Let ξ be a random measure on a Polish space S, such that $E\xi$ is locally finite. For every $G \in \mathcal{G}$, let $E(\xi\,|\,1_{G^c}\xi) = M^G \cdot E\xi$ on G, with an L^1-continuous, product-measurable* *density M^G. Then for fixed $s \in \operatorname{supp} E\xi$, and with G and I restricted to $\mathcal{G}$ and $\mathcal{I}$, we have*

(i) *the measures $\mathcal{L}(1_{G^c}\xi \,\|\, \xi)_s = M_s^G \cdot \mathcal{L}(\xi)$, $s \in G$, extend uniquely to a version Q_s of $\mathcal{L}(1_{\{s\}^c}\xi \,\|\, \xi)_s$,*

(ii) *when ξ is a.s. diffuse, the measures Q_r are tight and weakly continuous at $r = s$, iff*
$$\lim_{G\downarrow\{s\}} \sup_{r\in G} \liminf_{I\downarrow\{r\}} \frac{E(\xi I)(\xi G \wedge 1)}{E\xi I} = 0,$$

(iii) *when ξ is a simple point process, the measures Q_r are tight and weakly continuous at $r = s$, iff*
$$\lim_{G\downarrow\{s\}} \sup_{r\in G} \liminf_{I\downarrow\{r\}} \frac{E(\,\xi I;\, \xi G > 1)}{E\xi I} = 0.$$

Proof: For notational convenience, we may take ξ to be the identity mapping on $\Omega = \mathcal{M}_S$.

(i) Choose a sequence $G_n \downarrow \{s\}$ in $\mathcal{G}$. By Theorem 6.42 (iv), the associated Palm distributions $\mathcal{L}(1_{G_n^c}\xi \,\|\, \xi)_s = M^{G_n} \cdot P$ are consistent in n, and so the Daniell–Kolmogorov theorem yields a unique extension Q_s to $\mathcal{M}_{\{s\}^c}$, regarded as a probability measure on the space of σ-finite measures μ on S with $\mu\{s\} = 0$.

To see that Q_s is measurable, fix a countable basis $G_1, G_2, \ldots \in \mathcal{G}$ in S with $G_1 = S$, and put $\mathcal{G}_n = \{G_1, \ldots, G_n\}$ for $n \in \mathbb{N}$. For each $s \in S$, let $G_n(s)$ denote the intersection of the sets $G \in \mathcal{G}_n$ with $s \in G$, and note that
$$\mu\{B \setminus G_n(s)\} \uparrow \mu B, \quad s \in S \text{ a.e. } Q_s,\ \ B \in \hat{\mathcal{S}}.$$
By dominated convergence, we get for any $B_1, \ldots, B_n \in \hat{\mathcal{S}}$ and $r_1, \ldots, r_n \geq 0$
$$Q_s \bigcap_{k\leq n} \Big\{\mu\{B_k \setminus G_n(s)\} \leq r_k\Big\} \to Q_s \bigcap_{k\leq n} \{\mu B_k \leq r_k\}.$$
Here the measurability on the left carries over to the limit, and the required measurability of Q_s follows by a monotone-class argument.

A similar convergence holds for the reduced Palm measures $\mathcal{L}(1_{\{s\}^c}\xi \,\|\, \xi)_s$. Since the expressions on the left are the same as before, we get for $s \in S$ a.e. $E\xi$
$$Q_s \bigcap_{k\leq n} \{\mu B_k \leq r_k\} = \mathcal{L}\Big(\textstyle\bigcap_{k\leq n} \Big\{\xi(B_k \setminus \{s\}) \leq r_k\Big\} \Big\|\, \xi \Big)_s,$$
which extends a.e. to $Q_s = \mathcal{L}(1_{\{s\}^c}\xi \,\|\, \xi\,)_s$, by another monotone-class argument.

(ii) Since ξ is diffuse, Lemma 6.2 yields $E(\xi\{r\} \,\|\, \xi)_r = 0$ a.e. $E\xi$, and so by (i) we have for any $I \subset G$
$$\begin{aligned} E(\xi I)(\xi G \wedge 1) &= \int_I E\xi(dr)\, E(\,\xi G \wedge 1 \,\|\, \xi)_r \\ &= \int_I E\xi(dr) \int Q_r(d\mu)\, (\mu G \wedge 1). \end{aligned}$$

Thus, the stated condition is equivalent to

$$\lim_{G\downarrow\{s\}} \sup_{r\in G} \liminf_{I\downarrow\{r\}} \int_I \frac{E\xi(dp)}{E\xi I} \int Q_p(d\mu)\,(\mu G\wedge 1) = 0. \tag{26}$$

Fixing any $r\in G$, and letting $B\in\hat{\mathcal{S}}$ with $B\subset G$ and $r\notin\bar{B}$, we get by Theorem 6.42 (ii)

$$\int Q_r(d\mu)(\mu B\wedge 1) \le \liminf_{I\downarrow\{r\}} \int_I \frac{E\xi(dp)}{E\xi I} \int Q_p(d\mu)\,(\mu G\wedge 1).$$

Letting $B\uparrow G\setminus\{r\}$, we obtain the same relation with $B=G$, and so by (26),

$$\lim_{G\downarrow\{s\}} \sup_{r\in G} \int Q_r(d\mu)\,(\mu G\wedge 1) = 0, \tag{27}$$

which shows that the measures Q_r are uniformly tight at s.

Now choose any $G_k\downarrow\{s\}$ in $\mathcal{G}$. For fixed $k\in\mathbb{N}$, Theorem 6.42 (ii) yields $Q_r\to Q_s$ as $r\to s$ in $\operatorname{supp} E\xi$, in the sense of total variation, on the σ-field generated by the restriction to G_k^c. Since k was arbitrary, the stated convergence then holds weakly, for the vague topology on $\mathcal{M}_{\{s\}^c}$. Combining with (27), and using the implication (iii) $\Rightarrow$ (i) in Theorem 4.19, we obtain the required weak convergence, under the vague topology on $\mathcal{M}_S$.

Conversely, let Q_r be tight and weakly continuous at s. Let $G,B\in\mathcal{G}$, with $s\in G\subset B$ and $Q_s\{\mu\partial B>0\}=0$, and write

$$\sup_{r\in G} \int Q_r(d\mu)\,(\mu G\wedge 1) \le \sup_{r\in G} \int Q_r(d\mu)\,(\mu B\wedge 1).$$

Since $Q_r\xrightarrow{w} Q_s$ and Q_s is tight, we get as $G\downarrow\{s\}$ and then $B\downarrow\{s\}$

$$\lim_{G\downarrow\{s\}} \sup_{r\in G} \int Q_r(d\mu)\,(\mu G\wedge 1) \le \int Q_s(d\mu)\,(\mu\{s\}\wedge 1).$$

This proves (27), and (26) follows.

(iii) Here $P(\,\xi\{s\}=1\,\|\,\xi\,)_s=1$ a.e. $E\xi$, by Lemma 6.2, and so for any $I\subset G$,

$$\begin{aligned} E(\xi I;\,\xi G>1) &= \int_I E\xi(dr)\,P(\,\xi G>1\,\|\,\xi\,)_r \\ &= \int_I E\xi(dr)\,Q_r\{\mu;\,\mu G>0\}, \end{aligned}$$

which shows that the stated condition is again equivalent to (26). The proof may now be completed as before. □

We also have the following direct version, which leads to a useful kernel representation of random measures.

Theorem 6.44 *(Campbell measures and kernel representation) Consider a measure ρ on $S\times T$ and a random element η in T, such that $\rho(S\times\cdot)\ll\mathcal{L}(\eta)$, where T is measurable and S is Borel.*

(i) *If $T=\Omega$ and $\rho(S\times\cdot)\ll P$, then ρ is s-finite iff $\rho=C_\xi$ for some s-finite random measure ξ on S. Furthermore, ρ and ξ are simultaneously σ-finite, in which case ξ is a.s. unique.*

(ii) *In general, ρ is s-finite iff it equals $C_{\xi,\eta}$, for some s-finite random measure ξ on S. In that case, $\rho=C_{\zeta,\eta}$ with $\zeta=E(\xi|\,\eta)$, and $\zeta=\mu\circ\eta$ a.s. for some s-finite kernel $\mu\colon T\to S$.*

(iii) *An s-finite measure $\rho=C_{\xi,\eta}$ is σ-finite, iff $E(\xi Y|\,\eta)<\infty$ a.s. for some η-measurable process $Y>0$ on S. In that case, $Y_s=h(s,\eta)$ for some measurable function $h>0$, such that $E(\xi Y|\,\eta)=1\{E(\xi|\,\eta)\neq 0\}$ a.s.*

Proof: (i) This is clear from Theorem 1.23.

(ii) Let ρ be s-finite, so that $\rho=\sum_n\rho_n$ for some bounded measures $\rho_1,\rho_2,\dots$. By Theorem 1.23, we have $\rho_n=\mathcal{L}(\eta)\otimes\mu_n$ for some finite kernels $\mu_n: T\to S$, and so $\rho=\mathcal{L}(\eta)\otimes\mu$ for the s-finite kernel $\mu=\sum_n\mu_n$. Here $\xi=\mu\circ\eta$ is an s-finite random measure on S, and

$$\begin{aligned}C_{\xi,\eta}f &= E\int\xi(ds)\,f(s,\eta)\\ &= E\int(\mu\circ\eta)(ds)\,f(s,\eta)\\ &= \int P\{\eta\in dt\}\int\mu_t(ds)f(s,t)=\rho f,\end{aligned}$$

which shows that $\rho=C_{\xi,\eta}$.

Conversely, suppose that ξ is s-finite. Then Lemma 1.15 (iii) yields $\xi=\sum_n\xi_n$, for some random measures ξ_n with $\|\xi_n\|\le 1$. This gives $\|C_{\xi_n,\eta}\|\le 1$, and so $\rho=C_{\xi,\eta}=\sum_n C_{\xi_n,\eta}$ is again s-finite. As before, we have $\rho=\mathcal{L}(\eta)\otimes\mu$ for some s-finite kernel $\mu\colon T\to S$, and so

$$E(\,\xi f;\,\eta\in B)=\rho(f\otimes 1_B)=E\big\{(\mu\circ\eta)f;\,\eta\in B\big\},$$

which implies $E(\xi f|\,\eta)=(\mu\circ\eta)f$ a.s. Thus, $E(\xi|\,\eta)$ has the regular version $\mu\circ\eta$.

(iii) Let ρ be σ-finite, so that $\rho f\le 1$ for some measurable function $f>0$ on $S\times T$. Then the process $Y_s=f(s,\eta)$ is η-measurable with $E\xi Y\le 1$, and so $E(\xi Y|\,\eta)<\infty$ a.s. The process

$$\hat{Y}_s=Y_s\left(\frac{1\{E(\xi Y|\,\eta)>0\}}{E(\xi Y|\,\eta)}+1\{E(\xi Y|\,\eta)=0\}\right),\quad s\in S,$$

is again η-measurable and satisfies the stated conditions.

Conversely, if $E(\xi Y|\,\eta)<\infty$ a.s. for some η-measurable process Y, then the process $\hat{Y}$ above satisfies $E(\xi\hat{Y}|\,\eta)\le 1$ a.s. By FMP 1.13, we have

$\hat{Y}_s = h(s, \eta)$ for some measurable function $h > 0$ on $S \times T$, and ρ is σ-finite since $\rho h = E\xi\hat{Y} \le 1$. □

We conclude with an extension of Theorem 6.43, needed in Chapter 13.

Lemma 6.45 *(tightness and continuity) Let ξ be a diffuse random measure on $S = \mathsf{R}^d$, such that*

(i) *$E\xi^{\otimes n}$ is locally finite on $S^{(n)}$,*
(ii) *for any open set $G \subset S$, the kernel $\mathcal{L}(1_{G^c}\xi \,\|\, \xi^{\otimes n})_x$ has a version that is continuous in total variation in $x \in G^{(n)}$,*
(iii) *for any compact set $K \subset S^{(n)}$,*
$$\lim_{r\to 0} \sup_{x\in K} \liminf_{\varepsilon\to 0} \frac{E\,\xi^{\otimes n} B_x^\varepsilon(\xi U_x^r \wedge 1)}{E\xi^{\otimes n} B_x^\varepsilon} = 0.$$

Then the kernel $\mathcal{L}(\xi \,\|\, \xi^{\otimes n})_x$ has a tight, weakly continuous version on $S^{(n)}$, such that (ii) *holds for every G.*

Proof: Proceeding as before, we may construct a version of the kernel $\mathcal{L}(\xi \,\|\, \xi^{\otimes n})_x$ on $S^{(n)}$, satisfying the continuity property in (ii), for any open set G. Now fix any $x \in S^{(n)}$, and let $l(r)$ denote the "lim inf" in (iii). By Palm disintegration under condition (i), we may choose some points $x_k \to x$ in $S^{(n)}$, such that $E(\xi U_x^r \wedge 1 \,\|\, \xi^{\otimes n})_{x_k} \le 2l(r)$. Then for any open set $G \subset \mathsf{R}^d$ and point $x \in G^{(n)}$, we have by (ii)

$$E\Big\{\xi(U_x^r \cap G^c) \wedge 1 \,\Big\|\, \xi^{\otimes n}\Big\}_x = \lim_{k\to\infty} E\Big\{\xi(U_x^r \cap G^c) \wedge 1 \,\Big\|\, \xi^{\otimes n}\Big\}_{x_k} \le l(r).$$

Since G was arbitrary and $l(r) \to 0$ as $r \to 0$, we conclude that $\mathcal{L}(\xi \,\|\, \xi^{\otimes n})_x$ is tight at x. By a similar argument, using the full strength of (iii), we get the uniform tightness

$$\lim_{r\to 0} \sup_{x\in K} E\Big(\xi U_x^r \wedge 1 \,\|\, \xi^{\otimes n}\Big)_x = 0, \tag{28}$$

for any compact set $K \subset S^{(n)}$. For open G and $x_k \to x$ in $G^{(n)}$, we get by (ii)

$$\mathcal{L}(1_{G^c}\xi \,\|\, \xi^{\otimes n})_{x_k} \xrightarrow{w} \mathcal{L}(1_{G^c}\xi \,\|\, \xi^{\otimes n})_x\,.$$

Proceeding as in Theorem 4.19 and using (28), we see that the latter convergence remains valid with $1_{G^c}\xi$ replaced by ξ, as required. □

Chapter 7

Group Stationarity and Invariance

In Chapter 5, we considered random measures on R^d, stationary under arbitrary shifts. In particular, we saw how the invariance structure allows us to construct the Palm measure at the origin by a simple skew transformation. The general Palm measures of Chapter 6 can then be obtained by the obvious shifts.

We now consider a more general measurable group $(G, \mathcal{G})$ with identity element ι, acting measurably on some Borel spaces $(S, \mathcal{S})$ and $(T, \mathcal{T})$. Here the *measurability* of G means that the group operations $r \mapsto r^{-1}$ and $(r, r') \mapsto rr'$ are $\mathcal{G}$-measurable. An *action* of G on S is a mapping $(r, s) \mapsto rs$ from $G \times S$ to S, such that $\iota s = s$ and $r(r's) = (rr')s$ for all $r, r' \in G$ and $s \in S$. It is said to be *measurable*, if the defining map is a measurable function from $\mathcal{G} \otimes \mathcal{S}$ to $\mathcal{S}$. In particular, the group operation defines a measurable action $(r, r') \mapsto rr'$ of G onto itself.

The associated *shifts* θ_r and *projections* π_s are given by $rs = \theta_r s = \pi_s r$, so that $\theta_r \colon S \to S$ for all $r \in G$ whereas $\pi_s \colon G \to S$ for all $s \in S$. For any $s, s' \in S$, we write $s \sim s'$ if $s' = rs$ for some $r \in G$. The ranges $\pi_s G$, called the *orbits* of G, form a partition of S into equivalence classes, and the group action is said to be *transitive* if $s \sim s'$ for all $s, s' \in S$.

A measure μ on S is said to be *G-invariant*[1], if $\mu \circ \theta_r^{-1} = \mu$ for all $r \in G$. In particular, a *Haar measure* is defined as a σ-finite, left-invariant measure $\lambda \neq 0$ on G itself. It is known to exist when G is lcscH, in which case it is unique up to a normalization. We say that G acts *properly* on S, if there exists a measurable function $g > 0$ on S, such that $\lambda(g \circ \pi_s) < \infty$ for all $s \in S$. The measure $\lambda \circ \pi_s^{-1}$ is then σ-finite and G-invariant on the orbit $\pi_s G$ for every $s \in S$, and Theorem 7.3 shows that every G-invariant measure on S is essentially a unique mixture of such orbit measures.

A random measure ξ on S is said to be *G-stationary*, if $\xi \circ \theta_r^{-1} \stackrel{d}{=} \xi$ for every $r \in G$. Similarly, a random element η in T is said to be G-stationary if $r\eta \stackrel{d}{=} \eta$ for all $r \in G$. We also say that ξ and η are *jointly G-stationary* if $\theta_r(\xi, \eta) \stackrel{d}{=} (\xi, \eta)$, where the G-shifts on $\mathcal{M}_S$ are given by $\theta_r \mu = \mu \circ \theta_r^{-1}$. In the latter case, the Campbell measure $C_{\xi,\eta}$ is jointly G-invariant, and

[1] From this chapter on, we will constantly use the integral notation $\mu f = \int f d\mu$, $(f \cdot \mu)g = \mu(fg)$, and $(\mu \circ g^{-1})f = \mu(f \circ g)$.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_7

we may look for an invariant disintegration[2] $C_{\xi,\eta} = \nu \otimes \mathcal{L}(\eta \,\|\, \xi)$, in terms of an invariant supporting measure ν of ξ and a suitably invariant Palm kernel $\mathcal{L}(\eta \,\|\, \xi)$ from S to T. Such invariant disintegrations are studied most conveniently in an abstract setting.

For kernels $\mu\colon S \to T$, the appropriate notion of *G-invariance* is given by $\mu_{rs} = \mu_s \circ \theta_r^{-1}$, or more explicitly

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

For any G-invariant measure ν on S and G-invariant kernel $\mu\colon S \to T$, the composition $\nu \otimes \mu$ is a jointly G-invariant measure on $S \times T$. Conversely, Theorem 7.6 shows that, for any σ-finite and G-invariant measures ν on S and ρ on $S \times T$ with $\rho(\cdot \times T) \ll \nu$, there exists a G-invariant kernel $\mu\colon S \to T$ with $\rho = \nu \otimes \mu$. For given ρ, the projection $\nu = \rho(\cdot \times T)$ is automatically G-invariant, though it may fail to be σ-finite. However, a σ-finite supporting measure ν always exists when G acts properly on S.

When $S = G$, we may construct the desired disintegration by an elementary *skew transformation*, just as in Chapter 5. A similar construction applies to any product space $S = G \times S'$, where S' is Borel. For more general S, we may reduce to the previous case in various ways, such as by replacing the measures ν on S and ρ on $S \times T$ by the invariant measures $\lambda \otimes \nu$ on $G \times S$ and $\lambda \otimes \rho$ on $G \times S \times T$.

Another approach is based on the G-invariant *inversion kernel* $\gamma\colon S \to G$, which maps any G-invariant measures ν on S and ρ on $S \times T$ into invariant measures $\tilde\nu$ on $G \times A$ and $\tilde\rho$ on $G \times A \times T$, where A is an abstract measurable space, not subject to any G-action. In Theorem 7.14, we show that such a kernel exists when G acts properly on S. The present approach has the advantage of leading to some nice explicit formulas. We can also use the same kernel to map any stationary random measure ξ on S into a stationary random measure $\tilde\xi$ on $G \times A$.

In Chapter 5, we saw that a σ-finite measure ν is the Palm measure of a stationary point process ξ on R, iff it is cycle-invariant, in which case we can use a simple inversion formula to recover the associated distribution of ξ. We now consider the corresponding problem for an arbitrary measurable group, acting properly on S. In the simple version of Corollary 7.19, we take G to act directly on the basic probability space Ω, let ξ be a G-invariant random measure on G, and write τ for the identity map on Ω. Then a σ-finite measure $\tilde P$ on Ω has the form $P(\cdot \,\|\, \xi)_\iota$ for some σ-finite, G-invariant measure P on $\{\xi \neq 0\}$, iff $\tilde P\{\xi = 0\} = 0$ and

$$\tilde E \int \xi(ds)\, f(\tau, s) = \tilde E \int \xi(ds)\, f(\tilde s\tau, \tilde s)\, \Delta_s, \quad f \geq 0, \tag{1}$$

where Δ_s is the reciprocal modular function on G, and $\tilde s = s^{-1}$. In that case,

[2]Recall that the kernel composition is given by $(\nu \otimes \mu)f = \int \nu(ds) \int \mu_s(dt) f(s,t)$.

P is given uniquely by

$$Ef(\tau) = \tilde{E} \int \lambda(ds)\, f(s\tau)\, h(s\tau, s), \quad f \geq 0,$$

for a suitable *normalizing* function $h \geq 0$ on $\Omega \times G$.

Of the many *balance* and *transport* relations obtainable by the methods of duality and inversion, we mention only the celebrated *exchange formula* in Corollary 7.23. Once again, we assume P to be invariant under a group G, acting measurably directly on Ω. We further take ξ and η to be G-invariant kernels from Ω to G, making them into jointly G-stationary random measures on G. Defining τ as before, we show that

$$E\left\{ \int \xi(ds)\, f(\tau, s) \,\middle\|\, \eta \right\} = E\left\{ \int \eta(ds)\, f(\tilde{s}\tau, \tilde{s})\, \Delta_s \,\middle\|\, \xi \right\}.$$

This clearly reduces to (1) when $\xi = \eta$.

Turning to the subtle topics of stationary differentiation and disintegration, let ξ and η be jointly G-stationary random measures on a Borel space S, such that $\xi \ll \eta$ a.s. By a standard martingale argument, we may construct an associated jointly measurable density process $X \geq 0$ on S, so that[3] $\xi = X \cdot \eta$ a.s. We may need to find a version of X that is even G-stationary, jointly with ξ and η. In Corollary 7.36, we prove that such a density X exists when G is lcscH. More generally and under the same assumption, Theorem 7.35 shows that, for any jointly G-stationary random measures ξ on $S \times T$ and η on S with $\xi(\cdot \times T) \ll \eta$ a.s., there exists a random kernel $\zeta\colon S \to T$, such that the triple (ξ, η, ζ) is G-stationary and satisfies the disintegration formula $\xi = \eta \otimes \zeta$ a.s.

The latter results are based on some equally subtle properties of invariant differentiation and disintegration, proved in Theorem 7.24 and its Corollary 7.25. For the latter result, consider any locally finite measures $\mu \ll \nu$ on S. We may wish to find an associated product-measurable density function $\varphi \geq 0$ on $S \times \mathcal{M}_S^2$, which is *universally G-invariant*, in the sense that

$$\varphi(s, \mu, \nu) = \varphi\left\{ rs, (\mu, \nu) \circ \theta_r^{-1} \right\}, \quad r \in G,\ s \in S.$$

More generally, given any locally finite measures μ on $S \times T$ and ν on S such that $\mu(\cdot \times T) \ll \nu$, we may look for a kernel ψ from $S \times \mathcal{M}_{S\times T} \times \mathcal{M}_S$ to T, satisfying the disintegration $\mu = \nu \otimes \psi(\cdot, \mu, \nu)$, along with the universal invariance relation

$$\psi(s, \mu, \nu) \circ \theta_r^{-1} = \psi\left\{ rs, (\mu, \nu) \circ \theta_r^{-1} \right\}, \quad r \in G,\ s \in S.$$

Using the deep representation of locally compact groups in terms of projective limits of Lie groups, we prove that the desired function or kernel exists when G is lcscH. Note that, for fixed μ and ν, the invariance of ψ reduces to the elementary condition $\psi_{rs} = \psi_s \circ \theta_r^{-1}$.

[3]Note that $X \cdot \eta$ denotes the measure with η-density X, so that $(X \cdot \eta)f = \eta(fX)$.

So far, we have emphasized the role of invariant disintegrations to construct suitable versions of the Palm measures. The dual disintegration of the Campbell measure $C_{\xi,\eta}$ is equally important and leads to invariant representations of both ξ itself and its conditional intensities. Given a jointly G-stationary pair of a random measure ξ on S and a random element η in T, such that $C_{\xi,\eta}$ is σ-finite, we prove in Theorem 7.9 that $E(\xi|\,\eta) = \mu \circ \eta$ a.s. for some G-invariant kernel $\mu\colon T \to S$. In particular[4], we get $\xi = \mu \circ \eta$ a.s. when ξ is η-measurable.

We finally consider the related problem of establishing ergodic decompositions of $\mathcal{L}(\xi)$, when ξ is a G-stationary random measure on S. Here the idea is to write $\mathcal{L}(\xi) = E\mathcal{L}(\xi\,|\,\mathcal{I})$, where $\mathcal{I}$ denotes the σ-field of ξ-measurable events, invariant under the action of G. In Theorem 7.37 we show that, whenever G is lcscH, the conditional distribution $\mathcal{L}(\xi\,|\,\mathcal{I})$ has indeed a G-invariant, measure-valued version. This implies the corresponding properties for the conditional intensity $E(\xi\,|\,\mathcal{I})$.

7.1 Invariant Measures and Kernels

By a *measurable group* we mean a group G endowed with a σ-field $\mathcal{G}$, such that the group operations $r \mapsto r^{-1}$ and $(r, s) \mapsto rs$ on G are $\mathcal{G}$-measurable. It is said to be *Borel* if $(G, \mathcal{G})$ is a Borel space. Defining the left and right *shifts* θ_r and $\tilde{\theta}_r$ on G by $\theta_r s = \tilde{\theta}_s r = rs,\ r, s \in G$, we say that a measure λ on G is *left-invariant* if $\lambda \circ \theta_r^{-1} = \lambda$ for all $r \in G$. A σ-finite, left-invariant measure $\lambda \neq 0$ on G is called a *Haar measure* on G. Haar measures are known to exist when G is lcscH, in which case they are unique up to a normalization. If G is compact, then λ is bounded and may be normalized into a probability measure, which is also right-invariant. For any $f \in \mathcal{G}_+$, we put $\tilde{f}(r) = f(r^{-1})$, and define the corresponding right-invariant Haar measure $\tilde{\lambda}$ by $\tilde{\lambda} f = \lambda \tilde{f}$.

In the sequel, no topological or other assumptions are made on G, beyond the existence of a Haar measure. We show that the basic properties of Haar measures remain valid in our abstract setting. To simplify some subsequent proofs, we shall often consider *s-finite* (rather than σ-finite) measures, defined as countable sums of bounded measures. Note that the basic computational rules, including Fubini's theorem, remain valid for s-finite measures, and that s-finiteness has the advantage of being preserved by measurable transformations.

We list some basic properties of general Haar measures. The corresponding results for classical Haar measures and σ-finite invariant measures are of course well known.

Theorem 7.1 *(Haar measure) Let λ be a Haar measure on a measurable group G.*

[4]Such a representation has often been postulated as part of the definition, which bypasses the requirement of proof.

(i) *For any measurable space S and s-finite, G-invariant measure μ on $G \times S$, there exists a unique, s-finite measure ν on S, such that $\mu = \lambda \otimes \nu$. Here μ and ν are simultaneously σ-finite.*

(ii) *There exists a measurable homomorphism $\Delta : G \to (0, \infty)$, such that $\tilde{\lambda} = \Delta \cdot \lambda$ and $\lambda \circ \tilde{\theta}_r^{-1} = \Delta_r \lambda$ for all $r \in G$.*

(iii) *When $\|\lambda\| < \infty$, we have $\Delta \equiv 1$, and λ is also right-invariant with $\tilde{\lambda} = \lambda$.*

In particular, (i) shows that any s-finite, left-invariant measure on G equals $c\lambda$ for some $c \in [0, \infty]$. In the classical context, $\tilde{\Delta}$ is known as the *modular function* of G. The present s-finite versions lead to significant simplifications in subsequent arguments.

Proof: (i) Since λ is σ-finite, we may choose a measurable function $h > 0$ on G with $\lambda h = 1$, and define $\Delta_r = \lambda(h \circ \tilde{\theta}_r) \in (0, \infty]$ for $r \in G$. Using Fubini's theorem (three times), the G-invariance of μ and λ, and the definitions of Δ and $\tilde{\lambda}$, we get for any $f \geq 0$

$$\begin{aligned}
\mu(f\Delta) &= \iint \mu(dr\, ds)\, f(r,s) \int \lambda(dp)\, h(pr) \\
&= \int \lambda(dp) \iint \mu(dr\, ds)\, f(p^{-1}r, s)\, h(r) \\
&= \iint \mu(dr\, ds)\, h(r) \int \lambda(dp)\, f(p^{-1}, s) \\
&= \int \tilde{\lambda}(dp) \iint \mu(dr\, ds)\, h(r)\, f(p,s).
\end{aligned}$$

In particular, we may take $\mu = \lambda$ for a singleton S to get $\tilde{\lambda} f = \lambda(f\Delta)$, and so $\tilde{\lambda} = \Delta \cdot \lambda$. This gives $\lambda = \tilde{\Delta} \cdot \tilde{\lambda} = \tilde{\Delta}\Delta \cdot \lambda$, hence $\tilde{\Delta}\Delta = 1$ a.e. λ, and finally $\Delta \in (0, \infty)$ a.e. λ. Thus, for general S,

$$\mu(f\Delta) = \int \lambda(dp)\, \Delta(p) \iint \mu(dr\, ds)\, h(r)\, f(p,s).$$

Since $\Delta > 0$, we have $\mu(\cdot \times S) \ll \lambda$, and so $\Delta \in (0, \infty)$ a.e. $\mu(\cdot \times S)$, which yields the simplified formula

$$\mu f = \int \lambda(dp) \iint \mu(dr\, ds)\, h(r)\, f(p,s),$$

showing that $\mu = \lambda \otimes \nu$ with $\nu f = \mu(h \otimes f)$. If μ is σ-finite, then $\mu f < \infty$ for some measurable function $f > 0$ on $G \times S$, and Fubini's theorem yields $\nu f(r, \cdot) < \infty$ for $r \in G$ a.e. λ, which shows that even ν is σ-finite. The reverse implication is obvious.

(ii) For every $r \in G$, the measure $\lambda \circ \tilde{\theta}_r^{-1}$ is left-invariant, since θ_p and $\tilde{\theta}_r$ commute for all $p, r \in G$. Hence, (i) yields $\lambda \circ \tilde{\theta}_r^{-1} = c_r \lambda$ for some constants $c_r > 0$. Applying both sides to h gives $c_r = \Delta(r)$. The homomorphism property $\Delta(pq) = \Delta(p)\Delta(q)$ follows from the reverse semigroup property $\tilde{\theta}_{pq} = \tilde{\theta}_q \circ \tilde{\theta}_p$, and the measurability holds by the measurability of the group operation and Fubini's theorem.

(iii) From (ii) we get $\|\lambda\| = \Delta_r \|\lambda\|$ for all $r \in G$, and so $\Delta \equiv 1$ when $\|\lambda\| < \infty$. Then (ii) gives $\tilde{\lambda} = \lambda$ and $\lambda \circ \tilde{\theta}_r^{-1} = \lambda$ for all $r \in G$, which shows that λ is also right-invariant. □

An *action* of a group G on a space S is defined as a mapping $(r, s) \mapsto rs$ from $G \times S$ to S, such that $p(rs) = (pr)s$ and $\iota s = s$ for all $p, r \in G$ and $s \in S$, where ι denotes the identity element in G. If G and S are measurable spaces, we say that G acts *measurably* on S, if the action map is product-measurable. The *shifts* θ_r and *projections* π_s are defined by $rs = \theta_r s = \pi_s r$, and the *orbit* containing s is given by $\pi_s G = \{rs;\, r \in G\}$. The orbits form a partition of S, and we say that the action is *transitive* if $\pi_s G \equiv S$. For $s, s' \in S$, we write $s \sim s'$ if s and s' belong to the same orbit, so that $s = rs'$ for some $r \in G$. If G acts on both S and T, its *joint action* on $S \times T$ is given by $r(s, t) = (rs, rt)$.

When G is a group acting on a space S, we say that a set $B \subset S$ is *G-invariant*, if $B \circ \theta_r^{-1} = B$ for all $r \in G$, and a function f on S is *G-invariant* if $f \circ \theta_r = f$. When S is measurable with σ-field $\mathcal{S}$, the class of G-invariant sets in $\mathcal{S}$ is again a σ-field, denoted by $\mathcal{I}_S$. For a transitive group action, we have $\mathcal{I}_S = \{\emptyset, S\}$, and every invariant function is a constant. When the group action is measurable, a measure ν on S is said to be *G-invariant*, if $\nu \circ \theta_r^{-1} = \nu$ for all $r \in G$.

When G is a measurable group with Haar measure λ, acting measurably on S, we say that G acts *properly* on S, if there exists a *normalizing* function $g > 0$ on S, such that g is measurable with $\lambda(g \circ \pi_s) < \infty$ for all $s \in S$. We may then define a kernel φ on S by

$$\varphi_s = \frac{\lambda \circ \pi_s^{-1}}{\lambda(g \circ \pi_s)}, \quad s \in S. \tag{2}$$

We list some basic properties of the measures φ_s.

Lemma 7.2 *(orbit measures) The measures φ_s are σ-finite with $\varphi_s = \varphi_{s'}$ for $s \sim s'$, and $\varphi_s \perp \varphi_{s'}$ for $s \not\sim s'$. Furthermore,*

$$\varphi_s \circ \theta_r^{-1} = \varphi_s = \varphi_{rs}, \quad r \in G,\ s \in S. \tag{3}$$

Proof: The φ_s are σ-finite since $\varphi_s g \equiv 1$. In (3), the first equality holds by the left-invariance of λ, and the second one holds since $\lambda \circ \pi_{rs}^{-1} = \Delta_r \lambda \circ \pi_s^{-1}$, by Theorem 7.1 (ii). The latter equation yields $\varphi_s = \varphi_{s'}$ when $s \sim s'$. The orthogonality holds for $s \not\sim s'$, since φ_s is confined to the orbit $\pi_s(G)$, which is universally measurable by Corollary A1.2. □

We proceed to show that any s-finite, G-invariant measure on S is a unique mixture of the measures φ_s. An invariant measure ν on S is said to be *ergodic*, if $\nu I \wedge \nu I^c = 0$ for all $I \in \mathcal{I}_S$. It is further said to be *extreme*, if for any decomposition $\nu = \nu' + \nu''$ with invariant ν' and ν'', all three measures are proportional.

Theorem 7.3 *(invariant measures) Let G be a measurable group with Haar measure λ, acting properly on S, and define the kernel φ by (2), for some normalizing function $g > 0$ on S. Then an s-finite measure ν on S is G-invariant, iff $\nu = \int m\, \mu(dm)$ for some s-finite measure μ on the range $\varphi(S)$. The measure μ is then unique and satisfies*

(i) $\nu = \int \nu(ds)\, g(s)\, \varphi_s = \int m\, \mu(dm),$

(ii) $\mu f = \int \nu(ds)\, g(s)\, f(\varphi_s), \quad f \geq 0.$

Furthermore, μ and ν are simultaneously σ-finite, there exists a σ-finite, G-invariant measure $\tilde{\nu} \sim \nu$, and ν is ergodic or extreme iff μ is degenerate.

Proof: Let ν be s-finite and G-invariant. Using (2), Fubini's theorem, and Lemma 7.1 (ii), we get for any $f \in \mathcal{S}_+$

$$\begin{aligned}
\int \nu(ds)\, g(s)\, \varphi_s f &= \int \nu(ds)\, g(s)\, \frac{\lambda(f \circ \pi_s)}{\lambda(g \circ \pi_s)} \\
&= \int \lambda(dr) \int \nu(ds)\, \frac{g(s)\, f(rs)}{\lambda(g \circ \pi_s)} \\
&= \int \lambda(dr) \int \nu(ds)\, \frac{g(r^{-1}s)\, f(s)}{\lambda(g \circ \pi_{r^{-1}s})} \\
&= \int \lambda(dr)\, \Delta_r \int \nu(ds)\, \frac{g(r^{-1}s)\, f(s)}{\lambda(g \circ \pi_s)} \\
&= \int \lambda(dr) \int \nu(ds)\, \frac{g(rs)\, f(s)}{\lambda(g \circ \pi_s)} \\
&= \int \nu(ds)\, f(s) = \nu f,
\end{aligned}$$

which proves the first relation in (i). The second relation follows with μ as in (ii), by the substitution rule for integrals.

Conversely, suppose that $\nu = \int m\, \mu(dm)$ for some s-finite measure μ on $\varphi(S)$. For any $M \in \mathcal{M}_S$, we have $A \equiv \varphi^{-1} M \in \mathcal{I}_S$ by Lemma 7.2, and so

$$\begin{aligned}
\varphi_s(g; A) &= \frac{(\lambda \circ \pi_s^{-1})(1_A g)}{\lambda(g \circ \pi_s)} \\
&= 1_A(s) = 1_M(\varphi_s).
\end{aligned}$$

Hence,

$$\begin{aligned}
(g \cdot \nu) \circ \varphi^{-1} M &= \nu(g; A) = \int \mu(dm)\, m(g; A) \\
&= \int_M \mu(dm) = \mu M,
\end{aligned}$$

which shows that μ is uniquely given by (ii).

The measure μ is again s-finite by (ii), and μ and ν are simultaneously σ-finite by (i), (ii), and the facts that $g > 0$ and $\varphi_s \neq 0$. For general ν, we may choose a bounded measure $\hat{\nu} \sim \nu$, and define $\tilde{\nu} = \hat{\nu}\varphi$. Then $\tilde{\nu}$ is invariant by Lemma 7.2, and $\tilde{\nu} \sim (g \cdot \nu)\varphi = \nu$ by (ii). It is also σ-finite, since $\tilde{\nu}g = \hat{\nu}\varphi g = \hat{\nu}S < \infty$ by (i).

Since μ is unique, ν is extreme iff μ is degenerate. For any $I \in \mathcal{I}_S$, we have

$$\varphi_s I \wedge \varphi_s I^c = \frac{\lambda G}{\lambda(g \circ \pi_s)} \Big\{ 1_I(s) \wedge 1_{I^c}(s) \Big\} = 0, \quad s \in S,$$

which shows that the φ_s are ergodic. Conversely, let $\nu = \int m \, \mu(dm)$ for some non-degenerate μ, so that $\mu M \wedge \mu M^c \neq 0$ for some $M \in \mathcal{M}_S$. Then $I = \varphi^{-1} M \in \mathcal{I}_S$ with $\nu I \wedge \nu I^c \neq 0$, which shows that ν is non-ergodic. Hence, ν is ergodic iff μ is degenerate. $\Box$

We turn to a study of invariant disintegrations. The following result shows how densities and disintegration kernels are affected by shifts of the underlying measures.

Lemma 7.4 *(shifted integrals and compositions) Let G be a measurable group, acting measurably on S and T. Then*

(i) *for any measure μ and measurable function $h \geq 0$ on S,*

$$(h \cdot \mu) \circ \theta_r^{-1} = (h \circ \theta_{r^{-1}}) \cdot (\mu \circ \theta_r^{-1}), \quad r \in G,$$

(ii) *for any measure ν on S and kernel $\mu \colon S \to T$,*

$$(\nu \otimes \mu) \circ \theta_r^{-1} = (\nu \circ \theta_r^{-1}) \otimes (\mu_{r^{-1}(\cdot)} \circ \theta_r^{-1}), \quad r \in G.$$

Proof: (i) For any $r \in G$ and $f \in \mathcal{S}_+$,

$$\begin{aligned} \Big\{(h \cdot \mu) \circ \theta_r^{-1}\Big\} f &= (h \cdot \mu)(f \circ \theta_r) \\ &= \mu(f \circ \theta_r) h \\ &= (\mu \circ \theta_r^{-1})(h \circ \theta_{r^{-1}}) f \\ &= \Big\{(h \circ \theta_{r^{-1}}) \cdot (\mu \circ \theta_r^{-1})\Big\} f. \end{aligned}$$

(ii) For any $r \in G$ and $f \in (\mathcal{S} \otimes \mathcal{T})_+$,

$$\begin{aligned} \Big\{(\nu \otimes \mu) \circ \theta_r^{-1}\Big\} f &= (\nu \otimes \mu)(f \circ \theta_r) \\ &= \int \nu(ds) \int \mu_s(dt) \, f(rs, rt) \\ &= \int \nu(ds) \int (\mu_s \circ \theta_r^{-1})(dt) \, f(rs, t) \\ &= \int (\nu \circ \theta_r^{-1})(ds) \int (\mu_{r^{-1}s} \circ \theta_r^{-1})(dt) \, f(s, t) \\ &= \Big\{(\nu \circ \theta_r^{-1}) \otimes (\mu_{r^{-1}(\cdot)} \circ \theta_r^{-1})\Big\} f. \end{aligned}$$

$\Box$

For G-invariant ν and $\nu \otimes \mu$, relation (ii) reduces to

$$\nu \otimes \mu = \nu \otimes \left\{\mu_{r^{-1}(\cdot)} \circ \theta_r^{-1}\right\}, \quad r \in G,$$

which justifies that we define *G-invariance* of a kernel μ by the condition

$$\mu_{rs} = \mu_s \circ \theta_r^{-1}, \quad r \in G, \ s \in S.$$

The general formulas in (i) and (ii) yield more complicated invariance conditions, needed for the construction of stationary densities and disintegrations in Section 7.6.

If the measure ν on S and kernel $\mu\colon S \to T$ are both G-invariant, the composition $\nu \otimes \mu$ is easily seen to be a jointly G-invariant measure on $S \times T$. We turn to the converse problem of constructing invariant disintegrations of a jointly invariant measure on a product space $S \times T$. We begin with an elementary skew factorization, which applies whenever S contains G as a factor.

Theorem 7.5 *(skew factorization) Let G be a measurable group with Haar measure λ, acting measurably on S and T, where T is Borel. Then a σ-finite measure ρ on $G \times S \times T$ is jointly G-invariant, iff $\rho = \lambda \otimes \nu \otimes \mu$ for some G-invariant kernels $\nu\colon G \to S$ and $\mu\colon G \times S \to T$, in which case*

$$\nu_r = \hat{\nu} \circ \theta_r^{-1}, \quad \mu_{r,s} = \hat{\mu}_{r^{-1}s} \circ \theta_r^{-1}, \qquad r \in G, \ s \in S,$$

for some measure $\hat{\nu}$ on S and kernel $\hat{\mu}\colon S \to T$. The measure $\hat{\nu} \otimes \hat{\mu}$ on $S \times T$ is then unique.

Proof: When T is a singleton, we define on $G \times S$ the skew transformation ϑ and shifts θ_r and θ'_r by

$$\vartheta(r, s) = (r, rs), \qquad \theta_r(p, s) = (rp, rs), \qquad \theta'_r(p, s) = (rp, s),$$

for any $p, r \in G$ and $s \in S$. Then $\vartheta^{-1} \circ \theta_r = \theta'_r \circ \vartheta^{-1}$, and the joint G-invariance of ρ yields

$$\rho \circ \vartheta \circ \theta_r'^{-1} = \rho \circ \theta_r^{-1} \circ \vartheta = \rho \circ \vartheta, \quad r \in G,$$

where $\vartheta = (\vartheta^{-1})^{-1}$. Hence, $\rho \circ \vartheta$ is invariant under shifts in G only, and Theorem 7.1 yields $\rho \circ \vartheta = \lambda \otimes \nu$, for some σ-finite measure ν on S, which implies $\rho = (\lambda \otimes \nu) \circ \vartheta^{-1}$.

For general T, we conclude that $\rho = (\lambda \otimes \chi) \circ \vartheta^{-1}$ for some σ-finite measure χ on $S \times T$, where ϑ is now defined by $\vartheta(r, s, t) = (r, rs, rt)$, for all $(r, s, t) \in G \times S \times T$. Since T is Borel, Theorem 1.23 yields a further disintegration $\chi = \hat{\nu} \otimes \hat{\mu}$, in terms of some measure $\hat{\nu}$ on S and kernel

$\hat{\mu}: S \to T$. Defining ν and μ as stated, in terms of $\hat{\nu}$ and $\hat{\mu}$, we note that both are G-invariant. For any $f \in (\mathcal{G} \otimes \mathcal{S} \otimes \mathcal{T})_+$, we get

$$\begin{aligned}(\lambda \otimes \nu \otimes \mu)f &= \int \lambda(dr) \int \hat{\nu} \circ \theta_r^{-1}(ds) \int \hat{\mu}_{r^{-1}s} \circ \theta_r^{-1}(dt)\, f(r,s,t)\\ &= \int \lambda(dr) \int \hat{\nu}(ds) \int \hat{\mu}_s(dt)\, f(r,rs,rt)\\ &= \left\{(\lambda \otimes \hat{\nu} \otimes \hat{\mu}) \circ \vartheta^{-1}\right\} f = \rho f,\end{aligned}$$

which proves the required disintegration. Since $\lambda \otimes \hat{\nu} \otimes \hat{\mu} = \rho \circ \vartheta$, the measure $\hat{\nu} \otimes \hat{\mu}$ is unique by Theorem 7.1. □

The disintegration problem for general S and T is much harder. Given a σ-finite, jointly G-invariant measure ρ on $S \times T$, we say that ρ admits a G-invariant disintegration, if $\rho = \nu \otimes \mu$ for some σ-finite, G-invariant measure ν on S, and a G-invariant kernel $\mu: S \to T$. Under suitable conditions, we prove the existence of such a disintegration:

Theorem 7.6 *(invariant disintegration) Let G be a measurable group with Haar measure λ, acting measurably on S and T, where T is Borel. Consider a σ-finite, G-invariant measure ρ on $S \times T$, such that $\rho(\cdot \times T) \ll \nu$ for some σ-finite, G-invariant measure ν on S. Then $\rho = \nu \otimes \mu$, for some G-invariant kernel $\mu: S \to T$. The required measure ν exists in particular when G acts properly on S.*

In general, the σ-finiteness of a measure is not preserved by measurable maps. In particular, the S-projection $\rho(\cdot \times T)$ of a σ-finite measure ρ on $S \times T$ need not be σ-finite.

The following technical result will be useful, here and below.

Lemma 7.7 *(invariance sets) Let G be a measurable group with Haar measure λ, acting measurably on S and T. Fix a kernel $\mu : S \to T$, and a measurable, jointly G-invariant function f on $S \times T$. Then the following sets in S and $S \times T$ are G-invariant:*

$$\begin{aligned}A &= \left\{s \in S;\, \mu_{rs} = \mu_s \circ \theta_r^{-1},\ r \in G\right\},\\ B &= \left\{(s,t) \in S \times T;\, f(rs,t) = f(ps,t),\ (r,p) \in G^2 \ \text{a.e. } \lambda^2\right\}.\end{aligned}$$

Proof: When $s \in A$, we get for any $r, p \in G$

$$\begin{aligned}\mu_{rs} \circ \theta_p^{-1} &= \mu_s \circ \theta_r^{-1} \circ \theta_p^{-1}\\ &= \mu_s \circ \theta_{pr}^{-1} = \mu_{prs},\end{aligned}$$

which shows that even $rs \in A$. Conversely, $rs \in A$ implies $s = r^{-1}(rs) \in A$, and so $\theta_r^{-1} A = A$.

Now let $(s,t) \in B$. Then Lemma 7.1 (ii) yields $(qs,t) \in B$ for any $q \in G$, and the invariance of λ gives

$$f(q^{-1}rqs,t) = f(q^{-1}p\,qs,t), \quad (r,p) \in G^2 \text{ a.e. } \lambda^2,$$

which implies $(qs,qt) \in B$, by the joint G-invariance of f. Conversely, $(qs,qt) \in B$ implies $(s,t) = q^{-1}(qs,qt) \in B$, and so $\theta_q^{-1}B = B$. □

Proof of Theorem 7.6: Applying Theorem 7.5 to the G-invariant measures $\lambda \otimes \rho$ on $G \times S \times T$ and $\lambda \otimes \nu$ on $G \times S$, we obtain a G-invariant kernel $\varphi\colon G \times S \to T$ with $\lambda \otimes \rho = \lambda \otimes \nu \otimes \varphi$. Introducing the kernels $\varphi_r(p,s) = \varphi(rp,s)$, writing $f_r(p,s,t) = f(r^{-1}p,s,t)$ for measurable functions $f \geq 0$ on $G \times S \times T$, and using the invariance of λ, we get for any $r \in G$

$$\begin{aligned}(\lambda \otimes \nu \otimes \varphi_r)f &= (\lambda \otimes \nu \otimes \varphi)f_r \\ &= (\lambda \otimes \rho)f_r = (\lambda \otimes \rho)f,\end{aligned}$$

which shows that $\lambda \otimes \rho = \lambda \otimes \nu \otimes \varphi_r$. Hence, by the a.e. uniqueness,

$$\varphi(rp,s) = \varphi(p,s), \quad (p,s) \in G \times S \text{ a.e. } \lambda \otimes \nu, \quad r \in G.$$

Now let A be the set of all $s \in S$ satisfying

$$\varphi(rp,s) = \varphi(p,s), \quad (r,p) \in G^2 \text{ a.e. } \lambda^2,$$

and note that $\nu A^c = 0$ by Fubini's theorem. By Lemma 7.1 (ii), the defining condition is equivalent to

$$\varphi(r,s) = \varphi(p,s), \quad (r,p) \in G^2 \text{ a.e. } \lambda^2,$$

and so for any $g,h \in \mathcal{G}_+$ with $\lambda g = \lambda h = 1$,

$$(g \cdot \lambda)\varphi(\cdot,s) = (h \cdot \lambda)\varphi(\cdot,s) = \mu(s), \quad s \in A. \tag{4}$$

To extend this to an identity, we may redefine $\varphi(r,s) = 0$ when $s \in A^c$. That will not affect the disintegration of $\lambda \otimes \rho$, and by Lemma 7.7 it even preserves the G-invariance of φ. Fixing a $g \in \mathcal{G}_+$ with $\lambda g = 1$, we get for any $f \in (\mathcal{S} \otimes \mathcal{T})_+$

$$\begin{aligned}\rho f &= (\lambda \otimes \rho)(g \otimes f) \\ &= (\lambda \otimes \nu \otimes \varphi)(g \otimes f) \\ &= \{\nu \otimes (g \cdot \lambda)\varphi\}f \\ &= (\nu \otimes \mu)f,\end{aligned}$$

which shows that $\rho = \nu \otimes \mu$. Finally, we see from (4) and the G-invariance of φ and λ that

$$\begin{aligned}\mu_s \circ \theta_r^{-1} &= \int \lambda(dp)\, g(p)\, \varphi(p,s) \circ \theta_r^{-1} \\ &= \int \lambda(dp)\, g(p)\, \varphi(rp,rs) \\ &= \int \lambda(dp)\, g(r^{-1}p)\, \varphi(p,rs) = \mu_{rs},\end{aligned}$$

which shows that μ is G-invariant.

Now suppose that G acts properly on S. Since ρ is σ-finite and jointly G-invariant, the projection $\rho(\cdot \times T)$ is s-finite and G-invariant, and so by Theorem 7.3 there exists a σ-finite, G-invariant measure $\nu \sim \rho(\cdot \times T)$. □

We proceed with an invariant version of the Radon–Nikodym theorem.

Theorem 7.8 *(absolute continuity) Let G act measurably on S, and consider any s-finite, G-invariant measures μ and ν on S. Then these conditions are equivalent:*

(i) $\mu \ll \nu$ *on* $\mathcal{S}$,

(ii) $\mu \ll \nu$ *on* $\mathcal{I}_S$.

When μ and ν are σ-finite, it is also equivalent that

(iii) $\mu = h \cdot \nu$, *for some measurable, G-invariant function $h \in \mathcal{S}_+$.*

All conditions hold automatically, when the group action is transitive and $\nu \neq 0$.

Proof: The implications (iii) ⇒ (i) ⇒ (ii) being trivial, it suffices to prove the reverse statements.

(ii) ⇒ (i): Assume (ii), and fix any $B \in \mathcal{S}$ with $\nu B = 0$. Define $f(s) = \int 1_B(rs)\, \tilde{\lambda}(dr)$ and $A = \{s \in S;\, f(s) > 0\}$, and note that f and A are $\mathcal{S}$-measurable by Fubini's theorem. For any $r \in G$, we have $f(rs) = f(s)$ by the right-invariance of $\tilde{\lambda}$, and so $\theta_r^{-1} A = A$, which means that A is G-invariant and therefore belongs to $\mathcal{I}_S$. Next, use Fubini's theorem for s-finite measures, and the G-invariance of ν to get $\nu f = \|\tilde{\lambda}\|\, \nu B = 0$, which implies $f = 0$ a.e. ν, and therefore $\nu A = 0$. Then $\mu A = 0$ by hypothesis, and so $\mu f = 0$. Using the s-finite version of Fubini's theorem and the G-invariance of μ, we get $\|\tilde{\lambda}\|\, \mu B = 0$, which gives $\mu B = 0$, since $\tilde{\lambda} \neq 0$. This proves (i).

(i) ⇒ (iii): Let μ and ν be σ-finite with $\mu \ll \nu$. Introduce a singleton $T = \{t\}$, equipped with the trivial group action $(r, t) \mapsto t$, $r \in G$. The measure $\rho = \mu \otimes \delta_t$ on $S \times T$ is again σ-finite and jointly G-invariant, and $\rho(\cdot \times T) = \mu \ll \nu$. Hence, Theorem 7.6 yields a σ-finite, G-invariant kernel $\varphi\colon S \to T$, such that $\rho = \nu \otimes \varphi$. The function $h(s) = \varphi_s 1$ is finite, measurable, and invariant, and for any $f \in \mathcal{S}_+$,

$$\begin{aligned} \mu f &= \rho(f \otimes 1) = (\nu \otimes \varphi)(f \otimes 1) \\ &= \int \nu(ds)\, f(s)\, \varphi_s 1 = \nu(fh), \end{aligned}$$

which implies $\mu = h \cdot \nu$, thus proving (iii).

For transitive group actions, we have $\mathcal{I}_S = \{\emptyset, S\}$, and (ii) holds trivially when $\nu \neq 0$. □

7.2 Invariant Representations and Palm Kernels

Let G be a measurable group with Haar measure λ, acting measurably on some Borel spaces S and T. Given a random measure ξ on S and a random element η in T, we say that the pair (ξ,η) is *jointly G-stationary* if $\theta_r(\xi,\eta) \overset{d}{=} (\xi,\eta)$ for all $r \in G$, where $\theta_r\mu = \mu\circ\theta_r^{-1}$ and $\theta_r t = rt$ for any $\mu \in \mathcal{M}_S$ and $t \in T$. When $S = G$, the associated *Palm measure* $Q_{\xi,\eta}$ was defined in Section 5.1 by

$$Q_{\xi,\eta} f = E\int f(r^{-1}\eta)\, g(r)\, \xi(dr), \quad f \in \mathcal{T}_+,$$

for any measurable function $g \geq 0$ on G with $\lambda g = 1$. Since g is arbitrary, the formula extends to

$$(\lambda \otimes Q_{\xi,\eta}) f = E\int f(r, r^{-1}\eta)\, \xi(dr), \quad f \in (\mathcal{G}\otimes\mathcal{T})_+.$$

Putting $\vartheta(r,t) = (r, rt)$ for $r \in G$ and $t \in T$, we may write this relation as

$$(\lambda \otimes Q_{\xi,\eta})(f\circ\vartheta) = E\int f(r,\eta)\, \xi(dr) = C_{\xi,\eta} f, \tag{5}$$

where the last equality defines the *Campbell measure* $C_{\xi,\eta}$ of the pair (ξ,η). Note that (5) agrees with the skew factorization in Theorem 7.5, for the special case of a singleton S. A similar approach applies when S contains G as a factor.

For general S, we may base our definition of Palm measures on the invariant disintegrations in Theorem 7.6. Here we begin with a basic representation of stationary random measures, along with some elementary remarks.

Theorem 7.9 *(stationarity and invariance) Let G act measurably on S and T, where S is Borel. Consider a random measure ξ on S and a G-stationary random element η in T, such that $C_{\xi,\eta}$ is σ-finite. Then the conditions*

(i) *ξ and η are jointly G-stationary,*
(ii) *$E(\xi|\,\eta)$ and η are jointly G-stationary,*
(iii) *$C_{\xi,\eta}$ is jointly G-invariant,*
(iv) *$E(\xi|\,\eta) = \mu\circ\eta$ a.s., for some σ-finite, G-invariant kernel $\mu: T \to S$,*

are related by (i) $\Rightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv). *The kernel μ in* (iv) *is unique a.s. $\mathcal{L}(\eta)$, and the uniqueness holds identically when G acts transitively on T.*

The relation in (iv) may be written more explicitly as

$$E(\xi|\,\eta)\, f = E(\xi f|\,\eta) = \int \mu(\eta, ds)\, f(s), \quad f \in \mathcal{S}_+.$$

If ξ is η-measurable, then all four conditions are equivalent, and the representation in (iv) becomes $\xi = \mu\circ\eta$ a.s. When G acts directly on Ω, we get $\xi = \mu$ a.s., which means that ξ has a G-invariant version. If $T = G\times U$, for some measurable space U, the skew transformation in Theorem 7.5 yields an explicit representation of the kernel μ. Similar formulas can be obtained for general T, by means of the inversion kernel in Theorem 7.14 below.

Proof, (i) ⇒ (iii): If ξ and η are jointly G-stationarity, then for any $f \in (\mathcal{S} \otimes \mathcal{T})_+$ and $r \in G$,

$$\begin{aligned}\left(C_{\xi,\eta} \circ \theta_r^{-1}\right) f &= E \int \xi(ds)\, f(rs, r\eta)\\ &= E \int (\xi \circ \theta_r^{-1})(ds)\, f(s, r\eta)\\ &= E \int \xi(ds)\, f(s, \eta) = C_{\xi,\eta} f,\end{aligned}$$

which shows that $C_{\xi,\eta}$ is jointly G-invariant.

(ii) ⇒ (iii): The previous proof applies with ξ replaced by $\zeta = E(\xi | \eta)$, since the pairs (ξ, η) and (ξ, ζ) have the same Campbell measure.

(iii) ⇒ (iv): Since S is Borel and the measures $C_{\xi,\eta}$ and $\nu = \mathcal{L}(\eta)$ are jointly G-invariant with $C_{\xi,\eta}(S \times \cdot) \ll \nu$, Theorem 7.6 yields $C_{\xi,\eta} \cong \nu \otimes \mu$, for some σ-finite, G-invariant kernel $\mu: T \to S$. Thus,

$$\begin{aligned}E(\, \xi f;\, \eta \in B\,) &= C_{\xi,\eta}(f \otimes 1_B)\\ &= E\Big\{(\mu \circ \eta) f;\, \eta \in B\Big\},\end{aligned}$$

which implies $E(\xi f | \eta) = (\mu \circ \eta) f$ a.s. This shows that $E(\xi | \eta)$ has the regular version $\zeta = \mu \circ \eta$.

(iv) ⇒ (ii): Assuming $\zeta = \mu \circ \eta$, and using the G-stationarity of η and the G-invariance of μ, we get for any $r \in G$

$$\begin{aligned}\theta_r(\zeta, \eta) &= (\mu_\eta \circ \theta_r^{-1}, r\eta) = (\mu_{r\eta}, r\eta)\\ &\overset{d}{=} (\mu_\eta, \eta) = (\zeta, \eta),\end{aligned}$$

which shows that η and ζ are jointly G-stationary. □

Under further regularity conditions, we can disintegrate the invariant measure $C_{\xi,\eta}$ both ways, to obtain some dual, G-invariant disintegrations

$$C_{\xi,\eta} = \nu \otimes \mu \cong \nu' \otimes \mu'.$$

In explicit notation, this means that for any measurable functions $f \geq 0$ on $S \times T$,

$$\begin{aligned}E \int \xi(ds)\, f(s, \eta) &= \int \nu(ds) \int \mu_s(dt)\, f(s, t)\\ &= \int \nu'(dt) \int \mu'(ds)\, f(s, t).\end{aligned}$$

In particular, this holds when T is Borel and either $E\xi$ is σ-finite or G acts properly on S.

Both disintegrations are significant. Thus, if the intensity measure $E\xi$ is σ-finite, we may take $\nu = E\xi$, and choose the μ_s to be G-invariant versions of the Palm distributions $\mathcal{L}(\eta \,\|\, \xi)_s$. In general, we may assume G to act properly on S, in which case ξ has a σ-finite, G-invariant supporting measure ν, and we get a disintegration into a G-invariant family of Palm measures, again denoted by $\mathcal{L}(\eta \,\|\, \xi)_s$.

The previous methods also yield invariant disintegrations of moment measures. Here we consider a measurable group G with Haar measure λ, acting measurably on some Borel spaces S and T, and let ξ and η be jointly G-stationary random measures on S and T, respectively, such that the moment measure $E(\xi \otimes \eta)$ on $S \times T$ is σ-finite. Assuming that $E\xi$ is σ-finite or that G acts properly on S, we get a disintegration $E(\xi \otimes \eta) = \nu \otimes \mu$, in terms of a G-invariant measure ν on S and a G-invariant kernel $\mu\colon S \to T$. In explicit notation, we may write this for any $f \in (\mathcal{S} \otimes \mathcal{T})_+$ as

$$E\int \xi(ds) \int \eta(dt)\, f(s,t) = \int \nu(ds) \int \mu_s(dt)\, f(s,t).$$

Iterating this result, we obtain some n-fold disintegrations of the moment measure $E\xi^n$ on S^n.

Next, suppose that ξ is a point process on S. If ξ is G-stationary, then the factorial product measures $\xi^{(n)}$ are jointly G-stationary, by Lemma 1.12, and so the *factorial moment measures* $E\xi^{(n)}$ are jointly G-invariant. Thus, under suitable σ-finiteness conditions, we have an iterated disintegration of the form $E\xi^{(n)} = \mu_1 \otimes \cdots \otimes \mu_n$, or more explicitly

$$\begin{aligned} E\xi^{(n)} f = &\int \mu_1(ds_1) \int \mu_2(s_1, ds_2) \int \cdots \\ &\times \int \mu_n(s_1, \ldots, s_{n-1}, ds_n)\, f(s_1, \ldots, s_n), \quad f \in (\mathcal{S}^{\otimes n})_+. \end{aligned}$$

Similarly, the n-th order reduced Campbell measure $C_\xi^{(n)}$ is jointly G-invariant on $S^n \times \mathcal{N}_S$, and so, under appropriate regularity conditions, it admits a G-invariant disintegration $C_\xi^{(n)} = \nu \otimes \mu$. Here ν is a σ-finite, jointly G-invariant measure on S^n with $\nu \sim E\xi^{(n)}$, and $\mu\colon S \to T$ is a G-invariant version of the *reduced Palm kernel* $\mathcal{L}(\xi \,\|\, \xi^{(n)})$.

In all the mentioned cases, we may apply the various methods of skew factorization and measure inversion to derive some more explicit disintegration formulas. In the subsequent discussion, we will mostly focus our attention on the univariate case.

We proceed to characterize the invariant Palm kernels $\mathcal{L}(\eta \,\|\, \xi)$ of a G-stationary random element η with given distribution ν, here allowed to be unbounded but σ-finite.

Theorem 7.10 *(Palm criterion, Getoor, OK) Let G act measurably on some Borel spaces S and T. Fix some σ-finite, G-invariant measures μ on S and*

ν on T, along with a G-invariant kernel $\chi: S \to T$, and put $\rho = \mu \otimes \chi$. Then the conditions

(i) *$\chi_s \ll \nu$ on $\mathcal{I}_T$ for $s \in S$ a.e. μ,*
(ii) *$\rho(S \times \cdot) \ll \nu$ on $\mathcal{I}_T$,*
(iii) *$\rho = C_{\xi,\eta}$, for some G-stationary pair (ξ, η) with $\mathcal{L}(\eta) = \nu$,*

are related by (i) $\Rightarrow$ (ii) $\Leftrightarrow$ (iii). *They are further equivalent, under each of the conditions*

(iv) *G acts transitively on S or T,*
(v) *G acts properly on S, and μ is ergodic,*
(vi) *G acts properly on T, and ν is ergodic.*

In (iii)*, we may choose $\xi = \kappa \circ \eta$ for some G-invariant kernel $\kappa: T \to S$, and under* (iv)*, we may replace* (i) *by the same condition for a fixed $s \in S$.*

Proof: Each of the conditions (i) and (iii) clearly implies (ii). Now assume (ii). Since both sides are s-finite and G-invariant, the relation $\rho(S \times \cdot) \ll \nu$ extends to $\mathcal{T}$ by Theorem 7.8, and so Theorem 7.6 yields $\rho \overset{\sim}{=} \nu \otimes \kappa$ for some G-invariant kernel $\kappa: T \to S$. Choosing a random element η in T with $\mathcal{L}(\eta) = \nu$, and putting $\xi = \kappa \circ \eta$, we see from Theorem 7.9 that ξ and η are jointly G-stationary with $\rho = C_{\xi,\eta}$, which proves (iii). It remains to show that (ii) implies (i), under each of the conditions (iv)$-$(vi). Thus, we may henceforth assume (ii).

Since χ is G-invariant, $\chi_{rs} = \chi_s$ on $\mathcal{I}_T$ for all $r \in G$ and $s \in S$. If G acts transitively on S, we get $\chi_s = \chi_{s'}$ on $\mathcal{I}_T$ for all $s, s' \in S$. This gives $\rho = \mu \otimes \chi_s$ on $S \times \mathcal{I}_T$ for every s, and so (ii) implies (i) for all $s \in S$. If instead G acts transitively on T, we have $\mathcal{I}_T = \{\emptyset, T\}$, and $\chi_s \ll \nu$ holds trivially, since $\|\nu\| > 0$. This shows that (ii) and (iv) together imply (i) for every fixed s.

Next, assume (v). Choosing $g > 0$ to be measurable on S with $\lambda(g \circ \pi_s) < \infty$, we may define φ as in Lemma 7.2. Since μ is G-invariant and ergodic, Theorem 7.3 yields $\rho = c\,\varphi_a$, for some $a \in S$ and $c \geq 0$, and we note that $c = \rho g$. Since $\varphi_s = \varphi_a$ iff $s \sim a$, we get

$$\begin{aligned} O_a \equiv \pi_a G &= \Big\{ s \in S;\ \rho = (\rho\, g)\, \varphi_s \Big\} \\ &= \varphi^{-1}\{\varphi_a\} \in \mathcal{I}_S. \end{aligned}$$

The G-invariance of χ yields

$$(\rho\, g)\, \|\varphi_a\|\, \chi_s = \rho(S \times \cdot) \ll \nu \ \text{ on } \ \mathcal{I}_T, \quad s \in O_a,$$

which implies (i) on O_a. Furthermore, the G-invariance of O_a yields

$$\rho\, O_a^c = (\rho\, g)\, \varphi_a O_a^c = (\rho\, g)\, 1_{O_a^c}(a) = 0.$$

This shows that (ii) and (v) imply (i).

Finally, assume (vi), so that $\nu = c\varphi_b$ by Theorem 7.3, for some $b \in T$, where $c = \nu g$ as before. Again $O_b = \pi_b G$ lies in $\mathcal{I}_T$, and for every $A \in \mathcal{I}_T$ we have either $A \supset O_b$ or $A \subset O_b^c$. Since $\nu O_b > 0$ unless $\nu = 0$, it remains to show that $\chi_s Q_b^c = 0$ for $s \in S$ a.e. μ, which is obvious from (ii), since $\nu O_b^c = 0$. Thus, (ii) and (vi) imply (i). $\square$

If ξ is a random measure on G with bounded intensity measure $E\xi$, we can construct an associated *stationary* pseudo-random measure $\tilde{\xi}$ on G, and then form the associated Palm distribution Q. When G is unimodal, it is natural to interpret the latter as a suitably shifted Palm measure of ξ. This version has important applications in Section 13.10.

Theorem 7.11 *(centered Palm distribution) Let ξ be a random measure on a unimodal lcscH group G with Haar measure λ, such that $E\|\xi\| = 1$. Then the pseudo-random measure $\tilde{\xi}$ on G with distribution*

$$Ef(\tilde{\xi}) = E\int f(\theta_r\xi)\,\lambda(dr), \quad f \geq 0,$$

is stationary with intensity $E\tilde{\xi} = \lambda$ and Palm distribution

$$E^0 f(\tilde{\xi}) = E\int f(\theta_r^{-1}\xi)\,\xi(dr) = \int E\xi(dr)\,E\Big\{f(\theta_r^{-1}\xi)\Big\|\,\xi\Big\}_r.$$

Proof: To prove that $\tilde{\xi}$ is stationary, let $s \in G$ be arbitrary, and conclude from the left invariance of λ that

$$\begin{aligned} Ef(\theta_s\tilde{\xi}) &= E\int f(\theta_{sr}\xi)\,\lambda(dr) \\ &= E\int f(\theta_r\xi)\,(\theta_s\lambda)(dr) \\ &= E\int f(\theta_r\xi)\,\lambda(dr) = Ef(\tilde{\xi}). \end{aligned}$$

Furthermore, by Fubini's theorem and the right invariance of λ,

$$\begin{aligned} E\tilde{\xi}f &= E\int \lambda(dr)\int(\theta_r\xi)(ds)f(s) \\ &= E\int \lambda(dr)\int \xi(ds)\,f(rs) \\ &= E\int \xi(ds)\int \lambda(dr)\,f(rs) \\ &= E\|\xi\|\,\lambda f = \lambda f, \end{aligned}$$

and so $E\tilde{\xi} = \lambda$. Similarly, choosing $g \geq 0$ with $\lambda g = 1$, we obtain

$$\begin{aligned} E^0 f(\tilde{\xi}) &= E\int f(\theta_r^{-1}\tilde{\xi})\,g(r)\,\tilde{\xi}(dr) \\ &= E\int \lambda(ds)\int f\Big(\theta_r^{-1}\theta_s\xi\Big)\,g(r)\,(\theta_s\xi)(dr) \end{aligned}$$

$$
\begin{aligned}
&= E\int \lambda(ds) \int f\left(\theta_{sr}^{-1}\theta_s\xi\right) g(sr)\,\xi(dr)\\
&= E\int f(\theta_r^{-1}\xi)\,\xi(dr) \int \lambda(ds)\, g(sr)\\
&= E\int (\theta_r^{-1}\xi)\,\xi(dr),
\end{aligned}
$$

which proves the first expression for $E^0 f(\tilde{\xi})$. The second expression follows by Palm disintegration. □

We proceed with a general coupling property of Palm measures. Say that G acts *injectively* on S, if all projection maps π_s are injective.

Theorem 7.12 *(shift coupling, Thorisson, OK) Let G act measurably on some Borel spaces S and T. Consider a jointly G-stationary pair of a random measure ξ on S and a random element η in T, such that $E\xi$ is σ-finite and η is G-ergodic. Then each of the conditions*

(i) *G acts transitively on S,*
(ii) *$\|\lambda\| = 1$,*

implies $\mathcal{L}(\eta \,\|\, \xi)_s = \mathcal{L}(\tau_s\eta)$, $s \in S$ a.e. $E\xi$, for some random elements τ_s in G, $s \in S$. If G acts injectively on S, we can choose the τ_s to satisfy $\tau_{rs} = r\tau_s$, for all $r \in G$ and $s \in S$.

Proof: First assume (i), and fix any $A \in \mathcal{I}_T$. By Theorem 7.6, we may choose a G-invariant version of $\mu = \mathcal{L}(\eta \,\|\, \xi)$. Since the S-action is transitive and η is ergodic, $\mu_s A$ is independent of s, and $P\{\eta \in A\} \in \{0, 1\}$. For any $B \in \mathcal{S}$ and $s \in S$, we get

$$
\begin{aligned}
E(\xi B)\,\mu_s A &= E(\xi B;\, \eta \in A)\\
&= E(\xi B)\, P\{\eta \in A\},
\end{aligned}
$$

and so $\mu_s A = P\{\eta \in A\}$, which shows that $\mu_s = \mathcal{L}(\eta)$ on $\mathcal{I}_T$. Hence, $\mu_s = \mathcal{L}(\tau_s\eta)$ by Theorem A2.6, for some random elements τ_s in G.

Next assume (ii). Since η is G-stationary and ergodic, Theorem 7.3 yields $\mathcal{L}(\eta) = \varphi_b = \lambda \circ \pi_b^{-1}$ for some $b \in T$, where $\varphi_t = \lambda \circ \pi_t^{-1}$. Introduce the G-invariant set $O_b = \pi_b G = \varphi^{-1}\{\varphi_b\}$ and measures $\bar{\mu}_s = \int \lambda(dr)\, \mu_{rs}$. Using Theorem 7.3 again gives

$$
\begin{aligned}
\int E\xi(ds)\, \bar{\mu}_s O_b^c &= \int E\xi(ds) \int \lambda(dr)\, \mu_{rs} O_b^c\\
&= \int E\xi(ds)\, \mu_s O_b^c\\
&= E(\xi S;\, \eta \in O_b^c)\\
&\ll P\{\eta \in O_b^c\}\\
&= (\lambda \circ \pi_b^{-1}) O_b^c = 0,
\end{aligned}
$$

which shows that the probability measures $\bar{\mu}_s$ are a.e. restricted to O_b. Now a third application of Theorem 7.3 gives $\bar{\mu}_s = \lambda \circ \pi_b^{-1}$ a.e., and so $\mu_s A = \bar{\mu}_s A = P\{\eta \in A\}$ for any $A \in \mathcal{I}_T$, which shows that $\mu_s = \mathcal{L}(\eta)$ on $\mathcal{I}_T$ for all T. The required coupling now follows as before.

Now let G act injectively on S. For any $a \in S$, we may choose τ_a with $\mu_a = \mathcal{L}(\tau_a \eta)$. Since π_a is injective, any element $s \in \pi_a G$ has a unique representation $s = ra$, for some $r \in G$. Using the G-invariance of μ, we get for any $f \in \mathcal{T}_+$

$$Ef(r\tau_a\eta) = E\{f(r\eta) \| \xi\}_a = E\{f(\eta) \| \xi\}_{ra},$$

and so we may choose $\tau_{ra} = r\tau_a$. Then for any $s = pa$ and $r \in G$, we get $\tau_{rs} = rp\tau_a = r\tau_s$, which extends the relation to arbitrary $s \in \pi_a G$. To obtain the desired identity, it remains to choose one element a from each orbit. □

7.3 Measure Inversion

In the non-transitive case, we may base our representations and disintegrations on a suitable orbit selection. A measurable group action on S is said to be *σ-finite*, if there exists a measurable function $f > 0$ on S^2 such that $(\lambda \circ \pi_s^{-1}) f(s, \cdot) < \infty$ for all $s \in S$. It is further said to be *proper* if $f(s, \cdot)$ can be chosen to be independent of s, so that $\lambda(g \circ \pi_s) < \infty$ holds identically for some measurable function $g > 0$ on S. By a *selector of orbit elements* (or *orbit selector*) we mean a function $\alpha \colon S \to S$ such that $s \sim \alpha_s$, and $\alpha_s = \alpha_{s'}$ iff $s \sim s'$. Similarly, we define a *selector of orbit measures* as a function $\varphi \colon S \to \hat{\mathcal{M}}_S$, such that φ_s is supported by the orbit $\pi_s G$, and $\varphi_s = \varphi_{s'}$ iff $s \sim s'$.

A function f from a measurable space S to a Borel space T is said to be *universally measurable* (or *u-measurable* for short), if for any bounded measure ν on S, there exists a measurable function $f^\nu : S \to T$ satisfying $f = f^\nu$ a.e. ν, so that the set $\{f \neq f^\nu\}$ has outer ν-measure 0. Say that the group action on S is *u-proper*, if the defining property holds with both g and $g \cdot (\lambda \circ \pi_s^{-1})$ u-measurable. We prove that suitable orbit selectors exist when the group action is proper.

Theorem 7.13 *(orbit selection) Let G be a measurable group with Haar measure λ, acting σ-finitely on a Borel space S, and consider the conditions*

(i) *G acts properly on S,*

(ii) *there exists a measurable selector $\hat{\varphi}$ of orbit measures,*

(iii) *there exists a measurable selector α of orbit elements,*

along with their u-measurable counterparts (i*)–(iii*). *Then*

$$\text{(iii)} \Rightarrow \text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii}^*) \Rightarrow \text{(i}^*) \Rightarrow \text{(ii}^*).$$

This suggests that the listed properties are nearly equivalent.

Proof, (i) ⇒ (ii): Fix a measurable function $g > 0$ on S with $\lambda(g \circ \pi_s) < \infty$, and introduce on S the probability kernel

$$\hat{\varphi}_s = \frac{g \cdot (\lambda \circ \pi_s^{-1})}{\lambda(g \circ \pi_s)}, \quad s \in S,$$

which is measurable by Fubini's theorem and the measurability of the group action. For any $r \in G$ and $s \in S$, Theorem 7.1 (ii) yields $\lambda \circ \pi_{rs}^{-1} = \Delta_r(\lambda \circ \pi_s^{-1})$, and so $\hat{\varphi}_s = \hat{\varphi}_{s'}$ when $s \sim s'$. On the other hand, $\hat{\varphi}_s \perp \hat{\varphi}_{s'}$ when $s \not\sim s'$, since the orbits $\pi_s G$ are disjoint and u-measurable, by Corollary A1.2 (ii). Hence, $\hat{\varphi}_s = \hat{\varphi}_{s'}$ iff $s \sim s'$.

(i*) ⇒ (ii*): If $g > 0$ and $g \cdot (\lambda \circ \pi_s^{-1}) < \infty$ are only u-measurable, then $\hat{\varphi}$ has the same property, and the relation $\lambda \circ \pi_{rs}^{-1} = \Delta_r(\lambda \circ \pi_s^{-1})$ still yields $(\lambda \circ \pi_{rs}^{-1})fg = \Delta_r(\lambda \circ \pi_s^{-1})fg$. Now proceed as before.

(ii) ⇒ (iii*): Let $\hat{\varphi}$ be a measurable selector of orbit measures. Since $\hat{\mathcal{M}}_S$ is again Borel by Theorem 1.5, Corollary A1.2 (iii) yields a u-measurable function $f\colon \hat{\mathcal{M}}_S \to S$, such that $\hat{\varphi} \circ f \circ \hat{\varphi} = \hat{\varphi}$ on S. The function $\alpha = f \circ \hat{\varphi}$ is again u-measurable, and $s \sim s'$ implies $\alpha_s = f \circ \hat{\varphi}_s = f \circ \hat{\varphi}_{s'} = \alpha_{s'}$. The relation $\hat{\varphi} \circ \alpha = \hat{\varphi}$ also yields $\alpha_s \sim s$, and it shows that $\alpha_s = \alpha_{s'}$ implies $s \sim s'$.

(iii) ⇒ (i): Let α be a measurable selector of orbit elements, and fix a measurable function $f > 0$ on S^2, with $(\lambda \circ \pi_s^{-1})f(s, \cdot) < \infty$ for all $s \in S$. Then the function $g(s) = f(\alpha_s, s) > 0$ on S is again measurable. Choosing $p_s \in G$ with $\alpha_s = p_s s$, and using the G-invariance of α and Theorem 7.1 (ii), we get

$$\begin{aligned} \lambda(g \circ \pi_s) &= \int \lambda(dr)\, f(\alpha_s, rs) \\ &= \tilde{\Delta}(p_s)\, (\lambda \circ \pi_{\alpha_s}^{-1}) f(\alpha_s, \cdot) < \infty. \end{aligned} \tag{6}$$

(iii*) ⇒ (i*): If the selector α is only u-measurable, then so is $g(s) = f(\alpha_s, s)$. By the first equality in (6) and Fubini's theorem, $\lambda(g \circ \pi_s)$ is a measurable function of the pair (s, α_s), and hence u-measurable. Similarly, $\lambda(gh \circ \pi_s)$ is u-measurable for every measurable function $h \geq 0$. □

The projections $\pi_s\colon G \to S$ are not injective in general, and hence may have no inverses. However, we can always use an orbit selector α to construct a G-invariant probability kernel $\gamma\colon S \to G$, which may act as a measure-valued substitute. Recall that, for any $s \in S$, the associated *isotropy group* is given by $G_s = \{r \in G;\, rs = s\} = \pi_s^{-1}\{s\}$. For any $r \in G$ and $s \in S$, we define $\pi_s r = rs$ and $\tilde{\pi}_s r = (rs, r)$.

Theorem 7.14 *(inversion kernel) Let G be an lcscH group with Haar measure λ, acting properly on a Borel space S, and fix a u-measurable orbit* *selector α. Then there exists a unique, u-measurable, G-invariant probability kernel $\gamma\colon S \to G$, such that*

$$\nu(\lambda \circ \tilde{\pi}_\alpha^{-1}) = \nu(\lambda \circ \pi_\alpha^{-1}) \otimes \gamma, \quad \nu \in \hat{\mathcal{M}}_S. \tag{7}$$

For every $s \in S$, the group G_{α_s} has normalized Haar measure γ_{α_s}, both γ_s and λ are right G_{α_s}-invariant, and $\gamma_s \circ \pi_{\alpha_s}^{-1} = \delta_s$.

Equation (7) may be written more explicitly as

$$\int \nu(ds) \int \lambda(dr)\, f(r, r\alpha_s) = \int \nu(ds) \int \lambda(dr) \int \gamma_{\alpha_s}(dp)\, f(rp, r\alpha_s).$$

Proof: First, let G act transitively on S, so that $\alpha \in S$ is a constant, and (7) reduces to $\lambda \circ \tilde{\pi}_\alpha^{-1} = (\lambda \circ \pi_\alpha^{-1}) \otimes \gamma$, or in explicit notation,

$$\int \lambda(dr)\, f(r, r\alpha) = \int \lambda(dr) \int \gamma_\alpha(dp)\, f(rp, r\alpha). \tag{8}$$

Since G is Borel, and $\lambda \circ \tilde{\pi}_\alpha^{-1}$ is jointly G-invariant on $S \times G$ with σ-finite S-projection $\lambda \circ \pi_\alpha^{-1}$, Theorem 7.6 yields a unique, G-invariant kernel $\gamma\colon S \to G$ satisfying (8). The G-invariance gives $\gamma_\alpha \circ \theta_r^{-1} = \gamma_{r\alpha} = \gamma_\alpha$ for all $r \in G_\alpha$, which shows that γ_α is G_α-invariant. Writing $\hat{\gamma}_r = \gamma_{r\alpha}$, we get

$$\lambda\hat{\gamma}f = \int \lambda(dr) \int \gamma_\alpha(dq)\, f(rp\,q) = \Delta_p \lambda\hat{\gamma}f, \quad p \in G_\alpha.$$

Choosing f with $\lambda\hat{\gamma}f \in (0, \infty)$ gives $\Delta = 1$ on G_α, which shows that λ is right G_a-invariant. Since S is Borel, we may choose $f(r, s) = 1\{r\alpha \neq s\}$ in (8) to obtain $\|\lambda\|\gamma_\alpha G_\alpha^c = 0$, which implies $\gamma_\alpha G_\alpha^c = 0$. If instead $f(r, s) = g(s)\, 1\{r\alpha = s\}$ with $\lambda(g \circ \pi_\alpha) = 1$, we get $\|\gamma_\alpha\| = 1$, and so γ_α equals normalized Haar measure on G_α. For any $r \in G$, the G-invariance of γ yields

$$\begin{aligned}(\gamma_{r\alpha} \circ \pi_\alpha^{-1})f &= \int \gamma_\alpha(dp)\, f(rp\alpha) \\ &= \int \gamma_\alpha(dp)\, f(r\alpha) = f(r\alpha),\end{aligned}$$

and so $\gamma_s \circ \pi_\alpha^{-1} = \delta_s$ for all $s \in S$. Since γ_α is right G_α-invariant by Theorem 7.1 (iii), and θ_r and $\tilde{\theta}_p$ commute for all $r, p \in G$, we get for any $r \in G$ and $p \in G_\alpha$

$$\begin{aligned}\gamma_{r\alpha} \circ \tilde{\theta}_p^{-1} &= \gamma_\alpha \circ \theta_r^{-1} \circ \tilde{\theta}_p^{-1} \\ &= \gamma_\alpha \circ \tilde{\theta}_p^{-1} \circ \theta_r^{-1} \\ &= \gamma_\alpha \circ \theta_r^{-1} = \gamma_{r\alpha},\end{aligned}$$

which shows that the measures γ_s are right G_α-invariant.

In the non-transitive case, we may fix a measurable function $g > 0$ on S with $\lambda(g \circ \pi_s) < \infty$ for all $s \in S$, and put $f(s) = g(s)/\lambda(g \circ \pi_{\alpha_s}) > 0$. Since $\alpha(r\alpha_s) = \alpha_s$ for all $r \in G$ and $s \in S$, we have $\lambda(f \circ \pi_{\alpha_s}) = 1$ for all s, and so

$\nu(\lambda \circ \pi_\alpha^{-1})f = \|\nu\| < \infty$. Replacing f by an equivalent, strictly measurable function, which exists since ν and λ are σ-finite, we conclude that both sides of (7) are σ-finite. Then Theorem 7.6 yields a G-invariant kernel $\gamma^\nu : S \to G$ satisfying (7). Since α is G-invariant, we may replace ν in (7) by the measure $\hat{\nu} = \nu \circ \alpha^{-1}$, to obtain

$$\int \hat{\nu}(da)\,(\lambda \circ \tilde{\pi}_a^{-1})f = \int \hat{\nu}(da)\,\big\{(\lambda \circ \pi_a^{-1}) \otimes \gamma^\nu\big\}f.$$

Applying this to functions of the form $f(s,r) = g(s,r)\,h(\alpha_s)$, and using the G-invariance of α, we get the same relation with $\hat{\nu}$ replaced by $h \cdot \hat{\nu}$, and since h is arbitrary,

$$(\lambda \circ \tilde{\pi}_a^{-1})f = \big\{(\lambda \circ \pi_a^{-1}) \otimes \gamma^\nu\big\}f, \quad a \in S \text{ a.e. } \hat{\nu}.$$

Restricting f to a countable, measure-determining class, which exists since $S \times G$ is Borel, we get $\lambda \circ \tilde{\pi}_a^{-1} = (\lambda \circ \pi_a^{-1}) \otimes \gamma^\nu$ for $a \in S$ a.e. $\hat{\nu}$. Hence, the uniqueness in the transitive case yields $\gamma^\nu = \gamma^a$ for $a \in S$ a.e. $\hat{\nu}$, where $\gamma^s = \gamma^{\delta_s}$. Since $s \sim s'$ implies $\gamma^s = \gamma^{s'}$, by the G-invariance of α, we obtain $\gamma^\nu = \gamma^s$ for $s \in S$ a.e. ν. Thus, $\gamma_s^\nu = \gamma_s$ for $s \in S$ a.e. ν with $\gamma_s = \gamma_s^s$, which shows that $\gamma = (\gamma_s)$ is a u-measurable kernel from S to G. It is also G-invariant, since for any $r \in G$ and $s \in S$,

$$\begin{aligned}\gamma_{rs} &= \gamma_{rs}^{rs} = \gamma_{rs}^{s} \\ &= \gamma_s^s \circ \theta_r^{-1} = \gamma_s \circ \theta_r^{-1}.\end{aligned}$$

The remaining assertions follow from the corresponding statements in the transitive case. □

The inversion kernel allows us to extend the method of skew factorization to jointly G-invariant measures on more general product spaces $S \times T$. Here we highlight the basic construction. Let δ and δ_α denote the kernels $s \mapsto \delta_s$ and $s \mapsto \delta_{\alpha_s}$, and define the mappings ϑ_T and $\hat{\vartheta}_{A,T}$ on $G \times A \times T$ by

$$\vartheta_T(r,a,t) = (r,a,rt), \qquad \hat{\vartheta}_{A,T}(r,a,t) = (ra,rt).$$

Theorem 7.15 *(extended skew factorization) Let G be an lcscH group with Haar measure λ, acting properly on S and measurably on T, where S and T are Borel. Fix a u-measurable orbit selector α on S, with associated inversion kernel γ, and put $A = \alpha(S)$. Then an s-finite measure ρ on $S \times T$ is jointly G-invariant, iff*

$$\lambda \otimes \hat{\rho} = \rho(\gamma \otimes \delta_\alpha \otimes \delta) \circ \vartheta_T, \tag{9}$$

for some s-finite measure $\hat{\rho}$ on $A \times T$, in which case

$$\rho = (\lambda \otimes \hat{\rho}) \circ \hat{\vartheta}_{A,T}^{-1}. \tag{10}$$

Equations (9) *and* (10) *are equivalent, and provide a 1−1 correspondence between the sets of all s-finite, jointly G-invariant measures ρ on $S \times T$, and all s-finite measures $\hat{\rho}$ on $A \times T$. Furthermore, ρ and $\hat{\rho}$ are simultaneously σ-finite.*

The displayed equations are shorthand versions of

$$\begin{aligned}\int \lambda(dr) \iint \hat\rho(da\,dt)\, f(r,a,rt) &= \iint \rho(ds\,dt) \int \gamma_s(dr)\, f(r,\alpha_s,t),\\ \iint \rho(ds\,dt)\, f(s,t) &= \int \lambda(dr) \iint \hat\rho(da\,dt)\, f(ra,rt).\end{aligned}$$

For a singleton T, equations (9) and (10) reduce to the equivalent relations

$$\lambda \otimes \hat\nu = \nu(\gamma \otimes \delta_\alpha), \qquad \nu = (\lambda \otimes \hat\nu) \circ \hat\vartheta_A^{-1},$$

which exhibit a G-invariant measure ν on S as a mixture of G-invariant orbit measures. This is essentially equivalent to the more elementary representation in Theorem 7.3 above.

Proof: Let ρ be an s-finite, jointly G-invariant measure on $S \times T$, and define $\chi = \rho(\gamma \otimes \delta_\alpha \otimes \delta)$. Using the G-invariance of ρ and γ, we get for any $r \in G$

$$\begin{aligned}(\chi \circ \theta_r^{-1})f &= \iint \rho(ds\,dt) \int \gamma_s(dp)\, f(rp,\alpha_s,rt)\\ &= \iint \rho(ds\,dt) \int \gamma_s(dp)\, f(p,\alpha_s,t) = \chi f,\end{aligned}$$

which shows that χ is again jointly G-invariant. Hence, Theorem 7.5 yields $\chi = (\lambda \otimes \hat\rho) \circ \vartheta_T^{-1}$, for some s-finite measure $\hat\rho$ on $A \times T$, which proves (9).

Now let ρ be an s-finite measure on $S \times T$, satisfying (9) for some measure $\hat\rho$ on $A \times T$. Since $\gamma_s \circ \pi_{\alpha_s}^{-1} = \delta_s$ by Theorem 7.14, we obtain

$$\begin{aligned}(\chi \circ \hat\vartheta_A^{-1})f &= \chi(f \circ \hat\vartheta_A)\\ &= \iint \rho(ds\,dt) \int \gamma_s(dr)\, f(r\alpha_s,t) = \rho f, \qquad (11)\end{aligned}$$

and so by (9),

$$\begin{aligned}\rho &= \chi \circ \hat\vartheta_A^{-1}\\ &= (\lambda \otimes \hat\rho) \circ \vartheta_T^{-1} \circ \hat\vartheta_A^{-1}\\ &= (\lambda \otimes \hat\rho) \circ \vartheta_{A,T}^{-1},\end{aligned}$$

which proves (10).

Conversely, consider any s-finite measure $\hat\rho$ on $A \times T$, and define ρ by (10). Using the left-invariance of λ, we get for any $p \in G$ and $f \in (\mathcal{S} \otimes \mathcal{T})_+$

$$\begin{aligned}(\rho \circ \theta_p^{-1})f &= \iint \rho(ds\,dt)\, f(ps,pt)\\ &= \int \lambda(dr) \iint \hat\rho(da\,dt)\, f(p\,ra,p\,rt)\\ &= \int \lambda(dr) \iint \hat\rho(da\,dt)\, f(ra,rt)\\ &= \iint \rho(ds\,dt)\, f(s,t) = \rho f,\end{aligned}$$

which shows that ρ is jointly G-invariant. Using (10), the G-invariance of γ and α, and the right G_a-invariance of λ, we further obtain

$$\begin{aligned}\rho(\gamma\otimes\delta_\alpha\otimes\delta)f &= \iint \rho(ds\,dt)\int \gamma_s(dp)\,f(p,\alpha_s,t)\\ &= \int\lambda(dr)\iint\hat\rho(da\,dt)\int\gamma_{ra}(dp)\,f(p,\alpha_{ra},rt)\\ &= \int\lambda(dr)\iint\hat\rho(da\,dt)\int\gamma_a(dp)\,f(rp,a,rt)\\ &= \int\lambda(dr)\iint\hat\rho(da\,dt)\,f(r,a,rt)\\ &= \left\{(\lambda\otimes\hat\rho)\circ\vartheta_T^{-1}\right\}f,\end{aligned}$$

which proves (9).

Since $\hat\rho$ and $\lambda\otimes\hat\rho$ are simultaneously s-finite, so are ρ and $\hat\rho$. Similarly, $\hat\rho$ and $\chi=(\lambda\otimes\hat\rho)\circ\vartheta_T^{-1}$ are simultaneously σ-finite. If $\rho f<\infty$ for some measurable function $f>0$ on $S\times T$, then by (11) the function $g=f\circ\hat\vartheta_A>0$ satisfies $\chi g=\rho f<\infty$. Conversely, if $\chi g<\infty$ for some measurable function $g>0$ on $G\times A\times T$, then the function $f(s,t)=\int\gamma_s(dr)\,g(r,\alpha_s,t)>0$ is u-measurable with

$$\rho f=\iint\rho(ds\,dt)\int\gamma_s(dr)\,g(r,\alpha_s,t)=\chi g<\infty.$$

Choosing a measurable function $\tilde f>0$ on $S\times T$ with $\rho|f-\tilde f|=0$, we get $\rho\tilde f<\infty$. Thus, ρ and χ are simultaneously σ-finite. □

When the measures ρ and $\hat\rho$ of the last theorem are σ-finite, we may proceed to establish a 1−1 correspondence between the associated disintegrations. This will lead in the next section to some useful explicit formulas.

Theorem 7.16 *(equivalent disintegrations) For G, S, and T as in Theorem 7.15, let the measures ρ on $S\times T$ and $\hat\rho$ on $A\times T$ be σ-finite and related by* (9) *and* (10). *Then*

(i) *any G-invariant disintegration $\rho=\nu\otimes\mu$ determines a disintegration $\hat\rho=\hat\nu\otimes\hat\mu$, where*

$$\lambda\otimes\hat\nu=\nu(\gamma\otimes\delta_\alpha),\qquad \hat\mu_a\equiv\mu_a,$$

(ii) *any disintegration $\hat\rho=\hat\nu\otimes\hat\mu$ determines a G-invariant disintegration $\rho=\nu\otimes\mu$, where*

$$\nu=(\lambda\otimes\hat\nu)\circ\hat\vartheta_A^{-1},\qquad \mu=(\gamma\otimes\delta_\alpha)\hat\mu\circ\hat\vartheta_T^{-1}.$$

In (ii), *we may choose a version of $\hat\mu$, such that $\hat\mu_a$ is G_a-invariant for every $a\in A$. The mappings in* (i) *and* (ii) *are then equivalent, and provide a 1−1 correspondence between the disintegrations of ρ and $\hat\rho$.*

In explicit notation, the second formula in (ii) becomes

$$\mu_s f = \int \gamma_s(dr) \int \hat{\mu}_{\alpha_s}(dt)\, f(rt), \quad s \in S.$$

Proof: By Theorem 7.15, the relations involving ν are equivalent, and provide a $1-1$ correspondence between the sets of all σ-finite, G-invariant measures ν on S, and all σ-finite measures $\hat{\nu}$ on A.

(i) Given a G-invariant disintegration $\rho = \nu \otimes \mu$, define $\hat{\nu}$ and $\hat{\mu}$ as stated. Using the G-invariance of μ and γ, the right G_a-invariance of λ, and Fubini's theorem, we get

$$\begin{aligned}\rho(\gamma \otimes \delta_\alpha \otimes \delta)f &= \int \nu(ds) \int \mu_s(dt) \int \gamma_s(dp)\, f(p, \alpha_s, t) \\ &= \int \lambda(dr) \int \hat{\nu}(da) \int \mu_a(dt) \int \gamma_a(dp)\, f(rp, a, rt) \\ &= \int \lambda(dr) \int \hat{\nu}(da) \int \mu_a(dt)\, f(r, a, rt) \\ &= \left\{(\lambda \otimes \hat{\nu} \otimes \hat{\mu}) \circ \vartheta_T^{-1}\right\} f,\end{aligned}$$

and so $\hat{\rho} = \hat{\nu} \otimes \hat{\mu}$, by the uniqueness in Theorem 7.15.

(ii) Given a disintegration $\hat{\rho} = \hat{\nu} \otimes \hat{\mu}$, define ν and μ as stated. Then ν is G-invariant by Theorem 7.15. To see that even μ is G-invariant, we conclude from the G-invariance of γ that, for any $r \in G$ and $s \in S$,

$$\begin{aligned}\mu_{rs} f &= \int \gamma_s(dp) \int \hat{\mu}_{\alpha_s}(dt)\, f(rpt) \\ &= \mu_s(f \circ \theta_r) = (\mu_s \circ \theta_r^{-1}) f.\end{aligned}$$

To see that $\rho = \nu \otimes \mu$, we may use the G-invariance of μ, the right G_a-invariance of λ, and Fubini's theorem to get

$$\begin{aligned}(\nu \otimes \mu) f &= \int \lambda(dr) \int \hat{\nu}(da) \int \mu_a(dt)\, f(ra, rt) \\ &= \int \lambda(dr) \int \hat{\nu}(da) \int \gamma_a(dp) \int \hat{\mu}_a(dt)\, f(ra, rpt) \\ &= \int \lambda(dr) \int \hat{\nu}(da) \int \hat{\mu}_a(dt)\, f(ra, rt) \\ &= \left\{(\lambda \otimes \hat{\rho}) \circ \hat{\vartheta}_{A,T}^{-1}\right\} f = \rho f.\end{aligned}$$

To prove the last assertion, assume that $\hat{\rho} = \hat{\nu} \otimes \hat{\mu}$. Then $\rho = \nu \otimes \mu$, with μ as in (ii), and so by (i) we have $\hat{\rho} = \hat{\nu} \otimes \hat{\mu}'$ with $\hat{\mu}'_a \equiv \mu_a$. Here $\hat{\mu}' = \hat{\mu}$ a.e. $\hat{\nu}$, by the uniqueness in Theorem 1.23. Using the G-invariance of μ, we get for any $a \in A$ and $r \in G_a$

$$\begin{aligned}\hat{\mu}'_a \circ \theta_r^{-1} &= \mu_a \circ \theta_r^{-1} \\ &= \mu_{ra} = \mu_a = \hat{\mu}'_a,\end{aligned}$$

which shows that $\hat{\mu}'_a$ is G_a-invariant. Conversely, suppose that $\hat{\mu}_a$ is G_a-invariant for all $a \in A$. Since γ_a is restricted to G_a, we obtain

$$\begin{aligned}\hat{\mu}'_a f &= \mu_a f = \int \gamma_a(dp) \int \hat{\mu}_{\alpha_a}(dt)\, f(p\,t) \\ &= \int \gamma_a(dp) \int \hat{\mu}_a(dt)\, f(t) = \hat{\mu}_a f,\end{aligned}$$

which shows that $\hat{\mu}' = \hat{\mu}$.

Conversely, let $\rho = \nu \otimes \mu$ for some G-invariant ν and μ. Then $\hat{\rho} = \hat{\nu} \otimes \hat{\mu}$ with $\hat{\mu}$ as in (i), and so by (ii) we get $\rho = \nu \otimes \mu'$, where μ' is G-invariant with $\mu' = (\gamma \otimes \delta_\alpha)\hat{\mu} \circ \hat{\vartheta}_T^{-1}$. As before, we see that $\hat{\mu}_a$ is G_a-invariant for every $a \in A$, and so $\mu'_a = \hat{\mu}_a = \mu_a$ for all $a \in A$, which implies $\mu' = \mu$, since both kernels are G-invariant. □

We conclude with a probabilistic interpretation, involving a random measure on $G \times A$. As before, α is a uniformly measurable orbit selector with range $A \subset S$, and γ denotes the associated inversion kernel from S to G.

Theorem 7.17 *(random measure inversion) Let G be an lcscH group, acting properly on S and measurably on T, where S is Borel. Consider an s-finite random measure ξ on S and a random element η in T, and put $\rho = C_{\xi,\eta}$. For γ, α, and A as before, define a random measure $\hat{\xi} = \xi(\gamma \otimes \delta_\alpha)$ on $G \times A$. Then*

(i) $\xi = \hat{\xi} \circ \vartheta^{-1}$, *with* $\vartheta(r, a) = ra$,

(ii) ξ *and* $\hat{\xi}$ *are simultaneously σ-finite,*

(iii) (ξ, η) *and* $(\hat{\xi}, \eta)$ *are simultaneously jointly G-stationary,*

(iv) $(\hat{\xi}, \eta)$ *has Campbell measure* $\hat{\rho} = \rho(\gamma \otimes \delta_\alpha \otimes \delta)$,

(v) $E(\hat{\xi} \,|\, \eta) = E(\xi \,|\, \eta)(\gamma \otimes \delta_\alpha)$ *a.s.*

Proof: (i) Use Theorem 7.15, for fixed $\omega \in \Omega$ and singleton T.

(iii) Since γ and α are G-invariant, and ϑ commutes with all θ_r, we have

$$\begin{aligned}\hat{\xi} \circ \theta_r^{-1} &= (\xi \circ \theta_r^{-1})(\gamma \otimes \delta_\alpha), \\ \xi \circ \theta_r^{-1} &= (\hat{\xi} \circ \theta_r^{-1}) \circ \vartheta^{-1}.\end{aligned}$$

(iv) For any $f \in (\mathcal{G} \otimes \mathcal{A} \otimes \mathcal{T})_+$, define

$$\begin{aligned}\hat{f}(s,t) &= (\gamma_s \otimes \delta_{\alpha_s} \otimes \delta_t) f \\ &= \int \gamma_s(dr)\, f(r, \alpha_s, t),\end{aligned}$$

and note that

$$E\,\hat{\xi} f(\cdot,\cdot,\eta) = E\,\xi \hat{f}(\cdot,\eta) = \rho \hat{f} = \hat{\rho} f.$$

(ii) Use (iv) with $\eta(\omega) = \omega$, along with Theorems 6.44 (iii) and 7.15.

(v) Since ξ is s-finite, Lemma 1.15 (iii) allows us to assume $\|\xi\| = \|\hat{\xi}\| \leq 1$ a.s. Regarding $\hat{\xi}$ as a random measure on the Borel space $G \times S$, we see from Theorem 6.44 (ii) that a.s. $E(\xi \,|\, \eta) = \mu \circ \eta$ and $E(\hat{\xi} \,|\, \eta) = \hat{\mu} \circ \eta$, for some kernels $\mu \colon T \to S$ and $\hat{\mu} \colon T \to G \times S$. For fixed $f \in (\mathcal{G} \otimes \mathcal{S})_+$, we have a.s.

$$\begin{aligned}(\hat{\mu} \circ \eta)f &= E(\hat{\xi} f \,|\, \eta) = E\Big\{\xi(\gamma \otimes \delta_\alpha)f \,\Big|\, \eta\Big\} \\ &= (\mu \circ \eta)(\gamma \otimes \delta_\alpha)f.\end{aligned}$$

Applying this to a measure-determining sequence of functions f, which exists since $G \times S$ is Borel, we obtain $\hat{\mu} \circ \eta = (\mu \circ \eta)(\gamma \otimes \delta_\alpha)$ a.s. □

7.4 Duality and Mass Transport

For any measure ρ on $S \times T$, we define the reflected measure $\tilde{\rho}$ on $T \times S$ by $\tilde{\rho} f = \rho \tilde{f}$, for any $f \in (\mathcal{S} \otimes \mathcal{T})_+$, where $\tilde{f}(t,s) = f(s,t)$, and we write $\tilde{\rho} \cong \rho$. For any measures ν on S and ν' on T, and kernels $\mu : S \to T$ and $\mu' : T \to S$, the pairs (ν, μ) and (ν', μ') (along with their compositions and disintegrations) are said to be (mutually) *dual*, if $\nu \otimes \mu \cong \nu' \otimes \mu'$. Our first aim is to characterize duality. Given a σ-finite kernel $\mu' : T \to S$, we may choose a *normalizing function* $h \in (\mathcal{T} \otimes \mathcal{S})_+$, such that $\mu'_t h(t, \cdot) \equiv 1\{\mu'_t \neq 0\}$.

Theorem 7.18 *(duality) Let G be an lcscH group, acting properly on S and measurably on T, where S is Borel, and fix any G-invariant measure ν on S and kernels $\mu \colon S \to T$ and $\mu' \colon T \to S$. Then there exists a G-invariant measure ν' on T with $\nu \otimes \mu \cong \nu' \otimes \mu'$, iff these conditions are fulfilled:*

(i) $\quad \nu\mu\{\mu' = 0\} = 0,$

(ii) $\quad (\hat{\nu} \otimes \mu \otimes \mu')f = \int \hat{\nu}(da) \int \mu_a(dt) \int \mu'_t(ds) \int \gamma_s(dr)\, f(\alpha_s, \tilde{r}t, \tilde{r}a)\, \Delta_r.$

For any normalizing function h of μ', the measure ν' is then given, uniquely on $\{\mu' \neq 0\}$, by

(iii) $\quad \nu'(f;\, \mu' \neq 0) = \int \lambda(dr) \int \hat{\nu}(da) \int \mu_a(dt)\, f(rt)\, h(rt, ra).$

If G acts transitively on S, then $\alpha = a$ is a constant, and (ii) and (iii) reduce to

(ii′) $\quad (\mu_a \otimes \mu')f = \int \mu_a(dt) \int \mu'_t \gamma(dr)\, f(\tilde{r}t, \tilde{r}a)\, \Delta_r,$

(iii′) $\quad \nu'(f;\, \mu' \neq 0) = \int \lambda(dr) \int \mu_a(dt)\, f(rt)\, h(rt, ra).$

Specializing to $S = G$ and $T = \Omega$, we obtain the following probabilistic interpretation.

Corollary 7.19 *(Palm criterion and inversion, Mecke) Let G act measurably on Ω, consider a G-invariant random measure ξ on G, and write τ for the identity map on Ω. Fix any σ-finite measure $\tilde{P}$ on Ω with associated expectation $\tilde{E}$. Then there exists a σ-finite, G-invariant measure P on Ω with $\tilde{P} = P(\cdot \,\|\, \xi)_\iota$, iff these conditions are fulfilled:*

(i) $\tilde{P}\{\xi = 0\} = 0$,

(ii) $\tilde{E}\int \xi(ds)\, f(\tau, s) = \tilde{E}\int \xi(ds)\, f(\tilde{s}\tau, \tilde{s})\, \Delta_s$.

For any normalizing function h of ξ, the measure P is then given, uniquely on $\{\xi \neq 0\}$, by

(iii) $E\{f(\tau);\, \xi \neq 0\} = \tilde{E}\int \lambda(ds)\, f(s\tau)\, h(s\tau, s)$.

Proof of Theorem 7.18: Beginning with the transitive case, assume that $\nu' \otimes \mu' \cong (\lambda \circ \pi_a^{-1}) \otimes \mu$. Then (iii′) is an immediate consequence, and (i) holds by the G-invariance of μ and μ'. Now introduce the G-invariant kernels $\hat{\mu}_r = \mu_{ra}$ from G to T, and $\hat{\mu}'_t = \mu'_t\gamma$ from T to G, and note that $\mu'_t = \hat{\mu}'_t \circ \pi_a^{-1}$ by Theorem 7.14, and $\nu' \otimes \hat{\mu}' \cong \lambda \otimes \hat{\mu}$ as in Theorem 7.16. Choosing $g \in \mathcal{G}_+$ with $\lambda g = 1$, we claim that

$$\begin{aligned}
(\mu_a \otimes \mu')f &= \int \lambda(dr) \int \hat{\mu}_\iota(dt) \int \hat{\mu}'_t(dp)\, f(t, p\,a)\, g(r)\\
&= \int \nu'(dt) \int \hat{\mu}'_t(dr) \int \hat{\mu}'_t(dp)\, f(\tilde{r}t, \tilde{r}p\,a)\, g(r)\\
&= \int \lambda(dp) \int \hat{\mu}_\iota(dt) \int \hat{\mu}'_t(dr)\, f(\tilde{r}t, \tilde{r}a)\, g(p\,r)\\
&= \int \mu_a(dt) \int \hat{\mu}'_t(dr)\, f(\tilde{r}t, \tilde{r}a)\, \Delta_r,
\end{aligned}$$

which will prove (ii′). Here the first equality holds by the definition of $\hat{\mu}$, the $\mu'/\hat{\mu}'$-relation, and the choice of g. The second equality holds by the $\hat{\mu}/\hat{\mu}'$-duality and the G-invariance of $\hat{\mu}$. The third equality holds by Fubini's theorem, the $\hat{\mu}/\hat{\mu}'$-duality, and the G-invariance of $\hat{\mu}$ and $\hat{\mu}'$. The fourth equality holds by the definitions of $\hat{\mu}$ and Δ, Fubini's theorem, and the choice of g.

Conversely, assume (i) and (ii′), and define an s-finite measure ν' by the right-hand side of (iii′), for some normalizing function h of μ'. Writing $\hat{\mu}' = \mu'\gamma$ as before, we claim that

$$\begin{aligned}
(\nu' \otimes \mu')f &= \int \lambda(dr) \int \mu_a(dt) \int \hat{\mu}'_t(dp)\, f(rt, rpa)\, h(rt, ra)\\
&= \int \lambda(dr) \int \mu_a(dt) \int \hat{\mu}'_t(dp)\, h(r\tilde{p}t, r\tilde{p}a)\, f(r\tilde{p}t, ra)\, \Delta_p\\
&= \int \lambda(dr) \int \mu_a(dt) \int \mu'_{rt}(ds)\, f(rt, ra)\, h(rt, s)\\
&= \int \lambda(dr) \int \mu_a(dt)\, f(rt, ra)\\
&= \Big\{(\lambda \circ \pi_a^{-1}) \otimes \mu\Big\}\tilde{f},
\end{aligned}$$

which implies $\nu' \otimes \mu' \overset{\sim}{=} (\lambda \circ \pi_a^{-1}) \otimes \mu$. Here the first equality holds by the definition of ν', the $\mu'/\hat{\mu}'$-relation, and the G-invariance of $\hat{\mu}'$. The second equality holds by the definition of Δ and Fubini's theorem, the third equality holds by (ii′), the fourth equality holds by (i) and the choice of h, and the last equality holds by the G-invariance of μ. The σ-finiteness of ν' follows from the σ-finiteness of $\nu \otimes \mu$, and the G-invariance of ν' is clear from the uniqueness of the disintegration, and the G-invariance of λ, μ, and μ'.

In the non-transitive case, suppose that $\nu' \otimes \mu' \overset{\sim}{=} \nu \otimes \mu$. Then (i) and (iii) are again obvious. Now introduce the G-invariant kernels $\hat{\mu}_{r,a} = \mu_{ra}$ from $G \times A$ to T, and $\hat{\mu}'_t = \mu'_t(\gamma \otimes \delta_\alpha)$ from T to $G \times A$, so that $\mu'_t = \hat{\mu}'_t \circ \vartheta^{-1}$ by Theorem 7.15, and $\lambda \otimes \hat{\nu} \otimes \hat{\mu} \overset{\sim}{=} \nu' \otimes \hat{\mu}'$ as in Theorem 7.16. Then (ii) may be proved in the same way as (ii′), apart from the insertion of some further integral signs. Even the reverse implication may be proved as before, except that now we also need the inversion formula $\nu = (\lambda \otimes \hat{\nu}) \circ \vartheta^{-1}$. □

We may next extend some classical identities, known as *mass-transport principles.*

Theorem 7.20 *(balancing transports) Let G be an lcscH group, acting properly on the Borel spaces S, S', U and measurable on T. Consider some G-invariant measures ν, ν' on S, S' and kernels*

$$\mu\colon S \to T, \quad \mu'\colon S' \to T, \quad \chi\colon S \times T \to U, \quad \chi'\colon S' \times T \to U,$$

and define $\hat{\nu}$, $\hat{\nu}'$ as before. Fix an orbit selector α on U with associated inversion kernel γ. Then these conditions are equivalent:

(i) $(\nu \otimes \mu)(\chi \otimes \delta) = (\nu' \otimes \mu')(\chi' \otimes \delta)$ *on* $U \times T$,

(ii) $$\iint \hat{\nu}(\mu \otimes \chi)(dt\, du) \int \gamma_u(dr) f(\alpha_u, \tilde{r}t)\, \Delta_r = \iint \hat{\nu}'(\mu' \otimes \chi')(dt\, du) \int \gamma_u(dr) f(\alpha_u, \tilde{r}t)\, \Delta_r.$$

Here (i) is a compact form of the relation

$$\int \nu(ds) \int \mu_s(dt) \int \chi_{s,t}(du)\, f(u,t) = \int \nu'(ds') \int \mu'_{s'}(dt) \int \chi'_{s',t}(du)\, f(u,t).$$

When $S' = U$ and $\chi'_{s,t} = \delta_s$, (ii) reduces to the one-sided condition

$$(\hat{\nu}' \otimes \mu')f = \iint \hat{\nu}(\mu \otimes \chi)(dt\, ds) \int \gamma_s(dr)\, f(\alpha_s, \tilde{r}t)\, \Delta_r.$$

Specializing further to $S = S' = U = G$, we get the following version of a classical result.

Corollary 7.21 *(mass transport) Let G act measurably on T. Consider some random measures ξ and η on G, a random element τ in T, and a G-invariant kernel $\chi\colon G \times T \to G$, such that (ξ, η, τ) is G-stationary, and ξ and η are τ-measurable. Then $\xi(\chi \circ \tau) = \eta$ a.s., iff*

$$E\{f(\tau) \,\|\, \eta\} = E\Big\{ \int \chi_{\iota,\tau}(ds)\, f(\tilde{s}\tau)\, \Delta_s \Big\|\, \xi \Big\}, \quad f \in \mathcal{T}_+.$$

Proof: When $\nu = \nu' = \lambda$, $\mu = \mathcal{L}(\tau \,\|\, \xi)$, and $\mu' = \mathcal{L}(\tau \,\|\, \eta)$, condition (ii) of the last theorem reduces to the displayed formula. For convenience, we may take τ to be the identity mapping on $\Omega = T$. Then $\xi(\chi \circ \tau) = \eta$ a.s. implies

$$\begin{aligned}(\nu \otimes \mu)(\chi \otimes \delta) &\overset{\sim}{=} P \otimes \xi(\chi \circ \tau)\\ &= P \otimes \eta \overset{\sim}{=} \nu' \otimes \mu',\end{aligned}$$

which proves (i). Conversely, (i) gives $P \otimes \xi(\chi \circ \tau) = P \otimes \eta$, which implies $\xi(\chi \circ \tau)f = \eta f$ a.s. for every $f \in \mathcal{G}_+$, and hence $\xi(\chi \circ \tau) = \eta$ a.s. □

Proof of Theorem 7.20: Put $\rho = (\nu \otimes \mu)(\chi \otimes \delta)$ and $\hat{\rho} = \rho(\gamma \otimes \delta_\alpha \otimes \delta)$. By Fubini's theorem and the definition of Δ,

$$\begin{aligned}\hat{\rho} f &= \int \lambda(dp) \int \hat{\nu}(da) \int \mu_a(dt) \int \chi_{a,t}(du) \int \gamma_a(dr)\, f(p\, r, \alpha_u, p\, t)\\ &= \int \lambda(dp) \int \hat{\nu}(da) \int \mu_a(dt) \int \chi_{a,t}(du) \int \gamma_a(dr)\, f(p, \alpha_u, p\, \tilde{r} t)\, \Delta_r,\end{aligned}$$

and similarly with $\hat{\nu}$, μ, χ replaced by $\hat{\nu}'$, μ', χ'. Under (i), the integrands in the outer integrals on the right agree for $p \in G$ a.e. λ, and (ii) follows, as we replace f by the function $g(p, a, t) = f(a, \tilde{p}\, t)$.

Conversely, assume (ii). Replacing f by the functions $g_p(a, t) = f(p, a, p\, t)$ with $p \in G$, integrating in p with respect to λ, and reversing the previous calculation, we get $\hat{\rho} = \hat{\rho}'$, where $\hat{\rho}'$ is defined as $\hat{\rho}$, though in terms of ν', μ', χ'. Now (i) follows by Theorem 7.15. □

We continue with a similar relationship between dual disintegrations, again containing some classical statements as special cases.

Theorem 7.22 *(dual transports) Let G be an lcscH group, acting properly on S, S' and measurable on T, where S and S' are Borel. Consider some G-invariant measures ν, ν' on S, S' and kernels*

$$\mu \colon S \to T, \quad \mu' \colon S' \to T, \quad \chi \colon S \times T \to S', \quad \chi' \colon S' \times T \to S,$$

fix some orbit selectors α, α' on S, S', with associated inversion kernels γ, γ', and introduce the measures and kernels

$$\begin{aligned}\lambda \otimes \hat{\nu} &= \nu(\gamma \otimes \delta_\alpha), & \hat{\mu}_{r,a} &= \mu_{ra}, & \hat{\chi}_{r,a,t} &= \chi_{ra,t}(\gamma' \otimes \delta_{\alpha'}),\\ \lambda \otimes \hat{\nu}' &= \nu'(\gamma' \otimes \delta_{\alpha'}), & \hat{\mu}'_{r,a} &= \mu'_{ra}, & \hat{\chi}'_{r,b,t} &= \chi'_{rb,t}(\gamma \otimes \delta_\alpha).\end{aligned}$$

Then these conditions are equivalent:

(i) $\nu \otimes \mu \otimes \chi \overset{\sim}{=} \nu' \otimes \mu' \otimes \chi'$ *on* $S \times T \times S'$,

(ii) $\lambda \otimes \hat{\nu} \otimes \hat{\mu} \otimes \hat{\chi} \overset{\sim}{=} \lambda \otimes \hat{\nu}' \otimes \hat{\mu}' \otimes \hat{\chi}'$ *on* $G \times A \times T \times G \times A'$,

(iii) $(\hat{\nu} \otimes \mu \otimes \hat{\chi}) f = \int \hat{\nu}'(db) \int \mu'_b(dt) \iint \hat{\chi}'_{\iota,b,t}(dr\, da)\, f(a, \tilde{r} t, \tilde{r}, b)\, \Delta_r.$

In explicit notation, (i)−(iii) become respectively

(i′)
$$\int \nu(ds) \int \mu_s(dt) \int \chi_{s,t}(ds')\, f(s,t,s') = \int \nu'(ds') \int \mu'_{s'}(dt) \int \chi'_{s',t}(ds)\, f(s,t,s'),$$

(ii′)
$$\int \lambda(dr) \int \hat\nu(da) \int \hat\mu_{r,a}(dt) \iint \hat\chi_{r,a,t}(dr'\,da')\, f(r,a,t,r',a') = \int \lambda(dr') \int \hat\nu'(da') \int \hat\mu'_{r',a'}(dt) \iint \hat\chi'_{r',a',t}(dr\,da)\, f(r,a,t,r',a'),$$

(iii′)
$$\int \hat\nu(da) \int \mu_a(dt) \iint \hat\chi_{\iota,a,t}(dr\,db)\, f(a,t,r,b) = \int \hat\nu'(db) \int \mu'_b(dt) \iint \hat\chi'_{\iota,b,t}(dr\,da)\, f(a,\tilde r t,\tilde r,b)\, \Delta_r.$$

Taking $S = S' = G$, we get the following classical result:

Corollary 7.23 *(exchange formula, Neveu) Let G act measurably on T, and consider some random measures ξ, η on G and a random element τ in T, such that (ξ,η,τ) is G-stationary. Then*

$$E\Big\{\int \xi(ds)\, f(\tau,s) \Big\|\, \eta\Big\} = E\Big\{\int \eta(ds)\, f(\tilde s\tau,\tilde s)\, \Delta_s \Big\|\, \xi\Big\}.$$

Note that when $\xi = \eta$, the displayed formula reduces to Mecke's criterion in Corollary 7.19.

Proof: Take $\chi = \xi$ and $\chi' = \eta$, and put $\mu = P(\cdot \,\|\, \eta)$ and $\mu' = P(\cdot \,\|\, \xi)$. Then (i) and (ii) of the last theorem follow from the relation $\xi \otimes \eta \overset{\sim}{=} \eta \otimes \xi$. □

Proof of Theorem 7.22: Let $\rho = \nu \otimes \mu$, and define $\hat\rho = \rho(\gamma \otimes \delta_\alpha \otimes \delta)$. Given a function $f \geq 0$ on $G \times A \times T \times G \times A'$, put $g(r,a,t) = \hat\chi(r,a,t,\cdot,\cdot)$. Then Theorem 7.16 yields

$$\begin{aligned}(\nu \otimes \mu \otimes \chi)\big(\gamma \otimes \delta_\alpha \otimes \delta \otimes \gamma' \otimes \delta_{\alpha'}\big) f &= \rho(\gamma \otimes \delta_\alpha \otimes \delta) g = \hat\rho g \\ &= (\lambda \otimes \hat\nu \otimes \hat\mu) g \\ &= \big(\lambda \otimes \hat\nu \otimes \hat\mu \otimes \hat\chi\big) f.\end{aligned}$$

Combining with a similar formula for the measures on the right, we see that (i) implies (ii).

Next, we see as in Theorem 7.15 that $\hat\chi_{r,a,t} \circ \vartheta^{-1} = \chi_{ra,t}$ and $\hat\chi'_{r,b,t} \circ \vartheta^{-1} = \chi'_{rb,t}$. Writing $\varphi(r,a,t,p,b) = (ra,t,p\,b)$, and noting that $\nu = (\lambda \otimes \hat\nu) \circ \vartheta^{-1}$ by Theorem 7.15, we get

$$\begin{aligned}&\big(\lambda \otimes \hat\nu \otimes \hat\mu \otimes \hat\chi\big)(f \circ \varphi) \\ &\quad = \int \hat\nu(da) \int (\lambda \circ \pi_a^{-1})(ds) \int \mu_s(dt) \int \chi_{s,t}(ds')\, f(s,t,s') \\ &\quad = (\nu \otimes \mu \otimes \chi) f.\end{aligned}$$

Combining with the corresponding formula for $\nu' \otimes \mu' \otimes \chi'$, we see that even (ii) implies (i), and so the two conditions are equivalent.

Now assume (ii). Using the G-invariance of the various kernels, along with the definition of Δ, we obtain

$$\begin{aligned}
&\big(\lambda \otimes \hat\nu \otimes \hat\mu \otimes \hat\chi\big) f \\
&\quad = \int \lambda(dp) \int \hat\nu(da) \int \mu_a(dt) \int \hat\chi_{\iota,a,t}(dr\,db)\, f(p,a,p\,t,p\,r,b) \\
&\quad = \int \lambda(dp) \int \hat\nu'(db) \int \mu'_b(dt) \int \hat\chi'_{\iota,b,t}(dr\,da)\, f(p\,r,a,p\,t,p,b) \\
&\quad = \int \lambda(dp) \int \hat\nu'(db) \int \mu'_b(dt) \int \hat\chi'_{\iota,b,t}(dr\,da)\, f(p,a,p\,\tilde r t,p\,\tilde r,b)\, \Delta_r,
\end{aligned}$$

which implies (iii). Conversely, (ii) follows as we apply (iii) to the functions $g_p(a,t,r,b) = f(p,a,p\,t,p\,r,b)$ with $p \in G$, integrate in p with respect to λ, and use the definition of Δ. Thus, even (ii) and (iii) are equivalent. □

7.5 Invariant Disintegration

Given a measurable group G, acting measurably on some Borel spaces S and T, consider some locally finite measures μ on $S \times T$ and ν on S such that $\mu(\cdot \times T) \ll \nu$. Then Lemma 1.26 yields a disintegration of the form $\mu = \nu \otimes \varphi(\cdot, \mu, \nu)$, where φ is a kernel from $S \times \mathrm{M}_{S\times T} \times \mathcal{M}_S$ to T. When G is lcscH, we show that φ can be chosen to be *G-invariant*, in the sense that

$$\varphi(s,\mu,\nu) \circ \theta_r^{-1} = \varphi\Big\{rs, (\mu,\nu) \circ \theta_r^{-1}\Big\}, \quad r \in G, \ \ s \in S.$$

The result is crucial for our construction of G-stationary densities and disintegration kernels in the next section.

Theorem 7.24 *(invariant disintegration) For any lcscH group G, acting measurably on some Borel spaces S and T, there exists a G-invariant kernel $\varphi\colon S \times \mathcal{M}_{S\times T} \times \mathcal{M}_S \to T$, such that $\mu = \nu \otimes \varphi(\cdot,\mu,\nu)$ for all $\mu \in \mathcal{M}_{S\times T}$ and $\nu \in \mathcal{M}_S$ with $\mu(\cdot \times T) \ll \nu$.*

For invariant measures μ and ν, this essentially reduces to Theorem 7.6. When T is a singleton and G acts measurably on a Borel space S, the invariant disintegration reduces to an *invariant differentiation* $\mu = \varphi(\cdot,\mu,\nu)\cdot \nu$, where invariance is now defined, for any measurable functions $\varphi \ge 0$ on $S \times \mathcal{M}_S^2$, by

$$\varphi(s,\mu,\nu) = \varphi\Big\{rs, (\mu,\nu) \circ \theta_r^{-1}\Big\}, \quad r \in G, \ \ s \in S, \tag{12}$$

for arbitrary $\mu \le \nu$ in $\mathcal{M}_S$. When G acts transitively on S, it is clearly enough to require (12) for a single element s.

Corollary 7.25 *(invariant differentiation) For any lcscH group G, acting measurably on a Borel space S, there exists a measurable and G-invariant function $\varphi \geq 0$ on $S \times \mathcal{M}_S^2$, such that $\mu = \varphi(\cdot, \mu, \nu) \cdot \nu$ for all $\mu \ll \nu$ in $\mathcal{M}_S$.*

The main theorem will be proved in several steps. First we consider some partial versions of the density result, beginning with the case where $S = G$ is a Lie group.

Lemma 7.26 *(differentiation on Lie groups) An invariant density function exists on every Lie group G.*

Proof: By Theorem A5.3, any inner product on the basic tangent space generates a left-invariant Riemannian metric ρ on G, given in local coordinates $x = (x^1, \ldots, x^d)$ by a smooth family of symmetric, positive definite matrices $\rho_{ij}(x)$. By Theorem A5.1, the length of a smooth curve in G is obtained by integration of the length element ds along the curve, where

$$ds^2(x) = \sum\nolimits_{i,j} \rho_{ij}(x)\, dx^i dx^j, \quad x \in G,$$

and the distance between two points $x, y \in G$ is defined as the minimum length of all smooth curves connecting x and y. This determines an invariant metric on G, which in turn defines the open balls B_r^ε in G of radius $\varepsilon > 0$, centered at $r \in G$.

By Theorem A5.2, the Besicovitch covering Theorem A4.4 on R^d extends to any compact subset of a Riemannian manifold. Furthermore, the open ρ-balls in G form a differentiation basis for any measure $\nu \in \mathcal{M}_G$. Hence, for any measures $\mu \leq \nu \in \mathcal{M}_G$, we may define a measurable density function φ on G by

$$\varphi(r, \mu, \nu) = \limsup_{n\to\infty} \frac{\mu B_r^{1/n}}{\nu B_r^{1/n}}, \quad r \in G,$$

where $0/0$ is interpreted as 0. To see that φ is invariant, let ι denote the identity element in G, and note that

$$\begin{aligned} \varphi\Big\{r, (\mu, \nu) \circ \theta_r^{-1}\Big\} &= \limsup_{n\to\infty} \frac{(\mu \circ \theta_r^{-1}) B_r^{1/n}}{(\nu \circ \theta_r^{-1}) B_r^{1/n}} \\ &= \limsup_{n\to\infty} \frac{\mu B_\iota^{1/n}}{\nu B_\iota^{1/n}} = \varphi(\iota, \mu, \nu), \end{aligned}$$

since $r^{-1} B_r^\varepsilon = B_\iota^\varepsilon$, by the invariance of the metric ρ. □

Next, we choose G to be a *projective limit* of Lie groups, in the sense that every neighborhood of the identity ι contains a compact, invariant subgroup H, such that G/H is isomorphic to a Lie group.

Lemma 7.27 *(projective limits of Lie groups) An invariant density function exists on any projective limit G of Lie groups.*

Proof: By Theorem A5.4, there exist some compact, invariant subgroups $H_n \downarrow \{\iota\}$ of G, such that the quotient groups G/H_n are isomorphic to Lie groups. Since the projection maps $\pi_n : r \mapsto rH_n$ are continuous, we have $\mu \circ \pi_n^{-1} \in \mathcal{M}_{G/H_n}$, for any $\mu \in \mathcal{M}_G$. Choosing some invariant density functions ψ_n on G/H_n, as in Lemma 7.26, we introduce the functions

$$\varphi_n(r, \mu, \nu) = \psi_n\Big\{rH_n, (\mu, \nu) \circ \pi_n^{-1}\Big\}, \quad r \in G, \ n \in \mathsf{N},$$

which are ν-densities of μ on the σ-fields $\mathcal{G}_n$, generated by the coset partitions G/H_n of G, so that

$$\mu = \varphi_n(\cdot, \mu, \nu) \cdot \nu \ \text{ on } \ \mathcal{G}_n, \quad n \in \mathsf{N}.$$

Since $H_n \downarrow \{\iota\}$, the $\mathcal{G}_n$ are non-decreasing and generate the Borel σ-field $\mathcal{G}$ on G. Hence, for any fixed μ and ν, we get by martingale convergence

$$\varphi_n(r, \mu, \nu) \to \varphi(r, \mu, \nu), \quad r \in G \ \text{ a.e. } \nu,$$

where the limit φ is a density on $\mathcal{G}$. To ensure product measurability, we may define

$$\varphi(r, \mu, \nu) = \limsup_{n\to\infty} \varphi_n(r, \mu, \nu), \quad r \in G, \ \mu \le \nu \ \text{ in } \mathcal{M}_G. \tag{13}$$

To show that φ is invariant, we note that an arbitrary shift by $r \in G$ preserves the coset partitions G/H_n of G, since any coset aH_n is mapped into the coset $r(aH_n) = (ra)H_n$. Furthermore, the permutation of G/H_n induced by $r \in G$ agrees with a shift of G/H_n by the group element rH_n, since $(rH_n)(aH_n) = (ra)H_n = r(aH_n)$, by the definition of the group operation in G/H_n. Further note that $\pi_n \circ \theta_r = \theta_{rH_n} \circ \pi_n$ on G, since for any $r, s \in G$,

$$\begin{aligned}
\pi_n \theta_r s &= \pi_n(rs) = rsH_n \\
&= (rH_n)(sH_n) \\
&= \theta_{rH_n}(sH_n) = \theta_{rH_n}\pi_n s.
\end{aligned}$$

Using the G/H_n-invariance of each ψ_n, we get for any $r \in G$

$$\begin{aligned}
\varphi_n\Big\{r, (\mu, \nu) \circ \theta_r^{-1}\Big\} &= \psi_n\Big\{rH_n, (\mu, \nu) \circ \theta_r^{-1} \circ \pi_n^{-1}\Big\} \\
&= \psi_n\Big\{rH_n, (\mu, \nu) \circ \pi_n^{-1} \circ \theta_{rH_n}^{-1}\Big\} \\
&= \psi_n\Big\{H_n, (\mu, \nu) \circ \pi_n^{-1}\Big\} \\
&= \varphi_n(\iota, \mu, \nu),
\end{aligned}$$

which shows that the functions φ_n are G-invariant. The G-invariance of φ now follows by (13). □

We may now proceed to any locally compact group G.

Lemma 7.28 *(locally compact groups) An invariant density function exists on every lcscH group G.*

Proof: Theorem A5.5 shows that G has an open subgroup H that is a projective limit of Lie groups. Since the cosets of H are again open, and the projection map $\pi : r \mapsto rH$ is both continuous and open, the coset space G/H is countable and discrete. Choosing an invariant density function ψ on H, as in Lemma 7.27, we define

$$\varphi(r, \mu, \nu) = \psi\Big\{\iota, (\mu, \nu)_{rH} \circ \theta_{r^{-1}}^{-1}\Big\}, \quad r \in G, \ \mu \le \nu \text{ in } \mathcal{M}_G, \tag{14}$$

where $\mu_A = \mu(A \cap \cdot)$ denotes the restriction of μ to A.

By (14), the invariance property (12) for $s = \iota$ is equivalent to

$$\psi\{\iota, (\mu, \nu)_H\} = \psi\Big(\iota, \{(\mu, \nu) \circ \theta_r^{-1}\}_{rH} \circ \theta_{r^{-1}}^{-1}\Big), \quad r \in G.$$

This follows from the relation

$$\mu_H \circ \theta_r^{-1} = (\mu \circ \theta_r^{-1})_{rH}, \quad r \in G, \ \mu \in \mathcal{M}_G, \tag{15}$$

which holds since, for any $B \in \mathcal{G}$,

$$\begin{aligned} (\mu \circ \theta_r^{-1})_{rH} B &= (\mu \circ \theta_r^{-1})(rH \cap B) \\ &= \mu\{r^{-1}(rH \cap B)\} \\ &= \mu(H \cap r^{-1}B) \\ &= \mu_H(r^{-1}B) = (\mu_H \circ \theta_r^{-1})B. \end{aligned}$$

Using (14), (15), and the invariance and density properties of ψ, we get for any measurable function $f \ge 0$ on the coset aH

$$\begin{aligned} \int f(r)\, \varphi(r, \mu, \nu)\, \nu(dr) &= \int f(ar)\, \varphi(r, \mu, \nu)\, (\nu \circ \theta_{a^{-1}}^{-1})(dr) \\ &= \int f(ar)\, \psi\Big\{\iota, (\mu, \nu)_{arH} \circ \theta_{(ar)^{-1}}^{-1}\Big\}\, (\nu \circ \theta_{a^{-1}}^{-1})(dr) \\ &= \int f(ar)\, \psi\Big\{r, (\mu, \nu)_{aH} \circ \theta_{a^{-1}}^{-1}\Big\}\, (\nu \circ \theta_{a^{-1}}^{-1})(dr) \\ &= \int f(ar)\, \psi\Big(r, \Big\{(\mu, \nu) \circ \theta_{a^{-1}}^{-1}\Big\}_H\Big)\, (\nu \circ \theta_{a^{-1}}^{-1})(dr) \\ &= \int f(ar)\, (\mu \circ \theta_{a^{-1}}^{-1})(dr) = \int f(r)\, \mu(dr). \end{aligned}$$

Since G/H is countable and $a \in G$ was arbitrary, this proves the density property $\mu = \varphi(\cdot, \mu, \nu) \cdot \nu$. □

Next, we extend the density version to any product space $G \times S$.

Lemma 7.29 *(from G to $G \times S$) A G-invariant density function exists on $G \times S$, when G has no action on S.*

Proof: Fix a dissection system (I_{nj}) on S, and put $S' = G \times S$. For any $\mu \le \nu$ in $\mathcal{M}_{S'}$, the measures $\mu_{nj} = \mu(\cdot \times I_{nj})$ and $\nu_{nj} = \nu(\cdot \times I_{nj})$ belong to $\mathcal{M}_G$ and satisfy $\mu_{nj} \le \nu_{nj}$. Choosing an invariant density function ψ on G, as in Lemma 7.28, we obtain

$$\mu_{nj} = \psi(\cdot, \mu_{nj}, \nu_{nj}) \cdot \nu_{nj} \text{ on } \mathcal{G}, \quad n, j \in \mathsf{N}. \tag{16}$$

Now define some measurable functions φ_n on $S' \times \mathcal{M}^2_{S'}$ by

$$\varphi_n(r, s, \mu, \nu) = \sum\nolimits_j \psi(r, \mu_{nj}, \nu_{nj})\, 1\{s \in I_{nj}\}, \quad r \in G,\ \ s \in S,\ \ n \in \mathsf{N}.$$

Writing $\mathcal{I}_n$ for the σ-field on S, generated by the sets I_{nj} for fixed n, we conclude from (16) that

$$\mu = \varphi_n(\cdot, \mu, \nu) \cdot \nu \text{ on } \mathcal{G} \otimes \mathcal{I}_n, \quad n \in \mathsf{N}.$$

Since the σ-fields $\mathcal{G} \otimes \mathcal{I}_n$ are non-decreasing and generate $\mathcal{G} \otimes \mathcal{S}$, the φ_n form ν-martingales on bounded sets, and so they converge, a.e. ν, to a ν-density of μ on $\mathcal{G} \otimes \mathcal{S}$. Here we may choose the version

$$\varphi(r, s, \mu, \nu) = \limsup_{n \to \infty} \varphi_n(r, s, \mu, \nu), \quad r \in G,\ \ s \in S,$$

which inherits the product measurability and G-invariance from the corresponding properties of each φ_n. $\Box$

The next step is to go from densities to disintegration kernels.

Lemma 7.30 *(from singletons to general T) Let G act on S but not on T, and suppose that a G-invariant density function exists on S. Then there exists a G-invariant disintegration kernel from S to T.*

Proof: Define $\theta_r(s, t) = (rs, t)$ for $r \in G$, $s \in S$, and $t \in T$. Suppose that $\mu \in \mathcal{M}_{S \times T}$ and $\nu \in \mathcal{M}_S$ with $\mu(\cdot \times T) \ll \nu$. By hypothesis, there exists for every $B \in \hat{\mathcal{T}}$ a measurable function $\psi_B \ge 0$ on $S \times \mathcal{M}^2_S$, such that

$$\mu(\cdot \times B) = \psi_B(\cdot, \mu, \nu) \cdot \nu, \qquad B \in \hat{\mathcal{T}}, \tag{17}$$

$$\psi_B(s, \mu, \nu) = \psi_B\big\{rs, (\mu, \nu) \circ \theta_r^{-1}\big\}, \quad r \in G,\ \ s \in S. \tag{18}$$

Writing $\mathcal{C}$ for the class of functions $f : \hat{\mathcal{T}} \to \mathsf{R}_+$, we need to construct a measurable mapping $\Phi : \mathcal{C} \to \mathcal{M}_T$, such that the kernel $\varphi = \Phi(\psi)$ from $S \times \mathcal{M}^2_S$ to T satisfies

$$\varphi(s, \mu, \nu)B = \psi_B(s, \mu, \nu), \quad s \in S \text{ a.e. } \nu,\ \ B \in \hat{\mathcal{T}}, \tag{19}$$

for any $\mu, \nu \in \mathcal{M}_S$. Then by (17),

$$\mu(\cdot \times B) = \varphi(\cdot, \mu, \nu)B \cdot \nu, \quad B \in \hat{\mathcal{T}},$$

which shows that $\mu = \nu \otimes \varphi$. Since $\Phi(f)$ depends only on f, (18) is preserved by Φ, which means that φ is again G-invariant.

To construct Φ, we may embed T as a Borel set in R_+, and extend every $f \in \mathcal{C}$ to a function $\tilde{f}$ on $\hat{\mathcal{B}}_+$, by putting $\tilde{f}(B) = f(B \cap T)$ for all $B \in \hat{\mathcal{B}}_+$. If $\tilde{m}$ is a corresponding measure on R_+ with the desired properties, its restriction m to T has clearly the required properties on T. Thus, we may henceforth take $T = \mathsf{R}_+$.

For any $f \in \mathcal{C}$, we define $F(x) = \inf_{r>x} f[0,r]$ for all $x \geq 0$, with r restricted to Q_+, and note that F is non-decreasing and right-continuous. If it is also finite, there exists a locally finite measure m on R_+, such that $m[0,x] = F(x)$ for all $x \geq 0$. Writing A for the class of functions $f \in \mathcal{C}$ with $F(r) = f[0,r]$ for $r \in \mathsf{Q}_+$, we may put $\Phi(f) = m$ on A and $\Phi(f) = 0$ on A^c, which defines Φ as a measurable function from $\mathcal{C}$ to $\mathcal{M}_T$.

To verify (19), let $B_1, B_2 \in \hat{\mathcal{B}}$ be disjoint with union B. Omitting the arguments of ψ, we may write

$$\begin{aligned}\psi_B \cdot \nu = \mu(\cdot \times B) &= \mu(\cdot \times B_1) + \mu(\cdot \times B_2)\\ &= \psi_{B_1} \cdot \nu + \psi_{B_2} \cdot \nu\\ &= (\psi_{B_1} + \psi_{B_2}) \cdot \nu,\end{aligned}$$

which implies $\psi_{B_1} + \psi_{B_2} = \psi_B$ a.e. ν. In particular, $\psi_{[0,r]}$ is a.e. non-decreasing in $r \in \mathsf{Q}_+$. By a similar argument, $B_n \uparrow B$ implies $\psi_{B_n} \uparrow \psi_B$ a.e. ν, which shows that $\psi_{[0,r]}$ is a.e. right-continuous on Q_+. Hence, the kernel $\varphi = \Phi(\psi)$ satisfies $\varphi[0,r] = \psi_{[0,r]}$, a.e. for all $r \in \mathsf{Q}_+$. The extension to (19) now follows by a standard monotone-class argument. □

Combining Lemmas 7.29 and 7.30, we see that a G-invariant disintegration kernel exists from $G \times S$ to T, when G has no action on S or T. We turn to the case of a general group action.

Lemma 7.31 *(general group action) When G acts measurably on S and T, there exists a G-invariant disintegration kernel from $G \times S$ to T.*

Proof: Let $\mu(\cdot \times T) \ll \nu$ on $S' = G \times S$. Define $\vartheta(r,s,t) = (r, rs, rt)$ with inverse $\tilde{\vartheta} = \vartheta^{-1}$, and put

$$\begin{aligned}\tilde{\mu} &= \mu \circ \tilde{\vartheta}^{-1}, & \tilde{\nu} &= \nu \circ \tilde{\vartheta}^{-1},\\ \mu &= \tilde{\mu} \circ \vartheta^{-1}, & \nu &= \tilde{\nu} \circ \vartheta^{-1}.\end{aligned}$$

By Lemma 7.30, there exists a kernel $\psi \colon S' \times \mathcal{M}_{S' \times T} \times \mathcal{M}_{S'} \to T$ with

$$\begin{aligned}&\tilde{\mu} = \tilde{\nu} \otimes \psi(\cdot, \tilde{\mu}, \tilde{\nu}) \ \text{ on } \ G \times S \times T,\\ &\psi(p,s,\mu,\nu) = \psi\Big\{rp, s, (\mu,\nu) \circ \theta_r'^{-1}\Big\}, \quad r,p \in G, \ \ s \in S,\end{aligned}$$

where $\theta_r'(p,s,t) = (rp,s,t)$. Define a kernel φ between the same spaces by

$$\varphi(r,s,\mu,\nu) = \psi\Big(r, r^{-1}s, \tilde{\mu}, \tilde{\nu}\Big) \circ \theta_r^{-1}.$$

Using the definitions of $\tilde\mu$, ψ, $\tilde\nu$, $\tilde\vartheta$, and φ, we get for any $f \in (\mathcal{S}' \otimes \mathcal{T})_+$

$$\begin{aligned}
\mu f &= (\tilde\mu \circ \vartheta^{-1}) f = \tilde\mu (f \circ \vartheta) \\
&= \{\tilde\nu \otimes \psi(\cdot, \tilde\mu, \tilde\nu)\}(f \circ \vartheta) \\
&= \Big\{(\nu \circ \tilde\vartheta^{-1}) \otimes \psi(\cdot, \tilde\mu, \tilde\nu)\Big\}(f \circ \vartheta) \\
&= \iint (\nu \circ \tilde\vartheta^{-1})(dr\, ds) \int \psi(r, s, \tilde\mu, \tilde\nu)(dt)\, f(r, rs, rt) \\
&= \iint \nu(dr\, ds) \int \psi\Big(r,\, r^{-1}s,\, \tilde\mu,\, \tilde\nu\Big)(dt)\, f(r, s, rt) \\
&= \iint \nu(dr\, ds) \int \varphi(r, s, \mu, \nu)(dt)\, f(r, s, t) = (\nu \otimes \varphi) f,
\end{aligned}$$

which shows that $\mu = \nu \otimes \varphi$.

Next, we get for any $r, p \in G$, $s \in S$, and $t \in T$

$$\begin{aligned}
(\tilde\vartheta \circ \theta_r)(p, s, t) &= \tilde\vartheta(rp, rs, rt) \\
&= \Big\{rp,\, (rp)^{-1} rs,\, (rp)^{-1} rt\Big\} \\
&= \Big(rp,\, p^{-1}s,\, p^{-1}t\Big) \\
&= \theta'_r\Big(p,\, p^{-1}s,\, p^{-1}t\Big) \\
&= (\theta'_r \circ \tilde\vartheta)(p, s, t),
\end{aligned}$$

so that $\tilde\vartheta \circ \theta_r = \theta'_r \circ \tilde\vartheta$. Combining with the definitions of φ, $\tilde\mu$, $\tilde\nu$, and the G-invariance of ψ, we obtain

$$\begin{aligned}
\varphi(p, s, \mu, \nu) \circ \theta_r^{-1} &= \psi\Big(p,\, p^{-1}s,\, \tilde\mu,\, \tilde\nu\Big) \circ \theta_{rp}^{-1} \\
&= \psi\Big\{rp,\, p^{-1}s,\, (\tilde\mu, \tilde\nu) \circ \theta_r'^{-1}\Big\} \circ \theta_{rp}^{-1} \\
&= \psi\Big\{rp,\, p^{-1}s,\, (\mu, \nu) \circ \theta_r^{-1} \circ \tilde\vartheta^{-1}\Big\} \circ \theta_{rp}^{-1} \\
&= \varphi\Big\{rp,\, rs,\, (\mu, \nu) \circ \theta_r^{-1}\Big\},
\end{aligned}$$

which shows that φ is G-invariant, in the sense of action on both S and T. □

It remains to replace the product space $G \times S$ of the previous result by a general Borel space S.

Lemma 7.32 *(from $G \times S$ to S) When G acts measurably on S and T, there exists a G-invariant disintegration kernel from S to T.*

Proof: Write $S' = G \times S$. By Lemma 7.31, there exists a G-invariant kernel ψ from $S' \times \mathcal{M}_{S' \times T} \times \mathcal{M}_{S'}$ to T, such that for any $\mu \in \mathcal{M}_{S \times T}$ and $\nu \in \mathcal{M}_S$ with $\mu(\cdot \times T) \ll \nu$,

$$\lambda \otimes \mu = \lambda \otimes \nu \otimes \psi(\cdot, \mu, \nu) \;\text{ on }\; S' \times T, \tag{20}$$

where the G-invariance means that

$$\psi(p, s, \mu, \nu) \circ \theta_r^{-1} = \psi\Big\{rp,\ rs,\ (\mu, \nu) \circ \theta_r^{-1}\Big\}, \quad r, p \in G,\ \ s \in S. \tag{21}$$

Fixing μ and ν, and defining

$$\psi_r(p, s) = \psi(rp, s), \qquad f_r(p, s, t) = f(r^{-1}p,\ s,\ t),$$

for all $r, p \in G$, $s \in S$, and $t \in T$, we get by the invariance of λ

$$\begin{aligned}(\lambda \otimes \nu \otimes \psi_r) f &= (\lambda \otimes \nu \otimes \psi) f_r \\ &= (\lambda \otimes \mu) f_r = (\lambda \otimes \mu) f,\end{aligned}$$

which shows that (20) remains valid with ψ replaced by the kernels ψ_r. Hence, by uniqueness,

$$\psi(rp, s) = \psi(p, s), \quad (p, s) \in G \times S \ \text{ a.e. } \lambda \otimes \nu,\ \ r \in G.$$

Writing A for the set of triples (s, μ, ν) with

$$\psi(rp, s) = \psi(p, s), \quad (r, p) \in G^2 \ \text{ a.e. } \lambda^2,$$

and putting $A_{\mu,\nu} = \{s \in S;\ (s, \mu, \nu) \in A\}$, we get $\nu A^c_{\mu,\nu} = 0$ by Fubini's theorem. By Lemma 7.1 (ii), the defining condition for A is equivalent to

$$\psi(r, s) = \psi(p, s), \quad (r, p) \in G^2 \ \text{ a.e. } \lambda^2,$$

and so for any $g, g' \in \mathcal{G}_+$ with $\lambda g = \lambda g' = 1$,

$$\lambda\{g\, \psi(\cdot, s)\} = \lambda\{g' \psi(\cdot, s)\} \equiv \varphi(s), \quad s \in A_{\mu,\nu}. \tag{22}$$

To make this hold identically, we may redefine $\psi(\cdot, s) = 0$ for $s \notin A_{\mu,\nu}$, without affecting the validity of (20). Condition (21) is not affected either, since A is G-invariant by Lemma 7.7.

Fixing g as above, and using (20) and (22), we get for any $f \in (\mathcal{S} \otimes \mathcal{T})_+$

$$\begin{aligned}\mu f &= (\lambda \otimes \mu)(g \otimes f) \\ &= (\lambda \otimes \nu \otimes \psi)(g \otimes f) \\ &= \Big\{\nu \otimes (g \cdot \lambda)\psi\Big\} f = (\nu \otimes \varphi) f,\end{aligned}$$

which shows that $\mu = \nu \otimes \varphi$. By (21), (22), and the invariance of λ, we further obtain

$$\begin{aligned}\varphi(s, \mu, \nu) \circ \theta_r^{-1} &= \int \lambda(dp)\, g(p)\, \psi(p, s, \mu, \nu) \circ \theta_r^{-1} \\ &= \int \lambda(dp)\, g(p)\, \psi\Big\{rp,\ rs,\ (\mu, \nu) \circ \theta_r^{-1}\Big\} \\ &= \int \lambda(dp)\, g(r^{-1}p)\, \psi\Big\{p,\ rs,\ (\mu, \nu) \circ \theta_r^{-1}\Big\} \\ &= \varphi\Big\{rs,\ (\mu, \nu) \circ \theta_r^{-1}\Big\},\end{aligned}$$

which shows that φ is again G-invariant. □

This completes the proof of Theorem 7.24. In view of the complexity of the argument, it may be useful to note the following easy proofs when G is compact or countable:

Proof of Theorem 7.24 for compact G: By Lemma 1.26, there exists a kernel $\psi: S \times \mathcal{M}_{S\times T} \times \mathcal{M}_S \to T$ with $\mu = \nu \otimes \psi(\cdot, \mu, \nu)$. Applying this to the measures $\mu \circ \theta_r^{-1}$ and $\nu \circ \theta_r^{-1}$, we get

$$\mu \circ \theta_r^{-1} = (\nu \circ \theta_r^{-1}) \otimes \psi\Big\{\cdot, (\mu, \nu) \circ \theta_r^{-1}\Big\}, \quad r \in G,$$

and so by Lemma 7.4 (ii),

$$\mu = \nu \otimes \Big(\psi\Big\{r(\cdot), (\mu, \nu) \circ \theta_r^{-1}\Big\} \circ \theta_{r^{-1}}^{-1}\Big), \quad r \in G. \tag{23}$$

Assuming $\lambda G = 1$, we may introduce yet another kernel

$$\varphi(s, \mu, \nu) = \int \lambda(dr)\, \psi\Big\{rs, (\mu, \nu) \circ \theta_r^{-1}\Big\} \circ \theta_{r^{-1}}^{-1}, \quad s \in S,$$

and note that again $\mu = \nu \otimes \varphi(\cdot, \mu, \nu)$, by (23) and Fubini's theorem. Using the right invariance of λ, we get for any $p \in G$

$$\begin{aligned} \varphi(s, \mu, \nu) \circ \theta_p^{-1} &= \int \lambda(dr)\, \psi\Big\{rs, (\mu, \nu) \circ \theta_r^{-1}\Big\} \circ \theta_{pr^{-1}}^{-1} \\ &= \int \lambda(dr)\, \psi\Big\{rps, (\mu, \nu) \circ \theta_{rp}^{-1}\Big\} \circ \theta_{r^{-1}}^{-1} \\ &= \varphi\Big\{ps, (\mu, \nu) \circ \theta_p^{-1}\Big\}, \end{aligned}$$

which shows that φ is G-invariant. □

Proof of Theorem 7.24 for countable G: For ψ as in the compact case, we get by (23) and the a.e. uniqueness of the disintegration

$$\psi(s, \mu, \nu) = \psi\Big\{rs, (\mu, \nu) \circ \theta_r^{-1}\Big\} \circ \theta_{r^{-1}}^{-1}, \quad s \in S \text{ a.e. } \nu,\ r \in G. \tag{24}$$

Now let A be the set of triples (s, μ, ν) satisfying

$$\psi(s, \mu, \nu) \circ \theta_r^{-1} = \psi\Big\{rs, (\mu, \nu) \circ \theta_r^{-1}\Big\}, \quad r \in G.$$

Since G is countable, (24) yields $\nu\{s \in S;\ (s, \mu, \nu) \notin A\} = 0$ for all μ and ν, which justifies that we choose $\varphi = 1_A \psi$. Since A is G-invariant by Lemma 7.7, we have

$$(s, \mu, \nu) \in A \quad \Leftrightarrow \quad \Big\{rs, (\mu, \nu) \circ \theta_r^{-1}\Big\} \in A.$$

Thus, φ satisfies (24) identically, and is therefore G-invariant. □

7.6 Stationary Densities and Disintegration

For measurable processes and group actions, Theorem 2.11 yields a remarkable strengthening of the elementary notion of stationarity, defined in terms of finite-dimensional distributions.

Theorem 7.33 *(strong stationarity) Given some Borel spaces S, T, and a group G acting measurably on S, let X be a measurable process on S, taking values in T. Then X is G-stationary, in the sense of finite-dimensional distributions, iff for any measurable function $f \geq 0$ on T and measure μ on S,*

$$(\mu \circ \theta_r^{-1})(f \circ X) \stackrel{d}{=} \mu(f \circ X), \quad r \in G. \tag{25}$$

Proof: If X is G-stationary, in the sense of finite-dimensional distributions, then so is $f(X)$. Hence, $f \circ X \circ \theta_r \stackrel{d}{=} f \circ X$ for all $r \in G$, and (25) follows by Theorem 2.11, since X is measurable. Conversely, assume (25). Considering measures μ with finite support, and using the Cramér–Wold theorem, we see that $f(X)$ is G-stationary, in the sense of finite-dimensional distributions. Since T is Borel, the same property holds for X. □

The property of X in Theorem 7.33 will be referred to as *strong G-stationarity.* More generally, for any measurable space T, we say that a random measure ξ and a measurable process X on $S \times T$ are *jointly strongly G-stationary*, if

$$\left\{\xi \circ \theta_r^{-1}, \mu(f \circ X)\right\} \stackrel{d}{=} \left\{\xi, (\mu \circ \theta_r^{-1})(f \circ X)\right\}, \quad r \in G,$$

for any μ and f as above. Note that the shift of ξ by $r \in G$ corresponds to a shift of X by r^{-1}. (This is only a convention, justified by Lemma 7.4.) Under a suitable condition, the joint G-stationarity of ξ and X is preserved by integration:

Theorem 7.34 *(stationary integration) Given a group G, acting measurably on a Borel space S, consider a random measure ξ and a bounded, measurable process $X \geq 0$ on S, such that (ξ, X) is G-stationary and $\xi \ll E\xi$ a.s. Then $(\xi,\ X,\ X \cdot \xi)$ is strongly G-stationary.*

Proof: By Lemma 7.4, we have

$$(X \cdot \xi) \circ \theta_r^{-1} = (X \circ \theta_{r^{-1}}) \cdot (\xi \circ \theta_r^{-1}), \quad r \in G.$$

Since $\xi \ll E\xi$ a.s. and $(\xi \circ \theta_r^{-1}, X \circ \theta_{r^{-1}}) \stackrel{d}{=} (\xi, X)$, the assertion follows by Theorem 2.11. □

The last result fails without the hypothesis $\xi \ll E\xi$ a.s. For a counterexample, let ξ be a unit rate Poisson process on R with first point τ in R_+, and define $X_t \equiv 1\{\tau = t\}$. Then (ξ, X) is stationary, since $X_t = 0$ a.s. for every $t \in \mathsf{R}$, and yet stationarity fails for the process $X \cdot \xi = \delta_\tau$.

Using Theorem 7.24, we may construct some stationary densities and disintegration kernels, for jointly stationary random measures ξ and η. By a *random kernel* $\zeta\colon S\to T$, we mean a σ-finite kernel from $\Omega\times S$ to T. The triple (ξ,η,ζ) is said to be *strongly G-stationary*, if

$$(\xi,\eta,\mu\zeta)\circ\theta_r^{-1}\overset{d}{=}\Big\{\xi,\eta,(\mu\circ\theta_r^{-1})\zeta\Big\},\quad r\in G,\ \mu\in\mathcal{M}_S.$$

By Theorem 7.33, this is equivalent to ordinary G-stationarity. Here the stronger version follows by the same argument.

Theorem 7.35 *(stationary disintegration) For any lcscH group G, acting measurably on the Borel spaces S and T, consider some jointly G-stationary random measures ξ on $S\times T$ and η on S, with $\xi(\cdot\times T)\ll\eta$ a.s. Then there exists a (ξ,η)-measurable random kernel $\zeta\colon S\to T$, such that (ξ,η,ζ) is strongly G-stationary with $\xi=\eta\otimes\zeta$ a.s.*

When ξ and η are non-random, the stationarity becomes invariance, and the result essentially reduces to Theorem 7.6. For singleton T, we get conditions for the existence of a stationary density. Our proof simplifies only marginally in that case. In particular, we can then dispense with Lemma 7.30.

Corollary 7.36 *(stationary differentiation) For any lcscH group G, acting measurably on a Borel space S, consider some jointly G-stationary random measures $\xi\ll\eta$ on S. Then there exists a (ξ,η)-measurable process $X\geq 0$ on S, such that (ξ,η,X) is strongly G-stationary with $\xi=X\cdot\eta$ a.s.*

Proof of Theorem 7.35: Let φ be such as in Theorem 7.24, and define $\zeta_s=\varphi(s,\xi,\eta)$ for all $s\in S$, so that $\xi=\eta\otimes\zeta$. Since φ is G-invariant, we get for any $\mu\in\mathcal{M}_S$ and $r\in G$

$$\begin{aligned}\mu\zeta\circ\theta_r^{-1}&=\int\mu(ds)\,\varphi(s,\xi,\eta)\circ\theta_r^{-1}\\&=\int\mu(ds)\,\varphi\Big\{rs,(\xi,\eta)\circ\theta_r^{-1}\Big\}\\&=(\mu\circ\theta_r^{-1})\,\varphi\Big\{\cdot,(\xi,\eta)\circ\theta_r^{-1}\Big\},\end{aligned}$$

and so the G-stationarity of (ξ,η) yields

$$\begin{aligned}(\xi,\eta,\mu\zeta)\circ\theta_r^{-1}&=\Big\{(\xi,\eta)\circ\theta_r^{-1},(\mu\circ\theta_r^{-1})\,\varphi\Big(\cdot,(\xi,\eta)\circ\theta_r^{-1}\Big)\Big\}\\&\overset{d}{=}\Big\{\xi,\eta,(\mu\circ\theta_r^{-1})\,\varphi(\cdot,\xi,\eta)\Big\}\\&=\Big\{\xi,\eta,(\mu\circ\theta_r^{-1})\zeta\Big\},\end{aligned}$$

which shows that (ξ,η,ζ) is strongly G-stationary. □

If the random measure η of the last result is a.s. G-invariant, then by Theorem 2.12 it can be replaced by the random measure $E(\xi\,|\mathcal{I})$, where $\mathcal{I}$ denotes the associated invariant σ-field. Assuming ξ to be G-stationary, we show that $E(\xi\,|\mathcal{I})$, and indeed even $\mathcal{L}(\xi\,|\mathcal{I})$, has a G-invariant version.

Theorem 7.37 *(ergodic decomposition) For any lcscH group G, acting measurably on a Borel space S, let ξ be a G-stationary random measure on S, with associated G-invariant σ-field $\mathcal{I}$. Then $E(\xi\,|\mathcal{I}\,)$ and $\mathcal{L}(\xi\,|\mathcal{I}\,)$ have G-invariant, measure-valued versions.*

Proof: Given a measure-valued version $\Xi = \mathcal{L}(\xi\,|\mathcal{I}\,)$, we note that $E(\xi\,|\mathcal{I}\,)$ has the measure-valued version $\eta = \int m\,\Xi(dm)$. If Ξ is invariant under the shifts θ_r on $\mathcal{M}_S$, defined by $\theta_r\mu = \mu\circ\theta_r^{-1}$, then

$$\begin{aligned}\eta\circ\theta_r^{-1} &= \int (m\circ\theta_r^{-1})\,\Xi(dm)\\ &= \int m\,(\Xi\circ\theta_r^{-1})(dm)\\ &= \int m\,\Xi(dm) = \eta,\end{aligned}$$

which shows that η is invariant under shifts on S. It is then enough to prove the assertion for $\mathcal{L}(\xi\,|\mathcal{I}\,)$.

Since ξ is stationary and $\mathcal{I}$ is invariant, we have for any measurable function $f\ge 0$ on $\mathcal{M}_S$

$$E\Big\{f(\xi\circ\theta_r^{-1})\,\Big|\,\mathcal{I}\Big\} = E\{f(\xi)|\,\mathcal{I}\,\}\ \text{ a.s.},\quad r\in G. \tag{26}$$

Fixing a countable, measure-determining class $\mathcal{F}$ of functions f, and a countable, dense subset $G'\subset G$, we note that (26) holds simultaneously for all $f\in\mathcal{F}$ and $r\in G'$, outside a fixed P-null set. Since the space $\mathcal{M}_S$ is again Borel, we can choose a measure-valued version $\Xi = \mathcal{L}(\xi\,|\mathcal{I}\,)$.

For any non-exceptional realization Q of Ξ, we may choose a random measure η with distribution Q, and write the countably many relations (26), in probabilistic form, as

$$Ef(\eta\circ\theta_r^{-1}) = Ef(\eta),\quad f\in\mathcal{F},\ \ r\in G'.$$

Since $\mathcal{F}$ is measure-determining, we conclude that

$$\eta\circ\theta_r^{-1}\overset{d}{=}\eta,\quad r\in G'.$$

Using the invariance of Haar measure λ, and the measurability of the mapping $\mu\mapsto\lambda\otimes\mu$, we get for every $r\in G'$

$$(\lambda\otimes\eta)\circ\theta_r^{-1} = \lambda\otimes(\eta\circ\theta_r^{-1})\overset{d}{=}\lambda\otimes\eta, \tag{27}$$

where θ_r on the left denotes the joint shift $(p,s)\mapsto(rp,rs)$ on $G\times S$.

Next, we introduce on $G\times S$ the skew transformation $\vartheta(r,s) = (r,rs)$ with inverse $\tilde\vartheta$, and note as in Lemma 7.31 that

$$\tilde\vartheta\circ\theta_r = \theta_r'\circ\tilde\vartheta,\quad r\in G, \tag{28}$$

where $\theta'_r(p,s) = (rp,s)$ denotes the shift in $G \times S$, acting on G alone. Then (27) yields for any $r \in G'$

$$\begin{aligned}(\lambda \otimes \eta) \circ \tilde{\vartheta}^{-1} \circ \theta_r'^{-1} &= (\lambda \otimes \eta) \circ \theta_r^{-1} \circ \tilde{\vartheta}^{-1} \\ &\overset{d}{=} (\lambda \otimes \eta) \circ \tilde{\vartheta}^{-1},\end{aligned}$$

which means that the random measure $\zeta = (\lambda \otimes \eta) \circ \tilde{\vartheta}^{-1}$ on $G \times S$ is invariant in distribution, under shifts by $r \in G'$ in the component G alone.

To extend this to arbitrary $r \in G$, we may take $S = \mathsf{R}$, so that even $S' = G \times S$ becomes lcscH. Choosing a metrization of S', such that every bounded set is relatively compact, we see that the ε-neighborhood of a bounded set is again bounded. Then for any measures $\mu \in \mathcal{M}_{S'}$ and continuous functions $f \geq 0$ on S' with compact support, we get by dominated convergence, as $r \to \iota$ in G,

$$\begin{aligned}(\mu \circ \theta_r'^{-1})f &= \int \mu(dp\,ds)\, f(rp,s) \\ &\to \int \mu(dp\,ds)\, f(p,s) = \mu f.\end{aligned}$$

This gives $\mu \circ \theta_r'^{-1} \xrightarrow{v} \mu$, which shows that G acts continuously on $\mathcal{M}_{S'}$, under shifts in G alone. For any $r \in G$, we may choose some $r_n \to r$ in G', and conclude that

$$\zeta \overset{d}{=} \zeta \circ \theta_{r_n}'^{-1} \xrightarrow{v} \zeta \circ \theta_r'^{-1}.$$

Hence, $\zeta \circ \theta_r'^{-1} \overset{d}{=} \zeta$ for all $r \in G$, which means that ζ is stationary under shifts in G alone.

Reversing the skew transformation, and using (28) and the invariance of λ, we get for any $r \in G$

$$\lambda \otimes (\eta \circ \theta_r^{-1}) = (\lambda \otimes \eta) \circ \theta_r^{-1} \overset{d}{=} \lambda \otimes \eta.$$

Fixing any $B \in \mathcal{G}$ with $\lambda B \in (0,\infty)$, and using the measurability of the projection $\mu \mapsto \mu(B \times \cdot)$ on $G \times S$, we obtain

$$\begin{aligned}(\lambda B)\,(\eta \circ \theta_r^{-1}) &= \left\{\lambda \otimes (\eta \circ \theta_r^{-1})\right\}(B \times \cdot) \\ &\overset{d}{=} (\lambda \otimes \eta)(B \times \cdot) = (\lambda B)\,\eta,\end{aligned}$$

which implies $\eta \circ \theta_r^{-1} \overset{d}{=} \eta$. This shows that Q is G-invariant, and the a.s. G-invariance of Ξ follows. $\square$

Chapter 8

Exterior Conditioning

In Chapter 6, we saw that for any point process ξ on a localized Borel space S, the reduced Palm measures of different order can be obtained by disintegration of the universal *compound Campbell measure* C on $\hat{\mathcal{N}}_S \times \mathcal{N}_S$, given by $Cf = E\sum_{\mu\le\xi} f(\mu, \xi - \mu)$, where the summation extends over all bounded point measures $\mu \le \xi$. We now turn our attention to the equally important dual disintegration of C, needed in the context of exterior conditioning. For our present purposes, we need the associated supporting measure on $\mathcal{N}_S$ to be equal to $\mathcal{L}(\xi)$ itself.

A complete disintegration is possible only when $C(\hat{\mathcal{N}}_S \times \cdot) \ll \mathcal{L}(\xi)$. In general, we can only achieve a *partial disintegration*, as explained in Corollary 1.24. Thus, we may introduce the *maximal* kernel G from $\mathcal{N}_S$ to $\hat{\mathcal{N}}_S$ satisfying $\mathcal{L}(\xi) \otimes G \le C$, or in explicit notation,

$$\begin{aligned} E\,\Gamma f(\cdot, \xi) &= E\int G(\xi, d\mu)\, f(\mu, \xi) \\ &\le E\sum_{\mu\le\xi} f(\mu, \xi - \mu), \quad f \ge 0. \end{aligned} \tag{1}$$

Here the random measure $\Gamma = G(\xi, \cdot)$ on $\hat{\mathcal{N}}_S$, referred to below as the *Gibbs kernel* of ξ, has some remarkable properties, of great significance in statistical mechanics.

Most striking is perhaps the formula[1]

$$\mathcal{L}(1_B\xi \mid 1_{B^c}\xi) = \Gamma(\,\cdot \mid \mu B^c = 0) \text{ a.s. on } \{\xi B = 0\}, \quad B \in \hat{\mathcal{S}}, \tag{2}$$

which shows how all conditional distributions on the left can be obtained, by elementary conditioning, from the single random measure Γ. When ξ is a.s. bounded, the measure C becomes symmetric, and the dual disintegrations are essentially equivalent. This leads to a basic relationship between Gibbs and Palm kernels, stated in Theorem 8.5 as $\Gamma = Q_\xi / Q_\xi\{0\}$ a.s., where Q_μ denotes the reduced Palm distribution of ξ at the measure μ. For general ξ, the latter relation yields the equally remarkable formula

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

which may be regarded as the dual of (2).

[1] Recall that $1_B\xi$ denotes the restriction of ξ to B, so that $(1_B\xi)f = \int_B f d\xi$.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_8

Of special significance is the restriction of the Gibbs kernel Γ to point measures of total mass 1, which leads to the notion of *Papangelou kernel* η, defined by

$$\eta B = \Gamma\{\mu B = \mu S = 1\}, \quad B \in \hat{\mathcal{S}}.$$

Writing $1_B \xi = \delta_{\tau_B}$ when $\xi B = 1$, and putting $\hat{\eta}_B = 1_B \eta / \eta B$ when $\eta B > 0$, we get the remarkable formula

$$\mathcal{L}(\tau_B \,|\, 1_{B^c \xi},\, \xi B = 1) = \hat{\eta}_B \ \text{ a.s. on } \ \{\xi B = 0\},$$

which shows that all conditional distributions on the left are determined by the single random measure η.

Much of the indicated theory simplifies under suitable regularity conditions on the point process ξ. The basic hypothesis, traditionally denoted by (Σ), may be stated as

$$P(\xi B = 0 \,|\, 1_{B^c} \xi) > 0, \quad B \in \hat{\mathcal{S}}.$$

We also consider the stronger requirement $(\tilde{\Sigma})$, defined as condition (Σ) applied to a uniform randomization $\tilde{\xi}$ of ξ, regarded as a simple point process on $S \times [0,1]$. In Theorem 8.13, we show that (Σ) is equivalent to the celebrated *DLR equation*

$$E \int \Gamma(d\mu)\, f(\xi + \mu)\, 1\{\xi B = \mu B^c = 0\} = E f(\xi), \quad B \in \hat{\mathcal{S}},\ f \geq 0,$$

of significance in statistical mechanics. We further show that $(\tilde{\Sigma})$ is precisely the condition needed for the absolute continuity $C(\hat{\mathcal{N}}_S \times \cdot) \ll \mathcal{L}(\xi)$, required for the complete disintegration $C = \mathcal{L}(\xi) \otimes G$, along with the corresponding formula for the Papangelou kernel.

If ξ is a Poisson process on S with diffuse intensity $\lambda = E\xi$, we may easily verify that $\eta = \lambda$ a.s. Less obvious is the fact that, for simple point processes ξ satisfying (Σ), the stated property essentially characterizes the class of Poisson processes. In statistical mechanics, the latter may then be regarded as the basic point processes, representing total randomness. A weaker condition is the absolute continuity $\eta \ll \lambda$ a.s., which implies

$$\mathcal{L}(1_B \xi \,|\, 1_{B^c} \xi) \ll \Pi_{1_B \lambda} \ \text{ a.s.}, \quad B \in \hat{\mathcal{S}},$$

where Π_μ denotes the distribution of a Poisson process on S with intensity μ. The associated process $-\log(d\eta/d\lambda)$ is then interpreted as the *local energy* of the particle system ξ.

Of special importance is the case where η is λ-*invariant*, in the sense that[2], for any measurable mapping $f\colon S \to S$,

$$\lambda \circ f^{-1} = \lambda \quad \Rightarrow \quad \eta \circ f^{-1} = \eta \ \text{ a.s.}$$

[2]Recall that the measure $\lambda \circ f^{-1}$ is given by $(\lambda \circ f^{-1})g = \lambda(g \circ f)$.

This implies that ξ is λ-*symmetric*, in the sense that

$$\lambda \circ f^{-1} = \lambda \quad \Rightarrow \quad \xi \circ f^{-1} \stackrel{d}{=} \xi \text{ a.s.}$$

For unbounded λ, we may then conclude that ξ is a Cox process directed by η. This result, established in Theorems 8.19 and 8.20, along with a similar asymptotic statement, will be useful in Chapter 11 to analyze stationary line processes and particle systems.

The remainder of the chapter deals with some local limit theorems of great importance. For any K-marked point process ξ on S with locally finite S-projection $\bar{\xi}$, Theorem 8.22 provides the existence of a random measure π on $S \times K$ with purely atomic S-projection $\bar{\pi}$, such that a.s.

$$E\Big\{\xi(B \times \cdot);\, \bar{\xi}B = 1 \Big|\, 1_{B^c}\xi\Big\} \to \pi(\{s\} \times \cdot), \quad s \in S,$$

as $B \downarrow \{s\}$ along a fixed dissection system $\mathcal{I} \subset \hat{\mathcal{S}}$. We further show that ξ satisfies (Σ), iff a.s. $\bar{\pi}\{s\} < 1$ for all $s \in S$.

Now let η_d denote the restriction of the Papangeou kernel η to the set of all $s \in S$ with $\bar{\eta}\{s\} = 0$, and define the *external intensity* $\hat{\xi}$ of ξ as the random measure $\pi + \eta_d$. Letting (B_{nj}) be a dissection system on a subset $B \in \hat{\mathcal{S}}$, we prove in Theorem 8.25 that

$$\textstyle\sum_j E\Big\{\xi(B_{nj} \times \cdot);\, \bar{\xi}B_{nj} = 1 \Big|\, 1_{B^c_{nj}}\xi\Big\} \to \hat{\xi}(B \times \cdot) \text{ a.s.}$$

More generally, given a K-marked point process ξ on S, we may consider the associated external intensity $\hat{\zeta}$ of a random measure ζ on S with locally finite intensity $E\zeta$. In Theorem 8.26, we establish some limiting properties of the form

$$\textstyle\sum_j E\Big(\zeta B_{nj} \,\Big|\, 1_{B^c_{nj}}\xi\Big) \to \hat{\zeta}B \text{ a.s. and in } L^1.$$

We finally introduce the notions of *externally measurable* processes and random measures, relative to a marked point process ξ on S. In Theorem 8.27, we characterize the external intensity of a random measure ζ as the a.s. unique, externally measurable random measure $\hat{\zeta}$ on S, such that[3] $E\zeta Y = E\hat{\zeta} Y$ for every externally measurable process $Y \geq 0$ on S. Thus, $\hat{\zeta}$ may be regarded as the *dual external projection* of ζ, just as the compensator $\hat{\xi}$ of a random measure ξ on R_+ can be characterized as the *dual predictable projection* of ξ, with respect to the underlying filtration.

8.1 Gibbs and Papangelou Kernels

The Palm measures of a point process ξ may be thought of as providing the conditional distribution of the outer configuration, given the local behavior

[3]Note that $\zeta Y = \int Y d\zeta$.

of ξ in some bounded set. In this chapter we consider the reverse problem of describing the distribution of ξ in some bounded set, given the outer configuration. Such conditional distributions play an important role in statistical mechanics, where the points of ξ may be interpreted as particles in a suitable phase space, and the mentioned distributions are determined by the local interaction of particles, possibly combined with an external force field.

In Chapter 6, we saw how the whole family of Palm measures of different orders can be obtained by disintegration of the *compound Campbell measure* C, given by $Cf = E\sum_{\mu\le\xi} f(\mu, \xi-\mu)$, where the summation extends over all finite point measures $\mu \le \xi$. More precisely, assuming $\xi = \sum_{j\in J}\delta_{\tau_j}$, we consider the set of all measures $\mu = \sum_{i\in I}\delta_{\tau_i}$, for finite subsets $I \subset J$. The Palm disintegration may now be written[4] as $C = \nu \otimes Q$, for a suitable supporting measure ν on $\hat{\mathcal{N}}_S$, where Q is a kernel from $\hat{\mathcal{N}}_S$ to $\mathcal{N}_S$. When ξ is simple, we may think of Q_μ as the "reduced" conditional distribution of ξ, given that ξ supports the points of μ.

In this section, we prove the surprising fact that all external conditional distributions of the form $\mathcal{L}(1_B\xi \,|\, 1_{B^c}\xi)$ can essentially be constructed from a single random measure $\Gamma = G(\xi,\cdot)$ on $\hat{\mathcal{N}}_S$, called the *Gibbs kernel* of ξ, obtainable by dual disintegration of the same compound Campbell measure C. Since in this case the distribution $\mathcal{L}(\xi)$ itself must be chosen as our supporting measure, the required absolute continuity $C(\hat{\mathcal{N}}_S \times \cdot) \ll \mathcal{L}(\xi)$ may fail in general, so that only a partial disintegration is possible. For later needs, we will develop the entire theory in the presence of an auxiliary random element τ in a Borel space T. Then Γ is defined as the maximal random measure on $\hat{\mathcal{N}}_S \times T$, satisfying

$$E\iint \Gamma(d\mu\, dt)\, f(t,\mu,\xi) \le E\sum_{\mu\le\xi} f(\tau,\mu,\xi-\mu), \tag{3}$$

for any measurable function $f \ge 0$ on $T \times \hat{\mathcal{N}}_S \times \mathcal{N}_S$. The existence of such a disintegration kernel follows from Corollary 1.24.

Lemma 8.1 *(boundedeness and support) Let Γ be the Gibbs kernel of a pair (ξ,τ), where ξ is a point process on S, and τ is a random element in T. Then*

(i) $\Gamma\{\mu S = \infty\} = 0$ *and* $\Gamma\{\mu B = \mu S = n\} < \infty$ *a.s.,* $B \in \hat{\mathcal{S}}$, $n \in \mathsf{Z}_+$,

(ii) $\int \mu(\operatorname{supp}\xi)\,\Gamma(d\mu\, dt) = 0$ *a.s., when ξ is simple.*

Proof: (i) For any $B \in \hat{\mathcal{S}}$ and $m, n \in \mathsf{Z}_+$, we have

$$\begin{aligned} &E\Big(\Gamma\{\mu B = \mu S = n\};\, \xi B = m\Big) \\ &\qquad \le C\Big(\{\nu B = m\} \times \{\mu B = \mu S = n\}\Big) \\ &\qquad = E 1\{\xi B = m+n\} \sum_{\mu\le\xi} 1\{\mu B = \mu S = n\} \\ &\qquad = \binom{m+n}{n} P\{\xi B = m+n\} < \infty, \end{aligned}$$

[4] Recall that $(\nu\otimes\mu)f = \int \nu(ds)\int \mu_s(dt) f(s,t)$, where $\mu f = \int f d\mu$.

and so $\Gamma\{\nu B = \nu S = n\} < \infty$ a.s. on $\{\xi B = m\}$.

(ii) Using a dissection system on S to approximate $\operatorname{supp} m$, we see by dominated convergence that the function $f(m,\mu) = \mu(\operatorname{supp} m)$ on $\mathcal{N}_S \times \hat{\mathcal{N}}_S$ is product-measurable. Hence, for simple ξ, we get by (3)

$$E\int \mu(\operatorname{supp}\xi)\,\Gamma(d\mu\,dt) \le Cf = E\sum_{\mu\le\xi}\mu(\operatorname{supp}(\xi-\mu)) = 0,$$

and so the inner integral on the left vanishes a.s. □

The definition of Γ is justified by some remarkable formulas:

Theorem 8.2 *(Gibbs kernel) Let $\Gamma = G(\xi,\cdot)$ be the Gibbs kernel of a pair (ξ,τ), where ξ is a point process on S, and τ is a random element in T. Then for any $B \in \hat{\mathcal{S}}$ and measurable $f \ge 0$ on $\hat{\mathcal{N}}_S \times T$,*

(i) $$E\iint \Gamma(d\mu\,dt)\, f(\xi+\mu,t)\,1\{\xi B = \mu B^c = 0\} = E\Big\{f(\xi,\tau);\ P(\xi B = 0\,|\,1_{B^c}\xi) > 0\Big\},$$

(ii) $$\Gamma(\,\cdot\,;\ \mu B^c = 0) = \frac{\mathcal{L}(1_B\xi,\tau\,|\,1_{B^c}\xi)}{P(\,\xi B = 0\,|\,1_{B^c}\xi)} \quad \textit{a.s. on } \{\xi B = 0\},$$

(iii) $$\mathcal{L}(1_B\xi,\tau|\,1_{B^c}\xi) = \Gamma(\,\cdot\,|\,\mu B^c = 0) \quad \textit{a.s. on } \{\xi B = 0\}.$$

When ξ is simple, we see from (ii) and Lemma 8.1 (ii) that Γ is a.s. determined by the conditional distributions $\mathcal{L}(1_B,\tau|\,1_{B^c}\xi)$. Conversely, (iii) shows that the same conditional distributions are obtainable from the single random measure Γ. In subsequent proofs, we will often use the elementary fact that $P(A|\,\mathcal{F}) > 0$ a.s. on A, for any event A and σ-field $\mathcal{F}$, which holds since

$$P\Big\{P(A|\,\mathcal{F}) = 0;\ A\Big\} = E\Big\{P(A|\,\mathcal{F});\ P(A|\,\mathcal{F}) = 0\Big\} = 0.$$

Proof: (i) For any $B \in \hat{\mathcal{S}}$, define

$$M_0 = \Big\{\mu;\ \mu B = 0,\ P(\xi B = 0\,|\,1_{B^c}\xi)_\mu > 0\Big\},$$

and let $M \subset M_0$ be measurable. Then

$$\begin{aligned}
&C(M\times\{\mu B^c = 0\})\\
&\quad = E\sum_{\mu\le\xi} 1\Big\{\xi-\mu\in M,\ \mu B^c = 0\Big\}\\
&\quad = E\sum_{\mu\le\xi} 1\Big\{\xi-\mu\in M,\ P(\xi B = 0\,|\,1_{B^c}\xi) > 0,\ \mu B^c = 0\Big\}\\
&\quad = P\Big\{1_{B^c}\xi\in M,\ P(\xi B = 0\,|\,1_{B^c}\xi) > 0\Big\}\\
&\quad \ll E\Big\{P(\xi B = 0\,|\,1_{B^c}\xi);\ 1_{B^c}\xi\in M\Big\}\\
&\quad = P\Big\{\xi B = 0,\ 1_{B^c}\xi\in M\Big\} \le P\{\xi\in M\},
\end{aligned}$$

and so $C(\cdot \times \{\mu B^c = 0\}) \ll \mathcal{L}(\xi)$ on M_0. Noting that $P(\xi B = 0 \mid 1_{B^c}\xi) > 0$ a.s. on $\{\xi B = 0\}$, and using the maximality of G, we get for any $B \in \hat{\mathcal{S}}$ and measurable $f \geq 0$ on $\mathcal{N}_S$

$$\begin{aligned}
&E\int \Gamma(d\mu\, dt)\, f(\xi+\mu, t)\, 1\{\xi B = \mu B^c = 0\} \\
&\quad = E\int \Gamma(d\mu\, dt)\, f(\xi+\mu, t)\, 1\Big\{\xi B = \mu B^c = 0,\ P(\xi B = 0 \mid 1_{B^c}\xi) > 0\Big\} \\
&\quad = \iint C(d\mu\, d\nu\, dt)\, f(\mu+\nu, t)\, 1\Big\{\nu B = \mu B^c = 0,\ P(\xi B = 0 \mid 1_{B^c}\xi)_\nu > 0\Big\} \\
&\quad = E\sum_{\mu \leq \xi} f(\xi, \tau)\, 1\Big\{\xi B = \mu B,\ \mu B^c = 0,\ P(\xi B = 0 \mid 1_{B^c}\xi) > 0\Big\} \\
&\quad = E\Big\{f(\xi, \tau);\ P(\xi B = 0 \mid 1_{B^c}\xi) > 0\Big\}.
\end{aligned}$$

(ii) Replacing the function f in (i) by

$$g(\mu, t) = f(1_B\mu, t)\, 1_M(1_{B^c}\mu), \quad t \in T,\ \mu \in \mathcal{N}_S,$$

for measurable $M \subset \mathcal{N}_S$ and $f \geq 0$ on $T \times \hat{\mathcal{N}}_S$, we get

$$\begin{aligned}
&E\Big\{f(1_B\xi, \tau)\, 1_M(1_{B^c}\xi);\ P(\xi B = 0 \mid 1_{B^c}\xi) > 0\Big\} \\
&\qquad = E\iint \Gamma(d\mu\, dt)\, f(\mu, t)\, 1_M(\xi)\, 1\{\xi B = \mu B^c = 0\} \\
&\qquad = E\iint G(1_{B^c}\xi,\ d\mu\, dt)\, f(\mu, t)\, 1_M(1_{B^c}\xi)\, 1\{\xi B = \mu B^c = 0\}.
\end{aligned}$$

Since M was arbitrary, we obtain a.s.

$$\begin{aligned}
&E\Big\{f(1_B\xi, \tau) \mid 1_{B^c}\xi\Big\}\, 1\Big\{P(\xi B = 0 \mid 1_{B^c}\xi) > 0\Big\} \\
&\qquad = \iint G(1_{B^c}\xi,\ dt\, d\mu)\, f(\mu, t)\, 1\{\mu B^c = 0\}\, P(\xi B = 0 \mid 1_{B^c}\xi),
\end{aligned}$$

and since $P(\xi B = 0 \mid 1_{B^c}\xi) > 0$ a.s. on $\{\xi B = 0\}$, this implies

$$1_{\mathcal{N}_B}\Gamma f = \frac{E\Big\{f(1_B\xi, \tau)\,\Big|\, 1_{B^c}\xi\Big\}}{P(\xi B = 0 \mid 1_{B^c}\xi)} \ \text{ a.s. on } \{\xi B = 0\}.$$

The assertion now follows by Lemma 2.1.

(iii) By (ii) we have

$$\Gamma\{\mu B^c = 0\} = \Big\{P(\xi B = 0 \mid 1_{B^c}\xi)\Big\}^{-1} > 0 \ \text{ a.s. on } \ \{\xi B = 0\},$$

and the assertion follows by division. □

Of special importance is the restriction of the Gibbs kernel to the set of single point masses δ_s. Identifying δ_s with the point $s \in S$, we obtain a random measure $\eta = g(\xi, \cdot)$ on S, called the *(first order) Papangelou kernel* of ξ. It may also be defined directly by the maximum disintegration

$$E \int \eta(ds)\, f(s, \xi) \leq E \int \xi(ds)\, f(s, \xi - \delta_s), \tag{4}$$

for any measurable function $f \geq 0$ on $S \times \mathcal{N}_S$. Similarly, we define the *n-th order Papangelou kernel* $\eta_n = g_n(\xi, \cdot)$ as the random measure on S^n, given by the partial disintegration

$$E \int \eta_n(ds)\, f(s, \xi) \leq E \int \cdots \int \xi^{(n)}(ds)\, f\Big(s, \xi - \textstyle\sum_i \delta_{s_i}\Big).$$

Writing G_n for the image of $g_n/n!$ under the mapping $(s_1, \ldots, s_n) \mapsto \sum_{i \leq n} \delta_{s_i}$, we see that $\Gamma_n = G_n(\xi, \cdot)$ agrees with the restriction of Γ to the set of point measures μ with $\|\mu\| = n$. Defining $\Gamma_0 = G_0(\xi, \cdot) = \delta_0$ for consistency, we obtain $\Gamma = \sum_{n \geq 0} \Gamma_n$.

From Theorem 8.2, we can easily read off some basic properties of the first order Papangelou kernel. In particular, the conditional distributions in (ii) below are given by the single random measure η. For a striking formulation, we define the random element τ_B by $1_B \xi = \delta_{\tau_B}$, when $\xi B = 1$.

Corollary 8.3 *(Papangelou kernel) Let ξ be a point process on S. Then $\eta = g(\xi, \cdot)$ is a.s. locally finite, and for simple ξ, we have $\eta(\operatorname{supp} \xi) = 0$ a.s. Furthermore, we have for any $B \in \hat{\mathcal{S}}$*

(i) $$1_B \eta = \frac{E\Big(1_B \xi;\, \xi B = 1 \,\Big|\, 1_{B^c} \xi\Big)}{P(\,\xi B = 0 \,|\, 1_{B^c} \xi)} \quad \textit{a.s. on } \{\xi B = 0\},$$

(ii) $$\mathcal{L}\Big(\tau_B \,\Big|\, 1_{B^c} \xi,\, \xi B = 1\Big) = \frac{1_B \eta}{\eta B} \quad \textit{a.s. on } \{\xi B = 0,\ \eta B > 0\}.$$

Proof: The local finiteness is clear, since by Lemma 8.1 (i),

$$\eta B = \Gamma\{\mu B = \mu S = 1\} < \infty \text{ a.s.}, \quad B \in \hat{\mathcal{S}}.$$

When ξ is simple, part (ii) of the same lemma yields a.s.

$$\begin{aligned} \eta(\operatorname{supp} \xi) &= \Gamma\{\mu(\operatorname{supp} \xi) = \mu S = 1\} \\ &\leq \int \Gamma(d\mu)\, \mu(\operatorname{supp} \xi) = 0. \end{aligned}$$

(i) By Theorem 8.2 (ii), we get for any $C \subset B$ in $\hat{\mathcal{S}}$

$$\begin{aligned} \eta C &= \Gamma\{\mu C = \mu S = 1\} \\ &= \frac{P\Big(\xi C = \xi B = 1 \,\Big|\, 1_{B^c} \xi\Big)}{P(\xi B = 0 \,|\, 1_{B^c} \xi)} \text{ a.s. on } \{\xi B = 0\}. \end{aligned}$$

(ii) By (i), we have

$$\eta B = \frac{P(\xi B = 1 \,|\, 1_{B^c}\xi)}{P(\,\xi B = 0 \,|\, 1_{B^c}\xi)} \text{ a.s. on } \{\xi B = 0\}.$$

When this is positive, the assertion follows by division. □

We may supplement the elementary Lemma 8.1 (i) by a useful uniform bound, involving the linear order induced by a dissection system on S, defined as in Lemma 1.8.

Lemma 8.4 *(uniform bound) Let Γ be the Gibbs kernel of a pair (ξ, τ), where ξ is a K-marked point process on S, and τ is a random element in T. Fix a non-decreasing function $f \geq 0$ on $\mathcal{N}_{S\times K} \times T$, with $Ef(1_B\xi, \tau) < \infty$ for all $B \in \hat{\mathcal{S}}$, and let $\mathcal{I}$ be the class of bounded S-intervals, in the linear order induced by a dissection system on S. Then for any $J \in \mathcal{I}$,*

$$\sup_{I \in \mathcal{I} \cap J} \iint \Gamma(d\mu\, dt)\, f(\mu, t)\, 1\Big\{\bar{\xi} I = \bar{\mu} I^c = 0\Big\} < \infty \;\; a.s. \tag{5}$$

Proof: Fix any $J \in \mathcal{I}$. For any $[a, b] \in \mathcal{I} \cap J$ and $\mu \in \mathcal{N}$, let s_μ and t_μ be the first and last points of $\bar{\mu}$ in $[a, b]$, and put $I_\mu = (s_\mu, t_\mu)$, where $I_\mu = \emptyset$ when $\mu I = 0$. Writing $\xi' = 1_{[a,b]}\xi$, we get from (3)

$$\begin{aligned} E \iint \Gamma(d\mu\, dt)\, f(\mu, t)\, 1\Big\{\bar{\xi} I_\xi = \bar{\mu} I_\xi^c = 0\Big\} \\ \leq E \sum_{\mu+\nu=\xi'} f(\mu, \tau)\, 1\Big\{\bar{\mu} I_\nu^c = \bar{\nu} I_\nu = 0\Big\}. \end{aligned}$$

If $\bar{\xi}' \neq 0$, then $\bar{\nu}$ contains s_ξ and t_ξ, but no intermediate point of $\bar{\xi}$. If instead $\xi' = 0$, then trivially $\nu = 0$. In either case, the last sum contains exactly one term, and the left-hand side is bounded by $Ef(\mu, \tau) \leq Ef(\xi', \tau) < \infty$, which shows that

$$\iint \Gamma(d\mu, dt)\, f(\mu, t)\, 1\Big\{\bar{\xi} I_\xi = \bar{\mu} I_\xi^c = 0\Big\} < \infty \text{ a.s.} \tag{6}$$

Replacing I_μ by $I'_\mu = [a, t_\mu)$ or $I''_\mu = (s_\mu, b]$, we see that $\xi' \neq 0$ implies $\bar{\nu} = \delta_{t_\xi}$ or δ_{s_ξ}, respectively, and so (6) remains true for I'_ξ and I''_ξ. Since every connected component of $J \setminus \operatorname{supp} \xi$ is of the form I_ξ, I'_ξ, or I''_ξ for some interval $[a, b] \subset J$, and since the integral in (5) is a nondecreasing function of $I \subset I_\xi$, the asserted uniformity follows. □

Now let $Q(\mu, \cdot)$ be a version of the reduced Palm kernel of ξ, where $\mu \in \hat{\mathcal{N}}_S$. The following result gives the basic connection between Palm and Gibbs kernels.

Theorem 8.5 *(Gibbs/Palm connection) Let ξ be a point process on S with Gibbs kernel G.*

(i) *When* $\xi S < \infty$ *a.s., let* Q *be a reduced Palm kernel of* ξ. *Then* $Q(\xi, \{0\}) > 0$ *a.s., and*

$$\Gamma = G(\xi, \cdot) = \frac{Q(\xi, \cdot)}{Q(\xi, \{0\})} \ \text{ a.s.}$$

(ii) *Let the* Q_n *be reduced Palm kernels of the restrictions* $\xi_n = 1_{S_n}\xi$, *where* $S_n \uparrow S$ *in* $\hat{\mathcal{S}}$. *Then for any* $B \in \hat{\mathcal{S}}$,

$$\frac{Q_n(\xi_n, \cdot)}{Q_n(\xi_n, \{0\})} \to 1_{\mathcal{N}_B} G(\xi, \cdot) \ \text{ a.s. on } \{\xi B = 0\}.$$

Since the Palm kernels are only determined up to a normalization, (i) shows that the Palm and Gibbs kernels are essentially the same object. This may be surprising, since they represent dual forms of conditioning, from individual points of ξ to the outer configuration, and vice versa.

Proof: (i) When $\xi S < \infty$ a.s., the compound Campbell measure C is symmetric, in the sense that $C\tilde{f} = Cf$, where $\tilde{f}(s,t) = f(t,s)$. Since $\mathcal{L}(\xi) \ll C(\cdot \times \mathcal{N}_S)$, we may choose an $A \in \mathcal{B}_{\mathcal{N}_S}$ with $\xi \in A$ a.s., and $\mathcal{L}(\xi) \sim C(\cdot \times \mathcal{N}_S)$ on A. Then on $A \times \mathcal{N}_S$, we have the dual disintegrations

$$C = \mathcal{L}(\xi) \otimes G = \nu \otimes Q \ \text{ on } A \times \mathcal{N}_S,$$

where ν is the supporting measure on $\mathcal{N}_S$, associated with Q. By the uniqueness in Theorem 1.23, it follows that

$$\Gamma \equiv G(\xi, \cdot) = p(\xi)\, Q(\xi, \cdot) \ \text{ a.s.,} \tag{7}$$

for some measurable function $p > 0$ on A. To determine p, we may apply Theorem 8.2 (ii) to the sets $M = \{0\}$ and $B = \emptyset$, to get a.s.

$$\Gamma\{0\} = \frac{P(1_\emptyset \xi = 0 \,|\, 1_S \xi)}{P(\xi\emptyset = 0 \,|\, 1_S \xi)} = \frac{P(\Omega \,|\, \xi)}{P(\Omega \,|\, \xi)} = 1.$$

Inserting this into (7) yields $p(\xi)\, Q(\xi, \{0\}) = 1$ a.s., and the assertion follows.

(ii) Writing G_n for the Gibbs kernel of ξ_n, we get by (i) and Theorem 8.2 (ii)

$$\frac{Q_n(\xi_n, \cdot)}{Q_n(\xi_n, \{0\})} = 1_{\mathcal{N}_B} G_n(\xi_n, \cdot) = \frac{\mathcal{L}\Big(1_B \xi \,\Big|\, 1_{S_n \setminus B}\xi\Big)}{P\Big(\xi B = 0 \,\Big|\, 1_{S_n \setminus B}\xi\Big)},$$

a.s. on $\{\xi B = 0\}$. By the same theorem and martingale convergence, we also have for any $M \in \mathcal{B}_{\mathcal{N}_B}$

$$\frac{P\Big(1_B \xi \in M \,\Big|\, 1_{S_n \setminus B}\xi\Big)}{P\Big(\xi B = 0 \,\Big|\, 1_{S_n \setminus B}\xi\Big)} \to \frac{P(1_B \xi \in M \,|\, 1_{B^c}\xi)}{P(\xi B = 0 \,|\, 1_{B^c}\xi)} = G(\xi, M),$$

a.s. on the same set. The assertion now follows by combination. □

In Theorem 8.2 (iii), we saw how the conditional distributions $\mathcal{L}(1_B\xi|\, 1_{B^c}\xi)$ can be obtained from the Gibbs kernel by a suitable normalization. Using the last theorem, we can now show that the reverse conditional distributions $\mathcal{L}(1_{B^c}\xi \,|\, 1_B\xi)$ are obtainable from the reduced Palm kernel Q by an simple conditioning. This is another interesting case of duality.

Corollary 8.6 *(inner conditioning) Let ξ be a point process on S with reduced Palm kernel Q. Then for any $B \in \hat{\mathcal{S}}$, we have $Q_{1_B\xi}\{\mu B = 0\} > 0$, and*

$$\mathcal{L}(1_{B^c}\xi \,|\, 1_B\xi) = Q_{1_B\xi}(\,\cdot\,|\, \mu B = 0) \quad a.s.$$

Proof: Applying Theorem 8.2 (ii) with ξ, τ, and B^c replaced by $1_B\xi$, $1_{B^c}\xi$, and B, and using Theorem 8.5 with ξ replaced by $1_B\xi$, we get a.s.

$$\begin{aligned}
\mathcal{L}(1_{B^c}\xi \,|\, 1_B\xi) &= \Gamma^B_{1_B\xi}(1_{B^c}\mu \,|\, \mu B = 0)\\
&= Q^B_{1_B\xi}(1_{B^c}\mu \in \cdot \,|\, \mu B = 0)\\
&= Q_{1_B\xi}(1_{B^c}\mu \in \cdot \,|\, \mu B = 0),
\end{aligned}$$

where Γ^B and Q^B denote the Gibbs and Palm kernels of the restriction $1_B\xi$. □

8.2 Transformation Properties

Arguments involving the Gibbs kernel may often be simplified by a reduction to a simpler setting. First we note how the multiplicities of a general point process may be regarded as marks in N, so that a point process on S can be identified with an N-marked point process on S. This is formally achieved by a bi-measurable map $\rho\colon \mathcal{N}_S \to \mathcal{N}^*_{S\times\mathsf{N}}$ with inverse ρ^{-1}, given by

$$\rho(\mu) = \sum_{s,n} \delta_{s,n} 1\{\mu\{s\} = n\}, \qquad \rho^{-1}(\nu) = \iint n\, \delta_s\, \mu(ds\, dn),$$

where both summation and integration range over $S \times \mathsf{N}$.

Lemma 8.7 *(multiplicities as marks) For any point process ξ on S, form the N-marked point process $\xi' = \rho(\xi)$ on S. Then the associated Gibbs kernels Γ and Γ' are related by*

$$\Gamma' = \begin{cases} \Gamma \circ \rho^{-1} & a.s.\ on\ \{\, \bar{\mu}(\operatorname{supp}\xi) = 0\,\},\\ 0 & a.s.\ on\ \{\, \bar{\mu}(\operatorname{supp}\xi) > 0\,\}.\end{cases}$$

Proof: The expression on $\{\, \bar{\mu}(\operatorname{supp}\xi) = 0\}$ is obtained, most easily, from Theorem 8.2 (ii), whereas that on the complement may be proved in the same way as Lemma 8.1 (ii). □

Next, we show how a suitable conditioning yields a reduction to the case of bounded point processes. The statement may be compared with the Palm iteration in Theorem 6.22.

Lemma 8.8 *(conditioning) Given a simple point process ξ on S with Gibbs kernel G, and a set $B \in \hat{\mathcal{S}}$, let $G_B(\cdot, \cdot \,|\, \mu)$ be the Gibbs kernel of $1_B\xi$, under the conditional distribution $\mathcal{L}(1_B\xi \,|\, 1_{B^c}\xi)_\mu$. Then*

$$1_{\mathcal{N}_B} G(\xi, \cdot) = G_B(1_B\xi, \cdot \,|\, 1_{B^c}\xi) \;\; a.s.$$

The existence of a measurable version of G_B is clear from the martingale construction in Theorem 1.28.

Proof: For measurable $M \subset \{\mu B^c = 0\}$ and $M', M'' \subset \mathcal{N}_S$, we have

$$\begin{aligned} &C\Big(M \times \{1_B\mu \in M',\, 1_{B^c}\mu \in M''\}\Big) \\ &\quad = E\sum_{\mu \le \xi} 1\Big\{\mu \in M,\, 1_B\xi - \mu \in M',\, 1_{B^c}\xi \in M''\Big\} \\ &\quad = E\Big(E\Big\{\sum_{\mu \le 1_B\xi} 1\{\mu \in M,\, 1_B\xi - \mu \in M'\}\Big|\, 1_{B^c}\xi\Big\};\, 1_{B^c}\xi \in M''\Big) \\ &\quad = E\Big\{C^B(M \times M');\, 1_{B^c}\xi \in M''\Big\}, \end{aligned} \tag{8}$$

where C^B is the compound Campbell measure of $1_B\xi$, under conditioning on $1_{B^c}\xi$. Writing $\Gamma^B = G^B(1_B\xi, \cdot \,|\, 1_{B^c}\xi)$, we see that the right-hand side is bounded below by

$$\begin{aligned} &E\Big(E\Big\{\Gamma^B M;\, 1_B\xi \in M' \,\Big|\, 1_{B^c}\xi\Big\};\, 1_{B^c}\xi \in M''\Big) \\ &\qquad = E\Big(\,\Gamma^B M;\, 1_B\xi \in M',\, 1_{B^c}\xi \in M''\Big), \end{aligned}$$

which extends, as in Lemma 2.1 (i), to

$$C(M \times A) \ge E(\,\Gamma^B M;\, \xi \in A),$$

and so $\Gamma^B M \le \Gamma M$ a.s., by the maximality of G. Conversely, the left-hand side of (8) is bounded below by

$$\begin{aligned} &E\Big(\Gamma M;\, 1_B\xi \in M',\, 1_{B^c}\xi \in M''\Big) \\ &\qquad = E\Big(E\Big\{\Gamma M;\, 1_B\xi \in M' \,\Big|\, 1_{B^c}\xi\Big\};\, 1_{B^c}\xi \in M''\Big), \end{aligned}$$

which gives

$$C^B(M \times M') \ge E\Big(\Gamma M;\, 1_B\xi \in M' \,\Big|\, 1_{B^c}\xi\Big) \;\text{ a.s.},$$

and so $\Gamma M \leq \Gamma^B M$ a.s., by the maximality of G^B. Combining the two relations yields $\Gamma^B = \Gamma$ a.s. on $\{\mu B^c = 0\}$, as asserted. $\square$

We proceed to show how the Gibbs kernel is transformed by a randomization. When ξ is simple, we also show how a randomization may dissolve the possible atoms of the Papangelou kernels.

Theorem 8.9 *(randomization) Let ξ be a point process on S with ν-randomization $\tilde{\xi}$, for a probability kernel $\nu\colon S \to T$, and write $\tilde{\mu}$ for a ν-randomization of the fixed measure $\mu \in \mathcal{N}_S$. Then*

(i) *the associated Gibbs kernels Γ and $\tilde{\Gamma}$ are related by*

$$\tilde{\Gamma} f = \int \Gamma(d\mu)\, E f(\tilde{\mu}) \quad a.s., \qquad f \geq 0,$$

(ii) *when ξ is simple, the associated Papangelou kernels η_n and $\tilde{\eta}_n$ are related by*

$$\tilde{\eta}_n = \eta_n \otimes \nu^{\otimes n} \quad a.s., \qquad n \in \mathsf{N}.$$

Proof: (i) Define a random measure Γ' on $\hat{\mathcal{N}}_{S\times T}$ by the expression on the right. For measurable $f \geq 0$, we may apply (1) to the function $g(\mu, \mu') = Ef(\tilde{\mu}, \tilde{\mu}')$, to get

$$\begin{aligned} E\int \Gamma'(d\mu)\, f(\mu, \tilde{\xi}) &= E\int \Gamma(d\mu)\, E\big\{f(\tilde{\mu}, \tilde{\xi})\big|\, \xi\big\} \\ &= E\int \Gamma(d\mu)\, g(\mu, \xi) \\ &\leq Cg = E\sum_{\mu\leq\xi} g(\mu, \xi - \mu) \\ &= E\sum_{\mu\leq\xi} E\big\{f(\tilde{\mu}, (\xi-\mu)\tilde{}\,)\big|\, \xi, \mu\big\} \\ &= E\sum_{\tilde{\mu}\leq\tilde{\xi}} f(\tilde{\mu}, \tilde{\xi} - \tilde{\mu}) = \tilde{C} f, \end{aligned}$$

which shows that $\Gamma' \leq \tilde{\Gamma}$ a.s.

Next, we define a random measure Γ'' on $\hat{\mathcal{N}}_S$ by $\Gamma'' f = \int \tilde{\Gamma}(d\mu)\, f(\bar{\mu})$, where $\bar{\mu}$ denotes the projection $\mu(\cdot \times T)$. For measurable $f \geq 0$, we may apply (1) for $\tilde{\Gamma}$ and $\tilde{C}$ to the function $g(\mu, \mu') = f(\bar{\mu}, \bar{\mu}')$, to get

$$\begin{aligned} E\int \Gamma''(d\mu)\, f(\mu, \xi) &= E\int \tilde{\Gamma}(d\mu)\, f(\bar{\mu}, \xi) \\ &= E\int \tilde{\Gamma}(d\mu)\, g(\mu, \tilde{\xi}) \\ &\leq \tilde{C} g = E\sum_{\tilde{\mu}\leq\tilde{\xi}} g(\tilde{\mu}, \tilde{\xi} - \tilde{\mu}) \\ &= E\sum_{\mu\leq\xi} f(\mu, \xi - \mu) = Cf, \end{aligned}$$

which shows that $\Gamma'' \leq \Gamma$ a.s.

For any $n \in \mathsf{Z}_+$ and $B \in \hat{\mathcal{S}}$, we now define

$$M_{n,B} = \left\{\mu \in \hat{\mathcal{N}}_S;\ \mu B = \mu S \le n\right\},$$
$$\bar{M}_{n,B} = \left\{\mu \in \hat{\mathcal{N}}_{S\times T};\ \bar{\mu} B = \bar{\mu} S \le n\right\}.$$

Using the definitions of Γ' and Γ'', and the a.s. relations $\Gamma' \le \tilde{\Gamma}$ and $\Gamma'' \le \Gamma$, we get a.s. for measurable $M \subset \hat{\mathcal{N}}_{S\times T}$

$$\begin{aligned}
\tilde{\Gamma}(\bar{M}_{n,B}) &= \Gamma''(M_{n,B}) \le \Gamma(M_{n,B}) = \Gamma'(\bar{M}_{n,B}) \\
&= \Gamma'(M \cap \bar{M}_{n,B}) + \Gamma'(M^c \cap \bar{M}_{n,B}) \\
&\le \tilde{\Gamma}(M \cap \bar{M}_{n,B}) + \tilde{\Gamma}(M^c \cap \bar{M}_{n,B}) \\
&= \tilde{\Gamma}(\bar{M}_{n,B}).
\end{aligned}$$

Since $\Gamma(M_{n,B}) < \infty$ a.s. by Lemma 8.1, all relations reduce to a.s. equalities, and in particular

$$\tilde{\Gamma}(M \cap \bar{M}_{n,B}) = \Gamma'(M \cap \bar{M}_{n,B}) \ \text{ a.s.}, \quad n \in \mathsf{Z}_+,\ B \in \hat{\mathcal{S}}.$$

Letting $n \to \infty$, and then $B \uparrow S$, we conclude from Lemma 8.1 that a.s.

$$\begin{aligned}
\tilde{\Gamma}(M) &= \tilde{\Gamma}\{\mu \in M;\ \bar{\mu} S < \infty\} \\
&= \Gamma'\{\mu \in M;\ \bar{\mu} S < \infty\} = \Gamma'(M).
\end{aligned}$$

Since $\hat{\mathcal{N}}_{S\times T}$ is Borel by Theorem 1.5, we may use Lemma 2.1 to see that indeed $\tilde{\Gamma} = \Gamma'$ a.s.

(ii) Since the Gibbs kernel $\tilde{\Gamma}$ is restricted to the set $\{\bar{\mu}(\operatorname{supp}\xi) = 0\}$, the Papangelou kernel $\tilde{\eta}_n$ is restricted to $\{(\operatorname{supp}\xi)^c \times T\}^n$. The assertion now follows from Theorem 8.2, along with the fact that, for any $B \in \hat{\mathcal{S}}$, the restrictions $1_{B\times T}\tilde{\xi}$ and $1_{B^c\times T}\tilde{\xi}$ are conditionally independent, given ξ. □

8.3 Regularity Conditions

Much of the previous theory simplifies under suitable regularity conditions on the underlying point processes. Given a point process ξ on S, we write $\tilde{\xi}$ for a uniform randomization of ξ, regarded as a simple point process on $S \times [0,1]$. The basic conditions may then be stated as

(Σ): $P(\xi B = 0 \,|\, 1_{B^c}\xi) > 0$ a.s., $\quad B \in \hat{\mathcal{S}}$,

(Σ_1): $P(\xi B = 0 \,|\, 1_{B^c}\xi) > 0$ a.s. on $\{\xi B = 1\}$, $\quad B \in \hat{\mathcal{S}}$,

$(\tilde{\Sigma})$: $P(\tilde{\xi} B = 0 \,|\, 1_{B^c}\tilde{\xi}) > 0$ a.s., $\quad B \in \hat{\mathcal{S}} \otimes \mathcal{B}_{[0,1]}$.

To compare those and related conditions, we need the following basic consistency relation, which will also play a fundamental role in the sequel.

Lemma 8.10 *(consistency, Papangelou) Let ξ be a K-marked point process on S. Then for any $B \subset C$ in $\hat{\mathcal{S}}$, we have*

$$P(\,\cdot\,|\,1_{B^c}\xi) = \frac{P(\,\cdot\,,\,\xi B = \xi C\,|\,1_{C^c}\xi)}{P(\,\xi B = \xi C\,|\,1_{C^c}\xi)} \;\; \textit{a.s. on } \{\xi B = \xi C\}.$$

Proof: Write $\xi_+ = 1_{B^c}\xi$, $\xi_- = 1_{C^c}\xi$, and $N = \{\xi_B = \xi_C\}$. Since N is ξ_+-measurable, we get for any $A \in \mathcal{A}$

$$\begin{aligned} E\Big\{P(A \cap N|\,\xi_-);\, N\Big\} &= E\Big\{P(N|\,\xi_-);\, A \cap N\Big\} \\ &= E\Big\{P(A|\,\xi_+)\,P(N|\,\xi_-);\, N\Big\}. \end{aligned}$$

Replacing A by $A \cap \{\xi_+ \in M\}$ for measurable $M \subset \mathcal{N}_{S\times K}$, and noting that $\xi_+ = \xi_-$ on N, we obtain

$$E\Big\{P(A \cap N|\,\xi_-);\, \xi_+ \in M,\, N\Big\} = E\Big\{P(A|\,\xi_+)\,P(N|\,\xi_-);\, \xi_+ \in M,\, N\Big\},$$

which implies

$$P(A \cap N|\,\xi_-) = P(A|\,\xi_+)\,P(N|\,\xi_-) \;\text{ a.s. on } N.$$

The assertion now follows, since $P(N|\,\xi_-) > 0$ a.s. on N. □

We may now compare (Σ) and $(\tilde{\Sigma})$, for the mentioned processes. Here we may also identify a general point process $\xi = \sum_n \beta_n \delta_{\sigma_n}$ on S, with the marked point process $\xi' = \sum_n \delta_{\sigma_n,\beta_n}$ on $S \times \mathsf{N}$.

Lemma 8.11 *(regularity comparison) For any point process ξ on S,*

(i) $(\tilde{\Sigma}) \Rightarrow (\Sigma)$, *with equivalence when ξ is simple,*

(ii) (Σ) *holds simultaneously for ξ and ξ',*

(iii) $(\Sigma) \Rightarrow (\Sigma_1)$, *with equivalence when ξ is simple.*

At this point, we prove only (i)−(ii), whereas (iii) will be established in connection with Theorem 8.13 below.

Partial proof: (i) Writing $\bar{B} = B \times [0,1]$ for any $B \in \hat{\mathcal{S}}$, we get by conditional independence

$$P(\xi B = 0\,|\,1_{B^c}\xi) = P(\tilde{\xi}\bar{B} = 0\,|\,1_{\bar{B}^c}\tilde{\xi}), \quad B \in \hat{\mathcal{S}}, \tag{9}$$

which shows that $(\tilde{\Sigma}) \Rightarrow (\Sigma)$. Next, Lemma 8.10 yields for any $B, C \in \hat{\mathcal{S}}$

$$\Big\{P(\xi B = 0\,|\,1_{B^c}\xi) > 0\Big\} = \Big\{P(\xi C = 0\,|\,1_{C^c}\xi) > 0\Big\}, \\ \text{a.s. on } \{\xi(B\Delta C) = 0\}. \tag{10}$$

Now let ξ be a simple point process on S satisfying (Σ), fix any $B \in \hat{\mathcal{S}} \otimes \mathcal{B}_{[0,1]}$, and let $\mathcal{U} \subset \hat{\mathcal{S}}$ be a countable dissecting ring in S. Since $\tilde{\xi}B < \infty$, the dissecting property gives a.s. $\tilde{\xi}(B\Delta\bar{U}) = 0$ for some $U \in \mathcal{U}$. Applying (10) to $\tilde{\xi}$, and using (9) and (Σ), we get a.s.

$$\begin{aligned}\Big\{P(\tilde{\xi}B = 0 \,|\, 1_{B^c}\tilde{\xi}) > 0\Big\} &= \bigcup\nolimits_U \Big\{P(\tilde{\xi}B = 0 \,|\, 1_{B^c}\tilde{\xi}) > 0,\ \tilde{\xi}(B\Delta\bar{U}) = 0\Big\} \\ &= \bigcup\nolimits_U \Big\{P(\tilde{\xi}\bar{U} = 0 \,|\, 1_{\bar{U}^c}\tilde{\xi}) > 0,\ \tilde{\xi}(B\Delta\bar{U}) = 0\Big\} \\ &= \bigcup\nolimits_U \Big\{P(\xi U = 0 \,|\, 1_{U^c}\xi) > 0,\ \tilde{\xi}(B\Delta\bar{U}) = 0\Big\} \\ &= \bigcup\nolimits_U \Big\{\tilde{\xi}(B\Delta\bar{U}) = 0\Big\} = \Omega,\end{aligned}$$

which proves $(\tilde{\Sigma})$.

(ii) Proceed as in (i), with $\tilde{\xi}$ replaced by ξ'. □

Some further comparison is provided by the following lemma. Given a K-marked point process ξ on S with locally finite projection $\bar{\xi}$, we write (Σ') and (Σ'_1) for conditions (Σ) and (Σ_1), with B restricted to the class $\hat{\mathcal{S}} \times K$ of measurable cylinder sets.

Lemma 8.12 *(cylindrical conditions) Let ξ be a K-marked point process on S, with locally finite S-projection $\bar{\xi}$. Then (Σ) and (Σ_1) for ξ are equivalent to the cylindrical conditions (Σ') and (Σ'_1).*

Proof: Clearly $(\Sigma) \Rightarrow (\Sigma')$. Now assume (Σ'), let $B \in \hat{\mathcal{S}} \otimes \mathcal{K}$ be arbitrary, and fix a dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}$. For any $\mu \in \mathcal{N}_{S\times K}$, we may choose $U \in \mathcal{U}$ with $\mu\{B\Delta(U \times K)\} = 0$. Then (Σ') yields $P(\bar{\xi}U = 0 \,|\, 1_{U^c}\xi) > 0$ a.s. for every U, and so by Lemma 8.10

$$P(\xi'B = 0 \,|\, 1_{B^c}\xi') > 0 \ \text{ a.s. on } \Big\{\xi\{B\Delta(U \times K)\} = 0\Big\}, \quad U \in \mathcal{U}.$$

Here (Σ) follows, as we take the countable union over all $U \in \mathcal{U}$. The proof of the equivalence $(\Sigma_1) \Leftrightarrow (\Sigma'_1)$ is similar. □

The significance of (Σ) and $(\tilde{\Sigma})$ is clear from the following basic criteria. Further criteria and consequences will be stated throughout the remainder of the chapter.

Theorem 8.13 *(regularity criteria) Let ξ be a point process on S, and fix any $p = 1 - q \in (0, 1)$. Then (Σ) is equivalent to*

(i) $$E\int \Gamma(d\mu)\, f(\xi + \mu)\, 1\{\xi B = \mu B^c = 0\} = Ef(\xi), \quad B \in \hat{\mathcal{S}},\ f \geq 0,$$

whereas $(\tilde{\Sigma})$ is equivalent to each of the conditions

(ii) $$E\int \Gamma(d\mu)\, f(\xi + \mu)\, p^{\xi B} q^{\mu B} 1\{\mu B^c = 0\} = Ef(\xi), \quad B \in \hat{\mathcal{S}},\ f \geq 0,$$

(iii) $C_1(S \times \cdot\,) \ll \mathcal{L}(\xi)$,

(iv) $C(\hat{\mathcal{N}}_S \times \cdot\,) \ll \mathcal{L}(\xi)$.

Yet another lemma will be needed for the main proofs.

Lemma 8.14 *(absolute continuity)*

(i) *For any σ-finite measures μ, ν, and λ,*
$$\mu \ll \nu \quad \Rightarrow \quad \mu \otimes \lambda \ll \nu \otimes \lambda,$$

(ii) *For μ and ν as above, and for measurable functions f,*
$$\mu \ll \nu \quad \Rightarrow \quad \mu \circ f^{-1} \ll \nu \circ f^{-1}.$$

(iii) *When $\tilde{\xi}$ and $\tilde{\eta}$ are uniform randomizations of some point processes ξ and η, we have*
$$\mathcal{L}(\xi) \ll \mathcal{L}(\eta) \quad \Leftrightarrow \quad \mathcal{L}(\tilde{\xi}) \ll \mathcal{L}(\tilde{\eta}).$$

(iv) *Conditions* (iii)−(iv) *of Theorem 8.13 hold simultaneously for ξ and $\tilde{\xi}$.*

Proof: (i) If $\mu \ll \nu$, then $\mu = g \cdot \nu$ for some measurable function $g \geq 0$, and so $\mu \otimes \lambda = (g \otimes 1) \cdot (\nu \otimes \lambda)$, which implies $\mu \otimes \lambda \ll \nu \otimes \lambda$.

(ii) Assume $\mu \ll \nu$, and let B be such that $(\nu \circ f^{-1})B = 0$. Then $\nu(f^{-1}B) = 0$, and so $(\mu \circ f^{-1})B = \mu(f^{-1}B) = 0$, which shows that $\mu \circ f^{-1} \ll \nu \circ f^{-1}$.

(iii) When ξ is a point process on S, and $\vartheta \perp\!\!\!\perp \xi$ is $U(0,1)$, we may use Lemma 1.6 to construct a uniform randomization $\tilde{\xi}$ on $S \times [0,1]$, as a measurable function of (ξ, ϑ). Conversely, the projection $\tilde{\xi} \mapsto \xi$ is trivially measurable. The assertion now follows by combination with (i) and (ii).

(iv) Choosing a measurable enumeration σ_j, $j \in J$, of the points of ξ, and introducing some independent $U(0,1)$ random variables ϑ_j, we may define
$$\xi = \sum_{j \in J} \delta_{\sigma_j}, \quad \xi_I = \sum_{j \in J \setminus I} \delta_{\sigma_j}, \quad \tilde{\xi} = \sum_{j \in J} \delta_{\sigma_j, \vartheta_j}, \quad \tilde{\xi}_I = \sum_{j \in J \setminus I} \delta_{\sigma_j, \vartheta_j},$$
where $I \subset J$ is arbitrary. By Fubini's theorem,
$$\begin{aligned} \mathcal{L}(\tilde{\xi})f &= Ef(\tilde{\xi}) = EE\Big\{f(\tilde{\xi})\Big|\,\xi\Big\} \\ &= E\{Ef(\tilde{\nu})\}_{\nu=\xi}. \end{aligned}$$
Next, we note that
$$C(\hat{\mathcal{N}}_S \times \cdot)f = E\sum_{\mu \leq \xi} f(\xi - \mu) = E\sum_{I \subset J} f(\xi_I),$$
where the summations extend over all bounded point measures $\mu \leq \xi$, respectively over all finite index sets $I \subset J$. Writing $\tilde{C}$ for the compound Campbell measure of $\tilde{\xi}$, we further note that
$$\begin{aligned} \tilde{C}\Big(\hat{\mathcal{N}}_{S \times [0,1]} \times \cdot\Big)f &= E\sum_{\mu \leq \tilde{\xi}} f(\tilde{\xi} - \mu) = E\sum_{I \subset J} f(\tilde{\xi}_I) \\ &= E\sum_{I \subset J} E\Big\{f(\tilde{\xi}_I)\Big|\,\xi_I\Big\} \\ &= E\sum_{I \subset J} \{Ef(\tilde{\nu})\}_{\nu=\xi_I}. \end{aligned}$$

The assertion for C and $\tilde{C}$ now follows, by an s-finite version of (iii). The proof for C_1 and $\tilde{C}_1$ is similar. □

Proofs of Theorem 8.13 and Lemma 8.11 (iii), (i) ⇔ (Σ): Comparing with Theorem 8.2 (i), we see that (i) is equivalent to

$$E\Big\{f(\xi);\, P(\xi B = 0 \,|\, 1_{B_c}\xi) > 0\Big\} = Ef(\xi), \quad B \in \hat{\mathcal{S}},\ f \geq 0.$$

Taking $f = 1$, we see that this is equivalent to (Σ).

(ii) ⇔ (iv): For any $p \in (0,1)$, $B \in \hat{\mathcal{S}}$, and $f \geq 0$, we have

$$\begin{aligned} E\int \Gamma(d\mu)\, f(\xi+\mu)\, p^{\xi B} q^{\mu B} 1\{\mu B^c = 0\} \qquad & \\ \leq \iint C(d\mu\, d\nu)\, f(\mu+\nu)\, p^{\nu B} q^{\mu B} 1\{\mu B^c = 0\} & \\ = E\sum_{\mu \leq \xi} f(\xi)\, p^{\xi B - \mu B} q^{\mu B} 1\{\mu B^c = 0\} & \\ = E\, f(\xi) \sum_{\mu \leq 1_B \xi} p^{\xi B - \mu B} q^{\mu B} & \\ = E\, f(\xi) \sum_{k \leq \xi B} \binom{\xi B}{k} p^{\xi B - k} q^k & \\ = E\, f(\xi)\, (p+q)^{\xi B} = Ef(\xi), & \end{aligned}$$

with equality for all B and f, iff

$$C(\{\mu B^c = 0\} \times \cdot\,) \ll \mathcal{L}(\xi), \quad B \in \hat{\mathcal{S}}.$$

Summing over a localizing sequence $B_n \uparrow S$ in $\hat{\mathcal{S}}$, we see that the latter condition is equivalent to $C(\hat{\mathcal{N}}_S \times \cdot) \ll \mathcal{L}(\xi)$.

It remains to prove that ($\tilde{\Sigma}$) ⇔ (iii) ⇔ (iv). By Lemma 8.14, it is equivalent to replace (iii) and (iv) by the corresponding conditions for $\tilde{\xi}$. It is then enough to prove the equivalences (Σ) ⇔ (iii) ⇔ (iv), for any simple point process ξ. To complete the proof of Lemma 8.11 as well, we may show that

$$(\Sigma) \Rightarrow (\Sigma_1) \Rightarrow (\text{iii}) \Rightarrow (\Sigma) \Rightarrow (\text{iv}) \Rightarrow (\text{iii}).$$

The implications (Σ) ⇒ (Σ_1) and (iv) ⇒ (iii) being trivial, it remains to show that (Σ_1) ⇒ (iii) ⇒ (Σ) ⇒ (iv).

(Σ_1) ⇒ (iii): Assume (Σ_1), and let $M \subset \mathcal{N}_S$ be measurable with $P\{\xi \in M\} = 0$. First let $M \subset \{\mu B = 0\}$ for some $B \in \hat{\mathcal{S}}$. Then by (Σ_1),

$$\begin{aligned} C_1(B \times M) \;&=\; E\int_B 1\Big\{(\xi - \delta_s)B = 0,\, 1_{B^c}\xi \in M\Big\}\, \xi(ds) \\ &=\; P\{\xi B = 1,\, 1_{B^c}\xi \in M\} \\ &\ll\; E\Big\{P(\xi B = 0 \,|\, 1_{B^c}\xi);\, 1_{B^c}\xi \in M\Big\} \\ &=\; P\{\xi B = 0,\, 1_{B^c}\xi \in M\} \\ &\leq\; P\{\xi \in M\} = 0, \end{aligned}$$

which shows that $C_1(B \times M) = 0$.

Next assume $M \subset \{\mu B < m\}$. Given a dissection system (B_{nj}) in B, we get by the previous case and dominated convergence

$$\begin{aligned}
C_1(B \times M) &= \sum\nolimits_j C_1(B_{nj} \times M) \\
&= \sum\nolimits_j C_1\Big\{B_{nj} \times (M \cap \{\mu B_{nj} > 0\})\Big\} \\
&= \sum\nolimits_j E \int_{B_{nj}} 1\Big\{\xi - \delta_s \in M,\, \xi B_{nj} > 1\Big\}\, \xi(ds) \\
&\le \sum\nolimits_j E \int_{B_{nj}} 1\Big\{\xi B \le m,\, \xi B_{nj} > 1\Big\}\, \xi(ds) \\
&\le m\, E 1\{\xi B \le m\} \sum\nolimits_j 1\{\xi B_{nj} > 1\} \to 0,
\end{aligned}$$

and so again $C_1(B \times M) = 0$. The result extends by monotone convergence, first to general M, and then to $B = S$ for general M, proving (iii).

(iii) $\Rightarrow$ (Σ): Assume (iii). Then for any $n \in \mathsf{N}$, $B \in \hat{\mathcal{S}}$, and measurable $M \subset \mathcal{N}_S$,

$$\begin{aligned}
n\, P\Big\{\xi B = n,\, 1_{B^c}\xi \in M\Big\} &= E\Big(\xi B;\, \xi B = n,\, 1_{B^c}\xi \in M\Big) \\
&= C_1\Big(B \times \Big\{\mu B = n-1,\, 1_{B^c}\mu \in M\Big\}\Big) \\
&\ll P\Big\{\xi B = n-1,\, 1_{B^c}\xi \in M\Big\}.
\end{aligned}$$

Iterating this and summing over n, we obtain

$$P\{1_{B^c}\xi \in M\} \ll P\Big\{\xi B = 0,\, 1_{B^c}\xi \in M\Big\}.$$

In particular, we get for $M = \{\mu;\, P(\xi B = 0 \,|\, 1_{B_c}\xi)_\mu = 0\}$

$$\begin{aligned}
&P\Big\{P(\xi B = 0 \,|\, 1_{B^c}\xi) = 0\Big\} \\
&\qquad \ll P\Big\{\xi B = P(\xi B = 0 \,|\, 1_{B^c}\xi) = 0\Big\} \\
&\qquad = E\Big\{P(\xi B = 0 \,|\, 1_{B^c}\xi);\, P(\xi B = 0 \,|\, 1_{B^c}\xi) = 0\Big\} = 0,
\end{aligned}$$

and so $P(\xi B = 0 \,|\, 1_{B^c}\xi) > 0$ a.s., which proves (Σ).

(Σ) $\Rightarrow$ (iv): Fix a dissecting ring $\mathcal{U} \subset \hat{\mathcal{S}}$. Assuming ($\Sigma$), we get for measurable $M \subset \mathcal{N}_S$ with $P\{\xi \in M\} = 0$

$$\begin{aligned}
C(M \times \mathcal{N}_S) &= E \sum_{\mu \le \xi} 1\{\xi - \mu \in M\} \\
&\le E \sum_{\mu \le \xi} \sum_{B \in \mathcal{U}} 1\Big\{\xi - \mu \in M,\, (\xi - \mu)B = \mu B^c = 0\Big\} \\
&= \sum_{B \in \mathcal{U}} E\, 1\{1_{B^c}\xi \in M\} \sum_{\mu \le \xi} 1\Big\{(\xi - \mu)B = \mu B^c = 0\Big\} \\
&= \sum_{B \in \mathcal{U}} P\{1_{B^c}\xi \in M\}
\end{aligned}$$

$$\begin{aligned}
&\ll \sum_{B\in\mathcal{U}} E\Big\{P(\xi B = 0 \,|\, 1_{B^c}\xi);\, 1_{B^c}\xi \in M\Big\} \\
&= \sum_{B\in\mathcal{U}} P\Big\{\xi B = 0,\, 1_{B^c}\xi \in M\Big\} \\
&\leq \sum_{B\in\mathcal{U}} P\{\xi \in M\} = 0,
\end{aligned}$$

which proves (iv). □

For bounded ξ, we may also express (Σ) in terms of the Palm measures of ξ, as suggested by Theorem 8.5.

Corollary 8.15 *(Palm criterion) Assume $\xi S < \infty$ a.s., and let Q be a Palm kernel of ξ, with associated supporting measure ν on $\hat{\mathcal{N}}_S$. Then (Σ) holds iff $Q_\mu\{0\} > 0$ for $\mu \in \hat{\mathcal{N}}_S$ a.e. ν.*

Proof: Since $\xi S < \infty$ a.s., the measure C is symmetric on $\mathcal{N}_S^2$, and so for any $M \in \mathcal{B}_{\mathcal{N}_S}$,

$$\begin{aligned}
P\{\xi \in M\} &= C(\{0\} \times M) = C(M \times \{0\}) \\
&= \int_M \nu(d\mu)\, Q_\mu\{0\},
\end{aligned}$$

which shows that $\mathcal{L}(\xi) = q \cdot \nu$ with $q(\mu) = Q_\mu\{0\}$. Hence, $q > 0$ a.e. ν iff $\nu \sim \mathcal{L}(\xi)$, which is equivalent to (Σ) by Theorem 8.13. □

8.4 Recursion and Symmetry

Under (Σ), we show that the Papangelou kernels g_n of different orders may be constructed recursively from g_1. Even the Gibbs kernel G is then determined by g_1.

Theorem 8.16 *(kernel recursion) Let ξ be a point process on S with Papangelou kernels $g_1, g_2, \ldots$. Then*

$$g_{m+n}(\xi, ds\, dt) \geq g_m(\xi, ds)\, g_n\Big(\xi + \textstyle\sum_i \delta_{s_i}, dt\Big) \;\; a.s., \quad m, n \in \mathsf{N},$$

with equality when ξ is simple and satisfies (Σ).

In terms of the kernels G_n, the asserted relations become

$$\binom{m+n}{m} \int G_{m+n}(\xi, d\mu)\, f(\mu) \geq \int G_m(\xi, d\mu) \int G_n(\xi + \mu, d\nu)\, f(\mu + \nu),$$

for any measurable function $f \geq 0$ on $\hat{\mathcal{N}}_S$.

Proof: Let $m, n \in \mathsf{N}$, and consider a measurable function $f \geq 0$ on $S^{m+n} \times \mathcal{N}_S$. Using Lemma 1.12 (iii), the definition of C_n, relation (3), and Fubini's theorem, we get

$$\begin{aligned} C_{m+n} f &= E \iint \xi^{(m+n)}(ds\, dt)\, f\Big(s, t, \xi - \textstyle\sum_i \delta_{s_i} - \sum_k \delta_{t_k}\Big) \\ &= E \int \xi^{(n)}(dt) \int \Big(\xi - \textstyle\sum_i \delta_{t_i}\Big)^{(m)}(ds)\, f\Big(s, t, \xi - \sum_i \delta_{s_i} - \sum_k \delta_{t_k}\Big) \\ &= \iint C_n(dt\, d\mu) \int \mu^{(m)}(ds)\, f\Big(s, t, \mu - \textstyle\sum_i \delta_{s_i}\Big) \\ &\geq E \int g_n(\xi, dt) \int \xi^{(m)}(ds)\, f\Big(s, t, \xi - \textstyle\sum_i \delta_{s_i}\Big) \\ &= E \int \xi^{(m)}(ds) \int g_n(\xi, dt)\, f\Big(s, t, \xi - \textstyle\sum_i \delta_{s_i}\Big) \\ &= \iint C_m(ds\, d\mu) \int g_n\Big(\mu + \textstyle\sum_i \delta_{s_i}, dt\Big)\, f(s, t, \mu) \\ &\geq E \int g_m(\xi, ds) \int g_n\Big(\xi + \textstyle\sum_i \delta_{s_i}, dt\Big)\, f(s, t, \xi). \end{aligned}$$

Hence, the kernel on the right provides a partial disintegration of C_{m+n}, and the assertion follows by the maximality of g_{m+n}. Under (Σ), all relations become equalities, by Theorem 8.13. $\square$

We turn to the special case where all Papangelou kernels are absolutely continuous with respect to a fixed measure. The present result also provides some insight into the general case.

Theorem 8.17 *(density recursion) Let ξ be a point process on S with Papangelou kernels $g_1, g_2, \ldots$, such that*

$$g_n(\xi, \cdot) = r_n(\xi, \cdot) \cdot \lambda^{\otimes n} \quad a.s., \quad n \in \mathsf{N},$$

for a fixed measure $\lambda \in \mathcal{M}_S$ and some measurable functions r_n. Then for any $m, n \in \mathsf{N}$, and for $(s, t) \in S^{m+n}$ a.e. $\lambda^{\otimes(m+n)}$, we have a.s.

$$r_{m+n}(\xi, s, t)\, 1\Big\{r_m(\xi, s) > 0\Big\} = r_m(\xi, s)\, r_n\Big(\xi + \textstyle\sum_i \delta_{s_i}, t\Big).$$

When ξ is simple and subject to (Σ), we may omit the second factor on the left.

Proof: Randomizing as in Lemma 8.9, we may replace g_n by its product with n-dimensional Lebesgue measure $\lambda_1^{\otimes n}$, so that the hypotheses remain valid with λ replaced by $\lambda \otimes \lambda_1$. We may then take ξ to be a simple point process on S, and let λ be diffuse, so that the points $s_1, \ldots, s_m$ and $t_1, \ldots, t_n$ can be taken to be distinct. We may further restrict the latter to a fixed set $C \in \hat{\mathcal{S}}$ with $P(\xi C = 0 \,|\, 1_{C^c} \xi) > 0$, and by Lemma 8.8 we may even choose $C = S$, so that $P\{\xi S = 0\} > 0$. We may finally restrict attention to points $s_i \in B^c$ and $t_j \in B$, for some fixed set $B \in \mathcal{S}$.

Now define

$$A = \Big\{\mu = \textstyle\sum_{i\le m}\delta_{s_i};\ s\in (B^c)^m,\ r_m(0,s)>0\Big\}.$$

Writing Π_λ for the distribution of a Poisson process with intensity λ, we get on A

$$G(0; 1_{B^c}\mu\in\cdot) \ll \Pi_{1_{B^c}\lambda} \ll \Pi_\lambda \ll G(0,\cdot).$$

Letting $M\subset\{\mu B^c=0\}$ and $M'\subset A$ be measurable, we get

$$\begin{aligned}\frac{G\big(0;\ 1_B\mu\in M,\ 1_{B^c}\mu\in M'\big)}{G(0,M')} &= \frac{P\big\{1_B\xi\in M,\ 1_{B^c}\xi\in M'\big\}}{P\{\xi\in M'\}}\\ &= \frac{E\big\{P(1_B\xi\in M\,|\,1_{B^c}\xi);\ 1_{B^c}\xi\in M'\big\}}{E\big\{P(1_B\xi=0\,|\,1_{B^c}\xi);\ 1_{B^c}\xi\in M'\big\}}.\end{aligned}$$

Using the Radon-Nikodym theorem, along with the chain rule for densities $f\cdot\mu=(f/g)\cdot(g\cdot\mu)$, valid when $g>0$, we obtain for $\nu\in A$ a.e. $\mathcal{L}(\xi)$

$$\frac{G\big(0;\ 1_B\mu\in M,\ 1_{B^c}\mu\in d\nu\big)}{G(0,d\nu)} = \frac{P(1_B\xi\in M\,|\,1_{B^c}\xi)_\nu}{P(\xi B=0\,|\,1_{B^c}\xi)_\nu} = G(\nu,M),$$

since $\nu B=0$, by the choice of ν. To justify the last computation, we note that $P(\xi B=0\,|\,1_{B^c}\xi)>0$, a.s. on $\{\xi\in A\}\subset\{\xi B=0\}$.

The last equation may be written in integrated form as

$$G\big(0;\ 1_{B^c}\mu\in M',\ 1_B\mu\in M\big) = \int_{M'} G(0,d\nu)\,G(\nu,M),\quad M'\subset A.$$

Since $G(0,d\nu)=0$ for $\nu\in A^c$, this remains true for arbitrary $M'\subset\{\mu B=0\}$, with the set M' on the left replaced by $M'\cap A$. Using the symmetry of the kernels g_n, we conclude that, for any $s\in(B^c)^m$ and $t\in B^n$,

$$g_{m+n}(0;\ ds\,dt)\,1\{r_m(0,s)>0\} = g_m(0,ds)\,g_n\big(\textstyle\sum_i\delta_{s_i},dt\big),$$

which yields a similar relation between suitable versions of the corresponding densities r_n. □

Let Π_λ be the distribution of a Poisson process with intensity $\lambda\in\mathcal{M}_S$. Under (Σ), we show that the absolute continuity of the Papangelou kernel is equivalent to a suitable Poisson domination, for the external conditional distributions.

Corollary 8.18 *(Poisson domination, Matthes et al.) Let ξ be a simple point process on S satisfying (Σ), write η for the Papangelou kernel of ξ, and fix any $\lambda\in\mathcal{M}_S$. Then $\eta\ll\lambda$ a.s., iff*

$$\mathcal{L}(\,1_B\xi\,|\,1_{B^c}\xi\,)\ll\Pi_{1_B\lambda}\ \ a.s.,\quad B\in\hat{\mathcal{S}}. \tag{11}$$

Proof: Assuming (11), we get

$$\mathcal{L}\Big(1_B\xi;\, \xi B = 1 \,\Big|\, 1_{B_c}\xi\Big) \ll 1_B\lambda \le \lambda \ \text{a.s.}, \quad B \in \hat{\mathcal{S}},$$

and so, by Corollary 8.3, we have $1_B\eta \ll \lambda$ a.s. on $\{\xi B = 0\}$ for every $B \in \hat{\mathcal{S}}$. Restricting B to a dissection system $\mathcal{I}$, and noting that $(\operatorname{supp}\xi)^c = \bigcup\{I \in \mathcal{I};\, \xi I = 0\}$, we conclude that $\eta \ll \lambda$, a.s. on $(\operatorname{supp}\xi)^c$. Since ξ is simple, η vanishes a.s. on $\operatorname{supp}\xi$ by the same corollary, and so $\eta \ll \lambda$ a.s.

Conversely, suppose that $\eta \ll \lambda$ a.s. Since ξ is simple and satisfies (Σ), Theorem 8.16 gives $\eta_n \ll \lambda^{\otimes n}$, a.s. for all $n \in \mathsf{N}$, and so Theorem 3.4 gives $1_{\mathcal{N}_B}\Gamma \ll \Pi_{1_B\lambda}$, a.s. for every $B \in \hat{\mathcal{S}}$. Hence, Theorem 8.2 yields $\mathcal{L}(1_B\xi \,|\, 1_{B^c}\xi) \ll \Pi_{1_B\lambda}$, a.s. on $\{\xi B = 0\}$. The latter relation, here denoted by A, is clearly $1_{B^c}\xi$-measurable, and so by (Σ),

$$\begin{aligned} PA^c &\ll E\Big\{P(\xi B = 0 \,|\, 1_{B^c}\xi\,);\, A^c\Big\} \\ &= P(A^c;\, \xi B = 0) = 0. \end{aligned}$$

Hence, A extends a.s. to all of Ω, which proves (11). □

We proceed to show that the λ-symmetry of a marked point process ξ on S is essentially equivalent to λ-invariance of the associated Papangelou kernel η. When λ is diffuse and unbounded, ξ is then a Cox process directed by η. For any random measure ξ on S, we introduce the *remote σ-field* $\mathcal{R}_\xi = \mathcal{R}(\xi) = \bigcap_{B\in\hat{\mathcal{S}}} \bar{\sigma}(1_{B^c}\xi)$, where the bar denotes completion.

Theorem 8.19 *(invariance and Cox criteria, Papangelou, OK) Let ξ be a simple point process on $S \times T$ satisfying* (Σ), *write η for the Papangelou kernel of ξ, and fix any diffuse measure $\lambda \in \mathcal{M}_S$. Then these conditions are equivalent:*

(i) *ξ is λ-symmetric with* $\sup_s \xi(\{s\} \times T) \le 1$ *a.s.,*

(ii) *$\eta = \lambda \otimes \zeta$ a.s., for some random measure ζ on T.*

If also $\|\lambda\| = \infty$ or ζ is $\mathcal{R}_\xi$-measurable, then (i) *and* (ii) *imply*

(iii) *ξ is a Cox process directed by η.*

Proof, (i)$\Rightarrow$(ii): Under (i), Lemma 3.36 shows that ξ is either a λ-randomization of some point process β on T or a Cox process directed by $\lambda\otimes\nu$, for some random measure ν on T. In either case, $E(1_B\xi;\, \xi B = 1 \,|\, 1_{B^c}\xi)$ is a.s. λ-invariant, for every bounded, measurable rectangle $B \subset S\times T$, and so Corollary 8.3 shows that $1_B\eta$ is a.s. λ-invariant on $\{\xi B = 0\}$. Restricting B to a dissecting system in $S\times T$, we see that η is a.s. invariant on $(\operatorname{supp}\bar{\xi})^c\times T$, where $\bar{\xi} = \xi(\cdot \times T)$, hence a.s. of the form $\lambda \otimes \zeta$, for some random measure ζ on T. Since $\bar{\eta}$ and λ vanish a.s. on $\operatorname{supp}\bar{\xi}$, we obtain $\eta = \lambda\otimes\zeta$ a.s. on $S\times T$.

(ii)$\Rightarrow$(i): Assuming (ii), we see from Corollary 8.3 that $E(1_B\xi;\, \xi B = 1|$ $1_{B^c}\xi)$ is a.s. λ-invariant on $\{\xi B = 0\}$, for every bounded, measurable rectangle $B \subset S \times T$. Since ξ satisfies (Σ), the last statement extends a.s. to all of Ω, and since λ is diffuse, we obtain $\sup_s \xi(\{s\} \times T) \le 1$ a.s. Next, we see from (ii), (Σ), and Theorem 8.16 that $\eta_n = \lambda^{\otimes n} \otimes \zeta_n$ a.s., for some random measures ζ_n on T^n, $n \in \mathsf{N}$. Hence, $1_B \Gamma$ is a.s. λ-symmetric for every B as above, and so Theorem 8.2 yields the λ-symmetry of $\mathcal{L}(1_B \xi \,|\, 1_{B^c} \xi)$, a.s. on $\{\xi B = 0\}$ for every B. By (Σ), even the latter property extends to all of Ω. Taking expected values, we conclude that $\mathcal{L}(1_B \xi)$ is $1_B \lambda$-symmetric for all $B \in \hat{\mathcal{S}}$, which yields the asserted symmetry of ξ.

(iii) First suppose that $\|\lambda\| = \infty$. Then Theorem 3.37 shows that ξ is a Cox process directed by $\lambda \otimes \nu$, for some random measure ν on T. For every $B \in \hat{\mathcal{S}}$, the measures ζ and ν are both $1_{B^c}\xi$-measurable, the former by Corollary 8.3 and the latter by the law of large numbers. Since (4) holds with equality by Theorem 8.13, we get a.s. for any $B \in \hat{\mathcal{S}}$

$$\begin{aligned} 1_B \lambda \otimes \zeta &= E(1_B \eta \,|\, 1_{B^c} \xi) \\ &= E(1_B \xi \,|\, 1_{B^c} \xi) \\ &= E(1_B \xi \,|\, \nu) = 1_B \lambda \otimes \nu, \end{aligned}$$

which shows that $\nu = \zeta$ a.s.

Next, suppose that ζ is non-random. Then so is η, and hence $\eta_n = \eta^{\otimes n}$ for all $n \in \mathsf{N}$ by Theorem 8.16, which implies

$$P(1_B \xi \,|\, 1_{B^c} \xi) = \frac{1_{\mathcal{N}_B} \Gamma}{\Gamma(\mathcal{N}_B)} = \Pi_{1_B \eta} = \Pi_{1_B \lambda \otimes \zeta} \text{ a.s.}$$

Since the right-hand side is non-random, $1_B \xi$ is then Poisson $1_B \lambda \otimes \zeta$ for every $B \in \hat{\mathcal{S}}$, which shows that ξ is Poisson $\lambda \otimes \zeta$.

More generally, let η be $\mathcal{R}_\xi$-measurable. It is then $1_{B^c}\xi$-measurable for every $B \in \hat{\mathcal{S}}$, and so by Theorem 6.21 and Corollary 8.3, it remains a.s. the Papangelou kernel of ξ, conditionally on ζ. By the result for non-random η, it follows that ξ is conditionally Poisson $\lambda \otimes \zeta$, which means that ξ is Cox and directed by $\lambda \otimes \zeta$. □

We may also establish a corresponding asymptotic result:

Theorem 8.20 *(symmetry and Cox convergence) Let $\eta_1, \eta_2, \ldots$ be Papangelou kernels of some simple point processes $\xi_1, \xi_2, \ldots$ on $S \times T$ satisfying (Σ), where S and T are Polish, and assume that $\eta_n \xrightarrow{v} \lambda \otimes \zeta$ in L^1, where $\lambda \in \mathcal{M}_S$ is diffuse, and ζ is a random measure on T. Then*

(i) *the ξ_n are vaguely tight, and every limiting process ξ satisfies (Σ) and is λ-symmetric with $\sup_s (\xi\{s\} \times T) \le 1$ a.s.,*

(ii) *if ζ is $\bigcap_n \mathcal{R}(\xi_n)$-measurable, we have $\xi_n \xrightarrow{vd} \xi$, where ξ is Cox and directed by $\lambda \otimes \zeta$.*

Proof: (i) Since the ξ_n satisfy (Σ), and the η_n are locally uniformly integrable, we see from Theorem 8.13 and (4) that $E\xi_n = E\eta_n \xrightarrow{v} \lambda \otimes E\zeta$. In particular, the sequence $(E\xi_n)$ is vaguely relatively compact, which implies the two conditions in Theorem 4.2. Then the random sequence (ξ_n) satisfies both conditions in Theorem 4.10, and is therefore vaguely tight. Since even (η_n) is vaguely tight by Theorem 4.2, the same thing is true for the sequence of pairs (ξ_n, η), and so by Theorem 4.2 the latter is vaguely relatively compact in distribution. If $(\xi_n, \eta_n) \xrightarrow{d} (\xi, \eta)$ along a subsequence, we may assume that $\eta = \lambda \otimes \zeta$.

Applying (4) to each ξ_n, we get for any $B \in \hat{\mathcal{S}} \otimes \hat{\mathcal{T}}$ and $M \in \mathcal{B}_{\mathcal{N}_S}$

$$P\Big\{\xi_n B = 1,\, 1_{B^c}\xi_n \in M\Big\} = E\Big(\eta_n B;\, \xi_n B = 0,\, 1_{B^c}\xi_n \in M\Big), \quad n \in \mathsf{N},$$

and so by continuity and uniform integrability,

$$\begin{aligned} P\Big\{\xi B = 1,\, 1_{B^c}\xi \in M\Big\} &= E\Big(\eta B;\, \xi B = 0,\, 1_{B^c}\xi \in M\Big) \\ &= E\Big\{E(\eta B \,|\, \xi);\, \xi B = 0,\, 1_{B^c}\xi \in M\Big\}, \end{aligned}$$

as long as $\xi\partial B = 0$ a.s. and $P\{1_{B^c}\xi \in \partial M\} = 0$. For fixed B, the latter relation extends by a monotone-class argument to arbitrary M, and Lemma 8.10 yields

$$P(\xi B = 1 \,|\, 1_{B^c}\xi) = E(\eta B \,|\, \xi)\, P(\xi B = 0 \,|\, 1_{B^c}\xi).$$

In particular, $P(\xi B = 0 \,|\, 1_{B^c}\xi) > 0$ a.s. on $\{\xi B = 1\}$. If ξ is simple, then ξ satisfies (Σ) by Theorem 8.13, and Corollary 8.3 shows that ξ has Papangelou kernel $E(\eta \,|\, \xi) = \lambda \otimes E(\zeta \,|\, \xi)$. By Theorem 8.19, it follows that ξ is λ-symmetric.

It remains to show that any limiting process ξ is a.s. simple. Then repeat the previous argument for the uniform randomizations $\tilde{\xi}_n$, which have associated Papangelou kernels $\eta_n \otimes \nu$ by Lemma 8.9, where ν denotes Lebesgue measure on $[0,1]$. Noting that (Σ) extends to the processes $\tilde{\xi}_n$, that $\tilde{\xi}_n \xrightarrow{d} \tilde{\xi}$ for a uniform randomization $\tilde{\xi}$ of ξ, and that the random measures $\eta_n \otimes \nu$ remain locally uniformly integrable with $\eta_n \otimes \nu \xrightarrow{d} \lambda \otimes \nu \otimes \beta$, we conclude as before that $\tilde{\xi}$ is $(\lambda \otimes \nu)$-invariant. Since $\tilde{\xi}$ is a.s. simple, the same thing is true for ξ, by the representations in Lemma 3.36.

(ii) If ζ is non-random, then $E(\zeta \,|\, \xi) = \zeta$, and ξ is Poisson $\lambda \otimes \zeta$ by Theorem 8.19. Since the limit is independent of sub-sequence, the convergence $\xi_n \xrightarrow{d} \xi$ extends to the original sequence.

Now let ζ be measurable on $\mathcal{R}_{\xi_n}$ for every n. As before, the η_n then remain Papangelou kernels of ξ_n, conditionally on ζ. By Lemma 1.9, we can choose a countable dissecting semiring $\mathcal{I} \subset (\hat{\mathcal{S}} \otimes \hat{\mathcal{T}})_\eta$, approximated by a countable class $\mathcal{C} \subset \hat{C}_{S\times T}$, where the continuity property of $\mathcal{I}$ remains a.s. valid under conditioning on ζ. Now fix any sub-sequence $N' \subset \mathsf{N}$. The L^1-convergence $\eta_n \xrightarrow{v} \eta$ yields

$$\begin{aligned} EE(\eta f \,|\, \zeta) &= E\eta f < \infty, \\ EE\Big(|\eta_n f - \eta f| \,\Big|\, \zeta\Big) &= E|\,\eta_n f - \eta f| \to 0, \quad f \in \mathcal{C}, \end{aligned}$$

and so for every $f \in \mathcal{C}$, we have $E(\eta f | \zeta) < \infty$ and $E(|\eta_n f - \eta f| \,|\, \zeta) \to 0$ along a further sub-sequence $N'' \subset N'$, outside a fixed P-null set. By Corollary 4.9, it follows that $\eta_n \xrightarrow{v} \eta$ in L^1 along N'', conditionally on ζ, and so, under the same conditioning, the result for non-random ζ yields $\xi_n \xrightarrow{vd} \xi$ along N'', where ξ is Cox and directed by $\lambda \otimes \zeta$. The corresponding unconditional statement follows by dominated convergence, and since N' was arbitrary, the convergence remains valid along N. □

8.5 Local Conditioning and Decomposition

Here we prove a local limit theorem for marked point processes, which will play a crucial role in the next section. Given a random measure ξ on $S \times K$, we introduce the projections

$$\bar{\xi} = \xi(\cdot \times K), \qquad \xi_B = \xi(B \times \cdot), \qquad \xi_s = \xi(\{s\} \times \cdot),$$

so that $\bar{\xi}$ is a random measure on S, whereas ξ_B and ξ_s are random measures on K. Our first aim is to construct the purported limiting random measure.

Lemma 8.21 *(local intensity) For any K-marked point process ξ on S, there exists an a.s. unique random measure π on $S \times K$ with purely atomic projection $\bar{\pi}$, such that a.s. for each $B \in \hat{\mathcal{S}}$,*

$$\pi_s = \frac{E\Big(\xi_s;\, \bar{\xi}B = 1 \Big|\, 1_{B^c}\xi\Big)}{P(\,\xi_s = \xi_B \,|\, 1_{B^c}\xi\,)}, \quad s \in B \text{ with } \xi_s = \xi_B,$$

as long as the denominator is > 0, and otherwise

$$\pi_s = \frac{E\Big\{\xi(ds \times \cdot);\, \bar{\xi}B = 1 \Big|\, 1_{B^c}\xi\Big\}}{E\Big\{\bar{\xi}(ds);\, \bar{\xi}B = 1 \Big|\, 1_{B^c}\xi\Big\}}, \quad s \in B \text{ with } \xi_s = \xi_B.$$

Furthermore, π is a.s. related to the Papangelou kernel η by

$$\pi_s = \frac{\eta_s}{\bar{\eta}_s + 1}, \quad s \notin \operatorname{supp} \bar{\xi}.$$

Proof: By Lemma 8.10, the expressions for π_s are a.s. independent of B. To construct a specific version of π, we may proceed recursively, using only sets belonging to a fixed dissection system on S. The relation $P(\,\xi_s = \xi_B |\, 1_{B^c}\xi\,) = 0$ implies $P(\,\bar{\xi}B = 0 \,|\, 1_{B^c}\xi\,) = 0$, which is a.s. incompatible with $\bar{\xi}B = 0$. Thus, the second case may only occur when $s \in \operatorname{supp} \bar{\xi}$. The existence of a product-measurable version of π_s is then clear from Theorem 1.28. The a.s. uniqueness of the second expression follows from the fact that

$$E\Big(\bar{\xi}A;\, \bar{\xi}B = 1 \Big|\, 1_{B^c}\xi\Big) = P(\sigma_B \in A,\, \xi_{\sigma_B} = \xi_B \,|\, 1_{B^c}\xi) \text{ a.s.},$$

where σ_B is the unique atom site of $\bar{\xi}$ when $\bar{\xi}B = 1$.

Next, Corollary 8.3 yields for any $B \in \hat{\mathcal{S}}$

$$\eta_s = \frac{E\Big(\xi_s;\, \bar{\xi}B = 1\Big|\, 1_{B^c}\xi\Big)}{P\Big(\bar{\xi}B = 0\,\Big|\, 1_{B^c}\xi\Big)}, \quad s \in B, \;\text{ a.s. on } \{\bar{\xi}B = 0\},$$

and the last assertion follows, by combination with the first expression for π_s. This implies $\pi \le \eta$, a.s. on $(\operatorname{supp}\bar{\xi})^c$, which ensures that π is a locally finite, purely atomic random measure on S. □

With the *local intensity* π in place, we may now state the associated local limit theorem. Given a measure μ, we write $\mu \wedge 1$ for the set function $B \mapsto \mu B \wedge 1$. All subsequent limits should be understood in the setwise sense.

Theorem 8.22 *(local limits) Let ξ be a K-marked point process on S with local intensity π, and fix a dissection system $\mathcal{I} \subset \hat{\mathcal{S}}$. Then as $B \downarrow \{s\}$ along $\mathcal{I}$, we have a.s., simultaneously for all $s \in S$,*

(i) $E\Big(\xi_B;\, \bar{\xi}B = 1\Big|\, 1_{B^c}\xi\Big) \to \pi_s$,

(ii) $E(\,\xi_B \wedge 1 \,|\, 1_{B^c}\xi\,) \to \pi_s$,

(iii) $E(\,\xi_B \,|\, 1_{B^c}\xi\,) \to \pi_s$ *when* $E\bar{\xi} \in \mathcal{M}_S$.

In each case, the convergence is uniform over $\mathcal{K}$ whenever $\bar{\pi}_s < 1$.

Proof: Since $\mathcal{I}$ is countable, we may choose versions of the conditional distributions $\mathcal{L}(\xi|1_{B^c}\xi)$ with $B \in \mathcal{I}$, such that the consistency relations in Lemma 8.10 hold identically for sets in $\mathcal{I}$. Then for any $B \subset C$ in $\mathcal{I}$,

$$E\Big(\xi_B;\, \bar{\xi}B = 1\Big|\, 1_{B^c}\xi\Big) = \frac{E\Big(\xi_B;\, \bar{\xi}C = 1\Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_B = \xi_C \,|\, 1_{C^c}\xi\,)} \;\text{ on } \{\xi_B = \xi_C\}.$$

Letting $B \downarrow \{s\}$ for fixed s and C, we get by dominated convergence

$$E\Big(\xi_B;\, \bar{\xi}B = 1\Big|\, 1_{B^c}\xi\Big) \to \frac{E\Big(\xi_s;\, \bar{\xi}C = 1\Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_s = \xi_C \,|\, 1_{C^c}\xi\,)} \;\text{ on } \{\xi_s = \xi_C\},$$

as long as the denominator is > 0. Since

$$\begin{aligned}\Big\| E\Big(\xi_B;\, \bar{\xi}C = 1\Big|\, 1_{C^c}\xi\Big) - E\Big(\xi_s;\, \bar{\xi}C = 1\Big|\, 1_{C^c}\xi\Big)\Big\| \\ \le P(\,\xi_s \ne \xi_B|\, 1_{C^c}\xi\,) \to 0,\end{aligned}$$

the previous convergence is uniform over $\mathcal{K}$, and so (i) holds with π_s as in Lemma 8.21. Since also

$$\begin{aligned}P\Big(\bar{\xi}B > 1\Big|\, 1_{B^c}\xi\Big) &= \frac{P\Big(\bar{\xi}B = \bar{\xi}C > 1\Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_B = \xi_C|\, 1_{C^c}\xi\,)} \\ &\to \frac{P\Big(\bar{\xi}\{s\} = \bar{\xi}C > 1\Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_s = \xi_C|\, 1_{C^c}\xi\,)} = 0,\end{aligned}$$

even (ii) holds uniformly over $\mathcal{K}$. If $E\bar{\xi} \in \mathcal{M}_S$, then by dominated convergence,

$$\begin{aligned} E\Big(\bar{\xi}B;\, \bar{\xi}B > 1\Big|\, 1_{B^c}\xi\Big) &= \frac{E\Big(\bar{\xi}B;\, \bar{\xi}B = \bar{\xi}C > 1\Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_B = \xi_C|\, 1_{C^c}\xi\,)} \\ &\to \frac{E\Big(\bar{\xi}\{s\};\, \bar{\xi}\{s\} = \bar{\xi}C > 1\Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_s = \xi_C|\, 1_{C^c}\xi\,)} = 0, \end{aligned}$$

since $\|\,\bar{\xi}B - \bar{\xi}\{s\}\| \to 0$ as $B \downarrow \{s\}$, by the local finiteness of $\bar{\xi}$ and the dissecting property of $\mathcal{I}$. Thus, even (iii) follows by combination with (i), which completes the proof, when the denominator in Lemma 8.21 is positive.

If the denominator vanishes for some $B = C \in \mathcal{I}$, then $P(\bar{\xi}C = 0\,|\,1_{C^c}\xi) = 0$, which implies $\bar{\xi}C > 0$ a.s. Hence, it suffices to prove (i)−(iii) on the set $\{\bar{\xi}C = 1\}$, as $B \subset C$ decreases along $\mathcal{I}$ toward the unique atom site of $1_C\xi$. Assuming $E\bar{\xi} \in \mathcal{M}_S$, and using Lemmas 6.40, 8.10, and 8.21, along with the condition $P(\xi C = 0\,|\,1_{C^c}\xi) = 0$ a.s., we get a.s. as in Theorem 1.28

$$\begin{aligned} E(\xi_B\,|\,1_{B^c}\xi) &= \frac{E\Big(\xi_B;\, \bar{\xi}B = \bar{\xi}C\,\Big|\,1_{C^c}\xi\Big)}{P\Big(\bar{\xi}B = \bar{\xi}C\Big|\,1_{C^c}\xi\Big)} \\ &= \frac{E\Big(\xi_B;\, \bar{\xi}B = \bar{\xi}C > 0\,\Big|\,1_{C^c}\xi\Big)}{P\Big(\bar{\xi}B = \bar{\xi}C > 0\,\Big|\,1_{C^c}\xi\Big)} \\ &\sim \frac{E\Big(\xi_B;\, \bar{\xi}B = \bar{\xi}C = 1\Big|\,1_{C^c}\xi\Big)}{P\Big(\bar{\xi}B = \bar{\xi}C = 1\Big|\,1_{C^c}\xi\Big)} \\ &\to \frac{E\Big\{\xi(ds \times \cdot);\, \bar{\xi}B = \bar{\xi}C = 1\Big|\,1_{C^c}\xi\Big\}}{P\Big\{\bar{\xi}(ds) = \bar{\xi}C = 1\Big|\,1_{C^c}\xi\Big\}} = \pi_s, \end{aligned}$$

which proves (iii). A similar argument proves (i) and (ii) in this case. □

Given a simple point process $\zeta \le \xi$, we define the associated *partial Campbell measure* $C_{\xi,\zeta}$, for any measurable functions $f \ge 0$ on $S \times \mathcal{N}_S$, by

$$C_{\xi,\zeta}f = E\int \zeta(ds)f(s,\, \xi - \delta_s).$$

Say that ζ is *regular* or *singular* with respect to ξ, if $C_{\xi,\zeta}(S \times \cdot) \ll \mathcal{L}(\xi)$ or $\perp \mathcal{L}(\xi)$, respectively. The definitions for marked point processes are similar.

Theorem 8.23 *(regularity decomposition) Every K-marked point process ξ on S has an a.s. unique decomposition $\xi = \xi_r + \xi_s$, where ξ_r is regular and ξ_s is singular with respect to ξ. For π as in Lemma 8.21, we have a.s.*

$$\xi_r = 1\{\bar{\pi} < 1\} \cdot \xi, \qquad \xi_s = 1\{\bar{\pi} = 1\} \cdot \xi.$$

In particular, (Σ) holds iff a.s. $\bar{\pi}\{s\} < 1$ for all $s \in S$.

Proof: Suppose we can prove that ξ_r is regular and ξ_s is singular. Then for any $\xi' \le \xi$, we may write $\xi' = \xi' \wedge \xi_r + \xi' \wedge \xi_s$, where $\xi' \wedge \xi_r$ is regular and $\xi' \wedge \xi_s$ is singular, and so ξ' is regular iff $\xi' \le \xi_r$ a.s. and singular iff $\xi' \le \xi_s$ a.s. Thus, if $\xi = \xi'_r + \xi'_s$ with ξ'_r regular and ξ'_s singular, we have $\xi'_r \le \xi_r$ and $\xi'_s \le \xi_s$ a.s. By addition, both relations hold with equality, which proves the asserted a.s. uniqueness. The last assertion follows, since (Σ) is equivalent to the regularity of ξ, by Theorem 8.13.

We may assume that ξ is simple, the general case being similar. Fix any $B \in \hat{\mathcal{S}}$. When $\xi B = \xi_r B = 1$, there exists an $s \in B$ with $\xi_r\{s\} = 1$ and $\pi\{s\} < 1$, and so, by Lemma 8.21, we have a.s. on $\{\xi B = \xi_r B = 1\}$

$$\begin{aligned} &P(\xi B = 0 \,|\, 1_{B^c}\xi) \\ &\quad = P\Big(\xi\{s\} = \xi B \,\Big|\, 1_{B^c}\xi\Big) - P\Big(\xi\{s\} = \xi B = 1 \,\Big|\, 1_{B^c}\xi\Big) > 0. \end{aligned}$$

Now for any $M \in \mathcal{B}_{\mathcal{N}_S} \cap \{\mu B = 0\}$ and $\zeta \le \xi$,

$$C_{\xi,\zeta}(B \times M) = P\Big\{\xi B = \zeta B = 1,\, 1_{B^c}\xi \in M\Big\}. \tag{12}$$

Arguing as in the proof of (iii) $\Rightarrow$ (iv) in Theorem 8.13, we get $C_{\xi,\xi_r}(\cdot \times B) \ll \mathcal{L}(\xi)$, which shows that ξ_r is regular.

If ξ_s fails to be singular, there exist some $B \in \hat{\mathcal{S}}$ and $M \in \mathcal{B}_{\mathcal{N}_S}$ with

$$0 \ne C_{\xi,\xi_s}(B \times \cdot) \ll \mathcal{L}(\xi) \text{ on } M. \tag{13}$$

By monotone convergence, we may assume $M \subset \{\mu B < m\}$ for some $m \in \mathsf{N}$. For any dissection (B_{nj}) of B, we have

$$\begin{aligned} 0 &< C_{\xi,\xi_s}(B \times M) \\ &= \textstyle\sum_j C_{\xi,\xi_s}(B_{nj} \times M) \\ &= \textstyle\sum_j C_{\xi,\xi_s}\Big\{B_{nj} \times (M \cap \{\mu B_{nj} = 0\})\Big\} \\ &\quad + \textstyle\sum_j C_{\xi,\xi_s}\Big\{B_{nj} \times (M \cap \{\mu B_{nj} > 0\})\Big\}. \end{aligned}$$

Here the last sum tends to 0 as $n \to \infty$, by the argument for Theorem 8.13, and so the first sum on the right is positive for large enough n, and we may choose some n and j with $C_{\xi,\xi_s}\{B_{nj} \times (M \cap \{\mu B_{nj} = 0\})\} > 0$. Replacing B by B_{nj} and M by $M \cap \{\mu B_{nj} = 0\}$, we may henceforth assume that (13) holds with $M \subset \{\mu B = 0\}$.

Using (12) and (13), we get for any $N \subset M$ in $\mathcal{B}_{\mathcal{N}_S}$

$$\begin{aligned} P\Big\{\xi B = \xi_s B = 1,\, 1_{B^c}\xi \in N\Big\} &= C_{\xi,\xi_s}(B \times N) \\ &\ll P\{\xi \in N\} \\ &= P\Big\{\xi B = 0,\, 1_{B^c}\xi \in N\Big\}, \end{aligned}$$

which implies

$$P(\xi B = \xi_s = 1 \,|\, 1_{B^c}\xi) \ll P(\xi B = 0 \,|\, 1_{B^c}\xi) \text{ a.s. on } \{1_{B^c}\xi \in M\}.$$

On the other hand, Lemma 8.21 yields

$$P(\xi B = 0 \,|\, 1_{B^c}\xi) = 0 \ \text{ a.s. on } \{\xi B = \xi_s B = 1\}.$$

Hence, by combination,

$$\begin{aligned} C_{\xi,\xi_s}(B \times M) &= P\Big\{\xi B = \xi_s B = 1,\, 1_{B^c}\xi \in M\Big\} \\ &\ll E\Big\{P(\xi B = \xi_s B = 1 \,|\, 1_{B^c}\xi);\ \xi B = \xi_s B = 1,\, 1_{B^c}\xi \in M\Big\} \\ &\ll E\Big\{P(\xi B = 0 \,|\, 1_{B^c}\xi);\ \xi B = \xi_s B = 1,\, 1_{B^c}\xi \in M\Big\} = 0, \end{aligned}$$

contradicting (13). This shows that ξ_s is singular. □

The local convergence in Theorem 8.22 extends to more general random measures. The following version will be needed in the next section.

Theorem 8.24 *(local conditioning) For any K-marked point process ξ, and a random measure ζ on S with $E\zeta \in \mathcal{M}_S$, there exists a purely atomic random measure π_ζ on S, such that for any dissection system $\mathcal{I} \subset \hat{\mathcal{S}}$, we have a.s. as $B \downarrow \{s\}$ along $\mathcal{I}$, simultaneously for all $s \in S$,*

$$E(\zeta B \,|\, 1_{B^c}\xi) \to \pi_\zeta\{s\}, \quad s \in S.$$

Furthermore, for any fixed $B \in \hat{\mathcal{S}}$, we have a.s., simultaneously for all $s \in B$ with $\xi_s = \xi_B$,

$$\begin{aligned} \pi_\zeta\{s\} &= \frac{E\Big(\zeta_s;\, \bar\xi\{s\} = \bar\xi B \,\Big|\, 1_{B^c}\xi\Big)}{P(\,\xi_s = \xi_B \,|\, 1_{B^c}\xi\,)}, \qquad \bar\pi_s < 1, \\ \pi_\zeta\{s\} &= \frac{E\Big\{\,\zeta_s\, \bar\xi(ds);\, \bar\xi B = 1 \Big|\, 1_{B^c}\xi\Big\}}{E\Big\{\,\bar\xi(ds);\, \bar\xi B = 1 \Big|\, 1_{B^c}\xi\Big\}}, \qquad \bar\pi_s = 1. \end{aligned}$$

Proof: When $\bar\pi_s < 1$, we get a.s. on $\{\xi_s = \xi_C\}$, as $C \supset B \downarrow \{s\}$

$$\begin{aligned} E(\zeta B \,|\, 1_{B^c}\xi) &= \frac{E\Big(\zeta B;\, \bar\xi B = \bar\xi C \Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_B = \xi_C \,|\, 1_{C^c}\xi\,)} \\ &\to \frac{E\Big(\zeta\{s\};\, \bar\xi\{s\} = \bar\xi C \Big|\, 1_{C^c}\xi\Big)}{P(\,\xi_s = \xi_C \,|\, 1_{C^c}\xi\,)} = \pi_\zeta\{s\}. \end{aligned}$$

If instead $\bar\pi_s = 1$, then $E(\bar\xi B \,|\, 1_{B^c}\xi) = 1$ a.s., and so as $C \supset M \supset B \downarrow \{s\}$,

$$\begin{aligned} E(\zeta B \,|\, 1_{B^c}\xi) &\le E\Big(\zeta M;\, \bar\xi B > 0 \,\Big|\, 1_{B^c}\xi\Big) \\ &= \frac{E\Big(\zeta M;\, \bar\xi B = \bar\xi C > 0 \,\Big|\, 1_{C^c}\xi\Big)}{P\Big(\bar\xi B = \bar\xi C > 0 \,\Big|\, 1_{C^c}\xi\Big)} \\ &\to \frac{E\Big\{\,\zeta M;\, \bar\xi(ds) = \bar\xi C = 1 \Big|\, 1_{C^c}\xi\Big\}}{P\Big\{\,\bar\xi(ds) = \bar\xi C = 1 \Big|\, 1_{C^c}\xi\Big\}}, \end{aligned}$$

by Theorem 1.28 and its proof. Letting $M \downarrow \{s\}$ gives

$$\limsup_{B\downarrow\{s\}} E(\zeta B \mid 1_{B^c}\xi) \le \frac{E\Big\{\zeta\bar{\xi}(ds) = \bar{\xi}C > 0 \,\Big|\, 1_{C^c}\xi\Big\}}{P\Big\{\bar{\xi}(ds) = \bar{\xi}C = 1 \Big|\, 1_{C^c}\xi\Big\}} = \pi_\zeta\{s\}.$$

Similarly, we get as $C \supset B \downarrow \{s\}$

$$\begin{aligned} E(\zeta B \mid 1_{B^c}\xi) \;&\ge\; E\Big(\zeta\bar{\xi}B;\, \bar{\xi}B > 0 \,\Big|\, 1_{B^c}\xi\Big) \\ &= \frac{E\Big(\zeta\bar{\xi}B;\, \bar{\xi}B = \bar{\xi}C > 0 \,\Big|\, 1_{C^c}\xi\Big)}{P\Big(\bar{\xi}B = \bar{\xi}C > 0 \,\Big|\, 1_{C^c}\xi\Big)} \\ &\to \frac{E\Big\{\zeta\bar{\xi}B;\, \bar{\xi}(ds) = \bar{\xi}C = 1 \Big|\, 1_{C^c}\xi\Big\}}{P\Big\{\bar{\xi}(ds) = \bar{\xi}C = 1 \Big|\, 1_{C^c}\xi\Big\}} \to \pi_\zeta\{s\}, \end{aligned}$$

and so

$$\liminf_{B\downarrow\{s\}} E(\,\zeta B \mid 1_{B^c}\xi\,) \ge \pi_\zeta\{s\}.$$

Combining the upper and lower bounds, we obtain the desired convergence and expressions for π_ζ. □

8.6 External Intensity and Projection

Here we prove the convergence of sums of conditional probabilities, of the type considered in Theorem 8.22. Given a random measure η on $S \times K$, we write η_d for the restriction of η to the set where $\bar{\eta}\{s\} = 0$. For any K-marked point process ξ on S, we define the *external intensity* $\hat{\xi}$ as the random measure $\pi + \eta_d$, where π is the local intensity in Lemma 8.21, and η is the Papangelou random measure in Corollary 8.3.

Theorem 8.25 *(global limits, Papangelou, OK) Let ξ be a K-marked point process on S, and define $\hat{\xi} = \pi + \eta_d$. Then for any $B \in \hat{\mathcal{S}}$ with associated dissection system (B_{nj}), we have*

(i) $\sum_j E\Big(\xi_{B_{nj}};\, \bar{\xi}B_{nj} = 1 \Big|\, 1_{B_{nj}^c}\xi\Big) \to \hat{\xi}_B$ *a.s.*,

(ii) $\sum_j E\Big(\xi_{B_{nj}} \wedge 1 \Big|\, 1_{B_{nj}^c}\xi\Big) \to \hat{\xi}_B$ *a.s.*,

(iii) $\sum_j E\Big(\xi_{B_{nj}} \Big|\, 1_{B_{nj}^c}\xi\Big) \to \hat{\xi}_B$ *a.s. and in L^1, when $E\bar{\xi}B < \infty$.*

Under (Σ), *the convergence in* (i)$-$(iii) *is uniform over $\mathcal{K}$.*

Proof: (i) By Theorem 8.22 (i), we have a.s.

$$\begin{aligned} &\sum\nolimits_j \Big\{E\Big(\xi_{B_{nj}};\, \bar{\xi}B_{nj} = 1 \Big|\, 1_{B_{nj}^c}\xi\Big)\,;\, \bar{\xi}B_{nj} > 0\Big\} \\ &\qquad \to \int_B \bar{\xi}(ds)\, \pi_s = \hat{\xi}\,\Big\{(B \cap \operatorname{supp}\bar{\xi}) \times \cdot\Big\}. \end{aligned} \tag{14}$$

Now fix any finite union U of sets B_{nj}, and put $U' = U \setminus (\operatorname{supp} \bar\pi)$. Writing

$$\mathcal{J}_n(U) = \{j;\, B_{nj} \subset U\}, \quad B_n(s) = \bigcup\nolimits_j \{B_{nj};\, s \in B_{nj}\},$$

and using Lemma 8.10 and Theorems 8.3 and 8.22, we get a.s. on $\{\,\bar\xi U = 0\}$

$$\begin{aligned}
&\sum_{j\in\mathcal{J}_n(U)} E\Big(\xi_{B_{nj}};\, \bar\xi B_{nj} = 1\Big|\, 1_{B^c_{nj}}\xi\Big)\\
&\qquad = \sum_{j\in\mathcal{J}_n(U)} \frac{E\Big(\xi_{B_{nj}};\, \bar\xi U = 1\Big|\, 1_{U^c}\xi\Big)}{P\Big(\bar\xi B_{nj} = \bar\xi U\Big|\, 1_{U^c}\xi\Big)}\\
&\qquad = \int_U \frac{E\Big\{\xi(ds\times\cdot);\, \bar\xi U = 1\Big|\, 1_{U^c}\xi\Big\}}{P\Big\{\bar\xi B_n(s) = \bar\xi U\Big|\, 1_{U^c}\xi\Big\}}\\
&\qquad \uparrow \int_U \frac{E\Big\{\xi(ds\times\cdot);\, \bar\xi U = 1\Big|\, 1_{U^c}\xi\Big\}}{P(\,\xi_s = \xi_U \,|\, 1_{U^c}\xi)}\\
&\qquad = \sum_{s\in U} \frac{E\Big(\xi_s;\, \bar\xi U = 1\Big|\, 1_{U^c}\xi\Big)}{P(\,\xi_s = \xi_U \,|\, 1_{U^c}\xi)} + \frac{E\Big(\xi_{U'};\, \bar\xi U = 1\Big|\, 1_{U^c}\xi\Big)}{P(\,\xi_U = 0 \,|\, 1_{U^c}\xi)}\\
&\qquad = \pi_U + \eta_{U'} = \hat\xi_U,
\end{aligned}$$

uniformly over $\mathcal{K}$. Putting $\Delta U_n = U_n \setminus U_{n-1}$, where $U_0 = \emptyset$, and $U_n = \bigcup_j \{B_{nj};\, \bar\xi B_{nj} = 0\}$ for $n > 0$, we get a.s. by monotone convergence

$$\begin{aligned}
&\sum\nolimits_j \Big\{E\Big(\xi_{B_{nj}};\, \bar\xi B_{nj} = 1\Big|\, 1_{B^c_{nj}}\xi\Big);\, \bar\xi B_{nj} = 0\Big\}\\
&\qquad = \sum_{m\le n}\ \sum_{j\in\mathcal{J}_n(\Delta U_m)} E\Big(\xi_{B_{nj}};\, \bar\xi B_{nj} = 1\Big|\, 1_{B^c_{nj}}\xi\Big)\\
&\qquad \uparrow \sum_{m\ge 1} \hat\xi_{\Delta U_m} = \hat\xi\Big\{(B \setminus \operatorname{supp}\bar\xi) \times \cdot\Big\},
\end{aligned} \tag{15}$$

which again holds uniformly over $\mathcal{K}$, since as $1 \le k \to \infty$,

$$\begin{aligned}
\sum_{m\in(k,n]}\ \sum_{j\in\mathcal{J}_n(\Delta U_m)} P\Big(\bar\xi B_{nj} = 1\Big|\, 1_{B^c_{nj}}\xi\Big) &\le \sum_{m>k} \hat\xi(\Delta U_m \times K)\\
&= \hat\xi\Big(\{(B\setminus U_k) \setminus \operatorname{supp}\xi\} \times K\Big) \to 0.
\end{aligned}$$

Combining (14) and (15) yields (i).

(ii) By Theorem 8.22 (i), we have a.s. as $n \to \infty$

$$\sum\nolimits_j \Big\{P\Big(\bar\xi B_{nj} > 1\Big|\, 1_{B^c_{nj}}\xi\Big)\,;\, \bar\xi B_{nj} > 0\Big\} \to 0. \tag{16}$$

Using the linear order on S induced by (B_{nj}), in the sense of Lemma 1.8, we may decompose $B \setminus \operatorname{supp}\xi$ into connected components $I_1, \ldots, I_m$. Then Theorem 8.2 (iii) yields a.s.

$$\sum\nolimits_j \Big\{P\Big(\bar\xi B_{nj} > 1\Big|\, 1_{B^c_{nj}}\xi\Big);\, \bar\xi B_{nj} = 0\Big\}$$

$$\begin{aligned}
&\le \sum\nolimits_j \left\{ \frac{P\left(\bar\xi B_{nj} > 1 \,\middle|\, 1_{B_{nj}^c}\xi\right)}{P\left(\bar\xi B_{nj} = 0 \,\middle|\, 1_{B_{nj}^c}\xi\right)};\ \bar\xi B_{nj} = 0 \right\} \\
&= \sum\nolimits_j \left\{ \Gamma\{\,\bar\mu B_{nj} = \bar\mu S > 1\};\ \bar\xi B_{nj} = 0 \right\} \\
&\le \sum\nolimits_k \Gamma\Big\{ \max_{j\in\mathcal{J}(I_k)} \bar\mu B_{nj} > 1,\ \bar\mu I_k^c = 0 \Big\} \to 0,
\end{aligned}$$

where the convergence holds by dominated convergence as $n \to \infty$, since $\max_j \bar\mu(B_{nj} \vee 1) \to 1$, and $\Gamma\{\bar\mu I_k^c = 0\} < \infty$ a.s. for all k by Lemma 8.4. The assertion now follows by combination with (i) and (16).

(iii) Let $E\bar\xi B < \infty$. By Theorem 8.22 (i)$-$(ii), we have a.s. as $n \to \infty$

$$\sum\nolimits_j \left\{ E\left(\bar B_{nj};\ \bar\xi B_{nj} > 1 \,\middle|\, 1_{B_{nj}^c}\xi\right);\ \bar\xi B_{nj} > 0 \right\} \to 0. \tag{17}$$

Defining $I_1, \dots, I_m$ as before, and using Theorem 8.2 (iii), we next get a.s.

$$\begin{aligned}
&\sum\nolimits_j \left\{ E\left(\bar\xi B_{nj};\ \bar\xi B_{nj} > 1 \,\middle|\, 1_{B_{nj}^c}\xi\right);\ \bar\xi B_{nj} = 0 \right\} \\
&\qquad \le \sum\nolimits_j \left\{ \frac{E\left(\bar\xi B_{nj};\ \bar\xi B_{nj} > 1 \,\middle|\, 1_{B_{nj}^c}\xi\right)}{P\left(\bar\xi B_{nj} = 0 \,\middle|\, 1_{B_{nj}^c}\xi\right)};\ \bar\xi B_{nj} = 0 \right\} \\
&\qquad = \sum\nolimits_j \int \Gamma(d\mu)\, \bar\mu(B_{nj})\, 1\left\{ \bar\mu B_{nj} = \bar\mu S > 1,\ \bar\xi B_{nj} = 0 \right\} \\
&\qquad \le \sum\nolimits_k \int \Gamma(d\mu)\, \bar\mu(I_k)\, 1\Big\{ \max_{j\in\mathcal{J}(I_k)} \bar\mu B_{nj} > 1,\ \bar\mu I_k^c = 0 \Big\} \to 0,
\end{aligned}$$

where the convergence holds by dominated convergence as $n \to \infty$, since $\max_j \bar\mu(B_{nj} \vee 1) \to 1$, and

$$\int \Gamma(d\mu)\,\{\,\bar\mu(I_k);\ \bar\mu I_k^c = 0\,\} < \infty \ \text{ a.s.}, \quad k \ge 1,$$

by Lemma 8.4. The asserted a.s. convergence now follows, by combination with (i) and (17).

Next we note that

$$\begin{aligned}
\sum\nolimits_j \left\{ P\left(\bar\xi B_{nj} = 1 \,\middle|\, 1_{B_{nj}^c}\xi\right);\ \bar\xi B_{nj} > 0 \right\} &\le \sum\nolimits_j 1\{\,\bar\xi B_{nj} > 0\} \\
&\le \sum\nolimits_j \bar\xi B_{nj} = \bar\xi B,
\end{aligned}$$

which shows that the left-hand side is uniformly integrable. Furthermore,

$$\begin{aligned}
E\sum\nolimits_j \left\{ P\left(\bar\xi B_{nj} = 1 \,\middle|\, 1_{B_{nj}^c}\xi\right);\ \bar\xi B_{nj} = 0 \right\} &\le \sum\nolimits_j P\{\,\bar\xi B_{nj} = 1\} \\
&\le E\bar\xi B.
\end{aligned}$$

Since the sums on the left are nondecreasing in n, they too are uniformly integrable. Finally,

$$\begin{aligned}
E\sum\nolimits_j E\left(\bar\xi B_{nj};\ \bar\xi B_{nj} > 1 \,\middle|\, 1_{B_{nj}^c}\xi\right) &= \sum\nolimits_j E\left(\bar\xi B_{nj};\ \bar\xi B_{nj} > 1\right) \\
&\le E\left(\bar\xi B;\ \max\nolimits_j \bar\xi B_{nj} > 1\right) \to 0,
\end{aligned}$$

by dominated convergence. Combining the last three results, we conclude that the sums in (iii) are uniformly integrable, and so the convergence remains valid in L^1. In particular,

$$E\xi_B = E\sum\nolimits_j E\Big(\xi_{B_{nj}}\Big|\, 1_{B_{nj}^c}\xi\Big) \to E\hat{\xi}_B,$$

and so $E\xi_B = E\hat{\xi}_B$. Replacing B by any measurable subset, and applying a monotone-class argument, we obtain $E\xi = E\hat{\xi}$ on B. □

The last result may be extended to any random measure ζ with $E\zeta \in \mathcal{M}_S$, as long as the external conditioning remains defined with respect to a marked point process ξ. For any measures μ and ν on S, let $\mu\nu = \nu\mu$ denote the purely atomic measure $\int \mu\{s\}\,\nu(ds) = \int \nu\{s\}\,\mu(ds)$. Given a random measure ζ, write $E(\zeta\,|\,\xi)$ for a ξ-measurable and measure-valued version of the process $B \mapsto E(\zeta B\,|\,\xi)$, whose existence is guaranteed by Corollary 2.17.

Theorem 8.26 *(external intensity, van der Hoeven, OK) Consider a K-marked point process ξ on S, along with a random measure ζ with $E\zeta \in \mathcal{M}_S$. Then there exists an a.s. unique random measure $\hat{\zeta}$ on S with $E\hat{\zeta} = E\zeta$, such that for any $B \in \hat{\mathcal{S}}$ with associated dissection system (B_{nj}),*

$$\sum\nolimits_j E(\zeta B_{nj}\,|\,1_{B_{nj}^c}\xi) \to \hat{\zeta}B \quad \text{a.s. and in } L^1.$$

Proof: Writing $1_B\bar{\xi} = \sum_k \delta_{\tau_k}$, we get a.s. by Theorem 8.24

$$\sum\nolimits_j \Big\{E\Big(\zeta B_{nj}\,|\,1_{B_{nj}^c}\xi\Big);\, \xi B_{nj} > 0\Big\} \to \sum\nolimits_j \pi_\zeta\{\tau_k\} \equiv \chi_1 B.$$

To prove the a.s. convergence, it is then enough to sum over indices j with $\xi B_{nj} = 0$. Writing $\zeta' = E(\zeta\,|\,\xi)$, and using Lemma 8.10, we get a.s. on $\{\xi B_{nj} = 0\}$

$$\frac{E\Big(\zeta B_{nj};\, \bar{\xi}B_{nj} = 0\,\Big|\, 1_{B_{nj}^c}\xi\Big)}{P\Big(\bar{\xi}B_{nj} = 0\,\Big|\, 1_{B_{nj}^c}\xi\Big)} = E(\zeta B_{nj}\,|\,\xi) = \zeta' B_{nj}.$$

Hence, by Theorem 8.22 and dominated convergence, we have a.s.

$$\begin{aligned}
\sum\nolimits_j &\Big\{E\Big(\zeta B_{nj};\, \bar{\xi}B_{nj} = 0\,\Big|\, 1_{B_{nj}^c}\xi\Big);\, \bar{\xi}B_{nj} = 0\Big\} \\
&= \sum\nolimits_j \Big\{\zeta'(B_{nj})\, P\Big(\bar{\xi}B_{nj} = 0\,\Big|\, 1_{B_{nj}^c}\xi\Big);\, \bar{\xi}B_{nj} = 0\Big\} \\
&= \int_B \zeta'(ds)\, P\Big\{\bar{\xi}B_n(s) = 0\,\Big|\, 1_{B_n^c(s)}\xi\Big\}\, 1\{\bar{\xi}B_n(s) = 0\} \\
&\to \int_B \zeta'(ds)\,(1 - \bar{\pi}_s)(1 - \bar{\xi}_s) \equiv \chi_2 B.
\end{aligned}$$

Since the left-hand side is bounded by the integrable random variable $\tilde{\zeta}B$, the last convergence remains valid in L^1.

It remains to consider the terms $E(\zeta B_{nj};\, \bar\xi B_{nj} > 0 \,|\, 1_{B^c_{nj}}\xi)$ with $\bar\xi B_{nj} = 0$. Writing Γ for the Gibbs kernel of the pair (ξ, ζ), we note that

$$\begin{aligned}
&\textstyle\sum_j \Big\{E\Big(\zeta B_{nj};\, \bar\xi B_{nj} > 1\Big|\, 1_{B^c_{nj}}\xi\Big);\, \bar\xi B_{nj} = 0\Big\}\\
&\qquad \le \textstyle\sum_j \left\{\dfrac{E\Big(\zeta B_{nj};\, \bar\xi B_{nj} > 1\Big|\, 1_{B^c_{nj}}\xi\Big)}{P\Big(\bar\xi B_{nj} = 0\Big|\, 1_{B^c_{nj}}\xi\Big)};\, \bar\xi B_{nj} = 0\right\}\\
&\qquad = \textstyle\sum_j \displaystyle\iint \Gamma(d\nu\, d\mu)\, \bar\nu(B_{nj})\, 1\Big\{\bar\mu B_{nj} = \bar\mu S > 1,\, \bar\xi B_{nj} = 0\Big\}\\
&\qquad \le \textstyle\sum_k \displaystyle\iint \Gamma(d\nu\, d\mu)\, \nu(I_k)\, 1\Big\{\max_{j\in\mathcal{J}(I_k)} \bar\mu B_{nj} > 1,\, \bar\mu I^c_k = 0\Big\},
\end{aligned}$$

which tends a.s. to 0 by dominated convergence, since by Lemma 8.4,

$$\iint \Gamma(d\nu\, d\mu)\, \nu(I_k)\, 1\{\mu I^c_k = 0\} < \infty \ \text{ a.s.}$$

The convergence holds even in L^1, since the expectation on the left is bounded by $E(\zeta B;\, \max_j \bar\xi B_{nj} > 1)$, which tends to 0 by dominated convergence.

Now we need to consider only the terms $E(\zeta B_{nj};\, \bar\xi B_{nj} = 1 \,|\, 1_{B^c_{nj}}\xi)$ with $\bar\xi B_{nj} = 0$. Here we show that ζ can be replaced by $\zeta\bar\xi$. Then note as before that

$$\begin{aligned}
&\textstyle\sum_j \Big(E\Big\{(\zeta - \zeta\xi)B_{nj};\, \bar\xi B_{nj} > 0\Big|\, 1_{B^c_{nj}}\xi\Big\};\, \bar\xi B_{nj} = 0\Big)\\
&\qquad \le \textstyle\sum_k \displaystyle\iint \Gamma(d\nu\, d\mu) \sum_{j\in\mathcal{J}(I_k)} (\nu - \nu\bar\mu)(B_{nj})\, 1\{\bar\mu B_{nj} = \bar\mu S > 0\}\\
&\qquad \le \textstyle\sum_k \displaystyle\iint \Gamma(d\nu\, d\mu) \int_{I_k} (\nu - \nu\bar\mu)(ds)\, 1\Big\{\bar\mu B_n(s) > 0,\, \bar\mu I^c_k = 0\Big\},
\end{aligned}$$

which tends to 0, by Lemma 8.4 and dominated convergence. Indeed, writing $1_{I_k}\bar\mu = \sum_i \delta_{s_{ki}}$ and $B'_n(s) = B_n(s) \setminus \{s\}$, we get by dominated convergence

$$\int_{I_k} (\nu - \nu\bar\mu)(ds)\, 1\{\bar\mu B_n(s) > 0\} \le \textstyle\sum_i \nu B'_n(s_{ki}) \to 0.$$

Assuming $1_B\bar\xi = \sum_i \delta_{\tau_i}$, and using dominated convergence, we see that the expectation on the left is bounded by

$$\textstyle\sum_j E\Big\{(\zeta - \zeta\bar\xi)B_{nj};\, \bar\xi B_{nj} > 0\Big\} = E \sum_i \zeta B'_n(\tau_i) \to 0,$$

which shows that even the last convergence extends to L^1.

We are now left with the terms $E\{\, \zeta\bar\xi B_{nj};\, \bar B_{nj} = 1 \,|\, 1_{B^c_{nj}}\xi\,\}$ with $\bar\xi B_{nj} = 0$. Fixing any finite union U of dissecting sets, we get a.s. on $\{\bar\xi U = 0\}$

$$\begin{aligned}
&\sum_{j\in\mathcal{J}(U)} E\Big(\zeta\bar\xi B_{nj};\, \bar\xi B_{nj} = 1\Big|\, 1_{B^c_{nj}}\xi\Big)\\
&\qquad\qquad = \int_U \frac{E\Big\{\zeta\bar\xi(ds);\, \bar\xi U = 1\Big|\, 1_{U^c}\xi\Big\}}{P\Big\{\bar\xi B_n(s) = \bar\xi U\Big|\, 1_{U^c}\xi\Big\}} \uparrow \chi_3 U,
\end{aligned}$$

where the random measure χ_3 on U is given by

$$\chi_3 B = \int_U \frac{E\Big\{\zeta\bar{\xi}(ds);\, \bar{\xi}U = 1\Big|\, 1_{U^c}\xi\Big\}}{P\Big\{\bar{\xi}\{s\} = \bar{\xi}U\Big|\, 1_{U^c}\xi\Big\}}, \quad B \in \mathcal{S} \cap U.$$

By Lemma 8.10, χ_3 extends a.s. uniquely to a random measure on B with $\chi_3(\operatorname{supp}\xi) = 0$, and so by monotone convergence

$$\textstyle\sum_j \Big\{E\Big(\zeta\bar{\xi}B_{nj};\, \bar{\xi}B_{nj} = 1\Big|\, 1_{B_{nj}^c}\xi\Big);\, \bar{\xi}B_{nj} = 0\Big\} \uparrow \chi_3 B.$$

Since the expected value on the left is bounded by $E\zeta B < \infty$, even the last convergence holds in L^1.

In summary, we have $\sum_j E(\zeta B_{nj} \mid 1_{B_{nj}^c}\xi) \to \hat{\zeta}B$ a.s. with $\hat{\zeta} = \chi_1 + \chi_2 + \chi_3$, where the limit is clearly independent of the choice of dissection system (B_{nj}). To extend the convergence to L^1, we may write the last result as $\gamma_n \to \gamma$ a.s. If $\zeta_m \uparrow \zeta$ for some random measures ζ_m, we write the corresponding convergence as $\gamma_n^m \to \gamma^m$ a.s. By Fatou's lemma,

$$\begin{aligned} E(\gamma - \gamma^m) &\le \liminf_{n\to\infty} E(\gamma_n - \gamma_n^m) \\ &= E(\zeta - \zeta_m)B \to 0, \end{aligned}$$

and so it suffices to show that $E|\gamma_n^m - \gamma^m| \to 0$ as $n \to \infty$, for every fixed m. Choosing $\zeta_m = \zeta \cdot 1\{\zeta B \vee \bar{\xi}B \le m\}$, we get

$$\textstyle\sum_j \Big\{E\Big(\zeta_m B_{nj} \big|\, 1_{B_{nj}^c}\xi\Big);\, \bar{\xi}B_{nj} > 0\Big\} \le m^2,$$

which shows that the left-hand side is uniformly integrable. Since the same property has already been established for the remaining terms, the desired L^1-convergence follows. □

The random measures $\hat{\xi}$ and $\hat{\zeta}$ above may also be regarded as dual projections, in analogy with the compensator of a random measure on R_+. Fixing a K-marked point process ξ on S, we introduce the associated *external σ-field* $\mathcal{X}$ on $S \times \Omega$, generated by the families $B \times \bar{\sigma}(1_{B^c}\xi)$ with $B \in \hat{\mathcal{S}}$, where $\bar{\sigma}(\gamma)$ denotes the P-completed σ-field in Ω, generated by the random element γ. We say that a process X on S is *externally measurable* with respect to ξ, if it is $\mathcal{X}$-measurable as a function on $S \times \Omega$. A random measure ζ on S is said to be externally measurable, if it is ξ-measurable, and the process $X_s = \zeta\{s\}$ on S is $\mathcal{X}$-measurable. Let μ_d and μ_a denote the diffuse and purely atomic parts of a measure μ.

Theorem 8.27 *(external projection, van der Hoeven, OK) Consider a K-marked point process ξ on S with external σ-field $\mathcal{X}$, along with a random measure ζ with $E\zeta \in \mathcal{M}_S$. Then $\hat{\zeta}$ is the a.s. unique, externally measurable random measure on S, satisfying $E\zeta Y = E\hat{\zeta}Y$ for every $\mathcal{X}$-measurable* *process $Y \geq 0$ on S. The mapping $\zeta \mapsto \hat{\zeta}$ is a.s. linear, non-decreasing, continuous from below, and such that a.s.*

$$(\hat{\zeta_a}) \geq (\hat{\zeta})_a = \pi_\zeta, \qquad (\hat{\zeta})_d \geq (\hat{\zeta_d}) = E(\zeta_d |\, \xi). \tag{18}$$

Proof: The linearity and monotonicity are obvious. To prove (18), we refer to the proof of Theorem 8.26, where $\hat{\zeta}$ is exhibited as a sum $\chi_1+\chi_2+\chi_3$. If ζ is diffuse, then $\zeta\bar{\xi} = 0$, and so $\chi_3 = 0$ a.s. In this case, also $\pi_\zeta = 0$ a.s. by Theorem 8.24, which implies $\chi_1 = 0$ a.s. Hence, $\hat{\zeta} = \chi_2 = (1-\bar{\pi})(1-\bar{\xi}) \cdot \zeta'$ a.s., where $\zeta' = E(\zeta \,|\, \xi)$. Since ζ' is again a.s. diffuse, and $\bar{\pi} = \bar{\xi} = 0$ outside a countable set, we conclude that $\hat{\zeta} = \zeta'$ a.s. Hence, for general ζ, we have $(\hat{\zeta})_d \geq (\hat{\zeta_d}) = E(\zeta_d |\, \xi)$ a.s., and so by additivity $(\hat{\zeta})_a \leq (\hat{\zeta_a})$ a.s.

To identify $(\hat{\zeta})_a$, we note that $\chi_1 = \pi_\zeta$, and $\chi_2 = \chi_3 = 0$ a.s. on $\operatorname{supp}\bar{\xi}$. It remains to show that $\hat{\zeta} = \pi_\zeta$, a.s. on $(\operatorname{supp}\bar{\xi})^c$. Using Lemma 8.10 as before, we get a.s. on $\{\xi B = 0\}$, simultaneously for all $s \in B$,

$$\begin{aligned}
\chi_2\{s\} &= \zeta'\{s\}(1-\bar{\pi}_s) \\
&= \zeta'\{s\}\, \frac{P(\xi_B = 0 \,|\, 1_{B^c}\xi)}{P(\xi_s = \xi_B \,|\, 1_{B^c}\xi)} \\
&= \frac{E\Big(\zeta\{s\};\, \xi_B = 0 \,\Big|\, 1_{B^c}\xi\Big)}{P(\xi_s = \xi_B \,|\, 1_{B^c}\xi)}\,,
\end{aligned}$$

and so by Theorem 8.24 we have, a.s. on $\{\xi B = 0\}$ for all $s \in B$,

$$\hat{\zeta}\{s\} = \chi_2\{s\} + \chi_3\{s\} = \frac{E\Big(\zeta\{s\};\, \xi_s = \xi_B \,\Big|\, 1_{B^c}\xi\Big)}{P(\xi_s = \xi_B \,|\, 1_{B^c}\xi)} = \pi_\zeta\{s\}.$$

Since B was arbitrary, we obtain $(\hat{\zeta})_a = \pi_\zeta$ a.s.

Now fix a dissection system (I_{nj}) on S, and write $B_{nj} = B \cap I_{nj}$ for any $B \in \hat{\mathcal{S}}$. Putting

$$Z(s) = \hat{\zeta}\{s\}, \quad Z_n(s) = E\Big(\zeta I_n(s) \,\Big|\, 1_{I_n(s)}\xi\Big), \qquad s \in S,\ n \in \mathsf{N},$$

we see from (18) and Theorem 8.24 that $Z_n(s) \to Z(s)$, outside a fixed P-null set. Since the processes Z_n are obviously $\mathcal{X}$-measurable, the same thing is true for the limit Z. Since $\hat{\zeta}$ is also ξ-measurable, it is indeed externally measurable.

By the L^1-convergence in Theorem 8.26, we have for any $B \in \hat{\mathcal{S}}$ and $A \in \bar{\sigma}(1_{B^c}\xi)$

$$\begin{aligned}
E(\zeta B; A) &= E\textstyle\sum_j E\Big(\zeta B_{nj} \Big|\, 1_{B^c_{nj}}\xi\Big) 1_A \\
&\to E(\hat{\zeta} B; A),
\end{aligned}$$

which gives $E\zeta Y = E\hat{\zeta} Y$ for $Y = 1_{B\times A}$. This extends by a monotone-class argument to any $\mathcal{X}$-measurable indicator function Y, and then to general Y, by linearity and monotone convergence.

To prove the uniqueness, let ζ_1 and ζ_2 be locally integrable and externally measurable random measures on S, satisfying $E\zeta_1 Y = E\zeta_2 Y$ for all $\mathcal{X}$-measurable processes $Y \geq 0$. Putting $Z_i(s) = \zeta_i\{s\}$ for $i = 1, 2$, and choosing $Y = 1_B 1\{Z_1 > Z_2\}$ for arbitrary $B \in \hat{\mathcal{S}}$, we get $Z_1 \leq Z_2$ a.s. The symmetric argument gives $Z_2 \leq Z_1$ a.s., and so by combination $Z_1 = Z_2$ a.s., which shows that the atomic parts agree a.s. for ζ_1 and ζ_2. Subtracting those atomic components, we may assume that ζ_1 and ζ_2 are diffuse. Since $E(\zeta_1 B \,|\, 1_{B^c}\xi) = E(\zeta_2 B \,|\, 1_{B^c}\xi)$ a.s. for all $B \in \hat{\mathcal{S}}$, we have $\hat{\zeta}_1 = \hat{\zeta}_2$ a.s., and so by (18) we get $\zeta_1 = \hat{\zeta}_1 = \hat{\zeta}_2 = \zeta_2$ a.s., as required.

Finally, let $\zeta_n \uparrow \zeta$. Then $\hat{\zeta}_n \uparrow \chi$ a.s. by monotonicity, for some random measure χ, and so by monotone convergence, $E\zeta Y = E\chi Y$ for every $\mathcal{X}$-measurable process $Y \geq 0$. Since also $\hat{\zeta}_n\{s\} \uparrow \chi\{s\}$ for all s, χ is again externally measurable, and so by uniqueness $\eta = \hat{\zeta}$ a.s. This proves that $\hat{\zeta}_n \uparrow \hat{\zeta}$ a.s. $\square$

Chapter 9

Compensation and Time Change

The dynamical aspects of random measures or point processes evolving in time are captured by the associated compensators, along with the related machinery of filtrations, martingales, optional and predictable times[1], and progressive and predictable processes. The mentioned random times are among the basic building blocks of increasing and more general processes, codifying the underlying jump structure. Here it is often useful to identify a random time τ with the corresponding unit mass δ_τ.

The basic result of the subject is the celebrated *Doob–Meyer decomposition* of a local submartingale X into a sum of a local martingale M and a locally integrable, increasing and predictable process A, both a.s. unique. Here we give a complete proof, based on Dunford's weak compactness criterion and Doob's ingenious approximation of totally inaccessible times.

The mentioned decomposition applies in particular to any locally integrable and adapted random measure ξ on $(0,\infty)$, in which case the increasing process A with associated random measure $\hat{\xi}$ is refereed to as the *compensator* of ξ. More generally, we often need to consider random measures ξ on a product space $(0,\infty)\times S$, where the existence of a compensator $\hat{\xi}$ on the same space is ensured by Theorem 9.21. It is often useful to think of $\hat{\xi}$ as the *dual predictable projection* of ξ, in the sense that[2] $E\,\xi V = E\,\hat{\xi} V$, for any predictable process $V \geq 0$ on $(0,\infty)\times S$.

Another basic result is Theorem 9.22, which gives an a.s. unique decomposition of an adapted random measure ξ on $(0,\infty)$ into a sum $\xi^c + \xi^q + \xi^a$, where ξ^c is diffuse, while ξ^q and ξ^a are purely atomic, and ξ^q is ql-continuous[3] whereas ξ^a is accessible. For the special case of random measures $\xi = \delta_\tau$, this leads to important characterizations of predictable, accessible, and totally inaccessible times.

If ξ is a marked point process with independent increments, the associated compensator $\hat{\xi}$ is clearly a.s. non-random. The converse is also true, as shown by the fundamental Theorem 9.29. A related result is Theorem 9.28, which shows how a ql-continuous point process ξ on $(0,\infty)\times S$ can be reduced to Cox through a pathwise mapping, by means of a suitable predictable process

[1] the former also known as *stopping times*

[2] Note that $\xi V = \int V d\xi$.

[3] short for *quasi-leftcontinuous*

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_9

V. This leads in particular to the celebrated time-change reduction of any ql-continuous point processes $\xi_1, \ldots, \xi_n$ into independent Poisson processes. It also yields a similar time-change reduction of the integral process $(V \cdot X)_t = \int_0^t V dX$, where X is a strictly p-stable Lévy process and the integrand V is non-negative and predictable.

For a deeper result of this type, suppose that the random measure ξ on $[0,1] \times S$ or $\mathsf{R}_+ \times S$ is *exchangeable*[4], in the sense that[5]

$$\lambda \circ f^{-1} = \lambda \quad \Rightarrow \quad \xi \circ f^{-1} \stackrel{d}{=} \xi,$$

where λ denotes Lebesgue measure on $[0,1]$ or R_+. Then the same property holds for predictable processes V on $[0,1]$ or R_+, in the sense that

$$\lambda \circ V^{-1} = \lambda \ \text{ a.s.} \quad \Rightarrow \quad \xi \circ V^{-1} \stackrel{d}{=} \xi.$$

This is the *predictable mapping* Theorem 9.33, a continuous-time counterpart of the *predictable sampling theorem* in Lemma 9.34.

For an optional time $\tau > 0$ with distribution μ, the associated compensator η with respect to the induced filtration is given by an explicit formula, which can be inverted to retrieve μ as a function of η. For a general filtration $\mathcal{F}$, the same inversion formula yields a predictable random measure ζ, called the *discounted compensator* of τ. The associated tail process $Z_t = 1 - \zeta(0,t]$ is the a.s. unique solution to *Doléans' differential equation* $Z = 1 - Z_- \cdot \eta$, which leads to an explicit expression for ζ. A similar construction applies when τ has an attached mark χ, with values in a Borel space S.

Given an adapted pair (τ, χ) as above with discounted compensator ζ, consider a predictable process Y on $\mathsf{R}_+ \times S$, with values in a probability space (T, μ), such that $\zeta \circ Y^{-1} \le \mu$ a.s. Then, quite amazingly, the image random element $\sigma = Y(\tau, \chi)$ in T has distribution μ. In Theorem 9.44 we prove a multi-variate version of the same result, where the random pairs (τ_j, χ_j), $j = 1, \ldots, n$, are mapped into *independent* random elements $\sigma_1, \ldots, \sigma_n$ with prescribed distributions $\mu_1, \ldots, \mu_n$.

The mentioned results can be used to extend the previously mentioned time-change reduction of ql-continuous point processes ξ, to the case of possibly discontinuous compensators η. Apart from a logarithmic transformation at each atom of η, we now need a randomization at the joint atoms of ξ and η. The transformation formula leads to simple estimates of the deviation from Poisson, showing in particular that, if $\xi_1, \xi_2, \ldots$ are simple point processes with compensators $\eta_1, \eta_2, \ldots$ satisfying $\eta_n \xrightarrow{d} \lambda$, then ξ_n tends in distribution to a unit rate Poisson process ξ.

The deepest result of the chapter is Theorem 9.48, which describes the set of all distributions of triples (τ, χ, η), where (τ, χ) is an adapted pair, as before, with a general compensator η. The mentioned distributions form a

[4]same as λ-symmetric, when λ equals Lebesgue measure on $[0,1]$ or R_+
[5]Recall that the measure $\lambda \circ f^{-1}$ is given by $(\lambda \circ f^{-1})g = \lambda(g \circ f)$.

convex set, whose extreme points are the distributions P_μ, in the case where $\mathcal{L}(\tau,\chi) = \mu$, and η is the *natural* compensator of (τ,χ), obtained for the induced filtration. Our main result provides a unique integral representation

$$\mathcal{L}(\tau,\chi,\eta) = \int P_\mu \, \nu(d\mu),$$

in terms of a probability measure ν on the set of measures μ. This is equivalent to the existence of a random probability measure ρ on $(0,\infty)\times S$ with distribution ν, such that a.s.

$$\zeta = \pi_\tau \rho, \qquad \mathcal{L}(\tau,\chi \,|\, \rho) = \rho,$$

where ζ is the discounted compensator of (τ,χ), and $\pi_\tau\rho$ denotes the restriction of ρ to the set $(0,\tau]\times S$.

The chapter ends with a discussion of *tangential* processes, defined as pairs of semi-martingales with the same *local characteristics*. For local martingales X, the latter consist of the quadratic variation $[X^c]$ of the continuous martingale component X^c, along with the compensator $\hat\xi$ of the point process ξ, codifying the times and sizes of the jumps. For general processes X, we need in addition a suitable compensator A of the entire process X. For non-decreasing processes X, it is enough to consider the compensator $\hat\xi$ of the jump point process ξ, along with the continuous component X^c of X.

Though the local characteristics may not determine the distribution of the underlying process, two tangential processes do share some important asymptotic properties. A basic comparison is given by Theorem 9.57, which shows that, if X and Y are tangential and either non-decreasing or conditionally symmetric, then for any function φ of moderate growth,

$$E\,\varphi(X^*) \asymp E\,\varphi(Y^*).$$

Here $X^* = \sup_t |X_t|$, and the relation $\asymp$ means that the ratio between the two sides is bounded above and below by positive constants, depending only on φ. Furthermore, the *conditional symmetry* of a process X means that X and $-X$ are tangential.

Even without any symmetry assumption, we can still prove a one-sided comparison, of the form

$$E\,\varphi(X^*) \lesssim E\,\varphi(Y^*),$$

for any tangential semi-martingales X and Y, such that Y has conditionally independent increments. The two-sided version is also true, whenever φ is convex, and the processes X and Y are *weakly tangential* local martingales, in the sense that the associated quadratic variation processes $[X]$ and $[Y]$ are tangential, in the previous strict sense.

The two-sided comparison enables us to prove relations of the form

$$\{X^* < \infty\} = \{Y^* < \infty\} \text{ a.s.},$$

which is clearly stronger than the equivalence of $X^* < \infty$ a.s. and $Y^* < \infty$ a.s. For sequences of pairwise tangential processes $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$, we may further prove that

$$X_n^* \xrightarrow{P} 0 \quad \Leftrightarrow \quad Y_n^* \xrightarrow{P} 0.$$

Under suitable conditions, we can also establish some one-sided relations or implications of either kind.

In order to derive criteria for $X^* < \infty$ a.s. and related properties, we need to find a tangential reference process Y, for which the corresponding criteria are known, or can be established by elementary methods. In Theorem 9.55, we prove that any semi-martingale X has a tangential process $\tilde{X}$ with conditionally independent increments. In particular, any simple point process ξ with diffuse compensator η has a tangential Cox process $\tilde{\xi}$ directed by η, hence the significance of the map[6] $\xi \mapsto \tilde{\xi}$.

The mentioned result is more subtle than it may first appear, since it also involves an extension of the original filtration $\mathcal{F} = (\mathcal{F}_t)$. We say that the filtration $\mathcal{G} = (\mathcal{G}_t)$ is a *standard extension* of $\mathcal{F}$, if it preserves all adaptedness and conditioning properties, as expressed by the relations

$$\mathcal{F}_t \subset \mathcal{G}_t \perp\!\!\!\perp_{\mathcal{F}_t} \mathcal{F}, \quad t \geq 0. \tag{1}$$

We may now give a more precise statement of the tangential existence theorem: For any semi-martingale X with local characteristics Y, with respect to a filtration $\mathcal{F}$, there exists a tangential semi-martingale $\tilde{X} \perp\!\!\!\perp_Y \mathcal{F}$ with respect to a standard extension $\mathcal{G}$ of $\mathcal{F}$, such that $\tilde{X}$ has conditionally independent increments, given Y. Note that any standard extension of $\mathcal{F}$ preserves the local characteristics of X, which then remain Y under the new filtration $\mathcal{G}$.

Using the cited comparison theorems, we can derive some asymptotic properties of X from those of $\tilde{X}$. For the latter process, a simple conditioning reduces the analysis to the case of independent increments, where the classical methods of characteristic functions and Laplace transforms apply. In particular, convergence criteria for simple, ql-continuous point processes agree with those for suitable Poisson processes, where such results are elementary.

9.1 Predictable Times and Processes

Here we collect some basic definitions and results from the general theory of processes. Much of this material will be used frequently, without further comments, in subsequent sections.

A *filtration* $\mathcal{F}$ on a linearly ordered index set T is a non-decreasing family of sub-σ-fields $\mathcal{F}_t$ of the basic σ-field $\mathcal{A}$ in Ω. A process X on T is said to be

[6] somewhat jokingly referred to as *Coxification*

$\mathcal{F}$-*adapted*, if X_t is $\mathcal{F}_t$-measurable for every $t \in T$. When $T = \mathsf{R}_+$, we define $\mathcal{F}_\infty = \bigvee_n \mathcal{F}_n \equiv \sigma\{\mathcal{F}_t;\, t \in T\}$. A random time τ in the closure $\bar{T}$ is said to be $\mathcal{F}$-*optional* if the process $1\{\tau \le t\}$ is $\mathcal{F}$-adapted. Any optional time τ determines an associated σ-field

$$\mathcal{F}_\tau = \Big\{A \in \mathcal{A};\; A \cap \{\tau \le t\} \in \mathcal{F}_t,\; t \in \bar{T}\Big\}.$$

Lemma 9.1 *(optional times) For any optional times σ and τ,*

(i) *τ is $\mathcal{F}_\tau$-measurable,*

(ii) *$\mathcal{F}_\tau = \mathcal{F}_t$ on $\{\tau = t\}$, for all $t \in T$,*

(iii) *$\mathcal{F}_\sigma \cap \{\sigma \le \tau\} \subset \mathcal{F}_{\sigma\wedge\tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.*

In particular, (iii) shows that $\{\sigma \le \tau\} \in \mathcal{F}_\sigma \cap \mathcal{F}_\tau$, that $\mathcal{F}_\sigma = \mathcal{F}_\tau$ on $\{\sigma = \tau\}$, and that $\sigma \le \tau$ implies $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.

Proof: (i) This holds since

$$\{\tau \le s\} \cap \{\tau \le t\} = \{\tau \le s \wedge t\} \in \mathcal{F}_{s\wedge t} \subset \mathcal{F}_t.$$

(ii) By (i), we have $\{\tau = t\} \in \mathcal{F}_\tau \cap \mathcal{F}_t$. If $A \in \mathcal{F}_t$, then for any $s \ge 0$,

$$A \cap \{\tau = t\} \cap \{\tau \le s\} = A \cap \{\tau = t \le s\} \in \mathcal{F}_s,$$

showing that $\mathcal{F}_t \subset \mathcal{F}_\tau$ on $\{\tau = t\}$. Conversely, $A \in \mathcal{F}_\tau$ yields $A \cap \{\tau = t\} \in \mathcal{F}_\tau$, and so

$$A \cap \{\tau = t\} = A \cap \{\tau = t\} \cap \{\tau \le t\} \in \mathcal{F}_t.$$

(iii) For any $A \in \mathcal{F}_\sigma$ and $t \in T$, we have

$$\begin{aligned} & A \cap \{\sigma \le \tau\} \cap \{\tau \le t\} \\ & \quad = \Big(A \cap \{\sigma \le t\}\Big) \cap \{\tau \le t\} \cap \{\sigma \wedge t \le \tau \wedge t\} \in \mathcal{F}_t, \end{aligned}$$

since $\sigma \wedge t$ and $\tau \wedge t$ are both $\mathcal{F}_t$-measurable. Hence,

$$\mathcal{F}_\sigma \cap \{\sigma \le \tau\} \subset \mathcal{F}_\tau.$$

Replacing τ by $\sigma \wedge \tau$ yields the first relation. Next, we may replace (σ, τ) by the pairs $(\sigma \wedge \tau, \sigma)$ and $(\sigma \wedge \tau, \tau)$, to get $\mathcal{F}_{\sigma\wedge\tau} \subset \mathcal{F}_\sigma \cap \mathcal{F}_\tau$. Conversely, for any $A \in \mathcal{F}_\sigma \cap \mathcal{F}_\tau$ and $t \in T$,

$$A \cap \{\sigma \wedge \tau \le t\} = \Big(A \cap \{\sigma \le t\}\Big) \cup \Big(A \cap \{\tau \le t\}\Big) \in \mathcal{F}_t,$$

and so $A \in \mathcal{F}_{\sigma\wedge\tau}$, proving that $\mathcal{F}_\sigma \cap \mathcal{F}_\tau \subset \mathcal{F}_{\sigma\wedge\tau}$. □

Given a filtration $\mathcal{F}$ on R_+, we say that a random time τ is *weakly optional*, if the process $1\{\tau < t\}$ is $\mathcal{F}$-adapted. We also introduce the right-continuous filtration $\mathcal{F}_t^+ = \bigcap_{u>t} \mathcal{F}_t$, and put $\mathcal{F}_{\tau+} = \bigcap_{h>0} \mathcal{F}_{\tau+h}$.

Lemma 9.2 *(weakly optional times) A random time τ is weakly $\mathcal{F}$-optional, iff it is $\mathcal{F}^+$-optional, in which case*

$$\mathcal{F}_{\tau+} = \mathcal{F}_\tau^+ = \Big\{A \in \mathcal{A};\ A \cap \{\tau < t\} \in \mathcal{F}_t,\ t > 0\Big\}. \tag{2}$$

Proof: If $A \in \mathcal{F}_\tau^+$, then for any $t > 0$,

$$A \cap \{\tau < t\} = \bigcup_{r<t} \Big(A \cap \{\tau \le r\}\Big) \in \bigcup_{r<t} \mathcal{F}_r \subset \mathcal{F}_t,$$

with r restricted to Q_+. Conversely, if $A \cap \{\tau < t\} \in \mathcal{F}_t$ for all t, then for any $t \ge 0$,

$$A \cap \{\tau \le t\} = \bigcap_{h>0} \bigcap_{r\in(t,t+h)} \Big(A \cap \{\tau < r\}\Big) \in \bigcap_{h>0} \mathcal{F}_{t+h} = \mathcal{F}_t^+.$$

For $A = \Omega$, this yields the first assertion, and for general A it implies the second relation in (2). To prove the first relation, we note that $A \in \mathcal{F}_{\tau+}$ iff $A \cap \{\tau + h \le t\} \in \mathcal{F}_t$ for all $t \ge 0$ and $h > 0$, which is equivalent to $A \cap \{\tau \le t\} \in \mathcal{F}_{t+h}$ for all such t and h, hence to $A \cap \{\tau \le t\} \in \mathcal{F}_{t+}$ for all $t \ge 0$. This means that $A \in \mathcal{F}_t^+$. □

Lemma 9.3 *(closure properties) Consider any random times $\tau_1, \tau_2, \ldots$ and filtration $\mathcal{F}$.*

(i) *If the τ_n are $\mathcal{F}$-optional, then so is $\tau = \sup_n \tau_n$.*

(ii) *If the τ_n are weakly $\mathcal{F}$-optional, then so is $\sigma = \inf_n \tau_n$, and we have $\mathcal{F}_\sigma^+ = \bigcap_n \mathcal{F}_{\tau_n}^+$.*

Proof: To prove (i) and the first part of (ii), we may write respectively

$$\{\tau \le t\} = \bigcap\nolimits_n \{\tau_n \le t\} \in \mathcal{F}_t, \qquad \{\sigma < t\} = \bigcup\nolimits_n \{\tau_n < t\} \in \mathcal{F}_t.$$

To prove the second assertion in (ii), we see from Lemma 9.1 (iii) that

$$\mathcal{F}_\sigma^+ = \bigcap_{h>0} \mathcal{F}_{\sigma+h} \subset \bigcap_{h>0} \bigcap_{n\ge1} \mathcal{F}_{\tau_n+h} = \bigcap_{n\ge1} \mathcal{F}_{\tau_n}^+.$$

Conversely, Lemma 9.2 yields for any $A \in \bigcap_n \mathcal{F}_{\tau_n}^+$ and $t \ge 0$

$$A \cap \{\sigma < t\} = \bigcup\nolimits_n \Big(A \cap \{\tau_n < t\}\Big) \in \mathcal{F}_t,$$

and so $A \in \mathcal{F}_\sigma^+$ by the same lemma, which shows that $\bigcap_n \mathcal{F}_{\tau_n}^+ \subset \mathcal{F}_\sigma^+$. □

Corollary 9.4 *(hitting times) Let X be an adapted process in a metric space S, fix a Borel set $B \subset S$, and define $\tau = \inf\{t > 0;\ X_t \in B\}$.*

(i) *If X is left-continuous and B is closed, then τ is optional.*

(ii) *If X is right-continuous and B is open, then τ is weakly optional.*

Proof: Approximate τ from below or above, respectively. $\square$

The theory simplifies when the filtration $\mathcal{F}$ is right-continuous, since every weakly optional time is then optional. It is also useful to assume $\mathcal{F}$ to be *complete*, in the sense that each $\mathcal{F}_t$ contains all the P-null sets in the basic σ-field $\mathcal{A}$. Both properties can be achieved by a suitable augmentation of $\mathcal{F}$. For any σ-field $\mathcal{H}$, we write $\overline{\mathcal{H}}$ for the completion of $\mathcal{H}$ in the stated sense.

Lemma 9.5 *(augmented filtration) Every filtration $\mathcal{F}$ on R_+ has a smallest right-continuous and complete extension $\mathcal{G}$, given by*

$$\mathcal{G}_t = \overline{(\mathcal{F}_{t+})} = \overline{(\mathcal{F}_t)}_+, \quad t \geq 0.$$

Proof: It is clearly enough to prove the second equality. Then note that

$$\overline{(\mathcal{F}_{t+})} \subset \overline{(\overline{(\mathcal{F}_t)}_+)} \subset \overline{(\mathcal{F}_t)}_+, \quad t \geq 0.$$

Conversely, let $A \in \overline{(\mathcal{F}_t)}_+$ be arbitrary. Then $A \in \overline{(\mathcal{F}_{t+h})}$ for every $h > 0$, and so there exist some sets $A_h \in \mathcal{F}_{t+h}$ with $P(A\Delta A_h) = 0$. Putting $A' = \{A_{1/n} \text{ i.o.}\}$, we get $A' = \mathcal{F}_{t+}$ and $P(A\Delta A') = 0$, and so $A \in \overline{(\mathcal{F}_{t+})}$. This shows that $\overline{(\mathcal{F}_t)}_+ \subset \overline{(\mathcal{F}_{t+})}$. $\square$

A random time τ is said to be *predictable*, if it is *announced* by a sequence of optional times $\tau_n \uparrow \tau$, such that $\tau_n < \tau$ a.s. on $\{\tau > 0\}$ for all n. For any optional time τ, we introduce the σ-field $\mathcal{F}_{\tau-}$, generated by $\mathcal{F}_0$ and the families $\mathcal{F}_t \cap \{\tau > t\}$ for all $t > 0$.

Lemma 9.6 *(strict past) For any optional times σ and τ,*

(i) $\mathcal{F}_\sigma \cap \{\sigma < \tau\} \subset \mathcal{F}_{\tau-} \subset \mathcal{F}_\tau$,

(ii) $\{\sigma < \tau\} \in \mathcal{F}_{\sigma-} \cap \mathcal{F}_{\tau-}$, *when τ is predictable,*

(iii) $\bigvee_n \mathcal{F}_{\tau_n} = \mathcal{F}_{\tau-}$, *when τ is predictable and announced by (τ_n).*

Proof: (i) For any $A \in \mathcal{F}_\sigma$, we have

$$A \cap \{\sigma < \tau\} = \bigcup_{r \in \mathsf{Q}_+} \Big(A \cap \{\sigma \leq r\} \cap \{r < \tau\} \Big) \in \mathcal{F}_{\tau-},$$

since all intersections on the right are generators of $\mathcal{F}_{\tau-}$. This proves the first relation in (i), and the second one holds, since every generator of $\mathcal{F}_{\tau-}$ belongs to $\mathcal{F}_\tau$.

(ii) If τ is announced by $\tau_1, \tau_2, \ldots$, then (i) yields

$$\{\tau \leq \sigma\} = \{\tau = 0\} \cup \bigcap\nolimits_n \{\tau_n < \sigma\} \in \mathcal{F}_{\sigma-}.$$

(iii) For any $A \in \mathcal{F}_{\tau_n}$, we get by (i)

$$A = \Big(A \cap \{\tau_n < \tau\} \Big) \cup \Big(A \cap \{\tau_n = \tau = 0\} \Big) \in \mathcal{F}_{\tau-},$$

and so $\bigvee_n \mathcal{F}_{\tau_n} \subset \mathcal{F}_{\tau-}$. Conversely, (i) yields for any $t \geq 0$ and $A \in \mathcal{F}_t$

$$A \cap \{\tau > t\} = \bigcup\nolimits_n \Big(A \cap \{\tau_n > t\} \Big) \in \bigvee\nolimits_n \mathcal{F}_{\tau_n -} \subset \bigvee\nolimits_n \mathcal{F}_{\tau_n},$$

which shows that $\mathcal{F}_{\tau-} \subset \bigvee_n \mathcal{F}_{\tau_n}$. □

It is often convenient to regard a measurable process on R_+ as a function on $\Omega \times \mathsf{R}_+$. On the latter space, we introduce a basic σ-field.

Lemma 9.7 *(predictable σ-field) The following classes of sets or processes generate the same σ-field $\mathcal{P}$ in $\Omega \times \mathsf{R}_+$:*

(i) *the set of continuous, adapted processes,*
(ii) *the set of left-continuous, adapted processes,*
(iii) $\mathcal{F}_0 \times \mathsf{R}_+$*, and all sets* $A \times (t, \infty)$ *with* $A \in \mathcal{F}_t$, $t \geq 0$*,*
(iv) $\mathcal{F}_0 \times \mathsf{R}_+$*, and all intervals* (τ, ∞) *with* τ *optional.*

Proof: Let $\mathcal{P}_1, \ldots, \mathcal{P}_4$ be the σ-fields generated by the classes in (i)$-$(iv). Then trivially $\mathcal{P}_1 \subset \mathcal{P}_2$, and $\mathcal{P}_2 \subset \mathcal{P}_3$ holds since every left-continuous process X can be approximated by sums $X_0 1_{\{0\}}(t) + \sum_k X_{k/n} 1_{(k,k+1]}(nt)$. Next, $\mathcal{P}_3 \subset \mathcal{P}_4$, since the time $t_A = t \cdot 1_A + \infty \cdot 1_{A^c}$ is optional for any $t \geq 0$ and $A \in \mathcal{F}_t$. Finally, $\mathcal{P}_4 \subset \mathcal{P}_1$, since for any optional time τ, the process $1_{(\tau,\infty)}$ can be approximated by the continuous, adapted processes $X_t^n = n(t-\tau)_+ \wedge 1$. □

A process X on R_+ is said to be *predictable*, if it is measurable with respect to the σ-field $\mathcal{P}$ in Lemma 9.7. We further say that X is *progressively measurable* or simply *progressive*, if its restriction to $\Omega \times [0, t]$ is $(\mathcal{F}_t \otimes \mathcal{B}_{[0,t]})$-measurable for every $t \geq 0$. Every progressive process is clearly adapted. Conversely, approximating from the right, we see that any right-continuous, adapted process is progressive. In particular, any predictable process is progressive.

Lemma 9.8 *(progressive and predictable processes)*

(i) *For any optional time τ and progressive process X, the random variable $X_\tau 1\{\tau < \infty\}$ is $\mathcal{F}_\tau$-measurable.*
(ii) *For any optional time τ and predictable process X, the random variable $X_\tau 1\{\tau < \infty\}$ is $\mathcal{F}_{\tau-}$-measurable.*
(iii) *For any predictable time τ and $\mathcal{F}_{\tau-}$-measurable random variable α, the process $X_t = \alpha\, 1\{\tau \leq t\}$ is predictable.*

Proof: (i) Assuming X to take values in $(S, \mathcal{S})$, we need to show that

$$\{X_\tau \in B,\, \tau \leq t\} \in \mathcal{F}_t, \quad t \geq 0,\ B \in \mathcal{S},$$

which holds if $X_{\tau \wedge t}$ is $\mathcal{F}_t$-measurable for every t. We may then assume $\tau \leq t$, and prove instead that X_τ is $\mathcal{F}_t$-measurable. This is clear if we write

$X_\tau = X \circ \psi$ with $\psi(\omega) = (\omega, \tau(\omega))$, and note that ψ is measurable from $\mathcal{F}_t$ to $\mathcal{F}_t \otimes \mathcal{B}_{[0,t]}$, whereas X is measurable on $\Omega \times [0,t]$ from $\mathcal{F}_t \otimes \mathcal{B}_{[0,t]}$ to $\mathcal{S}$.

(ii) If $X = 1_{A\times(t,\infty)}$ for some $t > 0$ and $A \in \mathcal{F}_t$, then $X_\tau 1\{\tau < \infty\}$ is the indicator function of $A \cap \{t < \tau < \infty\} \in \mathcal{F}_{\tau-}$. This extends by a monotone-class argument to any predictable indicator function, and then by linearity and monotone convergence to the general case.

(iii) We may clearly assume α to be integrable. Given an announcing sequence (τ_n) of τ, we introduce the processes

$$X_t^n = E(\alpha \,|\, \mathcal{F}_{\tau_n}) \Big(1\{0 < \tau_n < t\} + 1\{\tau_n = 0\}\Big), \quad t \geq 0,\ n \in \mathsf{N},$$

which are left-continuous and adapted, hence predictable. Furthermore, $\bigvee_n \mathcal{F}_{\tau_n} = \mathcal{F}_{\tau-}$ by Lemma 9.6, and so $E(\alpha \,|\, \mathcal{F}_{\tau_n}) \to E(\alpha \,|\, \mathcal{F}_{\tau-}) = \alpha$ a.s., which implies $X^n \to X$ on R_+ a.s. Thus, X is again predictable. $\Box$

The *graph* of a random time τ is defined as the set

$$[\tau] = \Big\{(t,\omega) \in \mathsf{R}_+ \times \Omega;\ \tau(\omega) = t\Big\}.$$

An optional time τ is said to be *totally inaccessible*, if $[\tau] \cap [\sigma] = \emptyset$ a.s. for every predictable time σ, and *accessible*, if $[\tau] \cap [\sigma] = \emptyset$ a.s. for every totally inaccessible time σ. For any set $A \in \mathcal{F}_\tau$, the *restriction* $\tau_A = \tau\, 1_A + \infty \cdot 1_{A^c}$ of τ to A is again optional.

Lemma 9.9 *(decomposition of optional times) For any optional time τ, there exists an a.s. unique set $A \in \mathcal{F}_\tau \cap \{\tau < \infty\}$, such that τ_A is accessible and τ_{A^c} is totally inaccessible. Furthermore, $[\tau_A] \subset \bigcup_n [\tau_n]$ a.s., for some predictable times $\tau_1, \tau_2, \ldots$.*

Proof: Define

$$p = \sup P \bigcup\nolimits_n \{\tau = \tau_n < \infty\}, \tag{3}$$

where the supremum extends over all sequences of predictable times τ_n. Combining sequences where the probabilities in (3) approach p, we may construct a sequence where the supremum is attained. For such a maximal sequence, let A be the union in (3). Then τ_{A^c} is totally inaccessible, since we could otherwise improve on the value of p. Furthermore, $[\sigma] \cap [\tau_n] = \emptyset$ a.s. for any totally inaccessible time σ, and so $[\sigma] \cap [\tau_A] = \emptyset$ a.s., which shows that τ_A is accessible. To prove the a.s. uniqueness of A, let B be any other set with the stated properties. Then $\tau_{A\setminus B}$ and $\tau_{B\setminus A}$ are both accessible and totally inaccessible, and so $\tau_{A\setminus B} = \tau_{B\setminus A} = \infty$ a.s., which implies $A = B$ a.s. $\Box$

Justified by Lemma 9.5, and for technical convenience, we may henceforth assume the underlying filtration $\mathcal{F}$ to be both right-continuous and complete.

Lemma 9.10 *(approximation of inaccessible times, Doob) Given a totally inaccessible time τ, put $\tau_n = 2^{-n}[2^n\tau]$, and choose some right-continuous versions X_t^n of the processes $P(\tau_n \le t \,|\, \mathcal{F}_t)$. Then*

$$\lim_{n\to\infty} \sup_{t\ge 0} \left| X_t^n - 1\{\tau \le t\} \right| = 0 \ \ a.s.$$

Proof: Since $\tau_n \uparrow \tau$, we may assume that $X_t^1 \ge X_t^2 \ge \cdots \ge 1\{\tau \le t\}$ for all $t \ge 0$. Then $X_t^n = 1$ for $t \ge \tau$, and as $t \to \infty$,

$$\begin{aligned} X_t^1 \ &\le \ P(\tau < \infty \,|\, \mathcal{F}_t) \\ &\to 1\{\tau < \infty\} = 0 \ \text{ a.s. on } \{\tau = \infty\}. \end{aligned}$$

Thus, $\sup_n |X_t^n - 1\{\tau \le t\}| \to 0$ a.s. as $t \to \infty$. It remains to show that, for any $\varepsilon > 0$, the times

$$\sigma_n = \inf\Big\{ t \ge 0;\ X_t^n - 1\{\tau \le t\} > \varepsilon \Big\}, \quad n \in \mathsf{N},$$

tend to ∞ a.s. The σ_n are clearly optional and non-decreasing, say with limit σ, and we have either $\sigma \le \tau$ or $\sigma_n = \infty$ for all n.

By optional sampling, the disintegration theorem, and Lemma 9.1, we have

$$\begin{aligned} X_\sigma^n \, 1\{\sigma < \infty\} \ &= \ P(\,\tau_n \le \sigma < \infty \,|\, \mathcal{F}_\sigma\,) \\ &\to P(\,\tau \le \sigma < \infty \,|\, \mathcal{F}_\sigma\,) \\ &= \ 1\{\tau \le \sigma < \infty\}. \end{aligned}$$

Hence, $X_\sigma^n \to 1\{\tau \le \sigma\}$ a.s. on $\{\sigma < \infty\}$, and so by right-continuity, we have $\sigma_n < \sigma$ for large enough n. Thus, σ is predictable and announced by the times $\sigma_n \wedge n$.

Using the optional sampling and disintegration theorems, we get

$$\begin{aligned} \varepsilon\, P\{\sigma < \infty\} \ &\le \ \varepsilon\, P\{\sigma_n < \infty\} \\ &\le \ E(X_{\sigma_n}^n;\ \sigma_n < \infty) \\ &= \ P\{\tau_n \le \sigma_n < \infty\} \\ &= \ P\{\tau_n \le \sigma_n \le \tau < \infty\} \\ &\to P\{\tau = \sigma < \infty\} = 0, \end{aligned}$$

where the last equality holds, since τ is totally inaccessible. Hence, $\sigma = \infty$ a.s., and the assertion follows since ε was arbitrary. □

An integrable random measure ξ on $(0,\infty)$ is said to be *natural*, if $E\xi(\Delta M) = 0$ for every bounded martingale M.

Lemma 9.11 *(continuity at inaccessible times) Given a natural random measure ξ on $(0,\infty)$, and a totally inaccessible time τ, we have $\xi\{\tau\} = 0$ a.s. on $\{\tau < \infty\}$.*

Proof: By rescaling, we may assume that ξ is a.s. continuous at dyadic times. Define $\tau_n = 2^{-n}[2^n\tau]$. Since ξ is natural, we have

$$E\int P(\tau_n > t \,|\, \mathcal{F}_t)\,\xi(dt) = E\int P(\tau_n > t \,|\, \mathcal{F}_{t-})\,\xi(dt),$$

and since τ is totally inaccessible, Lemma 9.10 yields

$$\begin{aligned} E\,\xi(0,\tau) &= E\int 1\{\tau > t\}\,\xi(dt) \\ &= E\int 1\{\tau \geq t\}\,\xi(dt) = E\,\xi(0,\tau]. \end{aligned}$$

Hence, $E(\xi\{\tau\};\, \tau < \infty) = 0$, and the assertion follows. □

Lemma 9.12 *(constant martingales) A process M is a predictable martingale of integrable variation, iff $M_t \equiv M_0$ a.s.*

Proof: Let M be a predictable martingale of integrable variation, with $M_0 = 0$. On the predictable σ-field $\mathcal{P}$, we define a signed measure μ by

$$\mu B = E\int_0^\infty 1_B(t)\,dM_t, \quad B \in \mathcal{P}.$$

Since M is uniformly integrable, and hence extends to a martingale on $[0,\infty]$, we have $\mu\{F \times (t,\infty)\} = 0$ for all $t \geq 0$ and $F \in \mathcal{F}_t$, and so a monotone-class argument, based on Lemma 9.7, yields $\mu = 0$ on $\mathcal{P}$. The predictability of M extends to the process $\Delta M_t = M_t - M_{t-}$, and hence also to the sets $J_\pm = \{t > 0;\, \pm\Delta M_t > 0\}$. Then $\mu J_\pm = 0$, and so $E\sum_t |\Delta M_t| = 0$, which implies $\Delta M = 0$ a.s., and shows that M is a.s. continuous.

By localization, we may henceforth take M to be a continuous martingale of bounded variation. Using the martingale property and dominated convergence, we get for any $t > 0$

$$\begin{aligned} EM_t^2 &= E\sum_{k\leq n}\Big(M_{kt/n} - M_{(k-1)t/n}\Big)^2 \\ &\leq E\max_{j\leq n}\Big|M_{jt/n} - M_{(j-1)t/n}\Big| \sum_{k\leq n}\Big|M_{kt/n} - M_{(k-1)t/n}\Big| \to 0, \end{aligned}$$

which implies $M_t = 0$ a.s. By continuity, it follows that $M = 0$ a.s. □

We conclude with a simple technical lemma, which is easily established by standard monotone-class arguments.

Lemma 9.13 *(predictable random measures)*

(i) *For any predictable random measure ξ, and a predictable process $V \geq 0$ on $(0,\infty)\times S$, the random measure $V\cdot\xi$ is again predictable.*

(ii) *For any predictable process $V \geq 0$ on $(0,\infty)\times S$ and predictable, measure-valued process ρ on S, the process $Y_t = \rho_t V(t,\cdot)$ is again predictable.*

9.2 Doob–Meyer Decomposition and Compensation

All subsequent material in this chapter is based on the notion of *compensator*, whose very definition relies on the fundamental *Doob–Meyer decomposition* below. Here an *increasing process* is understood to be a non-decreasing, right-continuous, and adapted process A with $A_0 = 0$. It is said to be *integrable*, if $EA_\infty < \infty$. All sub-martingales are assumed to be right-continuous. Local sub-martingales and locally integrable processes are defined by the usual device of localization.

Theorem 9.14 *(sub-martingale decomposition, Meyer, Doléans) A process X is a local sub-martingale, iff $X = M + A$ for some local martingale M, and a locally integrable, increasing, predictable process A. The latter processes are then a.s. unique.*

Our proof will be based on the next two lemmas. Let (D) denote the class of measurable processes X, such that the family $\{X_\tau\}$ is uniformly integrable, where τ ranges over the set of all finite optional times τ.

Lemma 9.15 *(preliminary decomposition, Meyer) Every sub-martingale X of class (D) satisfies $X = M + A$, for some uniformly integrable martingale M, and a natural, increasing process A.*

Proof (Rao): We may assume that $X_0 = 0$. For any process Y, we introduce the differences $\Delta_k^n Y = Y_{t_{k+1}^n} - Y_{t_k^n}$, where $t_k^n = k2^{-n}$, $k \in \mathsf{Z}_+$. Define

$$A_t^n = \sum_{k<2^n t} E(\Delta_k^n X | \mathcal{F}_{t_k^n}), \quad t \ge 0, \ n \in \mathsf{N},$$

and note that $M^n = X - A^n$ is a martingale on the n-dyadic set $D_n = 2^{-n}\mathsf{Z}_+$.

Introducing the optional times $\tau_r^n = \sup\{t \in D_n;\ A_t^n \le r\}$, $n \in \mathsf{N}$, $r > 0$, we get by optional sampling, for any $t \in D_n$,

$$\begin{aligned}
\tfrac{1}{2}\, E\Big(A_t^n;\ A_t^n > 2r\Big) &\le E(A_t^n - A_t^n \wedge r) && (4)\\
&\le E(A_t^n - A_{\tau_r^n \wedge t}^n)\\
&= E(X_t - X_{\tau_r^n \wedge t})\\
&= E\Big(X_t - X_{\tau_r^n \wedge t};\ A_t^n > r\Big). && (5)
\end{aligned}$$

Furthermore, by the martingale property and uniform integrability,

$$r\, P\{A_t^n > r\} \le EA_t^n = EX_t \lesssim 1,$$

and so the probability on the left tends to 0 as $r \to \infty$, uniformly in t and n. Since the differences $X_t - X_{\tau_r^n \wedge t}$ are uniformly integrable by (D), so are the variables A_t^n by (4). In particular, the sequence (A_∞^n) is uniformly integrable, and each M^n is a uniformly integrable martingale.

By Dunford's Lemma A1.3, A_∞^n converges weakly in L^1, along a subsequence $N' \subset \mathbb{N}$, toward some random variable $\alpha \in L^1(\mathcal{F}_\infty)$. Defining

$$M_t = E(X_\infty - \alpha \,|\, \mathcal{F}_t), \quad A_t = X_t - M_t, \qquad t \geq 0,$$

we note that $A_\infty = \alpha$ a.s. For dyadic t and bounded random variables γ, the martingale and self-adjointness properties yield

$$\begin{aligned} E(A_t^n - A_t)\gamma &= E(M_t - M_t^n)\gamma \\ &= E E\Big(M_\infty - M_\infty^n \,\Big|\, \mathcal{F}_t\Big)\gamma \\ &= E(M_\infty - M_\infty^n)\, E(\gamma \,|\, \mathcal{F}_t) \\ &= E(A_\infty^n - \alpha)\, E(\gamma \,|\, \mathcal{F}_t) \to 0, \end{aligned}$$

as $n \to \infty$ along N', which shows that $A_t^n \to A_t$, weakly in L^1 for dyadic t. In particular, we get for dyadic $s < t$

$$\begin{aligned} 0 \;\leq\; & E\Big(A_t^n - A_s^n;\, A_t - A_s < 0\Big) \\ & \to E\{(A_t - A_s) \wedge 0\} \leq 0, \end{aligned}$$

and so the last expectation vanishes, which implies $A_t \geq A_s$ a.s. By right continuity, A is then a.s. non-decreasing. Furthermore, $A_0 = 0$ a.s., since $A_0^n = 0$ for all n.

For any bounded martingale N, we get by Fubini's theorem and the martingale properties of N and $A^n - A = M - M^n$

$$\begin{aligned} E\, N_\infty A_\infty^n &= \textstyle\sum_k E\, N_\infty\, \Delta_k^n A^n \\ &= \textstyle\sum_k E\, N_{t_k^n}\, \Delta_k^n A^n \\ &= \textstyle\sum_k E\, N_{t_k^n}\, \Delta_k^n A \\ &= E \textstyle\sum_k N_{t_k^n}\, \Delta_k^n A. \end{aligned}$$

Using weak convergence on the left and dominated convergence on the right, along with Fubini's theorem, and the martingale property of N, we get

$$\begin{aligned} E \int N_{t-}\, dA_t \;&=\; E\, N_\infty\, A_\infty = \textstyle\sum_k E\, N_\infty\, \Delta_k^n A \\ &= \textstyle\sum_k E\, N_{t_{k+1}^n} \Delta_k^n A \\ &= E \textstyle\sum_k N_{t_{k+1}^n} \Delta_k^n A \\ &\to E \int N_t\, dA_t, \end{aligned}$$

and so $E \int \Delta N_t\, dA_t = 0$, which shows that A is natural. □

Lemma 9.16 *(natural processes, Doléans) Every natural, increasing process is predictable.*

Proof: Fix any natural, increasing process A. Consider a bounded martingale M, and a predictable time $\tau < \infty$ announced by $\sigma_1, \sigma_2, \ldots$. Since $M^\tau - M^{\sigma_k}$ is again a bounded martingale, and A is natural, we get by dominated convergence $E\Delta M_\tau \Delta A_\tau = 0$. Taking $M_t = P(B \mid \mathcal{F}_t)$ with $B \in \mathcal{F}_\tau$, and using optional sampling, we obtain a.s. $M_\tau = 1_B$ and

$$M_{\tau-} \leftarrow M_{\sigma_n} = P(B \mid \mathcal{F}_{\sigma_k}) \to P(B \mid \mathcal{F}_{\tau-}).$$

Hence, $\Delta M_\tau = 1_B - P(B \mid \mathcal{F}_{\tau-})$ a.s., and so

$$\begin{aligned} E(\Delta A_\tau; B) &= E\Delta A_\tau P(B \mid \mathcal{F}_{\tau-}) \\ &= E\Big\{E(\Delta A_\tau \mid \mathcal{F}_{\tau-}); B\Big\}. \end{aligned}$$

Since $B \in \mathcal{F}_\tau$ was arbitrary, we get $\Delta A_\tau = E(\Delta A_\tau \mid \mathcal{F}_{\tau-})$ a.s., and so the process $A'_t = \Delta A_\tau 1\{\tau \le t\}$ is predictable, by Lemma 9.8 (iii). It is also natural, since $E \int \Delta M_t \, dA'_t = E\Delta M_\tau \Delta A_\tau = 0$ for any bounded martingale M.

An elementary construction gives $\{t > 0; \Delta A_t > 0\} \subset \bigcup_n [\tau_n]$ a.s., for some optional times $\tau_n < \infty$, and by Lemmas 9.9 and 9.11 we can choose the latter to be predictable. By the previous argument, the process $A^1_t = \Delta A_{\tau_1} 1\{\tau_1 \le t\}$ is both natural and predictable, and in particular $A - A^1$ is again natural. Proceeding by induction, we conclude that the jump component A^d of A is predictable. Since the remainder $A - A^d$ is continuous and adapted, hence predictable, the predictability of A follows. $\Box$

Proof of Theorem 9.14: The sufficiency is obvious, and the uniqueness holds by Lemma 9.12. To prove the necessity, we may assume, by Lemma A1.3, that X is of class (D). Then Lemma 9.15 yields $X = M + A$, for some martingale M and natural, increasing process A. Finally, A is predictable by Lemma 9.16. $\Box$

The two conditions in Lemma 9.16 are in fact equivalent.

Corollary 9.17 *(natural random measures, Doléans) An integrable random measure is natural iff it is predictable.*

Proof: It remains to prove that any integrable, increasing, predictable process A is natural. By Lemma 9.15, we have $A = M + B$ for some uniformly integrable martingale M, and a natural, increasing process B, and Lemma 9.16 shows that B is predictable. But then $A = B$ a.s. by Lemma 9.12, and so A is natural. $\Box$

This yields a basic connection between predictable times and processes.

Theorem 9.18 *(predictable times and sets, Meyer) For any optional time τ, these conditions are equivalent:*

(i) *τ is predictable,*

(ii) *the interval $[\tau,\infty)$ is predictable,*

(iii) *$E\Delta M_\tau = 0$, for every bounded martingale M.*

Proof (Chung & Walsh): Since (i) $\Rightarrow$ (ii) $\Leftrightarrow$ (iii) by Lemma 9.8 (iii) and Theorem 9.17, it remains to show that (iii) $\Rightarrow$ (i). Then introduce the martingale $M_t = E(e^{-\tau}|\,\mathcal{F}_t)$ and super-martingale

$$\begin{aligned} X_t &= e^{-\tau\wedge t} - M_t \\ &= E\Big(e^{-\tau\wedge t} - e^{-\tau}\Big|\,\mathcal{F}_t\Big) \geq 0, \end{aligned}$$

and note that $X_\tau = 0$ a.s. by optional sampling. Writing $\sigma = \inf\{t \geq 0;\ X_t X_{t-} = 0\}$, we get $\{t \geq 0;\ X_t = 0\} = [\sigma,\infty)$ a.s. (cf. FMP 7.31), and in particular $\sigma \leq \tau$. By optional sampling, $E(e^{-\sigma} - e^{-\tau}) = EX_\sigma = 0$, and so $\sigma = \tau$ a.s., which implies $X_t X_{t-} > 0$, a.s. on $[0,\tau)$. Finally, (iii) yields

$$\begin{aligned} EX_{\tau-} &= E(e^{-\tau} - M_{\tau-}) \\ &= E(e^{-\tau} - M_\tau) \\ &= EX_\tau = 0, \end{aligned}$$

and so $X_{\tau-} = 0$ a.s. But then τ is announced by the optional times $\tau_n = \inf\{t;\ X_t < n^{-1}\}$, and (i) follows. □

This leads in particular to a short proof of the following basic fact.

Corollary 9.19 *(restrictions of predictable times) For any predictable time τ and set $A \in \mathcal{F}_{\tau-}$, the restriction τ_A is again predictable.*

Proof: The process $1_A 1\{\tau \leq t\} = 1\{\tau_A \leq t\}$ is predictable by Lemma 9.8 (iii), and so the time τ_A is predictable by Theorem 9.18. □

We can also prove a partial extension of Lemma 9.12:

Theorem 9.20 *(predictable martingales) A local martingale is predictable, iff it is a.s. continuous.*

Proof: Let M be a predictable local martingale. For any optional time $\tau < \infty$, the process

$$M_t^\tau = M_t 1_{[0,\tau]}(t) + M_\tau 1_{(\tau,\infty)}(t), \quad t \geq 0,$$

is again predictable by Lemma 9.7, and so, by localization, we may assume that M is uniformly integrable. Now fix any $\varepsilon > 0$, and introduce the optional time $\tau = \inf\{t > 0;\ |\Delta M_t| > \varepsilon\}$. Since the process $\Delta M_t = M_t - M_{t-}$ is predictable, so is the set $A = \{t > 0;\ |\Delta M_t| > \varepsilon\}$, and then also the interval $[\tau,\infty) = A \cup (\tau,\infty)$. Thus, τ is predictable by Theorem 9.18. Choosing an announcing sequence (τ_n) for τ, and using optional sampling, martingale convergence, and Lemmas 9.6 (iii) and 9.8 (ii), we get a.s.

$$\begin{aligned} M_{\tau-} \leftarrow M_{\tau_n} &= E(M_\tau | \mathcal{F}_{\tau_n}) \\ &\to E(M_\tau | \mathcal{F}_{\tau-}) = M_\tau, \end{aligned}$$

which implies $\tau = \infty$ a.s. Since ε was arbitrary, it follows that M is a.s. continuous. □

Given a Borel space S, we say that a random measure ξ on $(0,\infty)\times S$ is *adapted, predictable*, or *locally integrable*, if the process $\xi_t B = \xi\{(0,t]\times B\}$ has the corresponding property for every $B \in \hat{\mathcal{S}}$. For the adaptedness or predictability, it is clearly equivalent to assume the relevant property for the measure-valued process ξ_t. A process V on $\mathsf{R}_+ \times S$ is said to be *predictable*, if it is $\mathcal{P}\otimes\mathcal{S}$ measurable, where $\mathcal{P}$ denotes the predictable σ-field on $\mathsf{R}_+\times\Omega$. The following fundamental result defines the *compensator* $\hat{\xi}$ of a random measure ξ on $(0,\infty)\times S$.

Theorem 9.21 *(compensation of random measure, Grigelionis, Jacod) Let ξ be a locally integrable, adapted random measure on $(0,\infty)\times S$, where S is Borel. Then there exists an a.s. unique, predictable random measure $\hat{\xi}$ on $(0,\infty)\times S$, such that $E\,\xi V = E\,\hat{\xi}V$ for every predictable process $V \ge 0$ on $\mathsf{R}_+\times S$.*

Proof: For singleton S, the process $X_t = \xi[0,t]$, $t \ge 0$, is locally integrable and increasing, hence a local sub-martingale. Hence, Theorem 9.14 yields an a.s. unique, locally integrable, increasing, predictable process A, such that $M = X - A$ is a local martingale. For fixed $\omega \in \Omega$, the process A determines a measure $\hat{\xi}$ on $(0,\infty)$ with $\hat{\xi}[0,t] = A_t$ for all $t \ge 0$, and we note that $\hat{\xi}$ is a predictable random measure.

Now let (τ_n) be a localizing sequence for M and A, so that for every n, the process $M_n = M^{\tau_n}$ is a uniformly integrable martingale, whereas $A_n = A^{\tau_n}$ is integrable and increasing. Let ξ_n and $\hat{\xi}_n$ denote the restrictions of ξ and $\hat{\xi}$ to the interval $[0,\tau_n]$. By optional sampling, $EM_n(\tau) = 0$ for every optional time τ, and so $E\,\xi_n[0,\tau] = E\,\hat{\xi}_n[0,\tau]$. Thus, $E\,\xi_n V = E\,\hat{\xi}_n V$ holds when $V = 1_{[0,\tau]}$, and a monotone-class argument, based on Lemma 9.7, yields the same relation for any predictable process $V \ge 0$. Finally, $E\,\xi V = E\,\hat{\xi}V$ for every V, by monotone convergence.

We turn to a general Borel space S. Since ξ is locally integrable, we may easily construct a predictable process $V > 0$ on $\mathsf{R}_+\times S$, such that $E\xi V < \infty$. If $\zeta = V\cdot\xi$ has compensator $\hat{\zeta}$, then ξ has compensator $\hat{\xi} = V^{-1}\cdot\hat{\zeta}$, by Lemma 9.13. Thus, we may henceforth assume that $E\|\xi\| = 1$. Writing $\eta = \xi(\cdot\times S)$, and applying Theorem 1.23 to the Campbell measures $P\otimes\eta$ and $P\otimes\xi$, restricted to the σ-fields $\mathcal{P}$ and $\mathcal{P}\otimes\mathcal{S}$, respectively, we obtain a $\mathcal{P}$-measurable probability kernel ρ from $\Omega\times\mathsf{R}_+$ to S, such that $P\otimes\xi = P\otimes\eta\otimes\rho$.

Letting $\hat{\eta}$ be the compensator of η, we define $\hat{\xi} = \hat{\eta} \otimes \rho$, which is again a predictable random measure, by Lemma 9.13.

Now consider any predictable process $V \geq 0$ on $\mathsf{R}_+ \times S$, and note that the process $Y_t = \rho_t V(\cdot, t)$ is again predictable, by Lemma 9.13 (ii). The compensation property for singletons S yields $E\,\hat{\xi}V = E\,\hat{\eta}Y = E\,\eta Y = E\,\xi V$. Finally, $\hat{\xi}$ is a.s. unique, by the uniqueness in Theorem 9.14. □

We can now extend the decomposition of optional times in Lemma 9.9 to random measures. An rcll process X or filtration $\mathcal{F}$ is said to be *quasi-leftcontinuous* (*ql-continuous* for short), if $\Delta X_\tau = 0$ a.s. on $\{\tau < \infty\}$ or $\mathcal{F}_\tau = \mathcal{F}_{\tau-}$, respectively, for every predictable time τ. We further say that X has *accessible jumps*, if $\Delta X_\tau = 0$ a.s. on $\{\tau < \infty\}$ for every totally inaccessible time τ. A random measure ξ on $(0, \infty)$ is said to be *ql-continuous*, if $\xi\{\tau\} = 0$ a.s. for every predictable time τ. It is further said to be *accessible*, if it is purely atomic with $\xi\{\tau\} = 0$ a.s. for every totally inaccessible time τ, where $\mu\{\infty\} = 0$ for measures μ on R_+.

Theorem 9.22 *(decomposition of random measures) Every adapted random measure ξ on $(0, \infty)$ has an a.s. unique decomposition $\xi = \xi^c + \xi^q + \xi^a$, where ξ^c is diffuse, and $\xi^q + \xi^a$ is purely atomic, ξ^q is ql-continuous, and ξ^a is accessible. Furthermore, ξ^a is a.s. supported by $\bigcup_n [\tau_n]$, for some predictable times $\tau_1, \tau_2, \ldots$ with disjoint graphs. If ξ is locally integrable with compensator $\hat{\xi}$, then $\xi^c + \xi^q$ has compensator $(\hat{\xi})^c$, and ξ is ql-continuous iff $\hat{\xi}$ is a.s. continuous.*

Proof: Subtracting the predictable component ξ^c, we may take ξ to be purely atomic. Put $\eta = \hat{\xi}$ and $A_t = \eta(0, t]$. Consider the locally integrable process $X_t = \sum_{s \leq t} (\Delta A_s \wedge 1)$ with compensator $\hat{X}$, and define

$$A_t^q = \int_0^{t+} 1\{\Delta \hat{X}_s = 0\}\, dA_s, \quad A_t^a = A_t - A_t^q, \qquad t \geq 0.$$

For any predictable time $\tau < \infty$, the graph $[\tau]$ is again predictable, by Theorem 9.18, and so by Theorem 9.21

$$\begin{aligned} E(\Delta A_\tau^q \wedge 1) &= E\Big(\Delta X_\tau;\, \Delta \hat{X}_\tau = 0\Big) \\ &= E\Big(\Delta \hat{X}_\tau;\, \Delta \hat{X}_\tau = 0\Big) = 0, \end{aligned}$$

which shows that A^q is ql-continuous.

Now let $\tau_{n0} = 0$ for $n \in \mathsf{N}$, and define recursively the random times

$$\tau_{nk} = \inf\Big\{t > \tau_{n,k-1};\, \Delta \hat{X}_t \in 2^{-n}(1, 2]\Big\}, \quad n, k \in \mathsf{N}.$$

They are all predictable by Theorem 9.18, and $\{t > 0;\, \Delta A_t^a > 0\} = \bigcup_{n,k} [\tau_{nk}]$ a.s., by the definition of A^a. Thus, for any totally inaccessible time τ, we have $\Delta A_\tau^a = 0$ a.s. on $\{\tau < \infty\}$, which shows that A^a has accessible jumps.

If $A = B^q + B^a$ is another decomposition with the stated properties, then $Y = A^q - B^q = B^a - A^a$ is ql-continuous with accessible jumps, and so by Lemma 9.9, $\Delta Y_\tau = 0$ a.s. on $\{\tau < \infty\}$ for any optional time τ, which means that Y is a.s. continuous. Since it is also purely discontinuous, we get $Y = 0$ a.s., which proves the asserted uniqueness.

When A is locally integrable, we may write instead $A^q = 1\{\Delta\hat{A} = 0\} \cdot A$, and note that $(\hat{A})^c = 1\{\Delta\hat{A} = 0\} \cdot \hat{A}$. For any predictable process $V \geq 0$, we get by Theorem 9.21

$$\begin{aligned} E\int V dA^q &= E\int 1\{\Delta\hat{A} = 0\}\, V dA \\ &= E\int 1\{\Delta\hat{A} = 0\}\, V d\hat{A} \\ &= E\int V d(\hat{A})^c, \end{aligned}$$

which shows that A^q has compensator $(\hat{A})^c$. □

The classification of optional times is clarified by the following characterizations in terms of compensators. By the *compensator* of an optional time τ, we mean the compensator of the associated random measure $\xi = \delta_\tau$ or process $X_t = 1\{\tau \leq t\}$.

Corollary 9.23 *(compensation of optional times) Let τ be an optional time with compensator η. Then*

(i) *τ is predictable, iff $\eta = \delta_\sigma$ a.s. for some random time σ,*
(ii) *τ is accessible, iff η is a.s. purely atomic,*
(iii) *τ is totally inaccessible, iff η is a.s. diffuse.*

In general, τ has accessible part τ_D, where $D = \{\eta\{\tau\} > 0\}$.

Proof: (i) If τ is predictable, then so is $\xi = \delta_\tau$ by Theorem 9.18, and so $\eta = \xi$ a.s. Conversely, if $\eta = \delta_\sigma$ for some optional time σ, then the latter is predictable by Theorem 9.18, and so by Theorem 9.21,

$$\begin{aligned} P\{\sigma = \tau < \infty\} &= E\,\xi\{\sigma\} = E\,\eta\{\sigma\} \\ &= P\{\sigma < \infty\} \\ &= E\|\eta\| = E\|\xi\| \\ &= P\{\tau < \infty\}, \end{aligned}$$

which implies $\tau = \sigma$ a.s. Hence, τ is predictable.

(ii) Accessibility of τ means that ξ is accessible, which holds by Theorem 9.22 iff $\eta = \eta^d$ a.s.

(iii) Total inaccessibility of τ means that ξ is ql-continuous, which holds by Theorem 9.22 iff $\eta = \eta^c$ a.s. □

The next result characterizes ql-continuity of both filtrations and martingales.

Theorem 9.24 *(quasi-left continuity, Meyer) For any filtration $\mathcal{F}$, these conditions are equivalent:*

(i) *Every accessible time is predictable,*

(ii) $\mathcal{F}_{\tau-} = \mathcal{F}_\tau$ *on* $\{\tau < \infty\}$, *for every predictable time* τ,

(iii) $\Delta M_\tau = 0$ *a.s. on* $\{\tau < \infty\}$, *for every martingale* M *and predictable time* τ.

If the basic σ-field in Ω is taken to be $\mathcal{F}_\infty$, then $\mathcal{F}_\tau = \mathcal{F}_{\tau-}$ on $\{\tau = \infty\}$ for every optional time τ, and the relation in (ii) extends to all of Ω.

Proof, (i) $\Rightarrow$ (ii): Fix a predictable time τ and a set $B \in \mathcal{F}_\tau \cap \{\tau < \infty\}$. Then $[\tau_B] \subset [\tau]$, and so τ_B is accessible. By (i), it is even predictable, and so the process $X_t = 1\{\tau_B \le t\}$ is predictable, by Theorem 9.18. Noting that $X_\tau 1\{\tau < \infty\} = 1_B$, we get $B \in \mathcal{F}_{\tau-}$ by Lemma 9.8 (ii), which shows that $\mathcal{F}_\tau = \mathcal{F}_{\tau-}$ on $\{\tau < \infty\}$.

(ii) $\Rightarrow$ (iii): Fix a martingale M, and a bounded, predictable time τ with announcing sequence (τ_n). By optional sampling and Lemma 9.6 (iii), we get by (ii) a.s.

$$\begin{aligned} M_{\tau-} \leftarrow M_{\tau_n} &= E(M_\tau \,|\, \mathcal{F}_{\tau_n}) \\ &\to E(M_\tau \,|\, \mathcal{F}_{\tau-}) \\ &= E(M_\tau \,|\, \mathcal{F}_\tau) = M_\tau, \end{aligned}$$

which implies $\Delta M_\tau = 0$ a.s.

(iii) $\Rightarrow$ (i): Let τ be accessible. Then Lemma 9.9 yields $[\tau] \subset \bigcup_n [\tau_n]$ a.s., for some predictable times τ_n. By (iii), we have $\Delta M_{\tau_n} = 0$ a.s. on $\{\tau_n < \infty\}$ for every martingale M, and so $\Delta M_\tau = 0$ a.s. on $\{\tau < \infty\}$. Hence, τ is predictable by Theorem 9.18. □

Next, we show how random measures and their compensators are affected by predictable maps.

Lemma 9.25 *(predictable mapping) Let ξ be an adapted random measure on $(0,\infty) \times S$ with compensator η, consider a predictable process $V : \mathbb{R}_+ \times S \to S$, and put $W_{t,x} = (t, V_{t,x})$. Suppose that the random measure $\xi \circ W^{-1}$ is a.s. locally finite. Then $\xi \circ W^{-1}$ is again adapted with compensator $\eta \circ W^{-1}$.*

Proof: Since the identity map on $\mathbb{R}_+$ is trivially predictable, the mapping W on $\mathbb{R}_+ \times S$ is again predictable. To see that $\xi \circ W^{-1}$ is adapted, let $t \ge 0$ and $B \in \mathcal{S}$ be arbitrary, and note that

$$\left(\xi \circ W^{-1}\right)([0,t] \times B) = \xi\Big\{(s,x);\, s \le t,\, V_{s,x} \in B\Big\}.$$

Since V is predictable and hence progressively measurable, the set on the right belongs to $\mathcal{F}_t \otimes \mathcal{B}_{[0,t]} \otimes \mathcal{S}$, and so the right-hand side is $\mathcal{F}_t$-measurable, by the adaptedness of ξ and Fubini's theorem.

To identify the compensator of $\xi \circ W^{-1}$, consider any predictable process $U \geq 0$ on R_+ and function $f \in \mathcal{S}_+$, and define

$$X_{t,x} = U_t f(x), \quad t \geq 0, \ x \in S.$$

Then the process

$$\begin{aligned} Y_{t,x} &= (X \circ W)(t,x) \\ &= X(t, V_{t,x}) = U_t f(t, V_{t,x}), \quad t \geq 0, \ x \in K, \end{aligned}$$

is again predictable on $\mathsf{R}_+ \times S$, and so by compensation,

$$\begin{aligned} E(\xi \circ W^{-1})X &= E\xi(X \circ W) = E\xi Y \\ &= E\eta Y = E\eta(X \circ W) \\ &= E(\eta \circ W^{-1})X, \end{aligned}$$

which extends by a monotone-class argument to any predictable process $X \geq 0$ on $\mathsf{R}_+ \times S$. It remains to note that the random measure $\eta \circ W^{-1}$ is again predictable, by Lemma 9.13 (i). □

At this point, we may insert a simple conditioning property, needed in the next chapter.

Lemma 9.26 *(conditional compensator) Let ξ be an $\mathcal{F}$-adapted, T-marked point process on $(0,\infty)$ with $\mathcal{F}$-compensator η. Then η remains an $\mathcal{F}$-compensator of ξ, conditionally on $\mathcal{F}_0$.*

Proof: It is enough to verify that $E(\xi V | \mathcal{F}_0) = E(\eta V | \mathcal{F}_0)$ a.s. for any $\mathcal{F}$-predictable process $V \geq 0$, which holds since the process $1_A V$ is again $\mathcal{F}$-predictable for any $A \in \mathcal{F}_0$. □

We conclude this section with a norm estimate for compensators.

Theorem 9.27 *(norm inequality, Garsia, Neveu) Let ξ be a locally integrable random measure on $(0,\infty) \times S$ with compensator $\hat{\xi}$, and consider an $\mathcal{F}$-predictable process $V \geq 0$ on $\mathsf{R}_+ \times S$. Then*

$$\|\hat{\xi} V\|_p \leq p \, \|\xi V\|_p, \quad p \geq 1.$$

Proof: By localization and monotone convergence, we may assume that $E\|\xi\| < \infty$ and V is bounded. Writing $X_t = (V \cdot \xi)_t$ and $A_t = (V \cdot \hat{\xi})_t$, and putting $X_\infty = \zeta$ and $A_\infty = \alpha$, we get a.s. by compensation

$$\begin{aligned} E(\alpha - A_t \,|\, \mathcal{F}_t) &= E(\zeta - X_t \,|\, \mathcal{F}_t) \\ &\leq E(\zeta \,|\, \mathcal{F}_t), \qquad t \geq 0. \end{aligned} \tag{6}$$

Now fix any $r > 0$, and define $\tau_r = \inf\{t;\, A_t \geq r\}$. Since A is right-continuous, we have $\{\tau_r \leq t\} = \{A_t \geq r\}$, and since the latter set is predictable, by the predictability of A, the time τ_r is predictable by Theorem

9.18. Hence, by optional sampling and Lemma 9.6, (6) remains true with t replaced by τ_r-. Since τ_r is $\mathcal{F}_{\tau_r-}$-measurable by the same lemma, we obtain

$$\begin{aligned} E(\alpha - r;\, \alpha > r) &\le E(\alpha - r;\, \tau_r < \infty) \\ &\le E\big(\alpha - A_{\tau_r-};\, \tau_r < \infty\big) \\ &\le E(\zeta;\, \tau_r < \infty) \\ &\le E(\zeta;\, \alpha \ge r). \end{aligned}$$

When $p > 1$, let $p^{-1} + q^{-1} = 1$, and conclude by elementary calculus, Fubini's theorem, and Hölder's inequality that

$$\begin{aligned} \|\alpha\|_p^p &= p^2 q^{-1} E \int_0^\alpha (\alpha - r)\, r^{p-2}\, dr \\ &= p^2 q^{-1} \int_0^\infty E(\alpha - r;\, \alpha > r)\, r^{p-2}\, dr \\ &\le p^2 q^{-1} \int_0^\infty E(\zeta;\, \alpha > r)\, r^{p-2}\, dr \\ &= p^2 q^{-1} E\, \zeta \int_0^\alpha r^{p-2}\, dr \\ &= p\, E\, \zeta \alpha^{p-1} = p\, \|\zeta\|_p\, \|\alpha\|_p^{p-1}. \end{aligned}$$

If $\|\alpha\|_p > 0$, we may divide both sides by $\|\alpha\|_p^{p-1}$ to get the required inequality $\|\alpha\|_p \le p\|\zeta\|_p$. The statements for $\|\alpha\|_p = 0$ or $p = 1$ are obvious. □

9.3 Predictable Invariance and Time Change

For any Borel space S, we define an *S-marked point process* on $(0, \infty)$ as a simple point process ξ on $(0, \infty) \times S$, such that a.s. $\xi(\{t\} \times S) \le 1$ for all $t > 0$. If ξ is adapted to a filtration $\mathcal{F}$, the associated compensator $\hat{\xi}$ exists, since ξ is locally integrable. We say that ξ is *ql-continuous*, if $\xi(\{\tau\} \times S) = 0$ a.s. for every predictable time τ. By Theorem 9.22, it is equivalent that $\hat{\xi}$ be a.s. continuous, in the sense that $\sup_t \hat{\xi}(\{t\} \times S) = 0$ a.s.

Theorem 9.28 *(Cox reduction) For any ql-continuous, S-marked point process ξ on $(0, \infty)$ with compensator $\hat{\xi}$, and a predictable process $V\colon \mathsf{R}_+ \times S \to T$, define some random measures $\zeta = \xi \circ V^{-1}$ and $\eta = \hat{\xi} \circ V^{-1}$ on T. Then ζ is $\mathcal{F}_0$-conditionally Poisson with intensity η, iff η is $\mathcal{F}_0$-measurable.*

Proof: The necessity is obvious. Now let η be $\mathcal{F}_0$-measurable. Fix any measurable function $f \ge 0$ on T with bounded support, define $\psi(x) = 1 - e^{-x}$, and introduce the process

$$X_t = \xi_t(f \circ V) - \hat{\xi}_t(\psi \circ f \circ V), \quad t \ge 0,$$

where $\xi_t g = (g \cdot \xi)_t$ and similarly for $\hat{\xi}_t$. Writing $M = e^{-X}$, and noting that for $\Delta\xi_s = \delta_{\tau_s}$,

$$\begin{aligned}\psi(\Delta X_s) &= \psi\{\Delta\xi_s(f\circ V)\}\\ &= \psi\circ f\circ V_s(\tau_s)\\ &= \Delta\xi_s(\psi\circ f\circ V),\end{aligned}$$

we get, by the substitution rule for functions of bounded variation,

$$\begin{aligned}M_t-1 &= -\sum_{s\le t} M_{s-}\psi(\Delta X_s) - \int_0^t M_{s-}dX_s^c\\ &= -\sum_{s\le t} M_{s-}\int(\psi\circ f\circ V_s)\,\Delta\xi_s + \int_0^t M_{s-}\int(\psi\circ f\circ V_s)\,d\hat{\xi}_s\\ &= \int_0^t M_{s-}\int(\psi\circ f\circ V_s)\,d(\hat{\xi}-\xi)_s.\end{aligned}$$

Since M_{s-} and V_s are predictable, and $\xi_s-\hat{\xi}_s$ is a martingale measure, we conclude that M is a local martingale. Furthermore,

$$\begin{aligned}0 < M_t &\le \exp\Big\{\hat{\xi}_t(\psi\circ f\circ V)\Big\}\\ &\le e^{\eta(\psi\circ f)} < \infty, \qquad t\ge 0,\end{aligned}$$

where the bounds are $\mathcal{F}_0$-measurable. Thus, M is conditionally a bounded martingale, and so $E(M_\infty\,|\,\mathcal{F}_0)=1$ a.s. Noting that

$$\begin{aligned}X_\infty &= \xi(f\circ V) - \hat{\xi}(\psi\circ f\circ V)\\ &= \zeta f - \eta(\psi\circ f),\end{aligned}$$

we get

$$E(e^{-\zeta f}\,|\,\mathcal{F}_0) = e^{-\eta(\psi\circ f)} \ \text{ a.s.},$$

and Lemma 3.1 (ii) shows that ζ is conditionally Poisson η. □

Choosing V in the last theorem to be the identity map, we see in particular that a ql-continuous, marked point process is Poisson, iff its compensator is a.s. non-random. Here we prove a more general statement. Say that a marked point process ξ has *$\mathcal{F}$-independent increments* if $\theta_t\xi \perp\!\!\!\perp \mathcal{F}_t$ for all t.

Theorem 9.29 *(independent increments, Watanabe, Jacod) Let ξ be an $\mathcal{F}$-adapted, T-marked point process on* $(0,\infty)$ *with $\mathcal{F}$-compensator $\hat{\xi}$. Then ξ has $\mathcal{F}$-independent increments iff $\hat{\xi}$ is a.s. non-random.*

Proof: The necessity is obvious. Now let $\hat{\xi}$ be a.s. non-random, hence of the form

$$\hat{\xi} = \alpha + \textstyle\sum_k(\delta_{s_k}\otimes\beta_k) = \alpha + \sum_k \nu_k,$$

for some measure α on $\mathsf{R}_+\times T$ with $\alpha(\{s\}\times T)\equiv 0$, some distinct points $s_1,s_2,\ldots>0$, and some measures β_k on T with $\|\beta_k\|\le 1$. Given a function $f\ge 0$ on $\mathsf{R}_+\times T$ with bounded support, we introduce the process

$$X_t = \xi_t f - \alpha_t(\psi\circ f) + \sum_{s_k\le t}\log\Big\{1-\beta_k(\psi\circ f_{s_k})\Big\}, \quad t\ge 0,$$

where $\xi_t f = (f \cdot \xi)_t$, $\alpha_t g = (g \cdot \alpha)_t$, and $f_s = f(s, \cdot)$. Writing $M = e^{-X}$, we get by the elementary substitution rule

$$M_t = 1 - \sum_{s \le t} M_{s-} \psi(\Delta X_s) + \int_0^t M_{s-} \int (\psi \circ f_s)\, d\alpha_s.$$

Now for $s = s_k$,

$$\begin{aligned}\psi(\Delta X_s) &= 1 - \exp\Big(-\Delta \xi_s f - \log\Big\{1 - \Delta\hat{\xi}_s(\psi \circ f)\Big\}\Big) \\ &= 1 - \frac{\exp(-\Delta\xi_s f)}{1 - \Delta\hat{\xi}_s(\psi \circ f)} \\ &= \frac{\Delta(\xi - \hat{\xi})_s(\psi \circ f)}{1 - \Delta\hat{\xi}_s(\psi \circ f)},\end{aligned}$$

which agrees with $\Delta\xi_s(\psi \circ f)$ when $\Delta\hat{\xi}_s = 0$. Hence,

$$M_t = 1 + \int_0^t \frac{M_{s-}}{1 - \Delta\hat{\xi}_s(\psi \circ f)} \int (\psi \circ f_s)\, d(\hat{\xi} - \xi)_s.$$

Since M is bounded, M_{s-} and $\Delta\xi_s$ are predictable, and $\hat{\xi} - \xi$ is a martingale measure, we conclude that M is a uniformly integrable martingale. In particular $EM_\infty = 1$, which gives

$$-\log E e^{-\xi f} = \alpha(\psi \circ f) - \textstyle\sum_k \log\Big\{1 - \beta_k(\psi \circ f_{s_k})\Big\},$$

and so Lemma 3.18 shows that ξ has independent increments. More generally, the martingale property yields $\theta_t \xi \perp\!\!\!\perp \mathcal{F}_t$ for all $t \ge 0$, which means that ξ has $\mathcal{F}$-independent increments. $\square$

Theorem 9.28 shows how any ql-continuous, simple point process on $(0, \infty)$ can be reduced to Poisson by a suitable random time change. Here we state a corresponding multi-variate result:

Corollary 9.30 *(time change to Poisson, Papangelou, Meyer) Let $\xi_1, \dots, \xi_n$ form a ql-continuous point process on $(0, \infty) \times \{1, \dots, n\}$ with a.s. unbounded compensators $A_1, \dots, A_n$. Then the processes $\zeta_k = \xi_k \circ A_k^{-1}$ are independent, unit rate Poisson processes on R_+.*

Proof: Note that $\hat{\xi}_k \circ A_k^{-1} = \lambda$ a.s. for all k, and apply Theorem 9.28. $\square$

The previous Poisson reduction is straightforward when the compensator is unbounded, but in general we need a rather subtle randomization. Recall from (1) that $\mathcal{G}$ is said to be a *standard extension* of the filtration $\mathcal{F}$, if $\mathcal{F}_t \subset \mathcal{G}_t \perp\!\!\!\perp_{\mathcal{F}_t} \mathcal{F}$ for all $t \ge 0$. This ensures that $\mathcal{G}$ will inherit all adaptedness and conditioning properties of $\mathcal{F}$.

Theorem 9.31 *(extended Poisson reduction) Let ξ be a simple, $\mathcal{F}$-adapted point process on R_+ with continuous compensator A, and define*

$$T_s = \inf\{t \geq 0;\, A_t > s\}, \quad \mathcal{G}_s = \mathcal{F}_{T_s}, \qquad s \geq 0.$$

Then there exist a standard extension $\mathcal{H}$ of $\mathcal{G}$ and a unit rate $\mathcal{H}$-Poisson process η on R_+, such that a.s. $\eta = \xi \circ A^{-1}$ on $[0, A_\infty]$, and $\xi = \eta \circ T_-^{-1}$ on R_+.

Proof: Consider a unit rate Poisson process $\zeta \perp\!\!\!\perp \mathcal{F}$ with induced filtration $\mathcal{Z}$, and put $\mathcal{H}_t = \mathcal{G}_t \vee \mathcal{Z}_t$. Since $\mathcal{G} \perp\!\!\!\perp \mathcal{Z}$, we note that $\mathcal{H}$ is a standard extension of both $\mathcal{G}$ and $\mathcal{Z}$, and in particular ζ remains $\mathcal{H}$-Poisson. Now define $\eta = \xi \circ A^{-1} + (A_\infty, \infty) \cdot \zeta$. Then clearly η has $\mathcal{H}$-compensator λ, and so η is unit rate $\mathcal{H}$-Poisson, by Theorem 9.29.

To see how the relation $\eta = \xi \circ A^{-1}$ on $[0, A_\infty]$ can be inverted into $\xi = \eta \circ T_-^{-1}$ a.s., we note that the random set

$$U = \bigcup_{h>0} \bigl\{t > 0;\, A_t = A_{t-h}\bigr\}$$

is predictable with $\hat{\xi} U = 0$ a.s. This gives $E\,\xi U = E\,\hat{\xi} U = 0$, and so $\xi U = 0$ a.s. Thus, if $\eta\{\sigma\} = 1$, the corresponding atom of ξ lies a.s. at $T_{\sigma-}$. □

The basic Poisson reduction also yields a similar property for predictable integrals with respect to a p-stable Lévy process. For simplicity, we consider only the case where $p < 1$.

Proposition 9.32 *(time change of stable integrals) Let X be a strictly p-stable Lévy process, where $p \in (0, 1)$, and consider a predictable process $V \geq 0$, such that the process $A = V^p \cdot \lambda$ is a.s. finite but unbounded. Then $(V \cdot X) \circ A^{-1} \overset{d}{=} X$.*

Here the process $(V \cdot X) \circ A^{-1}$ is defined by

$$\{(V \cdot X) \circ A^{-1}\}_t = \int_0^\infty 1\{A_s \leq t\}\, V_s\, dX_s, \quad t \geq 0.$$

When $X_t = \xi[0, t]$ a.s. for some purely discontinuous, signed random measure ξ, the assertion becomes $(V \cdot \xi) \circ A^{-1} \overset{d}{=} \xi$.

Proof: Let η be the jump point process of X on $\mathsf{R}_+ \times (\mathsf{R} \setminus \{0\})$, so that

$$X_t = \iint_0^{t+} x\, \eta(ds\, dx), \quad t \geq 0,$$

and note that η is Poisson with intensity and compensator $E\eta = \hat{\eta} = \lambda \otimes \nu$, where $\nu(dx) = c_\pm |x|^{-p-1} dx$ for $\pm x > 0$. Define a predictable mapping T on

$\mathsf{R}_+ \times \mathsf{R}$ by $T_{s,x} = (A_s,\, x\,V_s)$. Then for measurable $B \subset \mathsf{R}\setminus\{0\}$,

$$\begin{aligned}(\lambda\otimes\nu)\circ T^{-1}([0,t]\times B) &= (\lambda\otimes\nu)\Big\{(s,x);\, A_s \le t,\ x\,V_s \in B\Big\}\\ &= \int \nu\Big\{x;\, x\,V_s\in B\Big\}\,1\{A_s\le t\}\,ds\\ &= \nu B\int 1\{A_s \le t\}\,V_s^p\,ds = t\,\nu B,\end{aligned}$$

which shows that $(\lambda\otimes\nu)\circ T^{-1} = \lambda\otimes\nu$ a.s. Hence, Theorem 9.28 yields $\eta\circ T^{-1} \overset{d}{=} \eta$ a.s., and so

$$\begin{aligned}\Big\{(V\cdot X)\circ A^{-1}\Big\}_t &= \int_0^\infty 1\{A_s\le t\}\,V_s\,dX_s\\ &= \iint_0^\infty 1\{A_s\le t\}\,V_s\,x\,\eta(ds\,dx)\\ &= \iint_0^{t+} u\,(\eta\circ T^{-1})(dr\,du),\end{aligned}$$

which is a process with the same distribution as X. □

For $I = [0,1]$ or R_+, we say that a random measure ξ on $I\times S$ is $\mathcal{F}$-*exchangeable*, if for any $t\in I$ it is conditionally exchangeable on $(t,1]$ or (t,∞), respectively, given $\mathcal{F}_t$. In particular, any exchangeable random measure ξ on I is $\mathcal{F}$-exchangeable for the right-continuous and complete filtration $\mathcal{F}$ induced by ξ.

Theorem 9.33 *(predictable mapping) Let ξ be an $\mathcal{F}$-exchangeable random measure on $I\times S$, where $I=[0,1]$ or R_+ and S is Borel, and consider an $\mathcal{F}$-predictable process $Y\colon I\to I$, such that $\lambda\circ Y^{-1}=\lambda$ a.s. Then $\xi\circ Y^{-1}\overset{d}{=}\xi$.*

For $I=\mathsf{R}_+$, this is an easy consequence of Theorem 9.28. The proof for $I=[0,1]$ is more difficult and relies on the corresponding discrete-time version, where exchangeability is defined as before with respect to a discrete filtration $\mathcal{F}=(\mathcal{F}_n)$. An integer-valued random time τ is said to be $\mathcal{F}$-*predictable* if $\{\tau=n\}\in\mathcal{F}_{n-1}$ for all $n\ge 0$, where $\mathcal{F}_{-1}=\mathcal{F}_0$.

Lemma 9.34 *(predictable sampling) Let $\xi=(\xi_1,\xi_2,\dots)$ be a finite or infinite, $\mathcal{F}$-exchangeable random sequence in S indexed by I, and let $\tau_1,\dots,\tau_n$ be a.s. distinct, I-valued, $\mathcal{F}$-predictable times. Then*

$$(\xi_{\tau_1},\dots,\xi_{\tau_n}) \overset{d}{=} (\xi_1,\dots,\xi_n). \tag{7}$$

Proof: First let $I=\{1,\dots,n\}$, so that the τ_k form a random permutation of I. The inverse permutation $\alpha_1,\dots,\alpha_n$ is *predictable*, in the sense that

$$\{\alpha_j = k\} = \{\tau_k = j\}\in\mathcal{F}_{j-1},\quad j,k\in I.$$

For every $m \in \{0, \ldots, n\}$, we define an $\mathcal{F}_{m-1}$-measurable permutation α_1^m, $\ldots, \alpha_n^m$ of I by

$$(\alpha_1^m, \ldots, \alpha_m^m) = (\alpha_1, \ldots, \alpha_m), \qquad \alpha_{m+1}^m < \cdots < \alpha_n^m. \tag{8}$$

Fix any measurable functions $f_1, \ldots, f_n \geq 0$ on S. Using the first relation in (8), the $\mathcal{F}_m$-measurability of (α_j^{m+1}) and (α_j^m), the $\mathcal{F}$-exchangeability of ξ, and the disintegration theorem, and writing $E^{\mathcal{F}} = E(\cdot \mid \mathcal{F})$, we get for $0 \leq m < n$

$$\begin{aligned}
E \prod_{k \leq n} f_{\alpha_k^{m+1}}(\xi_k) &= E E^{\mathcal{F}_m} \prod_{k \leq n} f_{\alpha_k^{m+1}}(\xi_k) \\
&= E \prod_{k \leq m} f_{\alpha_k}(\xi_k)\, E^{\mathcal{F}_m} \prod_{k > m} f_{\alpha_k^{m+1}}(\xi_k) \\
&= E \prod_{k \leq m} f_{\alpha_k}(\xi_k)\, E^{\mathcal{F}_m} \prod_{k > m} f_{\alpha_k^{m}}(\xi_k) \\
&= E E^{\mathcal{F}_m} \prod_{k \leq n} f_{\alpha_k^m}(\xi_k) \\
&= E \prod_{k \leq n} f_{\alpha_k^m}(\xi_k).
\end{aligned}$$

Summing over m, and noting that $\alpha_k^n \equiv \alpha_k$ and $\alpha_k^0 \equiv k$, we get

$$E \prod_{k \leq n} f_k(\xi_{\tau_k}) = E \prod_{k \leq n} f_{\alpha_k}(\xi_k) = E \prod_{k \leq n} f_k(\xi_k),$$

and (7) follows.

Next let $I = \{1, \ldots, m\}$ for some $m > n$. Then extend (τ_k) to a random permutation of I with $\tau_{n+1} < \cdots < \tau_m$. The latter variables are predictable, since

$$\{\tau_r \leq k\} = \Big(\textstyle\sum_{i \leq n} 1\{\tau_i > k\} \geq r - k\Big) \in \mathcal{F}_{k-1}, \quad r > n,\ k \leq m.$$

Hence, the result in the previous case yields $(\xi_{\tau_1}, \ldots, \xi_{\tau_m}) \stackrel{d}{=} \xi$, and (7) follows.

Finally, let $I = \mathsf{N}$. Fixing any $m \in \mathsf{N}$, we introduce the random times

$$\tau_j^m = \tau_j 1\{\tau_j \leq m\} + (m + j) 1\{\tau_j > m\}, \quad j \leq n,$$

which are bounded by $m + n$. They are also predictable, since

$$\begin{aligned}
\{\tau_j^m \leq k\} &= \{\tau_j \leq k\} \in \mathcal{F}_{j-1}, \qquad\qquad k \leq m, \\
\{\tau_j^m = m + j\} &= \{\tau_j > n\} \in \mathcal{F}_{m-1} \subset \mathcal{F}_{m+j-1}.
\end{aligned}$$

As $m \to \infty$, we get a.s. $\tau_j^m \to \tau_j$ for all j, and so $\xi_{\tau_j^m} = \xi_{\tau_j}$ for large enough m. Here $(\xi_{\tau_j^m}) \stackrel{d}{=} \xi$, by the result in the previous case, and (7) follows. □

To state the next result, we say that a predictable set in I is *simple*, if it is a finite union of predictable intervals of the form $(\sigma, \tau]$, where $\sigma \leq \tau$ are optional times, taking values in a fixed, finite subset of Q_I.

Lemma 9.35 *(approximation) Let $A_1, \dots, A_m$ be disjoint, $\mathcal{F}$-predictable sets in $[0,1]$ with non-random, rational lengths. Then for any $n \in \mathsf{N}$ there exist some disjoint, simple, $\mathcal{F}$-predictable sets $U_{nj} \subset [0,1]$, $j \le m$, such that $\lambda U_{nj} \equiv \lambda A_j$ and $\sum_j \lambda(A_j \triangle U_{nj}) \xrightarrow{P} 0$.*

Proof: Let $\mathcal{C}$ be the class of simple, predictable intervals $(\sigma, \tau]$, and write $\mathcal{D}$ for the class of predictable sets $A \subset (0,1]$, admitting approximations $E\lambda(A \triangle U_n) \to 0$ by simple, predictable sets U_n. Then $\mathcal{C}$ is a π-system contained in the λ-system $\mathcal{D}$, and so a monotone-class argument yields $\mathcal{D} \supset \sigma(\mathcal{C}) = \mathcal{P}$, which means that every predictable set A admits the stated approximation.

Now let $A_1, \dots, A_m$ be such as stated, and proceed as above, to produce some simple, predictable sets U_{nj}, $j \le m$, such that $\sum_j E\lambda(A_j \triangle U_{nj}) \to 0$. Taking differences, if necessary, we may adjust the U_{nj} to become disjoint in j for fixed n. We can also adjust the lengths, so that $\lambda U_{nj} = \lambda A_j$ for all n and j. Since the required changes are of the same order as the original approximation errors, the previous approximation property remains valid, and the assertion follows. □

To justify the preceding approximations, we need to show that the corresponding approximation errors for the point process ξ are negligible.

Lemma 9.36 *(continuity) Consider an $\mathcal{F}$-exchangeable point process ξ on $[0,1]$, and some $\mathcal{F}$-predictable sets $A_1, A_2, \ldots \subset [0,1]$ with $\lambda A_n \xrightarrow{P} 0$. Then $\xi A_n \xrightarrow{P} 0$.*

Proof: Since $\xi(t,1]$ is invariant under measure-preserving transformations of $(t,1]$, we may assume $\kappa = \xi[0,1]$ to be $\mathcal{F}_0$-measurable, in which case we may take $\kappa = n$ to be constant. Then define $M_t = (n - \xi_t)/(1-t)$ for $t \in [0,1)$, so that $\xi_t = n - (1-t)M_t$. Integration by parts yields

$$\xi_t = n - \int_0^t (1-s)\, dM_s + \int_0^t M_s\, ds, \quad t \in [0,1], \quad t \in [0,1].$$

Here the first integral is an $\mathcal{F}$-martingale on $[0,1)$, since M is an $\mathcal{F}$-martingale, by the $\mathcal{F}$-exchangeability of ξ. Hence, ξ has $\mathcal{F}$-compensator $\hat{\xi} = M \cdot \lambda$, and in particular $E\,\lambda M = E\,\xi[0,1] = n$ a.s.

Now assume $\lambda A_n \xrightarrow{P} 0$, so that $E\,\lambda A_n \to 0$, by dominated convergence. Then for any subsequence $N' \subset \mathsf{N}$, we have $1_{A_n} \to 0$ a.e. $P \otimes \lambda$ along a further subsequence N''. Hence, by dominated convergence,

$$E\,\xi A_n = E\,\hat{\xi} A_n = E\,\lambda(1_{A_n} M) \to 0, \quad n \in N'',$$

and so $\xi A_n \xrightarrow{P} 0$ along N''. The assertion follows, since N' was arbitrary. □

We may finally express the predictable interval unions above in terms of predictable times.

Lemma 9.37 *(enumeration) Let A be a simple, $\mathcal{F}$-predictable set in* $[0,1]$ *or* R_+*, with interval endpoints in* Z_+/n*, and define* $\mathcal{G}_k = \mathcal{F}_{k/n}$*,* $k \le n$*. Then* $A = \bigcup_{j \le m}(\tau_{j-1}, \tau_j]/n$*, for some* $\mathcal{G}$*-predictable times* $\tau_0 < \cdots < \tau_m$*.*

Proof: The time $\sigma = \inf A$ satisfies $\{\sigma < t\} \in \mathcal{F}_t$ for all t and is therefore $\mathcal{F}$-optional, by the right-continuity of $\mathcal{F}$. Hence, $n\sigma = \tau_1 - 1$ is $\mathcal{G}$-optional, which means that τ_1 is $\mathcal{G}$-predictable. Furthermore, the stochastic interval $(\sigma, \sigma + n^{-1}]$ is $\mathcal{F}$-predictable, and so the same thing is true for the difference $A' = A \setminus (\sigma, \sigma + n^{-1}]$. The assertion now follows by induction. $\square$

Proof of Theorem 9.33, $I = \mathsf{R}_+$: By Theorem 3.37, it is enough to consider random measures of the form $\xi = \lambda \otimes \alpha + \eta$, for some Cox process η on $\mathsf{R}_+ \times S$ directed by $\lambda \otimes \nu$, where α and ν are suitable random measures on S with $\eta \perp\!\!\!\perp_\nu \alpha$. The latter are clearly ξ-measurable and invariant under measure-preserving transformations of ξ. Extending $\mathcal{F}$ if necessary, we may then assume α and ν to be $\mathcal{F}_0$-measurable. Such an extension clearly preserves the predictability of the processes Y. Conditioning on $\mathcal{F}_0$, we may finally reduce to the case where α and ν are non-random. The result is then an immediate consequence of Theorem 9.28.

$I = [0,1]$: Using Theorem 3.37, and proceeding as before, we may reduce to the case where ξ is a uniform randomization of some point process β on S, where the latter may be taken to be non-random. Now fix a predictable mapping Y on $[0,1]$ with $\lambda \circ V^{-1} = \lambda$ a.s. For any disjoint, rational intervals $I_1, \ldots, I_m \subset [0,1]$, the sets $A_j = V^{-1}I_j$ are clearly disjoint and predictable with $\lambda A_j = \lambda I_j$ a.s., $j \le m$. By Lemma 9.35, we may choose some simple, predictable sets U_{nj}, disjoint for fixed n, such that $\lambda U_{nj} = \lambda A_j = \lambda I_j$ a.s., and $\lambda(A_j \triangle U_{nj}) \xrightarrow{P} 0$ as $n \to \infty$ for fixed j. Then Lemma 9.36 yields $\xi(U_{nj} \times B) \xrightarrow{P} \xi(A_j \times B)$ for every $B \in \hat{\mathcal{S}}$, and so it suffices to show that, for every fixed $n \in \mathsf{N}$, the sets of random variables $\xi(U_{nj} \times B)$ and $\xi(I_j \times B)$ have the same joint distribution.

Dropping the subscript n, we may assume that $U_1, \ldots, U_m$ are disjoint, simple, and predictable with $\lambda U_j = \lambda I_j$. Let $n \in \mathsf{N}$ be such that each U_j is a finite union of intervals $I_{ni} = (i-1, i]/n$. By Lemma 9.37, we may write $U_j = \bigcup_{k \le h_j} I_{n,\tau_{jk}}$, where the times τ_{jk} are distinct and predictable with respect to the discrete filtration $\mathcal{G}_k = \mathcal{F}_{k/n}$. Since the increments $\xi(I_{ni} \times \cdot)$ are $\mathcal{G}$-exchangeable in i for fixed n, Lemma 9.34 yields

$$\Big\{\xi(U_1 \times \cdot), \ldots, \xi(U_m \times \cdot)\Big\} \stackrel{d}{=} \Big\{\xi(I_1 \times \cdot), \ldots, \xi(I_m \times \cdot)\Big\},$$

and the assertion follows. $\square$

Our next result clarifies how the compensator of a point process is affected by a randomization.

Lemma 9.38 *(randomization) Consider an $\mathcal{F}$-adapted point process ξ on $(0,\infty)\times S$ with compensator $\hat{\xi}$, and a probability kernel $\nu\colon \mathsf{R}_+\times S\to T$. Let η be a ν-randomization of ξ on $\mathsf{R}_+\times S\times T$, and let $\mathcal{G}$ be the right-continuous filtration generated by $\mathcal{F}$ and η. Then $\mathcal{G}$ is a standard extension of $\mathcal{F}$, and the $\mathcal{G}$-compensator of η equals $\hat{\eta}=\hat{\xi}\otimes\nu$.*

Proof: For any $t\ge 0$, let ξ_t and η_t denote the restrictions of ξ and η to $[0,t]\times S$ and $[0,t]\times S\times T$, respectively, and put $\eta_t'=\eta-\eta_t$. Then clearly

$$\eta \perp\!\!\!\perp_{\xi} \mathcal{F}, \qquad \eta_t \perp\!\!\!\perp_{\xi} \eta_t', \qquad \eta_t \perp\!\!\!\perp_{\xi_t} \xi.$$

Using the first relation, combining with the other two, and invoking the chain rule for conditional independence (FMP 6.8), we obtain

$$\eta_t \perp\!\!\!\perp_{\xi,\eta_t'} \mathcal{F}, \qquad \eta_t \perp\!\!\!\perp_{\xi_t} (\eta_t',\mathcal{F}),$$

and so

$$\eta_t \perp\!\!\!\perp_{\mathcal{F}_t} (\eta_t',\mathcal{F}), \qquad (\eta_t,\mathcal{F}_t) \perp\!\!\!\perp_{\mathcal{F}_t} (\eta_t',\mathcal{F}).$$

Approximating from the right in the last relation gives $\mathcal{G}_t \perp\!\!\!\perp_{\mathcal{F}_t} \mathcal{F}$, which shows that $\mathcal{G}$ is a standard extension of $\mathcal{F}$.

Using the latter property, the chain rule for conditional expectations, the relation $\eta \perp\!\!\!\perp_{\xi} \mathcal{F}$, Fubini's theorem, and the definitions of randomization and compensation, we get on $(t,\infty)\times S\times T$, for arbitrary $t\ge 0$

$$\begin{aligned}
E(\eta\,|\,\mathcal{G}_t) &= E(\eta\,|\,\mathcal{F}_t) = E\Big\{E(\eta\,|\mathcal{F}_t,\xi)\,\Big|\,\mathcal{F}_t\Big\}\\
&= E\Big\{E(\eta\,|\,\xi)\,\Big|\,\mathcal{F}_t\Big\} = E(\xi\otimes\nu\,|\,\mathcal{F}_t)\\
&= E(\xi\,|\,\mathcal{F}_t)\otimes\nu = E(\hat{\xi}\,|\,\mathcal{F}_t)\otimes\nu\\
&= E(\hat{\xi}\otimes\nu\,|\,\mathcal{F}_t) = E(\hat{\xi}\otimes\nu\,|\,\mathcal{G}_t).
\end{aligned}$$

Since $\hat{\eta}=\hat{\xi}\otimes\nu$ is $\mathcal{F}$-predictable and hence even $\mathcal{G}$-predictable, we conclude that $\hat{\eta}$ is indeed a $\mathcal{G}$-compensator of η. □

The last result has a partial converse, stated here under some simplifying assumptions.

Lemma 9.39 *(symmetry and independence) Consider an adapted point process ξ on $(0,\infty)$ with marks in $S\times T$, such that $\bar{\xi}=\xi(\cdot\times S\times T)$ is a.s. locally finite and unbounded, and let ξ have compensator $\hat{\xi}=\eta\otimes\nu$, for some predictable random measure η on $\mathsf{R}_+\times S$ and a fixed probability measure ν on T. Then $\zeta=\xi(\cdot\times T)$ is an S-marked point process on $(0,\infty)$ with compensator η, and the associated T-marks γ_k are i.i.d. ν. However, ζ and $\gamma=(\gamma_k)$ need not be independent.*

Proof: The assertion for ζ being obvious by projection, we turn to the one for γ. Letting $\bar{\xi}$ have supporting times $\tau_1<\tau_2<\cdots$, and putting $\tau_0=0$, we introduce the discrete filtration $\mathcal{G}_k=\mathcal{F}_{\tau_k}$, $k\in\mathsf{Z}_+$. Then γ is clearly

$\mathcal{G}$-adapted, and since $\bar{\xi}$ has exactly one point in each interval $(\tau_{k-1}, \tau_k]$, we see by optional sampling that γ has $\mathcal{G}$-compensator

$$\begin{aligned}\hat{\gamma}_k &= E\Big(\eta\{(\tau_{k-1}, \tau_k] \times S\}\Big|\, \mathcal{G}_{k-1}\Big)\nu \\ &= E\Big(\zeta\{(\tau_{k-1}, \tau_k] \times S\}\Big|\, \mathcal{G}_{k-1}\Big)\nu = \nu,\end{aligned}$$

a.s. for each $k \in \mathsf{N}$. It remains to apply the discrete-time version of Theorem 9.29.

To prove the last assertion, let ξ be a unit rate Poisson process on R_+, and attach some i.i.d. $U(0,1)$ marks $\gamma_1, \gamma_2, \ldots$ to the points $\tau_1, \tau_2, \ldots$ of ξ. Then form a new process ξ', by deleting all points of ξ in the interval $(\tau_1, \tau_1 + \gamma_1]$. The remaining marks γ'_k are again i.i.d. ν, e.g. by Lemma 9.34, but they are not independent of ξ', since the latter process depends on the first mark γ_1. □

9.4 Discounted Compensator and Predictable Maps

Here we extend the previous time-change theory to optional times and point processes with possibly discontinuous compensators. We begin with an elementary formula for compensators with respect to the induced filtration. When $0 \le s \le t$, let $\int_s^t$ denote the integral over $(s,t]$. For measures μ on $\mathsf{R}_+ \times S$, we write $\bar{\mu} = \mu(\cdot \times S)$ and $\mu_t = \mu([0,t] \times \cdot)$.

Lemma 9.40 *(induced compensator) Let (τ, χ) be a random pair in $(0, \infty] \times S$ with distribution μ, where S is Borel. Then the induced compensator η of (τ, χ) is given by*

$$\eta_t B = \int_0^{t\wedge\tau} \frac{\mu(dr \times B)}{1 - \bar{\mu}(0, r)}, \quad B \in \mathcal{S},\ t \ge 0. \tag{9}$$

Proof: The process $\eta_t B$ is clearly predictable for every $B \in \mathcal{S}$. Writing $\xi = \delta_{\tau,\chi}$, it remains to show that $M_t = \xi_t B - \eta_t B$ is a martingale, hence that $E(M_t - M_s; A) = 0$ for all $s < t$ and $A \in \mathcal{F}_s$. Since $M_t = M_s$ on the set $\{\tau \le s\}$, and $\{\tau > s\}$ is an a.s. atom of $\mathcal{F}_s$, it is enough to show that $E(M_t - M_s) = 0$ or $EM_t = 0$. Then conclude from Fubini's theorem that

$$\begin{aligned}E\,\eta_t B &= E\int_0^{t\wedge\tau} \frac{\mu(dr \times B)}{1 - \bar{\mu}(0, r)} \\ &= \int_0^\infty \bar{\mu}(dx) \int_0^{t\wedge x} \frac{\mu(dr \times B)}{1 - \bar{\mu}(0, r)} \\ &= \int_0^t \frac{\mu(dr \times B)}{1 - \bar{\mu}(0, r)} \int_{r-}^\infty \bar{\mu}(dx) \\ &= \mu_t B = E\,\xi_t B.\end{aligned}$$

□

We turn to the case of a general filtration.

Theorem 9.41 *(discounted compensator) For any adapted pair (τ, χ) in $(0, \infty) \times S$ with compensator η, there exists a unique, predictable random measure ζ on $(0, \tau] \times S$ with $\|\zeta\| \le 1$, satisfying* (9) *with μ replaced by ζ. Writing $Z_t = 1 - \bar{\zeta}_t$, we have a.s.*

(i) $\zeta = Z_- \cdot \eta$, *and Z is the unique solution of $Z = 1 - Z_- \cdot \bar{\eta}$,*

(ii) $Z_t = \exp(-\bar{\eta}^c_t) \prod_{s \le t} (1 - \Delta\bar{\eta}_s), \quad t \ge 0,$

(iii) $\Delta\bar{\eta} < 1$ *on* $[0, \tau)$, *and* ≤ 1 *on* $[\tau]$,

(iv) *Z is non-increasing and ≥ 0 with $Z_{\tau-} > 0$,*

(v) *$Y = Z^{-1}$ satisfies $dY_t = Y_t \, d\bar{\eta}_t$ on $\{Z_t > 0\}$.*

The process in (ii) is a special case of the *Doléans exponential.*

Proof: (iii) The time $\sigma = \inf\{t \ge 0;\ \Delta\bar{\eta}_t \ge 1\}$ is optional, by an elementary approximation. Hence, the random interval (σ, ∞) is predictable, and so the same thing is true for the graph $[\sigma] = \{\Delta\bar{\eta} \ge 1\} \setminus (\sigma, \infty)$. Thus,

$$P\{\tau = \sigma\} = E\,\bar{\xi}[\sigma] = E\,\bar{\eta}[\sigma].$$

Since also

$$1\{\tau = \sigma\} \le 1\{\sigma < \infty\} \le \bar{\eta}[\sigma],$$

by the definition of σ, we obtain

$$1\{\tau = \sigma\} = \bar{\eta}[\sigma] = \Delta\bar{\eta}_\sigma \ \text{ a.s. on } \{\sigma < \infty\},$$

which implies $\Delta\bar{\eta}_t \le 1$ and $\tau = \sigma$ a.s. on the same set.

(i)–(ii): If (9) holds with $\mu = \zeta$, for some random measure ζ on $(0, \tau] \times S$, then the process $Z_t = 1 - \bar{\zeta}_t$ satisfies $d\bar{\eta}_t = Z_{t-}^{-1} d\bar{\zeta}_t = -Z_{t-}^{-1} dZ_t$, and so $dZ_t = -Z_t d\bar{\eta}_t$, which implies $Z_t - 1 = -(Z \cdot \bar{\eta})_t$, or $Z = 1 - Z_- \cdot \bar{\eta}$. Conversely, any solution Z to the latter equation yields a measure solving (9) with $B = S$, and so a general solution to (9) is given by $\zeta = Z_- \cdot \eta$. Furthermore, Lemma A2.1 shows that the equation $Z = 1 - Z_- \cdot \bar{\eta}$ has the unique solution (ii). In particular, Z is predictable, and the predictability of ζ follows by Lemma 9.13.

(iv) Since $1 - \Delta\bar{\eta} \ge 0$ a.s. by (iii), (ii) shows that Z is a.s. non-increasing with $Z \ge 0$. Since also $\sum_t \Delta\bar{\eta}_t \le \bar{\eta}_\tau < \infty$ a.s., and $\sup_{t<\tau} \Delta\bar{\eta}_t < 1$ a.s. by (iii), we have $Z_{\tau-} > 0$.

(v) By elementary integration by parts, we get on the set $\{t \ge 0;\ Z_t > 0\}$

$$0 = d(Z_t Y_t) = Z_{t-}\, dY_t + Y_t\, dZ_t.$$

Using the chain rule for Stieltjes integrals, along with the equation $Z = 1 - Z_- \cdot \bar{\eta}$, we obtain

$$\begin{aligned} dY_t &= -Z_{t-}^{-1}\, Y_t\, dZ_t \\ &= Z_{t-}^{-1}\, Y_t\, Z_{t-}\, d\bar{\eta}_t = Y_t\, d\bar{\eta}_t. \end{aligned}$$

□

The following martingale property will play a key role in the sequel. Recall our convention $0/0 = 0$.

Lemma 9.42 *(fundamental martingale) For τ, χ, η, ζ, and Z as in Theorem 9.41, consider a predictable process V on $\mathsf{R}_+ \times S$ with $\zeta|V| < \infty$ a.s. and $\zeta V = 0$ a.s. on $\{Z_\tau = 0\}$, and define*

$$U_{t,x} = V_{t,x} + Z_t^{-1} \int_0^t \!\! \int V d\zeta, \qquad t \ge 0,\ x \in S, \tag{10}$$

$$M_t = U_{\tau,\chi} 1\{\tau \le t\} - \int_0^t \!\! \int U d\eta, \quad t \ge 0. \tag{11}$$

Then

(i) *M exists on $[0,\infty]$ and satisfies $M_\infty = V_{\tau,\chi}$ a.s.,*

(ii) *when $E|U_{\tau,\chi}| < \infty$, we have $EV_{\tau,\chi} = 0$, and M becomes a uniformly integrable martingale, with $\|M^*\|_p \lesssim \|V_{\tau,\chi}\|_p$ for every $p > 1$.*

Proof: Write $Y = Z^{-1}$. Using the conditions on V, the definition of ζ, and Theorem 9.41 (iv), we obtain

$$\eta|V| = \zeta(Y_-|V|) \le Y_{\tau-}\, \zeta|V| < \infty.$$

Next, we see from (10) and Theorem 9.41 (v) that

$$\eta|U - V| \le \zeta|V|\, \bar\eta Y = (Y_\tau - 1)\, \zeta|V| < \infty,$$

whenever $Z_\tau > 0$. If instead $Z_\tau = 0$, we have

$$\begin{aligned} \eta|U - V| &\le \zeta|V| \int_0^{\tau-} Y\, d\bar\eta \\ &= (Y_{\tau-} - 1)\, \zeta|V| < \infty. \end{aligned}$$

In either case, U is η-integrable and M is well defined.

Now let $t \ge 0$ with $Z_t > 0$. Using (10), Theorem 9.41 (v), Fubini's theorem, and the definition of ζ, we get for any $x \in S$

$$\begin{aligned} \int_0^t \!\! \int (U - V)\, d\eta &= \int_0^t Y_s\, d\bar\eta_s \int_0^s \!\! \int V d\zeta \\ &= \int_0^t dY_s \int_0^s \!\! \int V d\zeta \\ &= \int_0^t \!\! \int (Y_t - Y_{s-})\, V_{s,y}\, d\zeta_{s,y} \\ &= U_{t,x} - V_{t,x} - \int_0^t \!\! \int V d\eta. \end{aligned}$$

Simplifying, and combining with (10), we get

$$\int_0^t \!\! \int U d\eta = U_{t,x} - V_{t,x} = Y_t \int_0^t \!\! \int V d\zeta. \tag{12}$$

To extend to general t, suppose that $Z_\tau = 0$. Using the previous version of (12), the definition of ζ, and the conditions on V, we get

$$\begin{aligned}
\int_0^\tau \!\! \int U d\eta &= \int_0^{\tau-} \!\! \int U d\eta + \int_{[\tau]} \!\! \int U d\eta \\
&= Y_{\tau-} \int_0^{\tau-} \!\! \int V d\zeta + \int_{[\tau]} \!\! \int V d\eta \\
&= Y_{\tau-} \int_0^{\tau} \!\! \int V d\zeta = 0 \\
&= U_{\tau,x} - V_{\tau,x},
\end{aligned}$$

which shows that (12) is generally true. In particular,

$$\begin{aligned}
V_{\tau,\chi} &= U_{\tau,\chi} - \int_0^\tau \!\! \int U d\eta \\
&= M_\tau = M_\infty .
\end{aligned}$$

(ii) If $E\,|U_{\tau,\chi}| < \infty$, then $E\,\eta|U| < \infty$ by compensation, which shows that M is uniformly integrable. For any optional time σ, the process $1_{[0,\sigma]}\,U$ is again predictable and $E\eta$-integrable, and so by the compensation property and (11),

$$\begin{aligned}
EM_\sigma &= E\,(U_{\tau,\chi};\, \tau \le \sigma) - E \int_0^\sigma \!\! \int U d\eta \\
&= E\,\delta_{\tau,\chi}(1_{[0,\sigma]}\,U) - E\,\eta(1_{[0,\sigma]}\,U) = 0,
\end{aligned}$$

which shows that M is a martingale. Thus, in view of (i),

$$EV_{\tau,\chi} = EM_\infty = EM_0 = 0.$$

Furthermore, by Doob's inequality,

$$\|M^*\|_p \lesssim \|M_\infty\|_p = \|V_{\tau,\chi}\|_p, \quad p > 1. \qquad \square$$

We also need a multi-variate version of the last result. Say that the optional times $\tau_1, \dots, \tau_n$ are *orthogonal*, if they are a.s. distinct, and the atomic parts of their compensators have disjoint supports.

Corollary 9.43 *(product moments) For each $j \le m$, consider an adapted pair (τ_j, χ_j) in $(0, \infty) \times S_j$, and a predictable process V_j on $\mathsf{R}_+ \times S_j$, where the τ_j are orthogonal, and define Z_j, ζ_j, and U_j as in Theorem 9.41 and* (10). *Fix any $p_1, \dots, p_m > 0$ with $\sum_j p_j^{-1} \le 1$, and suppose that for every $j \le m$,*

$$\zeta_j|V_j| < \infty, \qquad \zeta_j V_j = 0 \;\; a.s. \text{ on } \{Z_j(\tau_j) = 0\},$$
$$E\,|U_j(\tau_j, \chi_j)| < \infty, \qquad E\,|V_j(\tau_j, \chi_j)|^{p_j} < \infty.$$

Then

$$E \prod_{j \le m} V_j(\tau_j, \chi_j) = 0.$$

Proof: Let $\eta_1,\dots,\eta_m$ be compensators of the pairs (τ_j,χ_j), and define the martingales $M_1,\dots,M_m$ as in (11). Fix any $i\neq j$ in $\{1,\dots,m\}$, and choose some predictable times $\sigma_1,\sigma_2,\dots$ as in Theorem 9.22, such that $\{t>0;\ \Delta\bar{\eta}_t>0\}=\bigcup_k[\sigma_k]$ a.s. By compensation and orthogonality, we have for any $k\in\mathsf{N}$

$$\begin{aligned} E\,\delta_{\sigma_k}[\tau_j] &= P\{\tau_j=\sigma_k\} \\ &= E\,\delta_{\tau_j}[\sigma_k]=E\,\eta_j[\sigma_k]=0. \end{aligned}$$

Summing over k gives $\eta_k[\tau_j]=0$ a.s., which shows that $\Delta M_i\,\Delta M_j=0$ a.s. for all $i\neq j$. Integrating repeatedly by parts, we conclude that $M=\prod_j M_j$ is a local martingale. Defining $p\geq 1$ by $p^{-1}=\sum_j p_j^{-1}$, we get by Hölder's inequality, Lemma 9.42 (ii), and the various hypotheses

$$\begin{aligned} \|M^*\|_1 &\leq \|M^*\|_p\leq \prod\nolimits_j \|M_j^*\|_{p_j} \\ &\lesssim \prod\nolimits_j \|V_{\tau_j,\chi_j}\|_{p_j}<\infty. \end{aligned}$$

Thus, M is a uniformly integrable martingale, and so by Lemma 9.42 (i),

$$\begin{aligned} E\prod\nolimits_j V_j(\tau_j,\chi_j) &= E\prod\nolimits_j M_j(\infty) \\ &= EM(\infty)=EM(0)=0. \end{aligned}$$

□

The following time-change property of adapted pairs may be regarded as an extension of Theorem 9.33. A more general conditional version will be given in Lemma 9.52 below.

Theorem 9.44 *(predictable mapping) For each $j\leq m$, consider an adapted pair (τ_j,χ_j) in $(0,\infty)\times S_j$ with discounted compensator ζ_j, and a predictable mapping Y_j of $\mathsf{R}_+\times S_j$ into a probability space (T_j,μ_j), such that the τ_j are orthogonal and $\zeta_j\circ Y_j^{-1}\leq\mu_j$ a.s. Then the images $\sigma_j=Y_j(\tau_j,\chi_j)$ are independent with distributions μ_j.*

Proof: For fixed $B_j\in\mathcal{T}_j$, consider the predictable processes

$$V_j(t,x)=1_{B_j}\circ Y_j(t,x)-\mu_j B_j,\quad t\geq 0,\ x\in S_j,\ j\leq m.$$

By definitions and hypotheses,

$$\begin{aligned} \int_0^t\!\!\int V_j\,d\zeta_j &= \int_0^t\!\!\int 1_{B_j}(Y_j)\,d\zeta_j-\mu_j B_j\,\{1-Z_j(t)\} \\ &\leq Z_j(t)\,\mu_j B_j. \end{aligned}$$

Replacing B_j by B_j^c affects only the sign of V_j, and so by combination,

$$-Z_j(t)\,\mu_j B_j^c\leq\int_0^t\!\!\int V_j\,d\zeta_j\leq Z_j(t)\,\mu_j B_j.$$

In particular, $|\zeta_j V_j| \leq Z_j(\tau_j)$ a.s., and so $\zeta_j V_j = 0$ a.s. on $\{Z_j(\tau_j) = 0\}$. Defining U_j as in (10), we get by the previous estimates

$$\begin{aligned} -1 &\leq -1_{B_j^c} \circ Y_j \\ &= V_j - \mu_j B_j^c \\ &\leq U_j \leq V_j + \mu_j B_j \\ &= 1_{B_j} \circ Y_j \leq 1, \end{aligned}$$

which implies $|U_j| \leq 1$. Letting $\emptyset \neq J \subset \{1, \ldots, m\}$, and applying Corollary 9.43 with $p_j = |J|$ for all $j \in J$, we see that

$$E \prod_{j \in J} \left\{ 1_{B_j}(\sigma_j) - \mu_j B_j \right\} = E \prod_{j \in J} V_j(\tau_j, \chi_j) = 0. \tag{13}$$

Proceeding by induction on $|J|$, we claim that

$$P \bigcap_{j \in J} \{\sigma_j \in B_j\} = \prod_{j \in J} \mu_j B_j, \quad J \subset \{1, \ldots, m\}. \tag{14}$$

For $|J| = 1$, this is immediate from (13). Now suppose that (14) holds for all $|J| < k$. Then for $|J| = k$, we may expand the product in (13), and apply the induction hypothesis to the terms involving at least one factor $\mu_j B_j$. Here all terms but one reduce to $\pm \prod_{j \in J} \mu_j B_j$, while the remaining term equals $P \bigcap_{j \in J} \{\sigma_j \in B_j\}$. Thus, (14) remains true for J, which completes the induction. By a monotone-class argument, the formula for $J = \{1, \ldots, m\}$ extends to $\mathcal{L}(\sigma_1, \ldots, \sigma_m) = \bigotimes_j \mu_j$. □

The last result allows us to extend the classical Poisson reduction in Theorem 9.31, beyond the ql-continuous case. To avoid distracting technicalities, we consider only unbounded point processes.

Theorem 9.45 *(time change to Poisson) Consider a simple, unbounded, $\mathcal{F}$-adapted point process $\xi = \sum_j \delta_{\tau_j}$ on $(0, \infty)$ with compensator η, along with some i.i.d. $U(0,1)$ random variables $\vartheta_1, \vartheta_2, \ldots \perp\!\!\!\perp \mathcal{F}$, and define*

$$\begin{aligned} \rho_t &= 1 - \textstyle\sum_j \vartheta_j \, 1\{t = \tau_j\}, \\ Y_t &= \eta_t^c - \displaystyle\sum_{s \leq t} \log\left(1 - \rho_s \, \Delta\eta_s\right), \qquad t \geq 0. \end{aligned}$$

Then $\tilde{\xi} = \xi \circ Y^{-1}$ is a unit rate Poisson process on R_+.

Proof: Write $\sigma_0 = 0$, and let $\sigma_j = Y_{\tau_j}$ for $j > 0$. We claim that the differences $\sigma_j - \sigma_{j-1}$ are i.i.d. and exponentially distributed with mean 1. Since the τ_j are orthogonal, it is enough to consider σ_1. Letting $\mathcal{G}$ be the right-continuous filtration induced by $\mathcal{F}$ and the pairs (σ_j, ϑ_j), we see from Lemma 9.38 that (τ_1, ϑ_1) has $\mathcal{G}$-compensator $\tilde{\eta} = \eta \otimes \lambda$ on $[0, \tau_1] \times [0, 1]$. Write ζ for the associated discounted version, with projection $\bar{\zeta}$ onto R_+, and put $Z_t = 1 - \bar\zeta(0,t]$. Define the $\mathcal{G}$-predictable processes U and $V = e^{-U}$ on $\mathsf{R}_+ \times [0,1]$ by

$$\begin{aligned} U_{t,r} &= \eta_t^c - \sum_{s<t} \log(1-\Delta\eta_s) - \log(1 - r\,\Delta\eta_t), \\ V_{t,r} &= \exp(-\eta_t^c) \prod_{s<t} (1-\Delta\eta_s)\,(1 - r\,\Delta\eta_t) \\ &= Z_{t-}\,(1 - r\Delta\eta_t), \end{aligned}$$

where the last equality holds by Theorem 9.41 (ii). For any random variable γ with distribution function F, we have

$$\begin{aligned} F_{t-} &= P\{F_\gamma < F_{t-}\} \\ &\le P\{F_\gamma \le F_t\} = F_t, \quad t \in \mathsf{R}. \end{aligned}$$

Thus, $\zeta \circ V^{-1} \le \lambda$ a.s. on $[0,1]$, and so $V(\tau_1, 1-\vartheta_1) = e^{-\sigma_1}$ is $U(0,1)$ by Theorem 9.44, which yields the required distribution for σ_1. □

The last result yields an asymptotic version of the basic Poisson characterization in Theorem 9.29.

Corollary 9.46 *(Poisson convergence, Brown) For every $n \in \mathsf{N}$, let ξ_n be a simple point process on $(0,\infty)$ with compensator η_n, and suppose that $\eta_n[0,t] \xrightarrow{d} t$ for all $t > 0$. Then $\xi_n \xrightarrow{vd} \xi$, where ξ is a unit rate Poisson process on R_+.*

Proof: For any sub-sequence $N' \subset \mathsf{N}$, we have a.s. $\eta_n[0,t] \to t$ for all $t \ge 0$ along a further sub-sequence N''. In particular, $\sup_{t\le c} \eta_n\{t\} \to 0$ a.s. for all $c > 0$. Writing $\xi_n = \sum_j \delta_{\tau_{nj}}$, and defining $\sigma_{nj} = Y_n(\tau_{nj})$ with Y_n as above, we get a.s. for every $j \in \mathsf{N}$

$$|\,\tau_{nj} - \sigma_{nj}| \le \Big|\tau_{nj} - \eta_n[0,\tau_{nj}]\Big| + \Big|\eta_n[0,\tau_{nj}] - \sigma_{nj}\Big| \to 0.$$

Then as $n \to \infty$ along N'',

$$\xi_n = \sum\nolimits_j \delta_{\tau_{nj}} \xrightarrow{vd} \sum\nolimits_j \delta_{\sigma_{nj}} \overset{d}{=} \xi,$$

which yields the required convergence. □

Using Theorem 9.44, we can also derive some useful bounds on the distributions of Z_τ, $Z_{\tau-}$, and η_τ.

Corollary 9.47 *(distributional bounds) Let τ be an $\mathcal{F}$-optional time with compensator η, and define Z as in Theorem 9.41. Then*

(i) $P\{Z_{\tau-} \le x\} \le x \le P\{Z_\tau < x\}, \quad x \in [0,1],$

(ii) $1 - x \le P\{\eta_\tau \ge x\} \le e^{1-x}, \quad x \ge 0,$

(iii) $\|\eta_\tau\|_p \lesssim 1, \quad p > 0.$

Proof: (i) Introduce a $U(0,1)$ random variable $\vartheta \perp\!\!\!\perp \mathcal{F}$, and let $\mathcal{G}$ be the right-continuous filtration generated by (τ, ϑ) and $\mathcal{F}$, so that (τ, ϑ) has $\mathcal{G}$-compensator $\eta \otimes \lambda$ by Lemma 9.38. Define a predictable mapping V from $\mathsf{R}_+ \times [0,1]$ to $[0,1]$ by

$$V_{t,x} = 1 - Z_{t-} - x\,\Delta Z_t, \quad t \geq 0,\ x \in [0,1], \tag{15}$$

and note as before that $(\zeta \otimes \lambda) \circ V^{-1} = \lambda$, a.s. on $[0, 1 - Z_\tau]$. Hence, $V_{\tau,\vartheta}$ is $U(0,1)$ by Theorem 9.44, and (15) yields

$$Z_\tau \leq 1 - V_{\tau,\vartheta} \leq Z_{\tau-}. \tag{16}$$

(ii) By Theorem 9.41 (ii), we have for any $t \geq 0$

$$\begin{aligned} -\log Z_t &= \eta_t^c - \sum_{s \leq t} \log\left(1 - \Delta\eta_t\right) \\ &\geq \eta_t^c + \sum_{s \leq t} \Delta\eta_t = \eta_t \\ &\geq 1 - Z_t. \end{aligned}$$

Hence, by (16) and Theorem 9.41 (iii),

$$\begin{aligned} V_{\tau,\vartheta} &\leq \eta_\tau \leq \eta_{\tau-} + 1 \\ &\leq 1 - \log Z_{\tau-} \\ &\leq 1 - \log\left(1 - V_{\tau,\vartheta}\right), \end{aligned}$$

and so

$$\begin{aligned} 1 - x &\leq P\{\eta_\tau \geq x\} \\ &\leq P\Big\{1 - \log\left(1 - V_{\tau,\vartheta}\right) \geq x\Big\} \\ &= P\Big\{1 - V_{\tau,\vartheta} \leq e^{1-x}\Big\} \leq e^{1-x}. \end{aligned}$$

(iii) The upper bound in (ii) yields

$$\begin{aligned} E\,|\eta_t|^p &= \int_0^\infty P\{|\eta_t|^p \geq r\}\,dr \\ &\leq \int_0^\infty \exp(1 - r^{1/p})\,dr \\ &= e\,p \int_0^\infty e^{-x} x^{p-1} dx \\ &= e\,p\,\Gamma(p) < \infty. \end{aligned}$$

□

9.5 Extended Compensator and Integral Representation

Consider an adapted pair (τ, χ) in $(0,\infty)\times S$, with distribution μ and induced compensator η as in Lemma 9.40, and define $P_\mu = \mathcal{L}(\tau, \chi, \eta)$. For a general filtration $\mathcal{F}$ with associated compensator η, we give an integral representation of $\mathcal{L}(\tau, \chi, \eta)$ in terms of the measures P_μ, allowing a simple probabilistic interpretation. For any measure μ on $\mathbb{R}_+ \times S$, write $\pi_t\mu$ for the restriction of μ to $[0,t] \times S$.

Theorem 9.48 *(integral representation and extended compensator) Consider an adapted pair (τ, χ) in $(0,\infty)\times S$, with compensator η and discounted compensator ζ, where S is Borel. Then*

(i) *there exists a probability measure ν on $\mathcal{P}_{(0,\infty)\times S}$ with*

$$\mathcal{L}(\tau, \chi, \eta) = \int P_\mu \, \nu(d\mu),$$

(ii) *there exists a random probability measure ρ on $(0,\infty) \times S$, such that a.s.*

$$\zeta = \pi_\tau \rho, \qquad \mathcal{L}(\tau, \chi \,|\, \rho) = \rho.$$

The measures ν and $\mathcal{L}(\rho)$ agree and are uniquely determined by $\mathcal{L}(\eta)$. Any measure ν as in (i) *may occur.*

The representation in (i) is equivalent to

$$\mathcal{L}(\tau, \chi, \zeta) = \int \tilde{P}_\mu \, \nu(d\mu) \tag{17}$$

with $\tilde{P}_\mu = \mathcal{L}(\tau, \chi, \pi_\tau\mu)$, for the case where $\mathcal{L}(\tau, \chi) = \mu$, and $\zeta = \pi_\tau\mu$ is the induced, discounted compensator of (τ, χ). Our proof requires several lemmas, beginning with a basic uniqueness property. To appreciate this result, note that the corresponding statement for general point processes is false.

Lemma 9.49 *(uniqueness) For any adapted pairs (τ, χ) in $(0,\infty) \times S$ with compensators η, the distribution $\mathcal{L}(\tau, \chi, \eta)$ is uniquely determined by $\mathcal{L}(\eta)$, regardless of the underlying filtration.*

Proof: Since the process $\eta_t = \eta([0,t] \times \cdot)$ is right-continuous and predictable, the times

$$\tau_r = \inf\big\{t \geq 0;\, \eta_t S \geq r\big\}, \quad r \geq 0,$$

are predictable by Theorem 9.18. Since η is predictable, its restriction $\pi_{\tau_r}\eta$ to $[0,\tau_r] \times S$ is $\mathcal{F}_{\tau_r-}$-measurable by Lemma 9.8 (ii), and so for any measurable sets $A \subset \mathcal{M}_{(0,\infty)\times S}$, the restriction of τ_r to $\{\pi_{\tau_r}\eta \in A\}$ is again predictable, by

Corollary 9.19. Hence, for measurable $B \subset (0,\infty) \times S$, we get by Theorems 9.18 and 9.21

$$\begin{aligned} & P\Big\{(\tau,\chi) \in B,\, \tau \geq \tau_r,\, \pi_{\tau_r}\eta \in A\Big\} \\ & \qquad = E\Big(\eta\Big\{(t,x) \in B;\, t \geq \tau_r\Big\};\, \pi_{\tau_r}\eta \in A\Big). \end{aligned}$$

Since the τ_r are η-measurable, the left-hand side is then determined by $\mathcal{L}(\eta)$, regardless of filtration. The events on the left form a π-system, and so the uniqueness extends by a monotone-class argument to the joint distribution of all random measures

$$1\Big\{(\tau,\chi) \in B,\, \tau \geq \tau_r\Big\}\, \pi_{\tau_r}\eta, \quad B \in \mathcal{R}_+ \otimes \mathcal{S},\ r \geq 0,$$

hence also to the distributions of

$$1_B(\tau,\chi)\eta = \sup_{r \in \mathbb{Q}_+} 1\Big\{(\tau,\chi) \in B,\, \tau \geq \tau_r\Big\}\, \pi_{\tau_r}\eta, \quad B \in \mathcal{R}_+ \otimes \mathcal{S}.$$

The uniqueness of $\mathcal{L}(\tau,\chi,\eta)$ now follows, by another monotone-class argument. □

Say that an $\mathcal{F}$-adapted random pair (τ,χ) in $(0,\infty) \times S$ is *pure*, if its $\mathcal{F}$-compensator agrees with the induced version in Lemma 9.40.

Lemma 9.50 *(conditional independence) For any pure, $\mathcal{F}$-adapted random pair (τ,χ) in $(0,\infty) \times S$, we have*

$$(\tau,\chi) \underset{\tau > t}{\perp\!\!\!\perp} \mathcal{F}_t, \quad t \geq 0 \ \textit{with}\ P\{\tau > t\} > 0.$$

Proof: Let η be the compensator of (τ,χ), so that $EV_{\tau,\chi} = E\,\eta V$ for every predictable process $V \geq 0$ on $\mathsf{R}_+ \times S$. For any $t \geq 0$ and $A \in \mathcal{F}_t$, the process $1_{A \times (t,\infty)}V$ remains predictable, and so

$$E\Big(V_{\tau,\chi};\, \tau > t\,\Big|\,\mathcal{F}_t\Big) = E\Big\{\eta(1_{(t,\infty)}V)\,\Big|\,\mathcal{F}_t\Big\}, \text{ a.s. on } \{\tau > t\},$$

which shows that (τ,χ) remains compensated by η, conditionally on $\mathcal{F}_t$ on the set $\{\tau > t\}$. Writing $\mathcal{L}\{\tau,\chi\,|\,\tau > t\} = \mu_t$, we conclude that (τ,χ) has conditionally discounted compensator $\pi_\tau\mu_t$. Since $\pi_\tau\mu_t \leq \mu_t$, Theorem 9.44 yields

$$\mathcal{L}\{\tau,\chi\,|\,\mathcal{F}_t\} = \mu_t \text{ a.s. on } \{\tau > t\}, \quad t \geq 0,$$

which remains true under conditioning on $\{\tau > t\}$. In particular,

$$\mathcal{L}\{\tau,\chi\,|\,\mathcal{F}_t\} = \mathcal{L}\{\tau,\chi\,|\,\tau > t\} \text{ a.s. on } \{\tau > t\}, \quad t \geq 0,$$

which is equivalent to the asserted relation. □

In the pure case, the last result allows certain adapted processes on $[0,\tau]$ to be extended beyond time τ. Define

$$\operatorname{ess\,sup}\tau = \sup\Big\{t \geq 0;\, P\{\tau > t\} > 0\Big\}.$$

Lemma 9.51 *(extension) Consider a pure, adapted pair* (τ,χ) *in* $(0,\infty)\times S$ *with* $\operatorname{ess\,sup}\tau=\infty$*, along with a left-continuous, adapted process* X *in* S*, where* S *is Polish. Then there exists a left-continuous process* $Y \perp\!\!\!\perp (\tau,\chi)$ *in* S*, with* $X=Y$ *a.s. on* $[0,\tau]$*.*

Proof: For any $t\geq 0$ and measurable $B\subset(t,\infty)\times S$, we get by Lemma 9.50

$$\begin{aligned}\mathcal{L}\Big\{\pi_t X;\,(\tau,\chi)\in B\Big\} &= E\Big(P\Big\{(\tau,\chi)\in B\,\Big|\,\mathcal{F}_t\Big\};\,\pi_t X\in\cdot\Big)\\ &= P\Big\{(\tau,\chi)\in B\,\Big|\,\tau>t\Big\}\,\mathcal{L}(\pi_t X;\,\tau>t)\\ &= P\{(\tau,\chi)\in B\}\,\mathcal{L}(\pi_t X\,|\,\tau>t\,).\end{aligned}\tag{18}$$

Choosing $B=(t,\infty)\times S$ with $s\leq t$, we get in particular

$$\mathcal{L}(\,\pi_s\pi_t X\,|\,\tau>t\,)=\mathcal{L}(\,\pi_s X\,|\,\tau>s\,),\quad 0\leq s\leq t.$$

Hence, the Daniell–Kolmogorov theorem yields a left-continuous process Y in S, satisfying

$$\mathcal{L}(\,\pi_t Y)=\mathcal{L}(\,\pi_t X\,|\,\tau>t\,),\quad t\geq 0.$$

Assuming $Y\perp\!\!\!\perp(\tau,\chi)$ and combining with (18), we get

$$\begin{aligned}\mathcal{L}\Big\{\pi_t X;\,(\tau,\chi)\in B\Big\} &= P\{(\tau,\chi)\in B\}\,\mathcal{L}(\,\pi_t Y)\\ &= \mathcal{L}\Big\{\pi_t Y;\,(\tau,\chi)\in B\Big\}.\end{aligned}\tag{19}$$

Now introduce the random times

$$\tau_n=\max\Big\{k2^{-n}<\tau;\,k\in\mathsf{Z}_+\Big\},\quad n\in\mathsf{N}.$$

For measurable $B\subset(0,\infty)\times S$, we get from (19) with $t=k2^{-n}$

$$\begin{aligned}&\mathcal{L}\Big\{\pi_{\tau_n}X;\,(\tau_n,\chi)\in B,\,\tau_n=k2^{-n}\Big\}\\ &\qquad=\mathcal{L}\Big\{\pi_{\tau_n}Y;\,(\tau_n,\chi)\in B,\,\tau_n=k2^{-n}\Big\}.\end{aligned}$$

Summing over k gives

$$(\tau_n,\,\chi,\,\pi_{\tau_n}X)\overset{d}{=}(\tau_n,\,\chi,\,\pi_{\tau_n}Y),\quad n\in\mathsf{N},$$

and so, by left continuity,

$$(\tau,\,\chi,\,\pi_\tau X)\overset{d}{=}(\tau,\,\chi,\,\pi_\tau Y).$$

By a transfer argument, there exists a left-continuous process $\tilde{Y}$ in S with

$$\pi_\tau X=\pi_\tau\tilde{Y}\ \text{ a.s.},\qquad (\tau,\chi,\tilde{Y})\overset{d}{=}(\tau,\chi,Y).$$

In particular, the latter relation yields $\tilde{Y}\perp\!\!\!\perp(\tau,\chi)$, and so the process $\tilde{Y}$ has the required properties. □

We also need the following refinement of Theorem 9.45.

Lemma 9.52 *(conditional tails) Let τ be an $\mathcal{F}$-optional time with discounted compensator ζ, put $Z_t = 1 - \zeta_t$, and define*

$$\begin{aligned} Y_{t,r} &= 1 - Z_{t-} - r\,\Delta Z_t, && t \ge 0,\ r \in [0,1], \\ L_s &= \inf\Big\{t \ge 0;\ 1 - Z_t \ge s\Big\}, && s \in [0,1). \end{aligned}$$

Choose $\vartheta \perp\!\!\!\perp \mathcal{F}$ to be $U(0,1)$, and define $\sigma = Y_{\tau,\vartheta}$. Then

$$\mathcal{L}(\,\sigma\,|\,\mathcal{F}_{L_s-},\,\sigma > s\,) = \frac{\lambda}{1-s} \ \text{a.s. on } (s,1], \quad s \in [0,1).$$

Proof: Fix any $s \in [0,1)$. Since Z is right-continuous and predictable, the time L_s is predictable by Theorem 9.18. For any $A \in \mathcal{F}_{L_s-}$, the restriction of L_s to A is again predictable by Corollary 9.19. This also applies to the set $A \cap \{L_s \le \tau\}$, which lies in $\mathcal{F}_{L_s-}$ by Lemma 9.6. The process $1_A 1\{Ł_s \le \tau \wedge t\}$ is then predictable by Theorem 9.18, and hence so is the process

$$V_{t,r} = \Big\{1_B(Y_{t,r}) - \lambda B\Big\}\, 1_A\, 1\{L_s \le \tau \wedge t\}, \quad t \ge 0,\ r \in [0,1],$$

for any Borel set $B \subset (s,1]$. Proceeding as in the proof of Theorem 9.44, we get

$$E\Big\{1_B(\sigma) - \lambda B;\ A, L_s \le \tau\Big\} = 0,$$

and since $\sigma \ge s$ implies $L_s \le \tau$, we have

$$P(\,\sigma \in B;\ A\,) = \lambda B \cdot P(\,A;\ L_s \le \tau\,).$$

Combining this with the formula for $B = (s,1]$, we obtain

$$\frac{P(\,\sigma \in B;\ A\,)}{P(\,\sigma > s;\ A\,)} = \frac{\lambda B}{1-s}, \quad B \in \mathcal{B}_{(s,1]},\ \ s \in [0,1),$$

which is equivalent to the asserted relation. $\Box$

We may now prove a weaker, preliminary version of Theorem 9.48 (ii). Write π for projection onto $(0,\infty)$.

Lemma 9.53 *(extended compensator) For any adapted pair (τ,χ) in $(0,\infty) \times S$ with discounted compensator ζ, there exists a random probability measure ρ on $(0,\infty) \times S$, such that a.s.*

$$\zeta = \pi_\tau \rho, \qquad \mathcal{L}(\tau|\,\rho\,) = \rho \circ \pi^{-1}. \tag{20}$$

Proof: Introduce the processes

$$\begin{aligned} &Y_t = 1 - Z_t = \bar{\zeta}_t = \zeta_t S = \zeta([0,t] \times S), && t \ge 0, \\ &L_s = \inf\Big\{t \ge 0,\ Y_t \ge s\Big\}, \quad D_s = (Y \circ L)_{s-}, && s \in [0,1), \\ &T_{t,y} = Y_{t-} + y\Delta Y_t, && t \ge 0,\ y \in [0,1]. \end{aligned}$$

Let $\vartheta \perp\!\!\!\perp \mathcal{F}$ be $U(0,1)$, and put

$$\sigma = T_{\tau,\vartheta}, \qquad \beta = (\zeta \otimes \lambda) \circ (T \otimes I)^{-1}, \tag{21}$$

where $(T \otimes I)_{t,x,y} = (T_{t,y}, x)$, which makes β a random measure on $[0,\sigma] \times S$ with $\beta \circ \pi^{-1} = \lambda$ on $[0,\sigma]$. Letting $\mathcal{G}$ be the filtration generated by the σ-fields $\mathcal{F}_{L_s-}$ and the process δ_σ, we see from Lemma 9.52 that σ is $U(0,1)$ and a pure $\mathcal{G}$-optional time.

The processes L_s, D_s, and β_s are non-decreasing and left-continuous, where the continuity of β is with respect to the weak topology on $\mathcal{M}_S$. Furthermore, L_s and D_s are $\mathcal{F}_{L_s-}$-measurable, in case of D by the predictability of Y. Even β_s is $\mathcal{F}_{L_s-}$-measurable, since for any $A \in \hat{\mathcal{S}}$

$$\beta_s A = \zeta_{L_s} A - \frac{(Y_{L_s} - s)}{(Y_{L_s} - Y_{L_s-})} \, \Delta\zeta_{L_s}(A) \, 1\{L_s < \infty\}.$$

Thus, L, D, and β are $\mathcal{G}$-adapted. Now Lemma 9.51 yields some left-continuous processes $\hat{L}$, $\hat{D}$, $\hat{\beta}$ independent of σ, such that a.s.

$$\hat{L}_s = L_s, \quad \hat{D}_s = D_s, \quad \hat{\beta}_s = \beta_s, \qquad s \in [0,\sigma]. \tag{22}$$

Conditioning on $\{\sigma \geq s\}$, for any $s \in [0,1]$, we see that $\hat{L}_s$, $\hat{D}_s$, and $\hat{\beta}_s$ are a.s. non-decreasing, and $\hat{\beta}_s$ is a.s. continuous. A routine extension yields a random measure $\hat{\beta}$ on $[0,1) \times S$, such that $\hat{\beta}_s A = \hat{\beta}([0,s] \times A)$ for all $s \in [0,1)$ and $A \in \hat{\mathcal{S}}$. Conditioning as before gives $\hat{\beta} \circ \pi^{-1} = \lambda$ a.s., and in particular $\hat{\beta}([0,1) \times S) = 1$ a.s.

We proceed to show that a.s.

$$L_s = \hat{L}_s = \tau, \quad \beta_s = \hat{\beta}_s, \qquad s \in [\sigma, Y_\tau], \tag{23}$$

$$L_s \wedge \hat{L}_s > \tau, \qquad s > Y_\tau. \tag{24}$$

Since τ is a.s. the last point of increase of Y, and $\vartheta > 0$ a.s., we have a.s. $L_s = \tau$ on $[\sigma, Y_\tau]$, and $L_s = \infty$ on (Y_τ, ∞). To prove the remaining relations, we first assume that $\Delta Y_\tau = 0$. Then (23) follows from (22), and for (24) it suffices to show that $\hat{L}_s > \tau$ for $s > Y_\tau$. Now $\Delta Y_\tau = 0$ implies $L_s = \hat{L}_s < L_\sigma = \tau$ a.s. for all $s < \sigma$. Thus, if $\hat{L}_s = \tau$ for some $s > \sigma$, then σ is the left endpoint of a constancy interval for $\hat{L}$. But this is a.s. excluded by the independence $\sigma \perp\!\!\!\perp \hat{L}$, since the paths of $\hat{L}$ have only countably many such points.

If instead $\Delta Y_\tau > 0$, we may condition on σ, and use (22) to get

$$D_{s+} = \begin{cases} \inf\big\{t \geq s;\, L_t > L_{s+}\big\}, & s \in [0,\sigma), \\ \inf\big\{t \geq s;\, \hat{L}_t > \hat{L}_{s+}\big\}, & s \in [0,1). \end{cases}$$

Hence, by comparison and (22),

$$\inf\big\{t \geq s;\, \hat{L}_t > L_{s+}\big\} = \inf\big\{t \geq s;\, L_t > L_{s+}\big\}, \quad s \in [0,\sigma).$$

Now $\Delta Y_\tau > 0$ implies $L_s = \tau$, for large enough $s < \sigma$ a.s., and so for such an s,

$$\inf\Big\{t \ge s;\ \hat{L}_t > \tau\Big\} = \inf\Big\{t \ge s;\ L_t > \tau\Big\} = Y_\tau \ \text{ a.s.},$$

which shows that $\hat{L}$ satisfies (23) and (24).

To prove the second part of (23) when $\Delta Y_\tau > 0$, we note that β_s increases linearly on every constancy interval of L. Conditioning on σ, we get the same property for $\hat{\beta}$ and $\hat{L}$. Since L is constant on $(Y_{\tau-}, Y_\tau]$, and since $\hat{L} = L$ a.s. on the same interval, by the relevant parts of (22) and (23), both χ_s and $\hat{\beta}$ increase linearly on the mentioned interval. Since also $\beta = \hat{\beta}$ a.s., on the a.s. non-empty interval $(Y_{\tau-}, \sigma]$, we get $\beta = \hat{\beta}$ a.s. on the entire interval. This completes the proof of (23) and (24).

Now define a random probability measure ρ on $\mathsf{R}_+\times S$ by

$$\rho = \hat{\beta} \circ (\hat{L} \otimes I)^{-1}, \tag{25}$$

where $(\hat{L} \otimes I)_{s,x} = (\hat{L}_s, x)$. Noting that

$$L \circ T_{t,y} = t, \qquad (t, y) \in \mathsf{R}_+\times [0,1] \ \text{ a.e. } \zeta \circ \pi^{-1} \otimes \lambda,$$

and using (21)$-$(25), we get for any $t \in [0,\tau]$ and $B \in \hat{\mathcal{S}}$

$$\begin{aligned}
\rho_t B &= \hat{\beta}\Big\{(s,x) \in [0,1) \times B;\ \hat{L}_s \le t\Big\}\\
&= \beta\Big\{(s,x) \in [0,1) \times B;\ L_s \le t\Big\}\\
&= (\chi \otimes \lambda)\Big\{(u,x,y) \in \mathsf{R}_+\times B \times [0,1];\ L \circ T_{u,y} \le t\Big\} = \chi_t B,
\end{aligned}$$

which shows that $\rho = \zeta$, a.s. on $[0,\tau] \times S$. Since $\zeta = 0$ on $(\tau, \infty) \times S$, we obtain $\zeta = \pi_\tau \rho$ a.s. Furthermore,

$$\begin{aligned}
\mathcal{L}(\tau \,|\, \rho\,) &= \mathcal{L}(\hat{L}_\sigma \,|\, \rho\,) = E\Big\{\mathcal{L}(\hat{L}_\sigma \,|\, \rho, \hat{L})\Big|\, \rho\Big\}\\
&= E(\,\lambda \circ \hat{L}^{-1} |\, \rho\,)\\
&= E\Big\{(\hat{\beta} \circ \pi^{-1}) \circ \hat{L}^{-1}\Big|\, \rho\Big\}\\
&= E(\,\rho \circ \pi^{-1} |\, \rho\,) = \rho \circ \pi^{-1},
\end{aligned}$$

which proves the asserted conditional property. □

Proof of Theorem 9.48 (ii): Let ρ be such as in Lemma 9.53, and choose (τ', χ') with $\mathcal{L}(\tau', \chi' \,|\, \rho) = \rho$ a.s. Comparing with (20) gives $(\tau', \rho) \overset{d}{=} (\tau, \rho)$, and so the transfer lemma yields a random element $\tilde{\chi}$ with $(\tau, \rho, \tilde{\chi}) \overset{d}{=} (\tau', \rho, \chi')$. In particular, (20) extends to

$$\zeta = \pi_\tau \rho, \qquad \mathcal{L}(\,\tau, \tilde{\chi} \,|\, \rho\,) = \rho \ \text{ a.s.} \tag{26}$$

Now introduce the right-continuous filtration

$$\mathcal{G}_t = \sigma\{\rho,\ \pi_t \delta_{\tau,\tilde{\chi}}\}, \quad t \ge 0. \tag{27}$$

Using the characterizations in Lemma 9.7 (i) and (iv), we see that the random measure ρ and set $[0,\tau]$ are both $\mathcal{G}$-predictable. Hence, so is ζ by Lemma 9.13, and so η is $\mathcal{G}$-predictable by Theorem 9.41. From Lemma 9.40 we see that $(\tau,\tilde{\chi})$ has $\mathcal{G}$-compensator η. Since η is also the $\mathcal{F}$-compensator of (τ,χ), Lemma 9.49 yields $(\tau,\tilde{\chi},\rho) \stackrel{d}{=} (\tau,\chi,\rho)$. In particular, (26) remains true with $\tilde{\chi}$ replaced by χ.

(i) $\Leftrightarrow$ (ii): First assume (ii), and put $\nu = \mathcal{L}(\rho)$. Then the disintegration theorem gives

$$\begin{aligned}
\mathcal{L}(\tau,\chi,\zeta) &= \mathcal{L}(\tau,\chi,\pi_\tau\rho)\\
&= E\,\mathcal{L}(\tau,\chi,\pi_\tau\rho\,|\,\rho)\\
&= E\,\rho\Big\{(t,x)\in \mathsf{R}_+\times S;\ (t,x,\pi_t\rho)\in\cdot\Big\}\\
&= \int \tilde{P}_\mu\,\nu(d\mu),
\end{aligned}$$

proving (i). Next assume (i), and choose τ', χ', and ρ' with $\mathcal{L}(\rho')=\nu$ and $\mathcal{L}(\tau',\chi'\,|\,\rho')=\rho'$ a.s. The same calculation gives $\mathcal{L}(\tau',\chi',\pi_{\tau'}\rho') = \int \hat{P}_\mu\,\nu(d\mu)$, and so $(\tau',\chi',\pi_{\tau'}\rho') \stackrel{d}{=} (\tau,\chi,\zeta)$. Hence, the transfer lemma yields a random measure ρ with $(\tau,\chi,\rho) \stackrel{d}{=} (\tau',\chi',\rho')$ and $\pi_\tau\rho=\zeta$ a.s. In particular, $\mathcal{L}(\tau,\chi\,|\,\rho)=\rho$ a.s., which proves (ii).

Existence: Given a probability measure ν on $\mathcal{P}_{(0,\infty)\times S}$, we may choose a random probability measure ρ with $\mathcal{L}(\rho)=\nu$. Next, we choose a random pair (τ,χ) with $\mathcal{L}(\tau,\chi\,|\,\rho)=\rho$ a.s., put $\zeta=\pi_\tau\rho$, and define η as in Lemma 9.40. Finally, let $\mathcal{G}$ be the filtration in (27). Then η is $\mathcal{G}$-predictable, as before, and we see as in Lemma 9.40 that (τ,χ) has $\mathcal{G}$-compensator η. This proves (ii), and (i) follows as before.

Uniqueness: Consider the processes $Y_t=\|\rho_t\|$ and $\hat{Y}_t=\|\zeta_t\|$, along with their left-continuous inverses

$$\begin{aligned}
L_r &= \inf\{t\ge 0;\ Y_t\ge r\},\\
\hat{L}_r &= \inf\{t\ge 0;\ \hat{Y}_t\ge r\}, \qquad r\in[0,1).
\end{aligned}$$

Let $\vartheta\perp\!\!\!\perp\rho$ be $U(0,1)$, and put $\tau'=L_\vartheta$. Then (ii) yields $(\tau',\rho)\stackrel{d}{=}(\tau,\zeta)$, and so the transfer lemma yields a random variable σ with $(\sigma,\rho)\stackrel{d}{=}(\vartheta,\rho)$ and $\tau=L_\sigma$. Here a.s.

$$\begin{aligned}
&P\Big(\sigma\ge s,\ \tau\in B\,\Big|\,\rho\Big)\\
&\quad = \lambda\Big\{x\in[s,1];\ L_x\in B\Big\}\\
&\quad = \int_B \bigg\{1\{s\le Y_{t-}\} + \frac{(Y_t-s)}{\Delta Y_t}\,1\{s\in(Y_{t-},Y_t]\}\bigg\}\lambda\{x;\ L_x\in dt\}\\
&\quad = E\bigg\{1\{s\le Y_{\tau-}\} + \frac{(Y_\tau-s)}{\Delta Y_\tau}\,1\{s\in(Y_{\tau-},Y_\tau]\};\ \tau\in B\,\bigg|\,\rho\bigg\}.
\end{aligned}$$

Since $Y = \hat{Y}$, a.s. on $[0, \tau]$, we get a.s.

$$P(\sigma \geq s \mid \tau, \rho) = 1\{s \leq Y_{\tau-}\} + \frac{(Y_\tau - s)}{\Delta Y_\tau} 1\Big\{s \in (Y_{\tau-}, Y_\tau]\Big\}.$$

Furthermore, $\pi_{L_s}\rho = \pi_{\hat{L}_s}\zeta$ a.s. on $\{s \leq \hat{Y}_\tau\}$. Since $\sigma \perp\!\!\!\perp \rho$, we have for any $s \in [0,1)$

$$\begin{aligned}
&(1-s)\,\mathcal{L}(\pi_{L_s}\rho) \\
&\qquad = \mathcal{L}(\pi_{L_s}\rho\,;\ \sigma \geq s) \\
&\qquad = E\Big\{P(\sigma \geq s \mid \rho, \tau);\ \pi_{L_s}\rho \in \cdot\Big\} \\
&\qquad = \mathcal{L}\Big\{\pi_{\hat{L}_s}\zeta\,;\ s \leq \hat{Y}_{\tau-}\Big\} + E\bigg\{\frac{\hat{Y}_\tau - s}{\Delta\hat{Y}_\tau}\,;\ s \in (\hat{Y}_{\tau-}, \hat{Y}_\tau],\ \pi_{\hat{L}_s}\zeta \in \cdot\bigg\}.
\end{aligned}$$

Since τ is a.s. the last point of increase of $\hat{Y}$, the right-hand side is uniquely determined by $\mathcal{L}(\zeta)$, and hence so is $\mathcal{L}(\pi_{L_s}\rho)$ for every $s \in [0,1)$. This remains true for $\nu = \mathcal{L}(\rho)$, since $\pi_{L_s}\rho \uparrow \rho$ as $s \uparrow 1$. □

9.6 Tangential Existence

Given a real semi-martingale X, we define the associated *jump point process* ξ on $\mathbb{R}_+ \times (\mathbb{R} \setminus \{0\})$ by

$$\xi_t B = \xi\{(0,t] \times B\} = \sum_{s \leq t} 1\{\Delta X_s \in B\}, \quad t > 0,\ \ B \in \mathcal{B}_{\mathbb{R}\setminus\{0\}}, \tag{28}$$

and write $\hat{\xi}$ for the compensator of ξ. If X is increasing with $X_0 = 0$, its *local characteristics* equal X^c and $\hat{\xi}$, where X^c denotes the continuous component of X. For a local martingale X with $X_0 = 0$, the local characteristics are instead $[X]^c$ and $\hat{\xi}$, where the former denotes the continuous component of the quadratic variation process $[X]$. For more general semi-martingales X, we need to add a suitable drift component $\hat{X}$ of the process X itself. The precise definition of $\hat{X}$ is not important for our purposes.

Two semi-martingales X and Y, adapted to a common filtration $\mathcal{F}$, are said to be *tangential*, if their local characteristics agree a.s. We also say that two local martingales X and Y are *weakly tangential*, if their quadratic variation processes $[X]$ and $[Y]$ are tangential in the previous sense. The latter property is clearly weaker than the strict tangential relation between X and Y. The main result of this section is Theorem 9.55, which shows that any semi-martingale X has a tangential process Y with conditionally independent increments.

The latter statement requires a suitable extension of the underlying filtration $\mathcal{F}$. Recall from (1) that, for any filtrations $\mathcal{F}$ and $\mathcal{G}$ on a common probability space, $\mathcal{G}$ is called a *standard extension* of $\mathcal{F}$, if

$$\mathcal{F}_t \subset \mathcal{G}_t \perp\!\!\!\perp_{\mathcal{F}_t} \mathcal{F}, \quad t \geq 0.$$

This ensures that the transition $\mathcal{F} \to \mathcal{G}$ will preserve all adaptedness and conditioning properties.

All major results in this and the next section have discrete-time versions, which follow as special cases of the continuous-time statements (but not the other way around). To see this, it suffices to embed the discrete time scale Z_+ into R_+, and extend the discrete filtration $\mathcal{F} = (\mathcal{F}_k)$ to R_+, by setting $\mathcal{F}_t = \mathcal{F}_{[t]}$ for all $t \geq 0$, where $[t]$ denotes the integral part of t. All tangential and prediction properties of random sequences $X = (X_k)$ extend immediately to their continuous-time extensions $X_t = X_{[t]}$.

For clarity, we begin with a preliminary result, of some independent interest. For any combination $\mathcal{X}$ of processes and filtrations, we define an $\mathcal{X}$-*compensator* as a compensator with respect to the filtration generated by $\mathcal{X}$. The definitions of $\mathcal{X}$-martingale, $\mathcal{X}$-adaptedness, etc., are similar.

Proposition 9.54 *(reduction to independence) Given a filtration $\mathcal{F}$ on R_+, consider a continuous local martingale M, and an S-marked, adapted point process ξ with compensator η. Form a Cox process $\zeta \perp\!\!\!\perp_\eta \mathcal{F}$ directed by η, and a Brownian motion $B \perp\!\!\!\perp (\zeta, \mathcal{F})$, put $N = B \circ [M]$, and let $\mathcal{G}$ denote the filtration generated by $\mathcal{F}$, ζ, and N. Then*

- (i) *$\mathcal{G}$ is a standard extension of $\mathcal{F}$,*
- (ii) *N is a continuous local $\mathcal{G}$-martingale with rate $[N] = [M]$ a.s., and both ξ and ζ have $\mathcal{G}$-compensator η,*
- (iii) *when η is continuous, it is both a (ξ, η)-compensator of ξ and a (ζ, η)-compensator of ζ.*

Proof: (i) Since $\zeta \perp\!\!\!\perp_\eta \mathcal{F}$ and $B \perp\!\!\!\perp_{(\zeta,\eta)} \mathcal{F}$, we have $(\zeta, B) \perp\!\!\!\perp_\eta \mathcal{F}$, by the chain rule for conditional independence (FMP 6.8), and so

$$(\zeta^t, N^t) \underset{\eta,\,[M]}{\perp\!\!\!\perp} \mathcal{F}, \quad t \geq 0.$$

Using the conditioning criterion in FMP 6.6, and the definition of ζ and N, we further note that

$$(\zeta^t, N^t) \underset{\eta^t,\,[M]^t}{\perp\!\!\!\perp} (\eta, [M]), \quad t \geq 0.$$

Combining those relations, and using the chain rule once more, we obtain

$$(\zeta^t, N^t) \underset{\eta^t,\,[M]^t}{\perp\!\!\!\perp} \mathcal{F}, \quad t \geq 0,$$

which implies $\mathcal{G}_t \perp\!\!\!\perp_{\mathcal{F}_t} \mathcal{F}$ for all $t \geq 0$.

(ii) Since $B \perp\!\!\!\perp (\zeta, \mathcal{F})$, we get

$$B \underset{[M],\,N^s}{\perp\!\!\!\perp} (\mathcal{F}, \zeta, N^s), \quad s \geq 0,$$

and so

$$\theta_s N \underset{[M],\, N^s}{\perp\!\!\!\perp} \mathcal{G}_s, \quad s \geq 0.$$

Combining with the relation $\theta_s N \perp\!\!\!\perp_{[M]} N^s$, and using the chain rule again, we obtain

$$\theta_s N \underset{[M]}{\perp\!\!\!\perp} \mathcal{G}_s, \quad s \geq 0.$$

Localizing if necessary, to ensure integrability, we get for any $s \geq t$ the desired martingale property

$$\begin{aligned} E(N_t - N_s \,|\, \mathcal{G}_s) &= E\Big\{ E\Big(N_t - N_s \,\Big|\, \mathcal{G}_s, [M]\Big) \,\Big|\, \mathcal{G}_s \Big\} \\ &= E\Big\{ E\Big(N_t - N_s \,\Big|\, [M]\Big) \,\Big|\, \mathcal{G}_s \Big\} = 0, \end{aligned}$$

and the associated rate property

$$\begin{aligned} E\Big(N_t^2 - N_s^2 \,\Big|\, \mathcal{G}_s\Big) &= E\Big\{ E\Big(N_t^2 - N_s^2 \,\Big|\, \mathcal{G}_s, [M]\Big) \,\Big|\, \mathcal{G}_s \Big\} \\ &= E\Big\{ E\Big(N_t^2 - N_s^2 \,\Big|\, [M]\Big) \,\Big|\, \mathcal{G}_s \Big\} \\ &= E\Big\{ E\Big([M]_t - [M]_s \,\Big|\, [M]\Big) \,\Big|\, \mathcal{G}_s \Big\}, \end{aligned}$$

proving the assertions for N.

Property (i) shows that η remains a $\mathcal{G}$-compensator of ξ. Next, the relation $\zeta \perp\!\!\!\perp_\eta \mathcal{F}$ implies $\theta_t \zeta \perp\!\!\!\perp_{(\eta,\zeta^t)} \mathcal{F}_t$. Combining this with the Cox property $\theta_t \zeta \perp\!\!\!\perp_\eta \zeta^t$, and using the chain rule, we get $\theta_t \zeta \perp\!\!\!\perp_\eta (\zeta^t, \mathcal{F}_t)$. Invoking the tower property of conditional expectations and the Cox property of ζ, we obtain

$$\begin{aligned} E(\theta_t \zeta \,|\, \mathcal{G}_t) &= E(\theta_t \zeta \,|\, \zeta^t, \mathcal{F}_t) \\ &= E\Big\{ E(\theta_t \zeta \,|\, \zeta^t, \eta, \mathcal{F}_t) \,\Big|\, \zeta^t, \mathcal{F}_t \Big\} \\ &= E\Big\{ E(\theta_t \zeta \,|\, \eta) \,\Big|\, \zeta^t, \mathcal{F}_t \Big\} \\ &= E(\theta_t \eta \,|\, \zeta^t, \mathcal{F}_t) = E(\theta_t \eta \,|\, \mathcal{G}_t). \end{aligned}$$

Since η remains $\mathcal{G}$-predictable, it is then a $\mathcal{G}$-compensator of ζ.

(iii) The martingale properties in (ii) extend to the filtrations generated by (ξ, η) and (ζ, η), respectively, by the tower property of conditional expectations. Since η is continuous and adapted to both filtrations, it is both (ξ, η)- and (ζ, η)-predictable. The assertions follow by combination of the mentioned properties. □

Using the previous ideas, we may construct tangential processes with conditionally independent increments.

Theorem 9.55 *(tangential existence) For any $\mathcal{F}$-semi-martingale X with local characteristics Y, there exist a standard extension $\mathcal{G}$ of $\mathcal{F}$, and a $\mathcal{G}$-tangential semi-martingale $\tilde{X} \perp\!\!\!\perp_Y \mathcal{F}$, such that $\tilde{X}$ has conditionally independent increments, given Y.*

Proof: Define the jump point process ξ of X as in (28), and let η denote the $\mathcal{F}$-compensator of ξ. Further, let M be the continuous martingale component of X, and let A be the predictable drift component of X, with respect to an arbitrarily fixed truncation function. When η is continuous, we may define ζ, N, and $\mathcal{G}$ as in Theorem 9.54, and construct an associated semi-martingale $\tilde{X}$ by compensating the jumps given by ζ. Then $\tilde{X}$ has the same local characteristics $Y = ([M], \eta, A)$ as X, so the two processes are $\mathcal{G}$-tangential. Since $\zeta \perp\!\!\!\perp_\eta \mathcal{F}$ and $B \perp\!\!\!\perp (\zeta, \mathcal{F})$, we have $\tilde{X} \perp\!\!\!\perp_Y \mathcal{F}$, and the independence properties of B and ζ show that $\tilde{X}$ has conditionally independent increments.

The tangential property may fail when η has discontinuities, since the projection $\bar{\zeta} = \zeta(\cdot \times S)$ may then have multiplicities, so that ζ is no longer a marked point process. We then need to replace the Cox process in Theorem 9.54 by a more general S-marked point process ζ on R_+, with conditionally independent increments, given η, and such that $E(\zeta \,|\, \eta) = \eta$ a.s. Note that a.s., the compensator η of ξ satisfies $\eta(\{t\} \times S) \le 1$ for all $t \ge 0$, since the discontinuity set is covered by countably many predictable times.

Just as in the construction of Cox processes, it is enough to consider a non-random measure $\mu = \eta$ with this property, and construct an S-marked point process ζ with independent increments and intensity μ. Then write as in (6)

$$\mu = \mu^c + \sum_{k \ge 1} (\delta_{t_k} \otimes \nu_k),$$

where μ^c is continuous, $t_1, t_2, \ldots$ are distinct times in $(0, \infty)$, and the ν_k are measures on S with $\|\nu_k\| \le 1$. Let α be a Poisson process on $\mathsf{R}_+ \times S$ with intensity μ^c, and choose some independent random elements β_k in S^Δ with distributions ν_k on S. As in case of Theorem 3.18, the S-marked point process

$$\zeta = \alpha + \sum_{k \ge 1} (\delta_{t_k} \otimes \delta_{\beta_k})$$

has clearly the desired properties. It is easy to check that the associated kernel is measurable in μ, which allows us to construct an associated S-marked point process with conditionally independent increments and conditional intensity η. The assertions of Theorem 9.54 remain valid in this case, and the associated semi-martingale $\tilde{X}$ has local characteristics Y, so that X and $\tilde{X}$ are indeed $\mathcal{G}$-tangential. □

We conclude with a lemma that is needed in the next section.

Lemma 9.56 *(tangential preservation) Given an $\mathcal{F}$-semi-martingale X with local characteristics ρ, let Y and Y' be $\mathcal{G}$-tangential to X with conditionally independent increments, where $\mathcal{G}$ is a standard extension of $\mathcal{F}$, and suppose that $\mathcal{F}$, Y, and Y' are conditionally independent, given ρ. Then $X - Y$ and $Y - Y'$ are symmetric $\mathcal{G}$-tangential.*

Proof: The only difficulty is to verify the tangential property for the associated jump point processes. Then let ξ, η, and η' be the jump processes of X, Y, and Y', and let $\hat{\xi}$ denote the associated compensator, so that $\mathcal{F}$, η, and η' are conditionally independent, given $\hat{\xi}$. On the continuity set A of $\hat{\xi}$, the processes η and η' are even conditionally Poisson $\hat{\xi}$, and so the discontinuity sets of the three processes are a.s. disjoint on A, and the compensators of the jump processes of $X - Y$ and $Y - Y'$ agree a.s. on A.

Since $\hat{\xi}$ is predictable, Theorem 9.22 shows that its discontinuity set A^c is covered by the graphs of countably many predictable times $\tau_1, \tau_2, \ldots$, and so it suffices to show that the two compensators agree a.s., at any predictable time $\tau < \infty$. Let τ be announced by the optional times $\sigma_1, \sigma_2, \ldots$. Since $1_B[\tau]$ is predictable for any $B \in \mathcal{G}_{\sigma_n}$, we note that $E(\Delta\hat{\xi}_\tau | \mathcal{G}_{\sigma_n}) = E(\Delta\xi_\tau | \mathcal{G}_{\sigma_n})$, which extends to $E(\Delta\hat{\xi}_\tau | \mathcal{G}_{\tau-}) = E(\Delta\xi_\tau | \mathcal{G}_{\tau-})$, by martingale convergence and Lemma 9.6. Since $\Delta\hat{\xi}_\tau$ is $\mathcal{G}_{\tau-}$-measurable by Lemma 9.8, we have a.s.

$$\begin{aligned} \Delta\hat{\xi}_\tau &= E(\Delta\xi_\tau | \mathcal{G}_{\tau-}) \\ &= E(\Delta\eta_\tau | \mathcal{G}_{\tau-}) = E(\Delta\eta'_\tau | \mathcal{G}_{\tau-}). \end{aligned} \tag{29}$$

Next, we note that η has compensator $\hat{\xi}$, with respect to the extended filtration $\mathcal{H}_t = \mathcal{G}_t \vee \mathcal{F}_\infty$. This implies $E\,\eta V = E\,\hat{\xi} V$, for any $\mathcal{H}$-predictable process $V \geq 0$. Given a predictable time τ as above, we may choose $V = 1_B 1_{[\tau]}$ for any $B \in \mathcal{H}_{\tau-}$, to get a.s.

$$E(\Delta\eta_\tau | \mathcal{H}_{\tau-}) = E(\Delta\hat{\xi}_\tau | \mathcal{H}_{\tau-}) = \Delta\hat{\xi}_\tau,$$

where the last equality holds since $\Delta\hat{\xi}_\tau$ is $\mathcal{G}_{\tau-}$-measurable. Since the same relations hold for the original filtration $\mathcal{G}$, we conclude that a.s.

$$E\Big(\Delta\eta_\tau \Big| \mathcal{G}_{\tau-}, \mathcal{F}_\infty\Big) = E\Big(\Delta\eta_\tau \Big| \mathcal{G}_{\tau-}\Big),$$

which means that $\Delta\eta_\tau$ and $\mathcal{F}_\infty$ are conditionally independent, given $\mathcal{G}_{\tau-}$. A similar argument applies to $\Delta\eta'_\tau$, and so

$$\Delta\xi_\tau \underset{\mathcal{G}_{\tau-}}{\perp\!\!\!\perp} \Delta\eta_\tau \underset{\mathcal{G}_{\tau-}}{\perp\!\!\!\perp} \Delta\eta'_\tau.$$

Combining with (29), we obtain a.s.

$$E\Big\{\Delta(\xi_\tau, \eta_\tau) \Big| \mathcal{G}_{\tau-}\Big\} = E\Big\{\Delta(\eta_\tau, \eta'_\tau) \Big| \mathcal{G}_{\tau-}\Big\},$$

and it follows easily that the compensators of the jump processes of $X - Y$ and $Y - Y'$ agree a.s. at τ. □

9.7 Tangential Comparison

Here we prove the basic comparison theorems for tangential processes, along with some related separation properties, leading to some general equivalences, for broad classes of tangential or semi-tangential processes.

Here and below, we write $X_t^* = \sup_{s \le t} |X_s|$ and $X^* = X_\infty^*$. For positive functions f and g, we mean by $f \asymp g$ that the ratio f/g is bounded above and below by positive constants. A nondecreasing, continuous function φ on R_+ with $\varphi_0 = 0$ is said to have *moderate growth*, if $\varphi(2x) \le c\,\varphi(x)$, $x > 0$, for some constant $c > 0$. This implies the existence of a function $h > 0$ on $(0, \infty)$, such that $\varphi(c\,x) \le h(c)\,\varphi(x)$ for all $c, x > 0$. The stated condition holds in particular for all power functions $\varphi(x) = |x|^p$ with $p > 0$, as well as for $\varphi(x) = x \wedge 1$ and $\varphi(x) = 1 - e^{-x}$. Note that the property of moderate growth is preserved under pairwise composition.

Theorem 9.57 *(tangential comparison) Let X and Y be tangential, increasing or conditionally symmetric processes. Then for any non-decreasing, continuous function φ on R_+, of moderate growth and with $\varphi_0 = 0$,*

$$E\,\varphi(X^*) \asymp E\,\varphi(Y^*),$$

where the domination constants depend only on φ. When φ is convex, this remains true for any weakly tangential local martingales X and Y.

For the proof, we begin with some tail estimates of independent interest.

Lemma 9.58 *(tail comparison) Let X and Y be tangential, increasing or conditionally symmetric processes. Then for any $c, x > 0$,*

(i) $P\{(\Delta X)^* > x\} \le 2\,P\{(\Delta Y)^* > x\}$,

(ii) $P\{X^* > x\} \le 3\,P\{Y^* > cx\} + 4c$.

Proof: (i) Let ξ and η be the jump point processes of X and Y. Fix any $x > 0$, and introduce the optional time

$$\tau = \inf\Big\{t > 0;\ |\Delta Y_t| > x\Big\}.$$

Since the set $(0, \tau]$ is predictable by Lemma 9.7, we get

$$\begin{aligned} P\{(\Delta X)^* > x\} &\le P\{\tau < \infty\} + E\,\xi\Big\{(0,\tau] \times [-x,x]^c\Big\} \\ &\le P\{\tau < \infty\} + E\,\eta\Big\{(0,\tau] \times [-x,x]^c\Big\} \\ &= 2\,P\{\tau < \infty\} \\ &= 2\,P\{(\Delta Y)^* > x\}. \end{aligned}$$

(ii) Fix any $c, x > 0$. For increasing X and Y, form $\hat{X}$ and $\hat{Y}$ by omitting all jumps greater than $c\,x$, which clearly preserves the tangential relation. If X and Y are instead conditionally symmetric, we may form $\hat{X}$ and $\hat{Y}$ by omitting all jumps of modulus greater than $2c\,x$. Then by FMP 26.5 and its proof, $\hat{X}$ and $\hat{Y}$ are local L^2-martingales with jumps a.s. bounded by $4cx$. They also remain tangential, and the tangential relation carries over to the quadratic variation processes $[\hat{X}]$ and $[\hat{Y}]$.

Now introduce the optional time

$$\tau = \inf\Big\{t > 0;\, |\hat{Y}_t| > c\,x\Big\}.$$

For increasing X and Y, we have

$$\begin{aligned} x\,P\{\hat{X}_\tau > x\} &\le E\hat{X}_\tau = E\hat{Y}_\tau \\ &= E\hat{Y}_{\tau-} + E\Delta\hat{Y}_\tau \le 3cx. \end{aligned}$$

If instead X and Y are conditionally symmetric, we may use the Bernstein–Lévy and Jensen inequalities, integration by parts (FMP 26.6 (vii)), and the tangential properties and bounds, to get

$$\begin{aligned} \Big(x\,P\{\hat{X}^*_\tau > x\}\Big)^2 &\le (E|\hat{X}_\tau|)^2 \le E\hat{X}^2_\tau \\ &= E[\hat{X}]_\tau = E[\hat{Y}]_\tau = E\hat{Y}^2_\tau \\ &\le E\Big(|\hat{Y}_{\tau-}| + |\Delta\hat{Y}_\tau|\Big)^2 \le (4cx)^2. \end{aligned}$$

Thus, in both cases $P\{\hat{X}^*_\tau > x\} \le 4c$.

Since $Y^* \ge \frac{1}{2}(\Delta Y)^*$, we have

$$\begin{aligned} \{(\Delta X)^* \le 2c\,x\} &\subset \{X = \hat{X}\}, \\ \{\tau < \infty\} \subset \{\hat{Y}^* > c\,x\} &\subset \{Y^* > c\,x\}. \end{aligned}$$

Combining with the previous tail estimate and (i), we obtain

$$\begin{aligned} P\{X^* > x\} &\le P\Big\{(\Delta X)^* > 2c\,x\Big\} + P\{\hat{X}^* > x\} \\ &\le 2\,P\Big\{(\Delta Y)^* > 2c\,x\Big\} + P\{\tau < \infty\} + P\{\hat{X}^*_\tau > x\} \\ &\le 3\,P\{Y^* > c\,x\} + 4c. \end{aligned}$$

□

Proof of Theorem 9.57: For any $c, x > 0$, we introduce the optional times

$$\begin{aligned} \tau &= \inf\Big\{t > 0;\, |X_t| > x\Big\}, \\ \sigma &= \inf\Big\{t > 0;\, P\{(\theta_t Y)^* > c\,x|\,\mathcal{F}_t\} > c\Big\}. \end{aligned}$$

Since $\theta_\tau X$ and $\theta_\tau Y$ remain conditionally tangential, given $\mathcal{F}_\tau$, Lemma 9.58 (ii) yields a.s. on $\{\tau < \sigma\}$

$$P\Big\{\,(\theta_\tau X)^* > x\Big|\,\mathcal{F}_\tau\Big\} \le 3\,P\Big\{\,(\theta_\tau Y)^* > c\,x\Big|\,\mathcal{F}_\tau\Big\} + 4c \le 7c,$$

and since $\{\tau < \sigma\} \in \mathcal{F}_\tau$ by Lemma 9.1, we get

$$\begin{aligned} P\Big\{X^* > 3x,\, (\Delta X)^* \le x,\, \sigma = \infty\Big\} & \\ \le P\Big\{(\theta_\tau X)^* > x,\, \tau < \sigma\Big\} & \\ = E\Big\{P\{(\theta_\tau X)^* > x|\,\mathcal{F}_\tau\};\, \tau < \sigma\Big\} & \\ \le 7c\,P\{\tau < \infty\} & \\ = 7c\,P\{X^* > x\}. & \end{aligned}$$

Furthermore, by Lemma 9.58 (i),

$$\begin{aligned} P\{(\Delta X)^* > x\} &\le 2\,P\{(\Delta Y)^* > x\} \\ &\le 2\,P\{Y^* > x/2\}, \end{aligned}$$

and the Bernstein–Lévy inequality gives

$$\begin{aligned} P\{\sigma < \infty\} &= P\Big\{\sup\nolimits_t P\{(\theta_t Y)^* > c\,x |\, \mathcal{F}_t\} > c\Big\} \\ &\le P\Big\{\sup\nolimits_t P(Y^* > c\,x/2 |\, \mathcal{F}_t) > c\Big\} \\ &\le c^{-1} P\{Y^* > c\,x/2\}. \end{aligned}$$

Combining the last three estimates, we obtain

$$\begin{aligned} P\{X^* > 3x\} &\le P\Big\{X^* > 3x,\, (\Delta X)^* \le x,\, \sigma = \infty\Big\} \\ &\qquad + P\{(\Delta X)^* > x\} + P\{\sigma < \infty\} \\ &\le 7c\,P\{X^* > x\} + 2\,P\{Y^* > x/2\} \\ &\qquad + c^{-1} P\{Y^* > c\,x/2\}. \end{aligned}$$

Since φ is non-decreasing with moderate growth, so that $\varphi(r\,x) \le h(r)\varphi(x)$ for some function $h > 0$, we obtain

$$\begin{aligned} (h_3^{-1} - 7c)\,E\,\varphi(X^*) &\le E\,\varphi(X^*/3) - 7c\,E\,\varphi(X^*) \\ &\le 2\,E\,\varphi(2Y^*) + c^{-1} E\,\varphi(2Y^*/c) \\ &\le \Big(2h_2 + c^{-1} h_{2/c}\Big)\,E\,\varphi(Y^*). \end{aligned}$$

Finally, we may choose $c < (7h_3)^{-1}$ to get $E\varphi(X^*) \lesssim E\varphi(Y^*)$.

Now let φ be a convex function of moderate growth, and note that the function $\varphi(x^{1/2})$ is again of moderate growth. For any weakly tangential local martingales X and Y, the processes $[X]$ and $[Y]$ are nondecreasing and strictly tangential, and so, by the previous result and a version of the BDG inequalities (FMP 26.12), we see that

$$\begin{aligned} E\,\varphi(X^*) &\asymp E\,\varphi\Big([X]_\infty^{1/2}\Big) \\ &\asymp E\,\varphi\Big([Y]_\infty^{1/2}\Big) \asymp E\,\varphi(Y^*), \end{aligned}$$

which proves the last assertion. □

We proceed with a similar one-sided result.

Theorem 9.59 *(one-sided comparison) Consider some tangential semi-martingales X and Y, where Y has conditionally independent increments. Then for any continuous, non-decreasing function φ of moderate growth,*

$$E\,\varphi(X^*) \lesssim E\,\varphi(Y^*),$$

where the domination constant depends only on φ.

Proof: Let ρ denote the local characteristics of X. Proceeding as in Theorem 9.55, we may construct two tangential processes Y' and Y'' with conditionally independent increments, and such that Y', Y'', and $\mathcal{F}$ are conditionally independent, given ρ. Since $Y' \overset{d}{=} Y'' \overset{d}{=} Y$, and the processes $X - Y'$ and $Y' - Y''$ are symmetric tangential by Lemma 9.56, the growth property of φ and Theorem 9.57 yield

$$\begin{aligned} E\,\varphi(X^*) &\lesssim E\,\varphi \circ (X - Y')^* + E\,\varphi(Y'^*) \\ &\lesssim E\,\varphi \circ (Y' - Y'')^* + E\,\varphi(Y'^*) \\ &\lesssim 2\,E\,\varphi(Y'^*) + E\,\varphi(Y''^*) \\ &= 3\,E\,\varphi(Y^*), \end{aligned}$$

where the domination constants depend only on φ. □

The previous results yield some useful comparisons of boundedness and convergence, for suitable pairs of strictly or weakly tangential processes. We begin with the following equivalences.

Theorem 9.60 *(boundedness and convergence) Let* (X, Y) *and* (X_n, Y_n), $n \in \mathsf{N}$, *be pairwise tangential, increasing or conditionally symmetric processes, or weakly tangential local martingales with uniformly bounded jumps. Then*

(i) $\{X^* < \infty\} = \{Y^* < \infty\}$ *a.s.,*

(ii) $X_n^* \xrightarrow{P} 0 \;\Leftrightarrow\; Y_n^* \xrightarrow{P} 0$.

Note that (i) is stronger than the equivalence

$$X^* < \infty \ \text{ a.s.} \quad \Leftrightarrow \quad Y^* < \infty \ \text{ a.s.},$$

which follows from (ii) with $X_n = X/n$ and $Y_n = Y/n$. Applying (ii) to the processes $X_n = \theta_n X - X(n)$ and $Y_n = \theta_n Y - Y(n)$, we further note that

$$X \ \text{converges a.s.} \quad \Leftrightarrow \quad Y \ \text{converges a.s.}$$

Proof: (i) First let X and Y be tangential, and increasing or conditionally symmetric. Fixing any $n \in \mathsf{N}$, we consider the optional time

$$\tau = \inf\Big\{t > 0;\ |X_t| > n\Big\}. \tag{30}$$

For any $a > 0$, we may apply Theorem 9.57 to the tangential, increasing or conditionally symmetric processes $a\,X^\tau$ and $a\,Y^\tau$, along with the function $\varphi(x) = x \wedge 1$, with associated domination constant $c > 0$, to get

$$\begin{aligned} P\Big\{X^* \le n,\, Y^* = \infty\Big\} &= P\Big\{\tau = \infty,\, Y_\tau^* = \infty\Big\} \\ &\le P\{Y_\tau^* = \infty\} \le E(a\,Y_\tau^* \wedge 1) \\ &\le c\,E(a\,X_\tau^* \wedge 1). \end{aligned}$$

Since $X_\tau^* \le n + |\Delta X_\tau| < \infty$ a.s., the right-hand side tends to 0 as $a \to 0$, by dominated convergence, and so

$$P\Big\{X^* < \infty,\, Y^* = \infty\Big\} \le \sum_{n\ge 1} P\Big\{X^* \le n,\, Y^* = \infty\Big\} = 0,$$

which means that $Y^* < \infty$ a.s. on $\{X^* < \infty\}$. By symmetry, we may interchange the roles of X and Y.

Next, let X and Y be weakly tangential local martingales with jumps bounded by b. For τ as above, the last statement of Theorem 9.57 yields $EX_\tau^* \asymp EY_\tau^*$, and so for any $a > 0$, we get as before

$$\begin{aligned} P\Big\{X^* \le n,\, Y^* = \infty\Big\} &= P\Big\{\tau = \infty,\, Y_\tau^* = \infty\Big\} \\ &\le P\{Y_\tau^* = \infty\} \le a\, E\, Y_\tau^* \\ &\le a\, c\, E\, X_\tau^*, \end{aligned}$$

for some constant $c > 0$. Since $a > 0$ was arbitrary, and $X_\tau^* \le n + b < \infty$, we conclude that the probability on the left equals 0. The proof may now be completed as before.

(ii) The statement for tangential, increasing or conditionally symmetric processes is immediate from Theorem 9.57 with $\varphi(x) = x \wedge 1$. Now let X_n and Y_n be weakly tangential local martingales with jumps bounded by b. Suppose that $X_n^* \xrightarrow{P} 0$, and introduce the optional times

$$\tau_n = \inf\Big\{t > 0;\, |X_n(t)| > 1\Big\}, \quad n \in \mathsf{N}.$$

Since $X_n^*(\tau_n) \le 1 + b$, and also $E\, X_n^*(\tau_n) \asymp E\, Y_n^*(\tau_n)$ by Theorem 9.57, we get

$$\begin{aligned} E(Y_n^* \wedge 1) &\le P\{\tau_n < \infty\} + E\, Y_n^*(\tau_n) \\ &\lesssim P\{X_n^* > 1\} + E\, X_n^*(\tau_n) \to 0, \end{aligned}$$

which shows that $Y_n^* \xrightarrow{P} 0$. By symmetry, we may again interchange the roles of X and Y. □

For any local martingale X and constant $r > 0$, we may write $X = X' + X''$, where X' is the sum of all compensated jumps in X of modulus $> r$. Recall from FMP 26.5 that X' has locally integrable variation, whereas X'' is a local martingale with jumps bounded by $2r$. The following preliminary result, of some independent interest, may sometimes enable us to separate the large and small jumps.

Lemma 9.61 *(separation) Let X and $X_1, X_2, \ldots$ be local martingales, with conditionally symmetric or independent increments, and write $X = X' + X''$ and $X_n = X_n' + X_n''$, where X' and X_n' are the sums of all compensated jumps in X and X_n of size > 1. Then*

(i) $\{X^* < \infty\} = \{X'^* + X''^* < \infty\}$ *a.s.*,
(ii) $X_n^* \xrightarrow{P} 0 \quad \Leftrightarrow \quad X_n'^* + X_n''^* \xrightarrow{P} 0.$

Since trivially $X^* \le X'^* + X''^*$ and $X_n^* \le X_n'^* + X_n''^*$, the interesting parts are the inclusion or implication to the right.

Proof: (i) First let X be a local martingale with conditionally independent increments. Conditioning on the local characteristics, we may reduce to the case where X is centered, with strictly independent increments. Since each side has then probability 0 or 1, by Kolmogorov's $0-1$ law, it suffices to show that $X^* < \infty$ a.s. implies $X'^* + X''^* < \infty$ a.s. Thus, we need to show that $X^* < \infty$ a.s. implies $X''^* < \infty$ a.s. Letting $\tilde{X}$ and $\tilde{X}''$ be symmetrizations of X and X'', respectively, we note that $X^* < \infty$ a.s. implies $\tilde{X}^* < \infty$ a.s., and so $\tilde{X}''^* < \infty$ a.s. by Theorem 3.27 (i). Since X'' has bounded jumps, Corollary 3.29 (i) shows that even $X''^* < \infty$ a.s.

If instead X has conditionally symmetric increments, then by Theorem 9.55, we may choose Y to be tangential to X with conditionally independent increments. Introducing the corresponding decomposition $Y = Y' + Y''$, we note that Y' and Y'' are tangential to X' and X'', respectively. Hence, the tangential case of Theorem 9.60 (i) yields a.s.

$$\{X^* < \infty\} = \{Y^* < \infty\},$$
$$\{X'^* + X''^* < \infty\} = \{Y'^* + Y''^* < \infty\},$$

which reduces the proof to the previous case of processes with conditionally independent increments.

(ii) If the X_n have conditionally symmetric increments, then by Theorem 9.55, we may choose $Y_1, Y_2, \ldots$ to be pairwise tangential to $X_1, X_2, \ldots$ with conditionally independent increments. Writing $Y_n = Y_n' + Y_n''$ as before, we note that Y_n' and Y_n'' are tangential to X_n' and X_n'', respectively. The tangential case of Theorem 9.60 (ii) yields

$$X_n^* \xrightarrow{P} 0 \quad \Leftrightarrow \quad Y_n^* \xrightarrow{P} 0,$$
$$X_n'^* + X_n''^* \xrightarrow{P} 0 \quad \Leftrightarrow \quad Y_n'^* + Y_n''^* \xrightarrow{P} 0,$$

which reduces the proof to the case of processes with conditionally independent increments. By the same argument, we may further assume that the processes X_n are conditionally independent, given the local characteristics of the whole family (X_n).

In the latter case, we may turn to sub-sequences to ensure a.s. convergence, and then condition on the set of local characteristics, to reduce to the case of processes with strictly independent increments. Then we introduce the symmetrizations $\tilde{X}_n$ and $\tilde{X}_n''$, and note as before that $X_n^* \xrightarrow{P} 0$ implies $\tilde{X}_n^* \xrightarrow{P} 0$, which implies $\tilde{X}_n''^* \xrightarrow{P}$ by Theorem 3.27 (ii). Then $X_n''^* \xrightarrow{P} 0$ by

Corollary 3.29 (ii), and so $X_n'^* + X_n''^* \le X_n^* + 2X_n''^* \xrightarrow{P} 0$. □

Two local martingales X and Y are said to be *semi-tangential*, if there exists a constant $r > 0$ with corresponding decompositions $X = X' + X''$ and $Y = Y' + Y''$, such that X' and Y' are tangential, while X'' and Y'' are weakly tangential. More generally, we say that two sequences of local martingales $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are pairwise *uniformly semi-tangential*, if X_n and Y_n are semi-tangential for every $n \in \mathsf{N}$, and the associated truncation level $r > 0$ can be chosen to be independent of n.

Using the previous lemma, we may supplement Theorem 9.60 by some one-sided implications.

Theorem 9.62 *(one-sided relations) Let (X, Y) and (X_n, Y_n), $n \in \mathsf{N}$, be pairwise uniformly semi-tangential local martingales, such that X and all X_n have conditionally symmetric increments. Then*

(i) $\{X^* < \infty\} \subset \{Y^* < \infty\}$ *a.s.,*

(ii) $X_n^* \xrightarrow{P} 0 \;\Rightarrow\; Y_n^* \xrightarrow{P} 0.$

This remains true when the processes (X, Y) and (X_n, Y_n) are pairwise tangential semi-martingales, such that X and all the X_n have conditionally independent increments.

Proof: (i) Write $X = X' + X''$ and $Y = Y' + Y''$, where X' and Y' are the sums of all compensated jumps in X and Y, respectively, of modulus $> r$, and assume that X' and Y' are strictly tangential, whereas X'' and Y'' are weakly tangential. Using Theorem 9.60 (i) and Lemma 9.61 (i), and noting that $Y^* \le Y'^* + Y'^*$, we get a.s.

$$\begin{aligned}
\{X^* < \infty\} &= \{X'^* < \infty\} \cap \{X''^* < \infty\} \\
&= \{Y'^* < \infty\} \cap \{Y''^* < \infty\} \\
&\subset \{Y^* < \infty\}.
\end{aligned}$$

(ii) Assume that $X_n^* \xrightarrow{P} 0$. Writing $X_n = X_n' + X_n''$ and $Y_n = Y_n' + Y_n''$ as before, we see from Lemma 9.61 (ii) that $X_n'^* \xrightarrow{P} 0$ and $X_n''^* \xrightarrow{P} 0$, and so by Theorem 9.60 (ii) we have $Y_n'^* \xrightarrow{P} 0$ and $Y_n''^* \xrightarrow{P} 0$, which implies $Y_n^* \xrightarrow{P} 0$.

To prove the last statement, use Theorem 9.59, and proceed as in the proof of Theorem 9.60. □

The preceding theory yields some important principles of *decoupling* and *symmetrization*. Thus, in order to determine, for a non-decreasing process or local martingale X, whether the maximum X^* is finite or infinite, we may assume X to have conditionally independent increments. By a further conditioning, we may reduce to the case of strictly independent increments, where the problem is elementary, and the solution is essentially known. The previous theory may also permit a further reduction to the case of symmetric jumps. Similar remarks apply to the criteria for convergence $X_n^* \xrightarrow{P} 0$ and the associated tightness.

When applying the previous results to stochastic integrals, we need to know which tangential and other properties are preserved by stochastic integration. Here we use the notation

$$(V \cdot X)_t = \int_0^t V\,dX, \quad t \geq 0.$$

Lemma 9.63 *(preservation properties) Consider some semi-martingales X and Y and a predictable process V, such that the stochastic integrals below exist. Then*

(i) *if X and Y are tangential, so are $V \cdot X$ and $V \cdot Y$,*

(ii) *if X and Y are centered, weakly tangential, so are $V \cdot X$ and $|V| \cdot Y$,*

(iii) *if X and Y are semi-tangential, so are $V \cdot X$ and $V \cdot Y$ for bounded V,*

(iv) *if X is conditionally symmetric, so is $V \cdot X$,*

(v) *if $\tilde{X}$ has conditionally independent increments, given $\mathcal{F}$, so has $V \cdot \tilde{X}$.*

Proof: (i) Recall that if ξ is a marked point process with compensator $\hat{\xi}$, and the process $V \geq 0$ is predictable, then $V \cdot \xi$ has compensator $V \cdot \hat{\xi}$. Further note that, if X has jump point process ξ, then the jump point process of $-X$ is the reflection of ξ, whose compensator is obtained by reflection of $\hat{\xi}$. Combining those facts, we see that, if the point processes ξ and η are tangential, then so are $V \cdot \xi$ and $V \cdot \eta$. For the continuous martingale component M of X, we have $[V \cdot M] = V^2 \cdot [M]$. We also note that, if X has predictable drift component A, then the process $V \cdot A$ is again predictable. The assertion for X and Y now follows easily by combination.

(ii) If X and Y are weakly tangential, then $[X]$ and $[Y]$ are strictly tangential, and hence so are $[V \cdot X] = V^2 \cdot [X]$ and $[|V| \cdot Y] = V^2 \cdot [Y]$ by part (i). Furthermore, the centering of X and Y is clearly preserved by stochastic integration.

(iii) If $|\Delta X| \leq b$ and $|V| \leq c$, then $|\Delta(V \cdot X)| \leq b\,c$, and similarly for Y. Now apply (i) and (ii).

(iv) It is enough to consider the jump point process ξ, where symmetry means that the compensator $\hat{\xi}$ is symmetric. The latter symmetry is clearly preserved by the integration $V \cdot \xi$ when $V \geq 0$, and also by a change of sign, since the reflected version of ξ has the same compensator $\hat{\xi}$. The assertion now follows by combination.

(v) Conditioning on $\mathcal{F}$, we may reduce to the case where $\tilde{X}$ has independent increments and V is non-random, in which case clearly $V \cdot \tilde{X}$ has again independent increments. $\square$

We may also examine the validity of the set inclusion

$$\{(V \cdot X)^* < \infty\} \subset \{(V \cdot Y)^* < \infty\} \text{ a.s.} \tag{31}$$

along with the corresponding a.s. equality. This is especially useful when X has conditionally independent increments, since precise criteria for the condition on the left can then be inferred from results in Section 3.4.

Corollary 9.64 *(boundedness) Consider some semi-martingales X and Y and a predictable process V, such that X has conditionally independent increments, and $V \cdot X$ and $V \cdot Y$ exist. Then* (31) *holds a.s., under each of the conditions*

(i) *X and Y are tangential,*

(ii) *X and Y are semi-tangential, X is conditionally symmetric, and V is bounded,*

whereas equality in (31) *holds a.s., under each of the conditions*

(iii) *X and Y are tangential increasing, and $V \geq 0$,*

(iv) *X and Y are tangential and conditionally symmetric,*

(v) *X and Y are centered, weakly tangential with bounded jumps, and V is bounded.*

Proof: (i) The processes $V \cdot X$ and $V \cdot Y$ are again tangential by Lemma 9.63 (i), and Lemma 9.63 (v) shows that $V \cdot X$ has again conditionally independent increments. Now (31) follows by Theorem 9.62 (i).

(ii) Here $V \cdot X$ and $V \cdot Y$ are again semi-tangential by Lemma 9.63 (iii), and Lemma 9.63 (iv) shows that $V \cdot X$ has again conditionally symmetric increments. Hence, (31) follows again by Theorem 9.62 (i).

(iii) Here Lemma 9.63 (i) shows that $V \cdot X$ and $V \cdot Y$ are again tangential increasing, and so equality holds in (31) by Theorem 9.60 (i).

(iv) Here $V \cdot X$ and $V \cdot Y$ are again tangential and conditionally symmetric by Lemma 9.63 (i) and (iv), and so equality holds in (31) by Theorem 9.60 (i).

(v) Here Lemma 9.63 (ii) shows that $V \cdot X$ and $V \cdot Y$ are again centered, weakly tangential with bounded jumps, and so we have again equality in (31), by Theorem 9.60 (i). $\square$

By similar arguments, here omitted, we may derive sufficient conditions for the implication

$$(V_n \cdot X_n)^* \xrightarrow{P} 0 \quad \Rightarrow \quad (V_n \cdot Y_n)^* \xrightarrow{P} 0, \tag{32}$$

as well as for the corresponding equivalence.

Corollary 9.65 *(convergence) Consider some semi-martingales X_n and Y_n and predictable processes V_n, such that the X_n have conditionally independent increments, and the integrals $V_n \cdot X_n$ and $V_n \cdot Y_n$ exist. Then* (32) *holds under each of these conditions:*

(i) *X_n and Y_n are tangential for each n,*
(ii) *the X_n and Y_n are uniformly semi-tangential, the X_n are conditionally symmetric, and the V_n are uniformly bounded,*

whereas equivalence holds in (32)*, under each of these conditions:*

(iii) *X_n and Y_n are increasing tangential, and $V_n \geq 0$ for each n,*
(iv) *the X_n and Y_n are tangential and conditionally symmetric,*
(v) *the X_n and Y_n are centered, weakly tangential with uniformly bounded jumps, and the V_n are uniformly bounded.*

Chapter 10

Multiple Integration

Multiple stochastic integration is a classical topic, going back to the pioneering work of Wiener and Itô on multiple integrals with respect to Brownian motion, along with the associated chaos representations. The study has since been broadened to encompass integrals with respect to many other processes, including stable and more general Lévy processes. Apart from some important applications, the subject has considerable independent interest, thanks to an abundance of deep and interesting results, and the great variety of methods that have been employed for their proofs.

The theory of Wiener–Itô integrals is well understood and covered by numerous textbooks and monographs. Here we take, as our starting point, the equally fundamental case of Poisson and related processes, which will later provide a key to the study of some more general cases. For any random measures $\xi_1, \dots, \xi_d$ or ξ on a Borel space S, we consider integrals of measurable functions $f \geq 0$ on S^d with respect to the product measures $\xi_1 \otimes \cdots \otimes \xi_d$ or $\xi^{\otimes d}$ on S^d, henceforth denoted as $\xi_1 \cdots \xi_d f$ and $\xi^d f$, respectively.

Most results for Poisson processes remain valid for any simple point processes with independent increments. Until further notice, we may then take ξ and the ξ_k to be independent processes of the latter kind, all with the same intensity $E\xi = \nu$. (For the moment, we may ignore the further extension to marked point processes with independent increments.) The integrand f is said to be *non-diagonal*, if it vanishes on all diagonal spaces where two or more coordinates $s_1, \dots, s_d$ agree. *Symmetry* means that f is invariant under all permutations of the arguments. When S is linearly ordered, we further say that f is *tetrahedral*, if it is supported by the set $\{s_1 > \cdots > s_d\}$.

A basic problem is to characterize the convergence

$$\xi_1 \cdots \xi_d f_n \xrightarrow{P} \xi_1 \cdots \xi_d f, \qquad \xi^d f_n \xrightarrow{P} \xi^d f,$$

or the corresponding a.s. convergence, where f and the f_n are non-negative, measurable functions on S^d. In Theorem 10.2 we show that, whenever the integrals are a.s. finite, the stated convergence is equivalent to

$$\xi_1 \cdots \xi_d |f_n - f| \xrightarrow{P} 0, \qquad \xi^d |f_n - f| \xrightarrow{P} 0.$$

This is a significant simplification, since it essentially reduces the original convergence problem to the case where $f = 0$. Criteria for the latter conver-

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_10

gence clearly yield conditions for both finiteness and tightness of the various integrals.

Existence and convergence criteria for the integrals $\xi^d f$ are more general, since they extend immediately to $\xi_1 \cdots \xi_d f$, by a simple "chess-board" argument. For non-diagonal integrands f and f_n, the two cases are in fact essentially equivalent, as shown by the fundamental Theorem 10.3, based on the tangential comparison in Theorem 9.57. This yields another significant simplification, since proofs for the integrals $\xi_1 \cdots \xi_d f_n$ are so much easier, owing to the independence of the underlying processes.

Our central results for Poisson and related integrals are the recursive criteria for integrability and convergence, in Theorems 10.5 and 10.9, stated in terms of finitely many Lebesgue-type integrals, involving the integrands f and f_n. Already from Lemma 3.26, we know that $\xi f < \infty$ a.s. iff $\nu(f \wedge 1) < \infty$, whereas $\xi f_n \xrightarrow{P} 0$ iff $\nu(f_n \wedge 1) \to 0$. To state our criteria for $\xi_1 \xi_2 f < \infty$ a.s., which apply even to the integrals $\xi^2 f$ with non-diagonal f, we may take $f \leq 1$, since the conditions for f and $f \wedge 1$ are clearly equivalent. Introducing the marginal integrals $f_1(x) = \nu f(x, \cdot)$ and $f_2(y) = \nu f(\cdot, y)$, we show that the stated integrability is equivalent to the three conditions

$$\nu\{f_1 \vee f_2 = \infty\} = 0, \qquad \nu\{f_1 \vee f_2 > 1\} < \infty,$$
$$\nu^2\big(f;\, \{f_1 \leq 1\} \times \{f_2 \leq 1\}\big) < \infty.$$

The higher-dimensional conditions are similar, but become increasingly complicated as d gets large. The same thing is true for the associated convergence criteria.

As another basic topic, we explore the effects of symmetrization. Given a simple point process ξ on S, we may create a *symmetric* version $\tilde{\xi}$ of ξ, by replacing each unit mass δ_{τ_k} of ξ by a signed point mass $\sigma_k \, \delta_{\tau_k}$, where the σ_k are independent random signs[1] with distribution $P\{\sigma_k = \pm 1\} = \frac{1}{2}$. Our objective is then to characterize the existence and convergence of the integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ or $\tilde{\xi}^d f$, in terms of corresponding properties of the positive integrals $\xi_1 \cdots \xi_d g$ or $\xi^d g$, for suitable integrands $g \geq 0$.

Since the sign sequences are independent of the underlying processes ξ and ξ_k, we may reduce by conditioning to the case of non-random ξ and ξ_k, where the integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ and $\tilde{\xi}^d f$ are equivalent to the *random multi-linear forms*

$$\sigma_1 \cdots \sigma_d \, A = \sum\nolimits_k a_k \, \sigma_{1,k_1} \cdots \sigma_{d,k_d}, \qquad \sigma^d A = \sum\nolimits_k a_k \, \sigma_{k_1} \cdots \sigma_{k_d},$$

for some infinite arrays $A = (a_k)$ of appropriate function values. To avoid complications arising from the diagonal terms, we may assume, in case of $\sigma^d A$, that A is non-diagonal and either symmetric or tetrahedral. Then our crucial Lemma 10.12 asserts that, for fixed $p > 0$,

$$\|\sigma_1 \cdots \sigma_d A\|_p \asymp \|\sigma^d A\|_p \asymp |A|_2,$$

[1] often called *Rademacher variables*

where $|A|_2$ denotes the l^2-norm of the set of entries of A, and the domination constants depend only on d and p. When A is tetrahedral and $p = 2$, both relations become equalities.

It is now relatively straightforward to analyze the symmetric integrals $\tilde\xi_1 \cdots \tilde\xi_d f$ and $\tilde\xi^d f$. In Theorem 10.11, we show that the former integral exists iff $\xi_1 \cdots \xi_d f^2 < \infty$ a.s., and similarly for the integral $\tilde\xi^d f$ when f is symmetric non-diagonal. Furthermore, Theorem 10.15 shows that $\tilde\xi_1 \cdots \tilde\xi_d f_n \xrightarrow{P} 0$ iff $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} 0$, and similarly for the integrals $\tilde\xi^d f_n$ when the integrands f_n are symmetric non-diagonal. In Theorem 10.16, we prove a Fubini-type reduction to single integrations, as in

$$\tilde\xi_1 \cdots \tilde\xi_d f = \tilde\xi_1(\tilde\xi_2(\cdots(\tilde\xi_d f)\cdots)) \text{ a.s.,}$$

and in Corollary 10.17 we establish, for any non-diagonal functions f_n, the decoupling property

$$\tilde\xi_1 \cdots \tilde\xi_d f_n \xrightarrow{P} 0 \quad \Rightarrow \quad \tilde\xi^d f_n \xrightarrow{P} 0.$$

Finally, we show in Theorem 10.18 that, whenever $\tilde\xi_1 \cdots \tilde\xi_d f_n$ converges in probability, the limit equals $\tilde\xi_1 \cdots \tilde\xi_d f$ a.s. for some function f, in which case $f_n \to f$ holds locally in measure.

A more difficult problem is to characterize the divergence $\xi_1 \cdots \xi_d f_n \xrightarrow{P} \infty$ or $|\tilde\xi_1 \cdots \tilde\xi_d f_n| \xrightarrow{P} \infty$, and similarly for the integrals $\xi^d f_n$ and $\tilde\xi^d f_n$, where the *escape* condition $\alpha_n \xrightarrow{P} \infty$ may be defined, for any random variables $\alpha_1, \alpha_2, \ldots$, by $e^{-\alpha_n} \xrightarrow{P} 0$. In Theorem 10.19, we show that, if ξ and $\xi_1, \xi_2, \ldots$ are independent Poisson processes on S with a common diffuse intensity ν, then for any non-diagonal, measurable functions f_n on S, the three conditions

$$|\tilde\xi_1 \cdots \tilde\xi_d f_n| \xrightarrow{P} \infty, \qquad \xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} \infty, \qquad \xi^d f_n^2 \xrightarrow{P} \infty,$$

are equivalent and follow from $|\tilde\xi^d f_n| \xrightarrow{P} \infty$. For symmetric f_n, the possible equivalence of all four conditions remains an open problem.

From the Poisson integrals, we can easily proceed to multiple integrals with respect to more general Lévy processes. The simplest case is when X is a pure jump type Lévy process in R_+^d with non-decreasing components $X_1, \ldots, X_d$. By Theorem 3.19, the associated jump point process η is Poisson with intensity measure $E\eta = \nu \otimes \lambda$, where ν is the *Lévy measure* of X, and the multiple integral $X_1 \cdots X_d f$ can be written as $\eta^d(Lf)$, for a suitable linear operator L. Thus, the basic criteria for existence and convergence of such integrals may be obtained from the previously mentioned results for Poisson integrals.

A more subtle case is when $X = (X_1, \ldots, X_d)$ is a symmetric, purely discontinuous Lévy process in R^d. Here we have again a representation of the form $X_1 \cdots X_d f = \zeta^d(Lf)$ a.s., where ζ is now a symmetric Poisson process on $\mathsf{R}_+ \times (\mathsf{R}^d \setminus \{0\})$, and the two integrals exist simultaneously. We can also use the decoupling properties of symmetric integrals $\tilde{\xi}^d f$, explored earlier, to obtain similar properties for symmetric Lévy integrals. Indeed, letting $X'_1, \ldots, X'_d$ be independent Lévy processes in R with $X'_k \overset{d}{=} X_k$ for all k, and restricting attention to tetrahedral integrands f and $f_1, f_2, \ldots$, we note in Corollary 10.28 that the integrals $X_1 \cdots X_d f$ and $X'_1 \cdots X'_d f$ exist simultaneously, and that $X_1 \cdots X_d f_n \xrightarrow{P} 0$ iff $X'_1 \cdots X'_d f_n \xrightarrow{P} 0$.

We may also mention the connection with the quadratic variation processes $[X_1], \ldots, [X_d]$, established in Corollary 10.29. Writing $[X_1 \cdots X_d] = \bigotimes_k [X_k]$, we show that $X_1 \cdots X_d f_n \xrightarrow{P} 0$ iff $[X_1 \cdots X_d] f_n^2 \xrightarrow{P} 0$, for any tetrahedral, measurable functions f_n, and similarly for the existence, tightness, and divergence to ∞. This clearly extends the relationship between the integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ and $\xi_1 \cdots \xi_d f^2$, mentioned earlier.

As we have seen, an analysis of multi-linear forms is helpful to understand the behavior of multiple integrals with respect to symmetric point processes. We may also proceed in the opposite direction, in using results for multiple stochastic integrals to analyze the behavior of suitable multiple sums. In particular, a very general class of multiple sums may be written in coded form as

$$\sum\nolimits_k f_k \circ \vartheta_k = \sum\nolimits_k f_k(\vartheta_{k_1}, \ldots, \vartheta_{k_d}),$$

where $\vartheta_1, \vartheta_2, \ldots$ are i.i.d. $U(0,1)$, and the f_k are measurable functions on $[0,1]^d$ indexed by N^d. The latter sum may also be written as a multiple stochastic integral $\eta^d f$, for a suitable function f, where η denotes the point process $\sum_k \delta_{\vartheta_k + k}$ on R_+. It then becomes natural to compare with the multiple integral $\xi^d f$, where ξ is a Poisson process on R_+ with Lebesgue intensity λ. In Theorem 10.32, we show that $\xi^d f_n \xrightarrow{P} 0$ iff $\eta^d f_n \xrightarrow{P} 0$, and similarly for the corresponding symmetrized versions $\tilde{\xi}^d f_n$ and $\tilde{\eta}^d f_n$.

The basic criteria for existence or convergence of the multiple stochastic integrals $\xi_1 \cdots \xi_d f$ or $\xi^d f$ were previously obtained, only in the special case of marked point processes with independent increments ξ or $\xi_1, \ldots, \xi_d$, or for the corresponding symmetrized processes. Our final challenge is to use the tangential existence and comparison results of the previous chapter to extend the mentioned criteria to more general processes. Assuming f to be non-diagonal, we may then write the multiple integral $\xi_1 \cdots \xi_d f$ as a sum of $d!$ tetrahedral components, each of which can be evaluated through a sequence of d single integrations, where in each step the integrand is predictable with respect to the underlying filtration.

The integrator in the last step can be replaced by a tangential version, which can be chosen to have conditionally independent increments. A problem arises when we try to continue the process recursively, since the predictability of the integrands would typically be destroyed by the required change in the order of integration. In Theorems 10.34 and 10.37, we find a way around those difficulties, leading to a comparison of the original integral $\xi_1 \cdots \xi_d f$ with the corresponding one for some *sequentially tangential* processes $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$, formed recursively by a construction involving successive changes of filtration.

The indicated construction leads to a sequence of *sequential compensators* $\eta_1, \dots, \eta_d$ of $\xi_1, \dots, \xi_d$, and our criteria for the existence $\xi_1 \cdots \xi_d f < \infty$ a.s. or convergence $\xi_1 \cdots \xi_d f_n \overset{P}{\to} 0$ agree with those for independent processes ξ_k with independent increments and intensities η_k. This solves the existence and convergence problems for positive point process integrals. In the more difficult signed case, only some partial results are available.

10.1 Poisson and Related Integrals

Throughout this section, we consider multiple integrals of the form

$$\xi_1 \cdots \xi_d f = \int \cdots \int \xi_1(dt_1) \cdots \xi_d(dt_d)\, f(t_1, \dots, t_d),$$

where $\xi_1, \dots, \xi_d$ are independent copies of a T-marked point process ξ on S with independent increments and intensity $E\xi = \nu$, and $f \geq 0$ is a measurable function on $\bar{S}^d \equiv (S \times T)^d$. We also consider the integrals $\xi^d f$, defined as above with $\xi_k = \xi$ for all k.

From Theorem 3.18, we know that ξ is Poisson iff $\nu(\{s\} \times T) \equiv 0$. The present generality, requiring little extra effort, covers both the classical case of multiple Poisson integrals, and the corresponding discrete-parameter case of random multi-linear forms, based on sequences of independent random variables. It is also needed for our subsequent extensions to multiple integrals with respect to general non-decreasing processes or local martingales, where the Poisson case alone would only allow extensions to ql-continuous processes. Most results extend to general Poisson processes by a simple projection.

Though the integrals $\xi_1 \cdots \xi_d f$ and $\xi^d f$ may be expressed as countable sums, their continuity and convergence properties are far from obvious. Here and below, measurability of the integrands is always understood, and we say that a function f on $\bar{S}^d$ is *non-diagonal*, if it vanishes on all diagonal spaces, where two or more S-coordinates agree. When S is linearly ordered, we also consider *tetrahedral* functions, supported by the set where $s_1 > \cdots > s_d$.

We begin our study with some simple moment relations, where we write $\psi(x) = 1 - e^{-x}$ for $x \geq 0$.

Lemma 10.1 *(moment relations) For non-diagonal functions $f \geq 0$ on $\bar{S}^d$, we have*

(i) $E\xi_1 \cdots \xi_d f = E\xi^d f = \nu^d f$,

(ii) $E(\xi_1 \cdots \xi_d f)^2 \leq \sum_{J \in 2^d} \nu^{J^c} (\nu^J f)^2$,

(iii) $E(\xi^d f)^2 \leq \sum_{k \leq d} k! \binom{d}{k}^2 \nu^{d-k} (\nu^k \bar{f})^2$.

When $\xi, \xi_1, \xi_2, \dots$ are Poisson, equality holds in (ii)–(iii), *and*

(iv) $E\psi(\xi_1 \cdots \xi_d f) \leq \psi(\nu\psi(\cdots \nu(\psi \circ f) \cdots))$.

Proof: (i) We may start with indicator functions of non-diagonal, measurable rectangles, and then extend by a monotone-class argument.

(ii) Writing η_j for the projection of ξ_j onto the diagonal in $\bar{S}^2$, and letting ζ_i denote the non-diagonal part of ξ_i^2, we have the diagonal decomposition

$$(\xi_1 \cdots \xi_d)^2 = \sum_{J \in 2^d} \bigotimes_{i \in J^c} \zeta_i \bigotimes_{j \in J} \eta_j.$$

Using the independence on the right, and noting that $E\zeta_i \le \nu^2$ and $E\eta_i = \nu_D$, the diagonal projection of ν, we get

$$\begin{aligned} E(\xi_1 \cdots \xi_d)^2 &= \sum_{J \in 2^d} \bigotimes_{i \in J^c} E\zeta_i \bigotimes_{j \in J} E\eta_j \\ &\le \sum_{J \in 2^d} (\nu^2)^{J^c} \otimes \nu_D^J, \end{aligned}$$

by extension from measurable rectangles, as before.

(iii) Proceed as in (ii), and use a combinatorial argument.

(iv) Equality holds for $d = 1$ by Lemma 3.25 (i). Now assume the truth in dimension $d - 1$. Then in d dimensions, we may condition on $\xi_2, \ldots, \xi_d$, and use Jensen's inequality, Fubini's theorem, and the monotonicity and concavity of ψ to get

$$\begin{aligned} E\psi(\xi_1 \cdots \xi_d f) &= E\psi(\nu\psi(\xi_2 \cdots \xi_d f)) \\ &\le \psi(\nu E\psi(\xi_2 \cdots \xi_d f)) \\ &\le \psi(\nu\psi(\nu\psi(\cdots \nu(\psi \circ f) \cdots))). \end{aligned}$$

□

Results for the integrals $\xi^d f$ can often be extended to $\xi_1 \cdots \xi_d f$ by a simple *chess-board* argument. Here we define $\tilde{\xi} = \sum_k \tilde{\xi}_k$, on the union $\bigcup_k \bar{S}_k$ of disjoint copies $\bar{S}_1, \ldots, \bar{S}_d$ of $\bar{S}$, with associated independent copies $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ of ξ. Writing $\tilde{f}$ for the periodic extension of f to $(\bigcup_k \bar{S}_k)^d$, and letting f' be the restriction of $\tilde{f}$ to $\bar{S}_1 \times \cdots \times \bar{S}_d$, we get

$$\{\xi_1 \cdots \xi_d f\} \stackrel{d}{=} \{\tilde{\xi}_1 \cdots \tilde{\xi}_d \tilde{f}\} \stackrel{d}{=} \{\tilde{\xi}^d f'\}.$$

We turn to a basic continuity property of the integrands, which also allows a reduction to the case of zero limits.

Theorem 10.2 *(integrand reduction and convergence) For any $\xi_1 \cdots \xi_d$-integrable functions $f, f_1, f_2, \ldots, g \ge 0$ on $\bar{S}^d$, the conditions*

(i) $\xi_1 \cdots \xi_d f_n \to \xi_1 \cdots \xi_d f$ *a.s.,*

(ii) $\xi_1 \cdots \xi_d |f_n - f| \to 0$ *a.s.,*

(iii) $f_n \to f$ *a.e.* ν^d,

are related by (i) ⇔ (ii) ⇒ (iii), *and they are all equivalent when $f_n \wedge 1 \le g$ for all n. Similar results hold for ξ^d-integrals, when $f, f_1, f_2, \ldots$ are symmetric non-diagonal, and all statements remain true for convergence in probability and locally in measure, respectively.*

Proof: It is enough to consider the integrals $\xi^d f_n$ and $\xi^d f$, since the statements for $\xi_1 \cdots \xi_d f_n$ and $\xi_1 \cdots \xi_d f$ then follow by a chess-board argument. Here we let $f, f_1, f_2, \ldots$ be symmetric, non-diagonal, and a.s. ξ^d-integrable. It is also enough to consider a.e. convergence, since the versions for convergence in probability and measure will then follow by a sub-sequence argument.

(i) $\Rightarrow$ (iii): Assume $\xi^d f_n \to \xi^d f$ a.s. Fix any disjoint sets $B_1, \ldots, B_d \in \hat{\mathcal{S}}$, and put $\bar{B}_k = B_k \times T$ and $R = \bar{B}_1 \times \cdots \times \bar{B}_d$. Since ξ has independent increments, conditioning on $\xi \bar{B}_k = 0$ is equivalent to restricting $f, f_1, f_2, \ldots$ to the set $A_k = (\bar{B}_k^c)^d$. Taking differences, and proceeding recursively in d steps, we conclude from the given convergence that

$$\xi^d(1_U f_n) \to \xi^d(1_U f) \text{ a.s.,} \tag{1}$$

where

$$U = \bigcap_{k \le d} A_k^c = \bigcup_{p \in \mathcal{P}_d} \Big\{ x \in \bar{S}^d;\, x \circ p \in R \Big\},$$

$\mathcal{P}_d$ being the set of permutations of $\{1, \ldots, d\}$.

Now put $S' = \{s \in S;\, \nu(\{s\} \times T) > 0\}$, and let $B_1, \ldots, B_d$ be disjoint, with either $B_k = \{s\}$ for some $s \in S'$, or $B_k \in \hat{\mathcal{S}} \cap S'^c$ with $\nu \bar{B}_k > 0$. Conditioning on the event $\bigcap_k \{\xi \bar{B}_k = 1\}$, so that ξ is supported in $\bar{B}_1, \ldots, \bar{B}_d$ by some unique points $\sigma_1, \ldots, \sigma_d$, we see from Theorem 3.4 and Lemma 3.18 that the distribution of $\sigma = (\sigma_1, \ldots, \sigma_d)$ is proportional to ν^d on R. The symmetry of $f, f_1, f_2, \ldots$ yields

$$\xi^d(1_U f) = \sum_{p \in \mathcal{P}_d} f(\sigma \circ p) = d!\, f(\sigma),$$

and similarly for $\xi^d(1_U f_n)$. Hence, $f_n(\sigma) \to f(\sigma)$ a.s. by (1), which means that $f_n \to f$ a.e. ν^d on R. Since $f, f_1, f_2, \ldots$ are non-diagonal, hence supported by countably many rectangles R, the convergence extends to all of $\bar{S}^d$.

(iii) $\Rightarrow$ (i): Let $f_n \to f$ a.e. and $f_n \wedge 1 \le g$. Noting that $\xi^d\{g \ge 1\} \le \xi^d g < \infty$ a.s., we have $\xi^d(f_n;\, g \ge 1) \to \xi^d(f;\, g \ge 1)$ a.s., since the integrals reduce to finite sums. On the other hand, $f_n \le g$ when $g < 1$, and so $\xi^d(f_n;\, g < 1) \to \xi^d(f;\, g < 1)$ a.s., by dominated convergence. Hence, by combination, $\xi^d f_n \to \xi^d f$ a.s.

(ii) $\Rightarrow$ (i): If $\xi^d|f_n - f| \to 0$ a.s., then $|\xi^d f_n - \xi^d f| \le \xi^d|f_n - f| \to 0$ a.s., which means that $\xi^d f_n \to \xi^d f$ a.s.

(i) $\Rightarrow$ (ii): Assume $\xi^d f_n \to \xi^d f$ a.s. Then $f_n \to f$ a.e. ν^d by (iii), and so $f_n \wedge f \to f$ a.e. ν^d. Noting that $|a - b| = a + b - 2(a \wedge b)$ and $a \wedge b \le b$ for any $a, b \in \mathsf{R}$, we get a.s. by (i)

$$\begin{aligned} \xi^d|f_n - f| &= \xi^d f_n + \xi^d f - 2\xi^d(f_n \wedge f) \\ &\to 2\xi^d f - 2\xi^d f = 0. \end{aligned}$$

□

Some basic properties hold simultaneously for the integrals $\xi_1 \cdots \xi_d f$ and $\xi^d f$. This will often be useful below, since the proofs for the former are usually much easier. For notational convenience, here and below, we consider only processes ξ with a fixed intensity ν. The results for differing intensities ν_k are similar and follow as simple corollaries.

Theorem 10.3 *(decoupling) For any non-diagonal, measurable functions $f, f_1, f_2, \ldots \geq 0$ on $\bar{S}^d$, we have*

(i) $\xi_1 \cdots \xi_d f < \infty$ *a.s.* $\Leftrightarrow$ $\xi^d f < \infty$ *a.s.,*

(ii) $\xi_1 \cdots \xi_d f_n \xrightarrow{P} 0 \;\Leftrightarrow\; \xi^d f_n \xrightarrow{P} 0$,

(iii) *the sequences $(\xi_1 \cdots \xi_d f_n)$ and $(\xi^d f_n)$ are simultaneously tight,*

(iv) $\xi_1 \cdots \xi_d f_n \xrightarrow{P} \xi_1 \cdots \xi_d f < \infty \;\Rightarrow\; \xi^d f_n \xrightarrow{P} \xi^d f < \infty$, *with equivalence when the f_n are symmetric.*

Our proof is based on a moment comparison of independent interest.

Lemma 10.4 *(moment decoupling) For any nondecreasing function $\varphi \geq 0$ of moderate growth, and for non-diagonal functions $f \geq 0$ on $\bar{S}^d$, we have*

$$E\varphi(\xi^d f) \asymp E\varphi(\xi_1 \cdots \xi_d f),$$

where the domination constants depend only on φ and d.

Proof: First we note that, for any $x_1, \ldots, x_n \geq 0$,

$$\begin{aligned} \varphi(x_1 + \cdots + x_n) &\leq \varphi(n \max_k x_k) \\ &\lesssim \varphi(\max_k x_k) \\ &= \max_k \varphi(x_k) \leq \textstyle\sum_k \varphi(x_k), \end{aligned}$$

uniformly for fixed n. To prove the asserted relation, we may take $S = \mathsf{R}_+$. It is then enough to consider tetrahedral functions f, since for general f with tetrahedral components f_k, $k \leq d!$, we then obtain

$$\begin{aligned} E\varphi(\xi^d f) &= E\varphi\Big(\textstyle\sum_k \xi^d f_k\Big) \asymp \textstyle\sum_k E\varphi(\xi^d f_k) \\ &\asymp \textstyle\sum_k E\varphi(\xi_1 \cdots \xi_d f_k) \\ &\asymp E\varphi\Big(\textstyle\sum_k \xi_1 \cdots \xi_d f_k\Big) \\ &= E\varphi(\xi_1 \cdots \xi_d f). \end{aligned}$$

This allows us to regard $\xi^d f$ and $\xi_1 \cdots \xi_d f$ as iterated predictable integrals.

The claim is trivial for $d = 1$. Assuming it is true in dimension $d-1$, we turn to the d-dimensional case. Choosing $\xi \perp\!\!\!\perp \xi_1$, we note that ξ_1 and ξ are tangential for the induced filtration $\mathcal{F}$, and that the inner integral $\xi^{d-1} f$ is $\mathcal{F}$-predictable. Using Fubini's theorem, the induction hypothesis, and Theorem 9.57, we get

$$
\begin{aligned}
E\varphi(\xi_1\cdots\xi_d f) &= E\varphi(\xi_2\cdots\xi_d\,\xi_1 f)\\
&\asymp E\varphi(\xi^{d-1}\xi_1 f)\\
&= E\varphi(\xi_1\,\xi^{d-1} f) \asymp E\varphi(\xi^d f).
\end{aligned}
\qquad \Box
$$

Proof of Theorem 10.3: (ii) Use Lemma 10.4.

(iii) This follows from (ii), since a sequence of random variables $\alpha_1, \alpha_2, \ldots$ is tight iff $c_n\alpha_n \xrightarrow{P} 0$ whenever $c_n \to 0$ (FMP 4.9).

(i) Use (iii) with $f_n \equiv f$.

(iv) Use (ii) and Theorem 10.2. $\Box$

Our next aim is to characterize the $\xi_1\cdots\xi_d$ -integrability of functions on $\bar{S}^d$. Given a measurable function $f \geq 0$ on $\bar{S}^J$, where $m = |J| < \infty$, we define recursively some functions $f_1, \ldots, f_m \geq 0$ on $\bar{S}^J$ by

$$
f_1 = f \wedge 1, \qquad f_{k+1} = f_k \prod_{|I|=k} 1\{\nu^I f_k \leq 1\}, \quad 1 \leq k < m, \tag{2}
$$

where the product extends over all sets $I \subset J$ with $|I| = k$, and ν^I denotes integration in the arguments indexed by I, so that $\nu^I f_k$ becomes a measurable function of the remaining arguments indexed by $J \setminus I$.

Now define recursively some classes $\mathcal{C}_d$ of measurable functions $f \geq 0$ on $\bar{S}^d$. Starting with $\mathcal{C}_0 = \{0, 1\}$, and assuming $\mathcal{C}_k$ to be known for all $k < d$, we define $\mathcal{C}_d$ by the conditions

$$
\begin{cases} \nu^{d-k}\{\nu^J f_k = \infty\} = 0, \\ 1\{\nu^J f_k > 1\} \in \mathcal{C}_{d-k}, \end{cases} \qquad J \in 2^d \text{ with } k = |J| > 0.
$$

We may now characterize the a.s. finiteness of the multiple integrals $\xi_1\cdots\xi_d f$ and $\xi^d f$.

Theorem 10.5 *(integrability criteria) For any measurable function $f \geq 0$ on $\bar{S}^d$, we have*

$$
P\{\xi_1\cdots\xi_d f < \infty\} = 1\{f \in \mathcal{C}_d\},
$$

and similarly for $\xi^d f$, when f is non-diagonal.

For $d = 1$ we get $\xi f < \infty$ a.s. iff $\nu(f \wedge 1) < \infty$, in agreement with Lemma 3.26. For $d = 2$, let $f \geq 0$ be non-diagonal on $\bar{S}^2$, and put

$$
f_1(x) = \nu\{f(x, \cdot) \wedge 1\}, \quad f_2(x) = \nu\{f(\cdot, x) \wedge 1\}, \qquad x \geq 0.
$$

Then the conditions for $\xi_1\xi_2 f < \infty$ or $\xi^2 f < \infty$ a.s. become

(i) $\nu\{f_1 \vee f_2 = \infty\} = 0$,

(ii) $\nu\{f_1 \vee f_2 > 1\} < \infty$,

(iii) $\nu^2\Big(f \wedge 1;\, \{f_1 \le 1\} \times \{f_2 \le 1\}\Big) < \infty.$

Our proof of Theorem 10.5 will be based on several lemmas.

Lemma 10.6 *(divergent series) For any random variables* $\alpha_1, \alpha_2, \ldots \ge 0$, *we have*

$$P\Big\{\sum\nolimits_k \alpha_k = \infty\Big\} \ge \lim_{\varepsilon \downarrow 0} \inf_k P\{\alpha_k > \varepsilon\}.$$

Proof: By Fatou's lemma, we have for any $\varepsilon > 0$

$$\begin{aligned} P\Big\{\sum\nolimits_k \alpha_k = \infty\Big\} &\ge P\{\alpha_k > \varepsilon \text{ i.o.}\} \\ &\ge \limsup_{k\to\infty} P\{\alpha_k > \varepsilon\} \\ &\ge \inf_k P\{\alpha_k > \varepsilon\}. \end{aligned}$$

It remains to let $\varepsilon \to 0$. □

Lemma 10.7 *(tails and integrability) Let* $f \ge 0$ *on* $\bar{S}^d$, *with* $\nu^J f \le 1$ *for all* $J \in 2^d \setminus \{\emptyset\}$. *Then*

(i) $\quad P\Big\{\xi_1 \cdots \xi_d f \ge r\nu^d f\Big\} \ge \dfrac{(1-r)^2}{1 + (2^d - 1)/\nu^d f}, \quad r \in (0,1),$

(ii) $\quad \xi_1 \cdots \xi_d f < \infty$ *a.s. iff* $\nu^d f < \infty$.

Proof: (i) Since $(\nu^J f)^2 \le \nu^J f$ for all $J \in 2^d \setminus \{\emptyset\}$, Lemma 10.1 yields

$$\begin{aligned} E(\xi_1 \cdots \xi_d f)^2 &\le (\nu^d f)^2 + (2^d - 1)\nu^d f \\ &= \Big\{1 + (2^d - 1)/\nu^d f\Big\}\,(E\xi_1 \cdots \xi_d f)^2, \end{aligned}$$

and the assertion follows by the Paley–Zygmund inequality (FMP 4.1).

(ii) Part (i) shows that $\nu^d f = \infty$ implies $\xi_1 \cdots \xi_d f = \infty$ a.s., and the converse holds by Lemma 10.1 (i). □

Lemma 10.8 *(partial integrability) For fixed* $k < d$, *let* $B \subset \bar{S}^{d-k}$, *and* $f \ge 0$ *on* $\bar{S}^d$ *with* $\nu^J f \le 1$ *for all* $J \in 2^k \setminus \{\emptyset\}$. *Then*

(i) *when* $\nu^k f = \infty$ *on* B,

$$\xi_1 \cdots \xi_d(f;\, \bar{S}^k \times B) < \infty \ \text{ a.s.} \ \Leftrightarrow \ \nu^{d-k} B = 0,$$

(ii) *when* $\nu^k f \in [1, \infty)$ *on* B,

$$\xi_1 \cdots \xi_d(f;\, \bar{S}^k \times B) < \infty \ \text{ a.s.} \ \Leftrightarrow \ \xi_{k+1} \cdots \xi_d B < \infty \ \text{ a.s.}$$

Proof: (i) By Fubini's theorem,

$$\xi_1 \cdots \xi_d(f;\, \bar{S}^k \times B) = \xi_{k+1} \cdots \xi_d(\xi_1 \cdots \xi_k f; B). \tag{3}$$

If $\nu^k f = \infty$ on B, then $\xi_1 \cdots \xi_k f = \infty$ a.s. on B, by Lemma 10.7 (ii), and so (3) is a.s. finite iff $\xi_{k+1} \cdots \xi_d B = 0$ a.s., which is equivalent to $\nu^{d-k} B = 0$, by Lemma 10.1 (i).

(ii) If $\nu^k f < \infty$ on B, then $\xi_1 \cdots \xi_k f < \infty$ a.s. on B, by Lemma 10.1 (i), and so (3) is a.s. finite when $\xi_{k+1} \cdots \xi_d B < \infty$ a.s. If also $\nu^k f \geq 1$ on B, then Lemma 10.7 (i) yields

$$P\{\xi_1 \cdots \xi_k f \geq r\} \geq 2^{-k}(1-r)^2 \text{ on } B, \quad r \in (0,1),$$

and so by (3), Fubini's theorem, and Lemma 10.6,

$$P\Big\{\xi_1 \ldots \xi_d(f;\, \bar{S}^k \times B) = \infty \,\Big|\, \xi_{k+1} \cdots \xi_d B = \infty\Big\} \geq 2^{-k},$$

which implies

$$P\Big\{\xi_1 \cdots \xi_d(f;\, \bar{S}^k \times B) = \infty\Big\} \geq 2^{-k} P\{\xi_{k+1} \cdots \xi_d B = \infty\}.$$

Thus, the a.s. finiteness in (3) implies $\xi_{k+1} \cdots \xi_d B < \infty$ a.s. □

Proof of Theorem 10.5: By Theorem 10.3 and Kolmogorov's $0-1$ law, it is enough to prove that $\xi_1 \ldots \xi_d f < \infty$ a.s. iff $f \in \mathcal{C}_d$. By induction on d, we then need to show that $\xi_1 \ldots \xi_d f < \infty$ a.s., iff

$$\begin{cases} \nu^{d-k}\{\nu^J f_k = \infty\} = 0, \\ \xi_{k+1} \cdots \xi_d\{\nu^J f_k > 1\} < \infty \text{ a.s.,} \end{cases} \quad J \in 2^d \text{ with } |J| = k > 0, \tag{4}$$

where the condition for $k = d$ is interpreted as $\nu^d f_d < \infty$. Since $\xi_1 \cdots \xi_d f < \infty$ iff $\xi_1 \cdots \xi_d f_1 < \infty$, and since (4) agrees for f and f_1, we may assume that $f \leq 1$.

For $J \in 2^d \setminus \{\emptyset\}$, we introduce in $\bar{S}^d$ the associated sets

$$B_J = \begin{cases} \big\{f = f_k,\ \nu^J f_k > 1\big\}, & k = |J| < d, \\ \{f = f_d\}, & |J| = d. \end{cases} \tag{5}$$

Since $\bigcup_J B_J = \bar{S}^d$, we have $\xi_1 \cdots \xi_d f < \infty$, iff $\xi_1 \cdots \xi_d(f; B_J) < \infty$ for all $J \in 2^d \setminus \{\emptyset\}$. Noting that $f = f_k$ on $\operatorname{supp} f_k$, we get

$$f 1_{B_J} = \begin{cases} f_k 1\{\nu^J f_k > 1\}, & k = |J| < d, \\ f_d, & |J| = d. \end{cases} \tag{6}$$

When $k = |J| < d$, we see from Lemma 10.8, with $f = f_k$ and $B = \{\nu^J f = \infty\}$ or $\{1 < \nu^J f < \infty\}$, that $\xi_1 \cdots \xi_d(f; B_J) < \infty$ a.s., iff (4) holds for this particular J. If instead $|J| = d$, we see from Lemma 10.7 (ii) with $f = f_d$

that $\xi_1 \cdots \xi_d(f; B_J) < \infty$ a.s. iff $\nu^d f_d < \infty$. The asserted equivalence follows by combination of the two results. $\square$

A similar argument yields recursive criteria for convergence in probability to 0. For notational convenience, we consider only the case of a common intensity ν. Given some measurable functions $f_n \geq 0$ on $\bar{S}^d$, we define the associated truncated functions $f_n^1, \ldots, f_n^d$ as in (2).

Theorem 10.9 *(convergence criteria) For any measurable functions $f_1, f_2, \ldots \geq 0$ on $\bar{S}^d$, we have $\xi_1 \cdots \xi_d f_n \xrightarrow{P} 0$, iff $\nu^d f_n^d \to 0$ and*

$$\xi_1 \cdots \xi_{d-k}\{\nu^J f_n^k > 1\} \xrightarrow{P} 0, \quad J \in 2^d \ \textit{with} \ k = |J| > 0,$$

and similarly for the integrals $\xi^d f_n$, when the f_n are non-diagonal.

Here the proof requires a sequential version of Lemma 10.8 (ii).

Lemma 10.10 *(partial convergence) For any $k < d$ and $n \in \mathsf{N}$, let $A_n \subset \bar{S}^{d-k}$ and $f_n \geq 0$ on $\bar{S}^d$ be such that $\nu^J f_n \leq 1$ for $J \in 2^k \setminus \{\emptyset\}$ and $\nu^k f_n > 1$ on A_n. Then*

$$\xi_1 \cdots \xi_d(f_n;\, \bar{S}^k \times A_n) \xrightarrow{P} 0 \quad \Leftrightarrow \quad \xi_{k+1} \cdots \xi_d A_n \xrightarrow{P} 0.$$

Proof: If $\xi_{k+1} \cdots \xi_d A_n \xrightarrow{P} 0$, then Fubini's theorem gives

$$P\Big\{\xi_1 \cdots \xi_d(f_n;\, \bar{S}^k \times A_n) > 0\Big\} \leq P\{\xi_{k+1} \cdots \xi_d A_n > 0\} \to 0,$$

and so $\xi_1 \cdots \xi_d(f_n;\, \bar{S}^k \times A_n) \xrightarrow{P} 0$.

Conversely, Lemma 10.1 (ii) yields on A_n

$$\begin{aligned} E(\,\xi_1 \cdots \xi_k f_n)^2 &= \sum_{J \in 2^k} \nu^{J^c} (\nu^J f_n)^2 \\ &\leq (\nu^k f_n)^2 + \sum_{J \in (2^k)'} \nu^{J^c} \nu^J f_n \\ &= (\nu^k f_n)^2 + (2^k - 1)\nu^k f_n \\ &\leq 2^k (\nu^k f_n)^2 = 2^k (E\xi_1 \cdots \xi_k f_n)^2. \end{aligned}$$

We also note that, on $\{\xi_{k+1} \cdots \xi_d A_n > 0\}$,

$$\begin{aligned} 1 &\leq \xi_{k+1} \cdots \xi_d A_n \\ &\leq \xi_{k+1} \cdots \xi_d(\nu^k f_n; A_n), \end{aligned}$$

and further that a.s.

$$E\Big\{\,\xi_1 \cdots \xi_d(f_n;\, \bar{S}^k \times A_n)\Big|\, \xi_{k+1}, \ldots, \xi_d\Big\} = \xi_{k+1} \cdots \xi_d(\nu^k f_n; A_n).$$

Applying the Paley–Zygmund inequality to the conditional distributions, given $\xi_{k+1},\dots,\xi_d$, we get

$$\begin{aligned}
P\Big\{\,\xi_1\cdots\xi_d(f_n;\,\bar S^k\times A_n)\ge r\Big\}
&\ge P\Big\{\xi_1\cdots\xi_d(f_n;\,\bar S^k\times A_n)\ge r,\ \xi_{k+1},\dots,\xi_d(\nu^k f_n;A_n)>0\Big\}\\
&\ge 2^{-k}(1-r)^2P\Big\{\xi_{k+1}\cdots\xi_dA_n>0\Big\},
\end{aligned}$$

as long as $\nu^k f_n<\infty$ on A_n. This extends to the general case, by a simple truncation. When $\xi_1\cdots\xi_d(f_n;\,\bar S^k\times A_n)\xrightarrow{P}0$, the probability on the left tends to 0, and we get $\xi_{k+1}\cdots\xi_dA_n\xrightarrow{P}0$. □

Proof of Theorem 10.9: By Corollary 10.3 (ii), it is enough to consider the integrals $\xi_1\cdots\xi_d f_n$. Since $\xi_1\cdots\xi_d$ is Z_+-valued, we have $\xi_1\cdots\xi_d f_n\xrightarrow{P}0$ iff $\xi_1\cdots\xi_d f_n^1\xrightarrow{P}0$, and so we may assume that $f_n\le 1$ for all n. Defining the sets $B_n^J\subset\bar S^d$ as in (5), we note as before that $\xi_1\cdots\xi_d f_n\xrightarrow{P}0$ iff $\xi_1\cdots\xi_d(f_n;B_n^J)\xrightarrow{P}0$ for all $J\in 2^d\setminus\{\emptyset\}$, and that the counterpart of (6) holds for the functions f_n and sets B_n^J. When $|J|<d$, we may apply Lemma 10.10, with $f_n=f_n^k$ and $A_n=\{\nu^J f_n>1\}$, to see that $\xi_1\cdots\xi_d(f_n;B_n^J)\xrightarrow{P}0$, iff $\xi_1\cdots\xi_{d-k}\{\nu^J f_n^k>1\}\xrightarrow{P}0$. If instead $|J|=d$, then $\nu^d f_n^d\to 0$ yields $\xi_1\cdots\xi_d(f_n;B_n^J)\xrightarrow{P}0$, and the converse holds by Lemma 10.7 (i). □

10.2 Symmetric Point-process Integrals

Given a simple point process ξ on S, we may form a symmetric randomization $\xi=\xi^+ + \xi^-$ on $S\times\{\pm1\}$, generated by the distribution $\nu=\frac12(\delta_1+\delta_{-1})$, and form the *symmetrization* $\tilde\xi=\xi^+-\xi^-$. Equivalently, we may fix any atomic representation $\xi=\sum_k\delta_{\tau_k}$, and define $\tilde\xi=\sum_k\sigma_k\delta_{\tau_k}$, where the σ_k are i.i.d. random signs independent of the τ_k with distribution $P\{\sigma_k=\pm1\}=\frac12$. Note that the distribution of $\tilde\xi$ is independent of the enumeration of atoms $\tau_1,\tau_2,\dots$. For the point processes $\xi_1,\dots,\xi_d$, we may choose the associated sign sequences $(\sigma_{i,k})$ to be mutually independent, to ensure the conditional independence, given $\xi_1,\dots,\xi_d$, of the associated symmetrized processes $\tilde\xi_1,\dots,\tilde\xi_d$. The latter will then be independent, iff independence holds between the ξ_k.

In this section, we consider multiple integrals of the form $\tilde\xi_1\cdots\tilde\xi_d f$ or $\tilde\xi^d f$ for suitable functions f on S^d, where $\tilde\xi$ and $\tilde\xi_1,\dots,\tilde\xi_d$ are conditionally independent symmetrizations of some simple point processes ξ and $\xi_1,\dots,\xi_d$ on S. Given a measurable function f on S^d, we define $\tilde\xi_1\cdots\tilde\xi_d f$ as the limit in probability of the elementary integrals $\tilde\xi_1\cdots\tilde\xi_d f_n$, for any functions $f_n\to f$ of bounded supports satisfying $|f_n|\le|f|$, whenever the limit exists and is independent of the choice of $f_1,f_2,\dots$. The definition of the integral $\tilde\xi^d f$ is similar.

In the special case where ξ and $\xi_1, \cdots, \xi_d$ are deterministic and equal to the counting measure on N, the symmetric integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ and $\tilde{\xi}^d f$ reduce to *multi-linear forms* in the associated sign sequences σ and $\sigma_1, \ldots, \sigma_d$, here written as

$$\sigma_1 \cdots \sigma_d\, A = \sum_{k \in \mathsf{N}^d} a_k\, \sigma_{1,k_1} \cdots \sigma_{d,k_d},$$
$$\sigma^d A = \sum_{k \in \mathsf{N}^d} a_k\, \sigma_{k_1} \cdots \sigma_{k_d},$$

in terms of some matrices $A = (a_k)$ indexed by N^d. They are said to *exist*, whenever the sums converge in L^2.

Returning to the general case, we have formally

$$\begin{aligned}
\tilde{\xi}_1 \cdots \tilde{\xi}_d f &= \sum_{k \in \mathsf{N}^d} f(\tau_{1,k_1}, \ldots, \tau_{d,k_d}) \prod_{i \le d} \sigma_{i,k_i} \\
&= (\sigma_1 \cdots \sigma_d) f(\xi_1 \cdots \xi_d), \\
\tilde{\xi}^d f &= \sum_{k \in \mathsf{N}^d} f(\tau_{k_1}, \ldots, \tau_{k_d}) \prod_{i \le d} \sigma_{k_i} = \sigma^d f(\xi^d),
\end{aligned}$$

where $f(\xi^d)$ denotes the random array with entries $f(\tau_{k_1}, \ldots, \tau_{k_d})$, and similarly for $f(\xi_1 \cdots \xi_d)$. The representations clearly depend on our choice of measurable enumerations of the atoms of ξ or $\xi_1, \ldots, \xi_d$. For a simple choice, we may reduce by a Borel isomorphism to the case where S is a Borel subset of R_+, and then enumerate the atoms in increasing order.

We may now state the basic existence and representation theorem for multiple symmetric integrals. Recall that a function f on S^d is said to be *non-diagonal*, if it vanishes on all diagonal spaces, where two or more coordinates agree.

Theorem 10.11 *(existence and representation) Let ξ and $\xi_1, \ldots, \xi_d$ be simple point processes on S with symmetrizations $\tilde{\xi}$ and $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$, generated by the independent sign sequences σ and $\sigma_1, \ldots, \sigma_d$, and fix a measurable function f on S^d. Then*

(i) *the integral $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ exists iff $\xi_1 \cdots \xi_d f^2 < \infty$ a.s., and similarly for the integrals $\tilde{\xi}^d f$ and $\xi^d f^2$, when f is symmetric and non-diagonal,*

(ii) *the following representations hold a.s., whenever either side exists:*

$$\tilde{\xi}_1 \cdots \tilde{\xi}_d f = (\sigma_1 \ldots \sigma_d) f(\xi_1 \cdots \xi_d), \qquad \tilde{\xi}^d f = \sigma^d f(\xi^d).$$

In other words, the two sides in (ii) exist simultaneously, in which case they are equal. By independence of the sign sequences, we may form the multi-linear forms on the right, by first conditioning on the processes ξ and $\xi_1, \ldots, \xi_d$. It is then clear from part (i) that the expressions on the right exist, iff $\xi_1 \cdots \xi_d f^2 < \infty$ or $\xi^d f^2 < \infty$ a.s., respectively.

For the proof, we begin with the special case of multi-linear forms, where more can be said. Given an array A indexed by N^d, and a constant $p > 0$, we define the associated l^p-norms $|A|_p$ by $|A|_p^p = \sum_k |a_k|^p$. Say that A is *non-diagonal*, if all diagonal elements vanish, and *tetrahedral*, if it is supported by the indices $k \in \mathsf{N}^d$ with $k_1 > \cdots > k_d$. The forms $\sigma_1 \cdots \sigma_d \, A$ or $\sigma^d A$ are said to be *tight*, if the associated families of finite partial sums are tight.

Proposition 10.12 *(random multi-linear forms, Krakowiak & Szulga) Consider some independent random sign sequences $\sigma, \sigma_1, \ldots, \sigma_d$, along with arrays $A = (a_k)$ indexed by N^d, tetrahedral in case of σ^d.*

(i) *If $|A|_2 < \infty$, then $\sigma_1 \cdots \sigma_d \, A$ and $\sigma^d A$ converge in L^p for every $p > 0$, and convergence holds a.s. along the sets $\{1, \ldots, n\}^d$. Furthermore, we have for every $p > 0$*

$$\|\sigma_1 \cdots \sigma_d \, A\|_p \asymp |A|_2, \qquad \|\sigma^d A\|_p \asymp |A|_2, \tag{7}$$

with equality for $p = 2$.

(ii) *When $|A|_2 = \infty$, the forms $\sigma_1 \cdots \sigma_d \, A$ and $\sigma^d A$ are not even tight.*

Here our proof relies on some elementary hyper-contraction criteria.

Lemma 10.13 *(hyper-contraction, Marcinkiewicz, Paley, Zygmund) For any class $\mathcal{C} \subset L^p(\Omega)$ with $p > 0$, these conditions are equivalent:*

(i) *There exists a constant $b > 0$ with*

$$P\{|\alpha| \geq b \, \|\alpha\|_p\} > b, \quad \alpha \in \mathcal{C},$$

(ii) *there exist some constants $q \in (0, p)$ and $c > 0$ with*

$$\|\alpha\|_p \leq c \, \|\alpha\|_q, \quad \alpha \in \mathcal{C},$$

(iii) *for every $q \in (0, p)$, there exists a constant $c = c_q > 0$, as in* (ii).

Proof: Clearly (iii) $\Rightarrow$ (ii).

(ii) $\Rightarrow$ (i): Assume (ii) for some $q \in (0, p)$ and $c > 0$, and write $q^{-1} = p^{-1} + r^{-1}$ and $q' = q \wedge 1$. Fixing any $t \in (0, 1)$, and putting $\hat{\alpha} = \alpha \, 1\{|\alpha| > t\|\alpha\|_q\}$, we get by Minkowski's inequality

$$\|\alpha\|_q^{q'} \leq \|\hat{\alpha}\|_q^{q'} + t^{\hat{q}} \|\alpha\|_q^{q'}.$$

Hence, by Hölder's inequality,

$$(1 - t^{q'})^{1/\hat{q}} \, \|\alpha\|_q \leq \|\hat{\alpha}\|_q \leq \|\alpha\|_p \left(P\Big\{|\alpha| > t \, \|\alpha\|_q\Big\}\right)^{1/r},$$

and so, for $\|\alpha\|_p > 0$, we get from (ii)

$$\begin{aligned} P\Big\{|\alpha| > t \, c^{-1} \, \|\alpha\|_p\Big\} &\geq P\Big\{|\alpha| > t \, \|\alpha\|_q\Big\} \\ &\geq (1 - t^{q'})^{r/q'} \left(\|\alpha\|_q / \|\alpha\|_p\right)^r \\ &\geq (1 - t^{q'})^{r/q'} c^{-r} > 0, \end{aligned}$$

which proves (i) with $b = tc^{-1} \wedge (1 - t^{q'})^{r/q'} c^{-r}$.

(i) ⇒ (iii): Assuming (i) for some $b > 0$, we get for any $q \in (0, p)$

$$\begin{aligned} E|\alpha|^q &\geq E\Big\{|\alpha|^q;\, |\alpha| \geq b\,\|\alpha\|_p\Big\} \\ &\geq b^q\,\|\alpha\|_p^q\,P\Big\{|\alpha| \geq b\,\|\alpha\|_p\Big\} \\ &\geq b^{q+1}\,\|\alpha\|_p^q\,, \end{aligned}$$

which proves the relation in (ii) with $c = b^{-1-1/q}$. □

Proof of Proposition 10.12: (i) First we prove (7) for arrays A with finite support. The general result will then follow by continuity, once the convergence claim is established. For $d = 1$ and $n \in \mathsf{N}$, the moment $E\prod_{j\leq 2n}\sigma_{k_j}$ vanishes, unless the indices k_j agree pairwise. Hence,

$$\begin{aligned} \|\sigma A\|_{2n}^{2n} &= E\Big(\sum\nolimits_k a_k\sigma_k\Big)^{2n} \\ &= \sum_{k_1,\dots,k_{2n}}\prod_{i\leq 2n} a_{k_i}\,E\prod_{j\leq 2n}\sigma_{k_j} \\ &\lesssim \sum_{k_1,\dots,k_n}\prod_{i\leq n} a_{k_i}^2 \\ &= \Big(\sum\nolimits_k a_k^2\Big)^n = |A|_2^{2n}, \end{aligned}$$

and so $\|\sigma A\|_{2n} \lesssim |A|_2$. Since equality holds for $n = 1$, we get $\|\sigma A\|_p \lesssim \|\sigma A\|_2$ for every $p \geq 2$, which implies $\|\sigma A\|_p \asymp |A|_2$ for all p, by Lemma 10.13.

The result for $d = 1$ may be extended to l^2-valued sequences $A = (a_{ik}) = (A^i) = (A_k)$. Then write $|\cdot|_2$ for the l^2-norm, and conclude from Minkowski's inequality and the result for real A that

$$\begin{aligned} \|\sigma A\|_{2n}^2 &= \Big\||\sigma A|_2^2\Big\|_n = \Big\|\sum\nolimits_i(\sigma A_i)^2\Big\|_n \\ &\leq \sum\nolimits_i\Big\|(\sigma A_i)^2\Big\|_n \\ &= \sum\nolimits_i\|\sigma A_i\|_{2n}^2 \\ &\lesssim \sum\nolimits_i|A_i|_2^2 = |A|_2^2, \end{aligned}$$

with equality for $n = 1$.

We proceed to the general relation $\|\sigma_1\cdots\sigma_d\,A\|_p \asymp |A|_2$. If this holds in dimension $d - 1$, then in d dimensions we get by conditioning on $\sigma = \sigma_1$

$$\begin{aligned} \|\sigma_1\cdots\sigma_d\,A\|_p^p &= E|\sigma_1\cdots\sigma_d\,A|^p \\ &= EE\Big(|\sigma_1\cdots\sigma_d\,A|^p\,\Big|\,\sigma_1\Big) \\ &\lesssim E\Big\{\sum\nolimits_k\Big(\sum\nolimits_i\sigma_i a_{ik}\Big)^2\Big\}^{p/2} \\ &= E\Big|\sum\nolimits_i\sigma_i a_i\Big|_2^p = \|\sigma A\|_p^p \\ &\lesssim \Big(\sum\nolimits_i|a_i|_2^2\Big)^{p/2} \\ &= \Big(\sum\nolimits_{i,k}a_{ik}^2\Big)^{p/2} = |A|_2^p, \end{aligned}$$

which again extends to $\|\sigma_1 \cdots \sigma_d A\|_p \asymp |A|_2$ for all $p > 0$.

Turning to the forms $\sigma^d A$, we note that for non-diagonal A, the random variables $\sigma_1^d A$ and $(\sigma_1 + \cdots + \sigma_d)^d A$ form a martingale. Furthermore, for distinct $k_1, \ldots, k_d \in \mathsf{N}$, we have by independence and symmetry

$$E(\, \sigma_{1,k_1} \cdots \sigma_{d,k_d} \,|\, \sigma_1 + \cdots + \sigma_d \,) = d^{-d} \prod_{j \le d} (\sigma_{1,k_j} + \cdots + \sigma_{d,k_j}).$$

Using Jensen's inequality and the relation for $\sigma_1 \cdots \sigma_d A$, we get for any $p \ge 1$

$$\begin{aligned}
\|\sigma_1^d A\|_p &\le \left\|(\sigma_1 + \cdots + \sigma_d)^d A\right\|_p \\
&= d^d \left\|E(\, \sigma_1 \cdots \sigma_d \, A \,|\, \sigma_1 + \cdots + \sigma_d \,)\right\|_p \\
&\le d^d \, \|\, \sigma_1 \cdots \sigma_d \, A \,\|_p \asymp |A|_2,
\end{aligned}$$

which extends as before to the general case.

Now let $|A|_2 < \infty$, and consider any finite subarrays $A_n \to A$. Then by (7), the sequence $\sigma_1 \cdots \sigma_d \, A_n$ is Cauchy in L^p for every $p > 0$, and so it converges in L^p to some limit $\sigma_1 \cdots \sigma_d \, A$, which is clearly independent of p and (A_n). The proof for $\sigma^d A$ is similar. Next write A_n for the restriction of A to $\{1, \ldots, n\}^d$. Then $\sigma_1 \cdots \sigma_d \, A_n$ and $\sigma^d A_n$ are L^2-bounded martingales, which converge a.s. and in L^2.

(ii) Let $|A|_2 = \infty$, and choose $A_n \to A$ as before, so that $|A_n|_2 \to \infty$. Then by (7) and Lemma 10.13, there exists a constant $b > 0$, such that

$$P\Big\{\, |\sigma_1 \cdots \sigma_d \, A_n| \ge b \, |A_n|_2 \Big\} > b, \quad n \in \mathsf{N}, \tag{8}$$

and so the sequence $\{\sigma_1 \cdots \sigma_d \, A_n\}$ fails to be tight. Again a similar argument applies to $\sigma^d A$. □

To prove the main result, we need some conditional moment relations for symmetric point process integrals, in the special case of integrands with bounded support, where the results are immediate consequences of Lemma 10.12. Once the main theorem is established, the general version will follow easily by Fubini's theorem.

Corollary 10.14 *(conditional moments) For fixed $p > 0$ and measurable functions f on S^d with $\xi_1 \cdots \xi_d f^2 < \infty$ or $\xi^d f^2 < \infty$ a.s., we have*

(i) $E\Big(|\tilde{\xi}_1 \cdots \tilde{\xi}_d f|^{2p} \,\Big|\, \xi_1, \ldots, \xi_d\Big) \asymp (\xi_1 \cdots \xi_d f^2)^p$ *a.s.,*

(ii) $E\Big(|\tilde{\xi}^d f|^{2p} \,\Big|\, \xi\Big) \asymp (\xi^d f^2)^p$ *a.s., when f is symmetric and non-diagonal.*

Equality holds in (i) *when $p = 1$, and also in* (ii) *when S is ordered and f is tetrahedral.*

Proof of Theorem 10.11: (i) First assume that $\xi_1 \cdots \xi_d f^2 < \infty$ a.s., and consider some functions $f_n \to f$ on S^d with bounded supports, satisfying $|f_n| \le |f|$. By Corollary 10.14 (i) with $p = 1$, the sequence $(\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n)$ is conditionally Cauchy in L^2, given $\xi_1, \ldots, \xi_d$. By dominated convergence, it remains unconditionally Cauchy for the metric $\rho(\alpha, \beta) = E(|\alpha - \beta| \wedge 1)$, and the desired convergence follows, since ρ is complete in L^0. To see that the limit is a.s. unique, consider another approximating sequence (g_n) with the same properties, and apply the previous result to the alternating sequence $f_1, g_1, f_2, g_2, \ldots$.

To prove the necessity of the condition $\xi_1 \cdots \xi_d f^2 < \infty$ a.s., choose any functions f_n as above. By Lemma 10.13 and Corollary 10.14, there exists a constant $b > 0$, such that a.s.

$$P\Big\{(\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n)^2 \ge b^2\, \xi_1 \cdots \xi_d f_n^2 \,\Big|\, \xi_1, \ldots, \xi_d\Big\} > b, \quad n \in \mathsf{N}.$$

Taking expected values and noting that $\xi_1 \cdots \xi_d f_n^2 \to \infty$ a.s. on the set $A = \{\xi_1 \cdots \xi_d f^2 = \infty\}$, we see that $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n$ fails to be tight when $PA > 0$, which excludes the convergence in L^p for any $p > 0$.

(ii) Fix any bounded sets $B_n \uparrow S$, and let f_n denote the restrictions of f to B_n^d. Then, trivially,

$$\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n = (\sigma_1 \ldots \sigma_d) f_n(\xi_1 \cdots \xi_d), \quad n \in \mathsf{N}.$$

Assuming $\xi_1 \cdots \xi_d f^2 < \infty$ a.s., and letting $n \to \infty$, we get

$$\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} \tilde{\xi}_1 \cdots \tilde{\xi}_d f,$$

by the definition of the integral, whereas

$$(\sigma_1 \ldots \sigma_d) f_n(\xi_1 \cdots \xi_d) \xrightarrow{P} (\sigma_1 \ldots \sigma_d) f(\xi_1 \cdots \xi_d),$$

by Proposition 10.12 and dominated convergence.

The assertion for $\tilde{\xi}^d f$ follows by the same argument, as long as f is symmetric. For general f, we may reduce by a Borel isomorphism to the case where $S = \mathsf{R}_+$. Then decompose f into its tetrahedral components f_k, $k \le d!$, and note as before that the integrals $\tilde{\xi}^d f_k$ exist, iff $\xi^d f_k^2 < \infty$ a.s. for all k, which is equivalent to $\xi^d f^2 < \infty$ a.s. Since trivially $\tilde{\xi}^d f$ exists whenever $\tilde{\xi}^d f_k$ exists for every k, it remains to prove the reverse implication.

Then assume that $\tilde{\xi}^d f$ exists. By symmetry, it is enough to show that even $\tilde{\xi}^d g$ exists, where g is the tetrahedral component of f supported by the set $\Delta = \{t_1 > \cdots > t_d\}$. For any $n \in \mathsf{N}$, let f_n and g_n denote the restrictions of f and g to the cube $[0, n]^d$, and define $f_{m,n} = f_m + g_n - g_m$ for $m < n$. As $m, n \to \infty$, we have $\tilde{\xi}^d f_{m,n} \xrightarrow{P} \tilde{\xi}^d f$ by hypothesis, and so

$$\begin{aligned}
\tilde{\xi}^d g_n - \tilde{\xi}^d g_m &= \tilde{\xi}^d (g_n - g_m) \\
&= \tilde{\xi}^d (f_{m,n} - f_{m,m}) \\
&= \tilde{\xi}^d f_{m,n} - \tilde{\xi}^d f_{m,m} \xrightarrow{P} 0,
\end{aligned}$$

which shows that the sequence $(\tilde{\xi}^d g_n)$ is Cauchy in L^0, and hence converges in probability as $n \to \infty$. In particular it is tight, which ensures the existence of $\tilde{\xi}^d g$. □

We turn to some basic convergence and tightness criteria.

Theorem 10.15 *(symmetrization) For processes $\xi_1, \dots, \xi_d$ and $\tilde{\xi}_1, \dots, \tilde{\xi}_d$ as in Theorem 10.11, and for any measurable functions $f_1, f_2, \dots$ on S^d with $\xi_1 \cdots \xi_d f_n^2 < \infty$ a.s., we have*

(i) $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} 0 \;\Leftrightarrow\; \xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} 0$,

(ii) $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \to 0$ *a.s.* $\Rightarrow\; \xi_1 \cdots \xi_d f_n^2 \to 0$ *a.s.*,

(iii) *the sequences $(\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n)$ and $(\xi_1 \cdots \xi_d f_n^2)$ are simultaneously tight,*

and similarly for the integrals $\tilde{\xi}^d f_n$ and $\xi^d f_n^2$, when the f_n are symmetric non-diagonal.

Proof: The proofs for $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n$ and $\tilde{\xi}^d f_n$ being similar, we consider only the former.

(ii) If $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \to 0$ a.s., the same convergence holds conditionally on $\xi_1, \dots, \xi_d$, which reduces the claim to the corresponding statement for multilinear forms. Then assume $|a_n|_2 \not\to 0$, so that $|a_n|_2 > \varepsilon > 0$ along a subsequence $N' \subset \mathsf{N}$. By (8), we have

$$P\Big\{ |\sigma_1 \cdots \sigma_d \, a_n| \ge b\,\varepsilon \Big\} > b, \quad n \in N',$$

which contradicts the a.s. convergence $\sigma_1 \cdots \sigma_d \, a_n \to 0$.

(i) By Corollary 10.14 (i) with $p = 1$,

$$E\Big\{(\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n)^2 \wedge 1\Big\} \le E\Big\{\xi_1 \cdots \xi_d f_n^2 \wedge 1\Big\}, \quad n \in \mathsf{N},$$

which shows that $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} 0$ implies $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} 0$. The reverse implication follows from (ii), applied to a.s. convergent sub-sequences.

(iii) This follows from (i), since the sequence $(\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n)$ is tight, iff $t_n \tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} 0$ for any constants $t_n \to 0$, and similarly for the sequence $(\xi_1 \cdots \xi_d f_n^2)$. □

The following Fubini-type theorem shows how symmetric multiple integrals can be formed recursively by repeated single integrations. Given some representations of $\tilde{\xi}_1, \dots, \tilde{\xi}_d$, in terms of independent sign sequences $\sigma_1, \dots, \sigma_d$, we define the integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d Y$ and $\tilde{\xi}^d Y$, for measurable processes Y on S^d with $\xi_1 \cdots \xi_d Y^2 < \infty$ or $\xi^d Y^2 < \infty$, respectively, by the a.s. relations

$$\tilde{\xi}_1 \cdots \tilde{\xi}_d Y = (\sigma_1 \dots \sigma_d)\, Y(\xi_1 \cdots \xi_d), \qquad \tilde{\xi}^d Y = \sigma^d Y(\xi^d).$$

Theorem 10.16 *(recursion) For $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ as in Theorem 10.11, and any measurable function f on S^d, the following relation holds a.s., whenever either side exists:*

$$\tilde{\xi}_1 \cdots \tilde{\xi}_d f = \tilde{\xi}_1(\tilde{\xi}_2(\cdots(\tilde{\xi}_d f)\cdots)).$$

Proof: By induction on d, it is enough to prove that

$$\tilde{\xi}_1 \cdots \tilde{\xi}_d f = \tilde{\xi}_1(\tilde{\xi}_2 \cdots \tilde{\xi}_d f) \text{ a.s.,}$$

whenever either side exist. First we give a proof for multi-linear forms, where the relation reduces to

$$\sigma_1 \cdots \sigma_d \, A = \sigma_1(\sigma_2 \cdots \sigma_d \, A) \text{ a.s.} \tag{9}$$

The general case then follows easily by conditioning on $\xi_1, \ldots, \xi_d$, using the representation in Theorem 10.11.

By definition of the iterated sum, we have for suitable arrays $A = (a_{j,k_2,\ldots,k_d})$

$$\begin{aligned}\sigma_1(\sigma_2 \cdots \sigma_d \, A) &= \textstyle\sum_j \sigma_{1j} \sum_k (\sigma_2 \cdots \sigma_d)_k \, a_{jk} \\ &= \textstyle\sum_j \sigma_{1j}(\sigma_2 \cdots \sigma_d \, A)_j,\end{aligned}$$

where the summations extend over all $j \in \mathsf{N}$ and $k = (k_2, \ldots, k_d) \in \mathsf{N}^{d-1}$. Assuming $|A|_2 < \infty$, we get by Fubini's theorem, along with the L^2/l^2-isometry from Theorem 10.12 in dimensions 1 and $d-1$

$$\begin{aligned}E\Big\{\sigma_1(\sigma_2 \cdots \sigma_d \, A)\Big\}^2 &= E\, E\Big(\Big\{\textstyle\sum_i \sigma_{1i}(\sigma_2 \cdots \sigma_d \, A)_i\Big\}^2 \Big| \, \sigma_2, \ldots, \sigma_d\Big) \\ &= E \textstyle\sum_i (\sigma_2 \cdots \sigma_d \, A)_i^2 \\ &= \textstyle\sum_i \sum_k a_{ik}^2 = |A|_2^2,\end{aligned}$$

which shows that

$$\Big\|\sigma_1(\sigma_2 \cdots \sigma_d \, A)\Big\|_2 = |A|_2 = \|\sigma_1 \cdots \sigma_d \, A\|_2.$$

In particular, both sides of (9) exist when $|A|_2 < \infty$. The a.s. relation (9) holds trivially when A has bounded support, and it extends to general A with $|A|_2 < \infty$, by the L^2-continuity of both sides. If instead $|A|_2 = \infty$, the existence fails on the right, since at least the outer sum then fails to exist, in view of the previous calculation. □

We now specialize to the case of marked point processes with independent increments. First we extend Corollary 10.3 to the symmetric case.

Corollary 10.17 *(decoupling) Let $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ be independent copies of a symmetric, T-marked point process $\tilde{\xi}$ on S with independent increments. Then for any non-diagonal, measurable functions $f, f_1, f_2, \ldots$ on $\bar{S}^d$ with $\xi_1 \cdots \xi_d f_n^2 < \infty$ a.s. for all $n \in \mathsf{N}$, we have*

(i) *$\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ and $\tilde{\xi}^d f$ exist simultaneously,*

(ii) $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} 0 \;\Rightarrow\; \tilde{\xi}^d f_n \xrightarrow{P} 0,$

(iii) *if* $(\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n)$ *is tight, then so is* $(\tilde{\xi}^d f_n)$,

with equivalence throughout when the f_n *are symmetric.*

Proof: (i) The integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ and $\tilde{\xi}^d f$ exist by Theorem 10.11, iff a.s. $\xi_1 \cdots \xi_d f^2 < \infty$ or $\xi^d f^2 < \infty$, respectively, where the latter conditions are equivalent by Corollary 10.3 (i).

(ii) Without loss, we may assume that $S = \mathsf{R}_+$. Then decompose each f_n into its tetrahedral components f_{nk}, $k \leq d!$. By Theorem 10.15, (ii) we have $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} 0$ iff $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} 0$, which holds by Corollary 10.3 (ii) iff $\xi^d f_n^2 \xrightarrow{P} 0$. This is clearly equivalent to $\xi^d f_{nk}^2 \xrightarrow{P} 0$, as $n \to \infty$ for fixed k. By Theorem 10.15 (ii), the latter condition holds iff $\tilde{\xi}^d f_{nk} \xrightarrow{P} 0$ for all k, which implies $\tilde{\xi}^d f_n \xrightarrow{P} 0$. The last implication clearly reduces to an equivalence when the f_n are symmetric.

(iii) This follows from (ii), as in case of Corollary 10.3 (iii). □

We conclude with a version of Theorem 10.2 for symmetric integrals.

Theorem 10.18 *(convergence and closure) Let* $\tilde{\xi}, \tilde{\xi}_1, \ldots, \tilde{\xi}_d$ *be i.i.d., symmetric,* T*-marked point processes on* S *with independent increments, and let* $f_1, f_2, \ldots, g$ *be* $\tilde{\xi}_1 \cdots \tilde{\xi}_d$*-integrable functions on* $\bar{S}^d$.

(i) *If* $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n$ *converges a.s., the limit equals* $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ *a.s. for some function* f *on* $\bar{S}^d$, *in which case* $f_n \to f$ *a.e.*

(ii) *If* $f_n \to f$ *in* ν^d_{loc} *with* $f_n^2 \wedge 1 \leq g^2$ *a.e., then* $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} \tilde{\xi}_1 \cdots \tilde{\xi}_d f$.

Similar results hold for the integrals $\tilde{\xi}^d f_n$, *when the* f_n *are symmetric, non-diagonal. Part* (i) *remains true for convergence in probability and locally in measure, respectively.*

Proof: We consider only the integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n$, the proof for $\tilde{\xi}^d f_n$ being similar.

(i) If $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n$ converges a.s., then as $m, n \to \infty$ we have $(\tilde{\xi}_1 \cdots \tilde{\xi}_d)(f_m - f_n) \to 0$ a.s., and so $(\xi_1 \cdots \xi_d)(f_m - f_n)^2 \to 0$ a.s. by Theorem 10.15 (i). Hence, Theorem 10.2 (i) yields $f_m - f_n \to 0$ a.e. ν^d, and so $f_n \to f$ a.e. ν^d for some function f on $\bar{S}^d$. By Fatou's lemma,

$$\begin{aligned}(\xi_1 \cdots \xi_d)(f_n - f)^2 &= (\xi_1 \cdots \xi_d) \lim_{m \to \infty} (f_n - f_m)^2 \\ &\leq \liminf_{m \to \infty} (\xi_1 \cdots \xi_d)(f_n - f_m)^2,\end{aligned}$$

which tends to 0 as $n \to \infty$. In particular, $\xi_1 \cdots \xi_d f^2 < \infty$ a.s., and so $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ exists by Theorem 10.11. Furthermore, Theorem 10.15 (ii) yields $(\tilde{\xi}_1 \cdots \tilde{\xi}_d)(f_n - f) \xrightarrow{P} 0$, which implies $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} \tilde{\xi}_1 \cdots \tilde{\xi}_d f$. The same convergence holds a.s. by hypothesis. The case of convergence in probability follows by a sub-sequence argument.

(ii) Taking differences, we may reduce to the case $f = 0$, so that $f_n^2 \to 0$ in $\nu_{\rm loc}^d$. Since $f_n^2 \wedge 1 \le g^2$ and $\xi_1 \cdots \xi_d \, g^2 < \infty$ a.s. by Theorem 10.11, we get $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} 0$ by Theorem 10.2 (i), and so $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} 0$ by Theorem 10.15 (ii). □

The closure property in Theorem 10.18 (i) has no counterpart for positive integrals $\xi_1 \cdots \xi_d f_n$. For a counter-example, let $\xi_1, \ldots, \xi_d$ be independent, unit rate Poisson processes on R_+, and choose $f_n = n^{-d} 1_{[0,n]^d}$, so that $\xi_1 \cdots \xi_d f_n \to 1$ a.s., by the law of large numbers. To see that the limit 1 has no representation as an integral $\xi_1 \cdots \xi_d f$, we may reduce by conditioning to the case of $d = 1$, and conclude from Lemma 3.25 that $\mathrm{Var}(\xi f) = \lambda f^2$, which can only vanish when $f = 0$ a.e. λ.

10.3 Escape Criteria

Here we consider conditions for divergence to infinity, for multiple integrals of the form $\xi_1 \cdots \xi_d f_n$ or $\xi^d f_n$, along with their symmetric counterparts $\tilde{\xi}_1, \cdots \tilde{\xi}_d f_n$ and $\tilde{\xi}^d f_n$. To avoid some technical and notational complications, we take all processes ξ and $\xi_1, \cdots, \xi_d$ to be Poisson with a common diffuse intensity ν, defined on an arbitrary Borel space S. All functions f and f_n are assumed to be measurable.

Some basic equivalences and implications are given below:

Theorem 10.19 *(decoupling and symmetrization) For any non-diagonal, measurable functions $f_1, f_2, \ldots$ on $\bar{S}^d$, with $\xi_1 \cdots \xi_n f_n^2 < \infty$ a.s. for all n, the conditions*

(i) $|\, \tilde{\xi}_1 \cdots \tilde{\xi}_d f_n| \xrightarrow{P} \infty$,

(ii) $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} \infty$,

(iii) $\xi^d f_n^2 \xrightarrow{P} \infty$,

are equivalent and follow from

(iv) $|\, \tilde{\xi}^d f_n| \xrightarrow{P} \infty$.

We conjecture that all four conditions are equivalent when the f_n are symmetric. Our proof relies on several lemmas, beginning with some scaling properties of independent interest.

Lemma 10.20 *(scaling) For fixed $c_1, \ldots, c_d > 0$, let $\eta_1, \ldots, \eta_d$ be independent Poisson processes on S with intensities $c_1 \nu, \ldots, c_d \nu$. Then for any measurable functions $f_1, f_2, \ldots$ on S,*

(i) $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} \infty \;\Leftrightarrow\; \eta_1 \cdots \eta_d f_n^2 \xrightarrow{P} \infty$,

(ii) $|\, \tilde{\xi}_1 \cdots \tilde{\xi}_d f_n| \xrightarrow{P} \infty \;\Leftrightarrow\; |\, \tilde{\eta}_1 \cdots \tilde{\eta}_d f_n| \xrightarrow{P} \infty$.

Once the main result is proved, a similar property will follow for the integrals $\xi^d f_n^2$.

Proof: (i) By iteration, we may assume that $\xi_i = \eta_i$ for $1 < i \le d$. To prove the implication to the right, it is enough to consider sub-sequences with $\xi_1 \cdots \xi_d f_n \to \infty$ a.s. Conditioning on $\xi_2, \dots, \xi_d$, we may next reduce to $d = 1$, in which case the assertion follows by Theorem 3.26 (iv).

(ii) Proceed as before, now using Theorem 3.27 instead of 3.26. □

Lemma 10.21 *(scaling and thinning) For fixed $c > 1$ and $p \in (0,1]$, let η be a Poisson process on S with intensity $c\nu$, and let η_p be a p-thinning of η. Then for any non-diagonal, measurable functions $f_1, f_2, \ldots \ge 0$ on S,*

$$\xi^d f_n \xrightarrow{P} \infty \quad \Rightarrow \quad \eta_p \, \eta^{d-1} f_n \xrightarrow{P} \infty.$$

Proof: The statement is obvious for $p = c^{-1}$, since we may then take ξ to be a p-thinning of η. Using a sub-sequence argument, we may strengthen the result to $\eta_p \, \eta^{d-1} f_n \to \infty$ a.s., and then reduce by conditioning to the case of non-random η. The assertion for general $p \in (0,1)$ now follows by Corollary 3.33. □

A crucial role is played by the following decoupling property for one-dimensional integrals, of some independent interest.

Lemma 10.22 *(decoupling in escape conditions) For every $n \in \mathsf{N}$, consider a T-marked point process ξ_n on $(0,\infty)$ with compensator η_n and p-thinnings ξ_n^p, a predictable process V_n on $(0,\infty) \times T$, and a ξ_n-tangential process $\zeta_n \perp\!\!\!\perp_{\eta_n} V_n$ with conditionally independent increments. Let $\tilde\xi_n$ and $\tilde\zeta_n$ be symmetric versions of ξ_n and ζ_n. Then*[2]

(i) $\xi_n^p V_n^2 \xrightarrow{P} \infty,\ p \in (0,1] \ \Rightarrow\ \zeta_n V_n^2 \xrightarrow{P} \infty,$

(ii) $\big(V_n \cdot \tilde\xi_n^p\big)^* \xrightarrow{P} \infty,\ p \in (0,1] \ \Rightarrow\ \big(V_n \cdot \tilde\zeta_n\big)^* \xrightarrow{P} \infty.$

Proof: (i) Put $U_n = V_n^2$ and $\hat U_n = U_n \wedge 1$. For fixed $r > 0$, consider the optional times

$$\tau_n = \inf\Big\{t \ge 0;\ (\hat U_n \cdot \eta_n)_t > r\Big\}, \quad n \in \mathsf{N}.$$

Using Chebyshev's inequality, the conditional independence $\xi_n^p \perp\!\!\!\perp_{\xi_n} U_n$, and the compensation property of η_n, we get

$$\begin{aligned} P\Big\{(U_n \cdot \xi_n^p)_{\tau_n} \ge 1\Big\} &= P\Big\{(\hat U_n \cdot \xi_n^p)_{\tau_n} \ge 1\Big\} \\ &\le E\,(\hat U_n \cdot \xi_n^p)_{\tau_n} \\ &= p\,E\,(\hat U_n \cdot \xi_n)_{\tau_n} \\ &= p\,E\,(\hat U_n \cdot \eta_n)_{\tau_n} \le p\,(r+1), \end{aligned}$$

[2]Recall that $V \cdot \tilde\xi$ denotes the process $\int_0^t V_s \, \tilde\xi(ds)$, and $X^* = \sup_t |X_t|$.

and so

$$\begin{aligned}
P\{\eta_n \hat{U}_n \le r\} &= P\{\tau_n = \infty\} \\
&\le P\{\xi_n^p U_n < 1\} + P\Big\{(U_n \cdot \xi_n^p)_{\tau_n} \ge 1\Big\} \\
&\le P\{\xi_n^p U_n < 1\} + p\,(r+1).
\end{aligned}$$

Since $\xi_n U_n \xrightarrow{P} \infty$, we obtain $P\{\eta_n \hat{U}_n \le r\} \to 0$, as $n \to \infty$ and then $p \to 0$, and r being arbitrary, we get $\eta_n \hat{U}_n \xrightarrow{P} \infty$. Then for every sub-sequence $N' \subset \mathsf{N}$, we have $\eta_n \hat{U}_n \to \infty$ a.s. along a further sub-sequence N''. Since ζ_n has conditionally independent increments and intensity η_n, given (η_n, U_n), Theorem 3.26 (iv) yields $\zeta_n U_n \xrightarrow{P} \infty$, conditionally along N''. The corresponding unconditional property follows by dominated convergence, and the convergence extends to N, since N' was arbitrary.

(ii) Here we put instead $\hat{V}_n = 2V_n/(|V_n| \vee 2)$, and consider for every $r > 0$ the optional times

$$\tau_n = \inf\Big\{t \ge 0;\ (\hat{V}_n^2 \cdot \eta_n)_t > r\Big\}, \quad n \in \mathsf{N}.$$

Using a BDG-inequality and proceeding as before, we get

$$\begin{aligned}
P\Big\{(V_n \cdot \tilde{\xi}_n^p)^*_{\tau_n} \ge 1\Big\} &= P\Big\{(\hat{V}_n \cdot \tilde{\xi}_n^p)^*_{\tau_n} \ge 1\Big\} \\
&\le E\,(\hat{V}_n \cdot \tilde{\xi}_n^p)^{*2}_{\tau_n} = p^2 E\,(\hat{V}_n \cdot \tilde{\xi}_n)^{*2}_{\tau_n} \\
&\lesssim p^2 E\,(\hat{V}_n^2 \cdot \xi_n)_{\tau_n} = p^2 E\,(\hat{V}_n^2 \cdot \eta_n)_{\tau_n} \\
&\le p^2 (r+4),
\end{aligned}$$

and so

$$\begin{aligned}
P\{\eta_n \hat{V}_n^2 \le r\} &= P\{\tau_n = \infty\} \\
&\le P\Big\{(V_n \cdot \xi_n^p)^* < 1\Big\} + P\Big\{(V_n \cdot \xi_n^p)^*_{\tau_n} \ge 1\Big\} \\
&\lesssim P\Big\{(V_n \cdot \xi_n^p)^* < 1\Big\} + p^2(r+4),
\end{aligned}$$

which implies $\eta_n \hat{V}_n^2 \xrightarrow{P} \infty$. We may now complete the proof as before, using Theorem 3.27 instead of 3.26. □

For functions f on S^d, we define the *symmetrization* $\tilde{f}$ by

$$\tilde{f}(s_1, \dots, s_d) = E f(s_{\pi_1}, \dots, s_{\pi_d}), \quad s_1, \dots, s_d \in S,$$

where $\pi = (\pi_1, \dots, \pi_d)$ denotes an exchangeable permutation of $(1, \dots, d)$.

Lemma 10.23 *(symmetrization) For functions $f_1, f_2, \ldots \ge 0$ on S^d with symmetrizations $\tilde{f}_1, \tilde{f}_2, \ldots$, we have*

$$\xi_1 \cdots \xi_d f_n \xrightarrow{P} \infty \quad \Leftrightarrow \quad \xi_1 \cdots \xi_d \tilde{f}_n \xrightarrow{P} \infty.$$

Proof: Since $f_n \le d!\,\tilde f_n$, it suffices to prove the implication to the left. Then assume $\xi_1 \cdots \xi_d \tilde f_n \xrightarrow{P} \infty$. By Lemma 10.20 (i), we have also $\eta_1 \cdots \eta_d \tilde f_n \xrightarrow{P} \infty$, where the η_k are i.i.d. Poisson with intensity $c\nu$. Writing $\tilde\pi(s) = s \circ \pi$ for $s \in S^d$ and $\pi \in \mathcal{P}_d$, the set of permutations on $(1, \dots, d)$, we get for any $r > 0$

$$\begin{aligned} P\Big\{\eta_1 \cdots \eta_d \tilde f_n > r\Big\} &\le P \bigcup_{\pi \in \mathcal{P}_d} \Big\{\eta_1 \cdots \eta_d (f_n \circ \tilde\pi) > r\Big\} \\ &\le \sum_{\pi \in \mathcal{P}_d} P\Big\{\eta_1 \cdots \eta_d (f_n \circ \tilde\pi) > r\Big\} \\ &= d!\, P\{\eta_1 \cdots \eta_d f_n > r\}, \end{aligned}$$

by the symmetry of $\eta_1 \cdots \eta_d$. Now write $\xi_j = \xi_{j1}^m + \cdots + \xi_{jm}^m$ for arbitrary $m \in \mathsf{N}$, where the ξ_{jk}^m are i.i.d. Poisson with intensity ν/m. Applying the previous estimate to $\eta_j = \xi_{j1}^m$, we get for any $r > 0$

$$\begin{aligned} P\{\xi_1 \cdots \xi_d f_n \le r\} &\le P \bigcap_{k \le m} \Big\{\xi_{1k}^m \cdots \xi_{dk}^m f_n \le r\Big\} \\ &= \Big(P\Big\{\xi_{1k}^m \cdots \xi_{dk}^m f_n \le r\Big\}\Big)^m \\ &\le \Big(1 - P\Big\{\xi_{1k}^m \cdots \xi_{dk}^m \tilde f_n > r\Big\}\,/d!\Big)^m. \end{aligned}$$

Letting $n \to \infty$ and then $m \to \infty$, we get $P\{\xi_1 \cdots \xi_d f_n \le r\} \to 0$, and $\xi_1 \cdots \xi_d f_n \xrightarrow{P} \infty$ follows, since $r > 0$ was arbitrary. □

We further list some elementary probability estimates, where we write $P^{\mathcal{F}} = P(\cdot \mid \mathcal{F})$ for convenience.

Lemma 10.24 *(conditioning) For any event A, σ-fields $\mathcal{F} \perp\!\!\!\perp \mathcal{G}$, and constants $a, b \ge 0$, we have*

(i) $PA \ge 2a \;\Rightarrow\; P\{P^{\mathcal{F}}A \ge a\} \ge a,$

(ii) $P\{P^{\mathcal{F}}A \ge 1-a\} \ge 1-b \;\Rightarrow\; PA \ge 1-a-b,$

(iii) $P\{P^{\mathcal{F}}A \ge 1-a\} \ge b \;\Rightarrow\; P\{P^{\mathcal{G}}A \ge b/2\} \ge 1-2a.$

Proof: (i) If $P\{P^{\mathcal{F}}A \ge a\} < a$, then

$$\begin{aligned} PA &\le P\{P^{\mathcal{F}}A \ge a\} + E\Big\{P^{\mathcal{F}}A;\, P^{\mathcal{F}}A < a\Big\} \\ &< a + a = 2a. \end{aligned}$$

(ii) If $P\{P^{\mathcal{F}}A \ge 1-a\} \ge 1-b$, then $P\{P^{\mathcal{F}}A^c > a\} \le b$, and so

$$PA^c \le P\{P^{\mathcal{F}}A^c > a\} + E\Big\{P^{\mathcal{F}}A^c;\, P^{\mathcal{F}}A^c \le a\Big\} \le b + a.$$

(iii) Put $B = \{P^{\mathcal{F}}A \geq 1 - a\}$ and $Q = P(\cdot\,|\,B) = E(\cdot\,|\,B)$, so that $PB \geq b$ and $QA^c \leq a$. By Chebyshev's inequality,

$$\begin{aligned} Q\{Q^{\mathcal{G}}A \geq \tfrac{1}{2}\} &= 1 - Q\{Q^{\mathcal{G}}A^c > \tfrac{1}{2}\} \\ &\geq 1 - 2\,Q\,Q^{\mathcal{G}}A^c \\ &= 1 - 2\,QA^c \geq 1 - 2a. \end{aligned}$$

Using the independence $B \perp\!\!\!\perp \mathcal{G}$ twice, we further obtain

$$\begin{aligned} Q\{Q^{\mathcal{G}}A \geq \tfrac{1}{2}\} &= P\{Q^{\mathcal{G}}A \geq \tfrac{1}{2}\} \\ &= P\Big\{P^{\mathcal{G}}(A \cap B) \geq \tfrac{1}{2}PB\Big\} \\ &\leq P\{P^{\mathcal{G}}A \geq b/2\}. \end{aligned}$$

The assertion follows by combination of the two estimates. □

Proof of Theorem 10.19, (i) ⇒ (ii): By a sub-sequence argument followed by conditioning on $\xi_1, \dots, \xi_d$, we may take the latter to be non-random. If (ii) fails, then $\xi_1 \cdots \xi_d f_n^2$ is bounded along a sub-sequence $N' \subset \mathbb{N}$, and so $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n$ is tight along N' by Theorem 10.12 (iv), which contradicts (i).

(iv) ⇒ (iii): Proceed as above, except that before using Theorem 10.12, we may take $S = \mathbb{R}_+$, and divide the functions f_n^2 into tetrahedral components.

(ii) ⇒ (iii): Let $\eta_1, \dots, \eta_d$ be i.i.d. Poisson processes with rate d^{-1}. Assuming $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} \infty$, we get by Lemma 10.20

$$\begin{aligned} \xi^d f_n^2 &\overset{d}{=} (\eta_1 + \cdots + \eta_d)^d f_n^2 \\ &\geq \eta_1 \cdots \eta_d f_n^2 \xrightarrow{P} \infty. \end{aligned}$$

(iii) ⇒ (ii): By a Borel isomorphism, the identity $\xi^d f_n = \xi^d \tilde{f}_n$, and Lemma 10.23, we may take $S = \mathbb{R}_+$, and let the f_n be tetrahedral. The assertion is trivial for $d = 1$. Assuming its truth in dimension $d - 1$, let $\xi^d f_n \xrightarrow{P} \infty$. Then Lemma 10.21 gives $\eta_p\, \eta^{d-1} f_n \xrightarrow{P} \infty$ for all $p \in (0,1)$, where η is Poisson with intensity $c\nu > \nu$, and η_p is a p-thinning of η. Now Lemma 10.22 yields $\eta_1\, \eta^{d-1} f_n \xrightarrow{P} \infty$, where η_1 is an independent copy of η. By a sub-sequence argument, we may assume that even $\eta_1\, \eta^{d-1} f_n \to \infty$ a.s., which remains conditionally true, given η_1. Now the induction hypothesis yields $\eta_1 \cdots \eta_d f_n \xrightarrow{P} \infty$, conditionally on η_1, where $\eta_1, \dots, \eta_d$ are independent copies of η. By dominated convergence, the latter statement remains unconditionally true, and (ii) follows by Lemma 10.20.

(ii) ⇒ (i): Assume $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} \infty$. For fixed $m \in \mathbb{N}$, let $\eta_1, \dots, \eta_d$ be independent Poisson processes with intensity ν/m, and note that even

$\eta_1 \cdots \eta_d f_n^2 \xrightarrow{P} \infty$, by Lemma 10.20. By Corollary 10.14 and the Paley–Zygmund inequality, there exists an absolute constant $c_d > 0$, such that for large enough $n_0 = n_0(t)$,

$$P\Big\{|\tilde\eta_1 \cdots \tilde\eta_d f_n| > t\Big\} \ge c_d, \quad n \ge n_0. \tag{10}$$

By Lemma 10.24 (i), we get for $n \ge n_0$

$$P\Big\{P\Big(|\,\tilde\eta_1 \cdots \tilde\eta_d f_n| \ge t \,\Big|\, \tilde\eta_2, \ldots, \tilde\eta_d\Big) \ge c_d/2\Big\} \ge c_d/2,$$

and so by Theorems 10.11, 10.16, and A1.6,

$$P\Big\{P\Big(|\,\tilde\xi_1 \tilde\eta_2 \cdots \tilde\eta_d f_n| \ge t \,\Big|\, \tilde\eta_2, \cdots, \tilde\eta_d\Big) \ge 1 - c\,(mc_d/2)^{-1/2}\Big\} \ge c_d/2,$$

for some absolute constant $c > 0$. Then Lemma 10.24 (iii) yields

$$P\Big\{P\Big(|\,\tilde\xi_1 \tilde\eta_2 \cdots \tilde\eta_d f_n| \ge t \,\Big|\, \tilde\xi_1\Big) \ge c_d/4\Big\} \ge 1 - 2c\,(mc_d/2)^{-1/2}. \tag{11}$$

Applying the same argument to the conditional probability in (11), and using Theorem 10.16 again, we obtain

$$\begin{aligned} P\Big\{P\Big(P\Big\{|\,\tilde\xi_1 \tilde\xi_2 \tilde\eta_3 \cdots \tilde\eta_d f_n| \ge t \,\Big|\, \tilde\xi_1, \tilde\xi_2\Big\} \ge c_d\, 4^{-2} \Big|\, \tilde\xi_1\Big) \ge 1 - 2c\,(mc_d/8)^{-1/2}\Big\} \\ \ge 1 - 2c\,(mc_d/2)^{-1/2}, \end{aligned}$$

and so by Lemma 10.24 (ii),

$$P\Big\{P\Big(|\,\tilde\xi_1 \tilde\xi_2 \tilde\eta_3 \cdots \tilde\eta_d f_n| \ge t \,\Big|\, \tilde\xi_1, \tilde\xi_2\Big) \ge c_d\, 4^{-2}\Big\} \ge 1 - 6c\,(mc_d/2)^{-1/2}.$$

Continuing recursively in d steps, we obtain

$$\begin{aligned} P\Big\{|\,\tilde\xi_1 \cdots \tilde\xi_d f_n| \ge t\Big\} &= P\Big\{P\Big(|\,\tilde\xi_1 \cdots \tilde\xi_d f_n| \ge t \,\Big|\, \tilde\xi_1, \ldots, \tilde\xi_d\Big) \ge c_d\, 4^{-d}\Big\} \\ &\ge 1 - c\,(mc_d/2)^{-1/2}\,(2 + 2^2 + \cdots + 2^d) \\ &\ge 1 - c\, 2^{d+1} (mc_d/2)^{-1/2}. \end{aligned}$$

Fixing $\varepsilon, t > 0$, we may choose m so large that the right-hand side is $> 1 - \varepsilon$, and then choose n_0 accordingly to make (10) fulfilled. Then $P\Big\{|\tilde\xi_1 \cdots \tilde\xi_d f_n| < t\Big\} < \varepsilon$ for all $n \ge n_0$, as required. □

We turn to some useful truncation criteria.

Theorem 10.25 *(truncation) For any measurable sets $A_n \subset S^d$ and functions f_n on S^d, non-diagonal in case of* (ii′)*, we have*

(i) $\xi_1 \cdots \xi_d A_n \xrightarrow{P} \infty \;\Leftrightarrow\; \xi_1 \cdots \xi_d A_n \wedge 1 \xrightarrow{P} 1$,

(ii) $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} \infty \;\Leftrightarrow\; \xi_1 \cdots \xi_d (f_n^2 \wedge 1) \xrightarrow{P} \infty$,

(ii′) $\xi^d f_n^2 \xrightarrow{P} \infty \;\Leftrightarrow\; \xi^d (f_n^2 \wedge 1) \xrightarrow{P} \infty$,

(iii) $|\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n| \xrightarrow{P} \infty \;\Leftrightarrow\; \left|\tilde{\xi}_1 \cdots \tilde{\xi}_d(|f_n| \wedge 1)\right| \xrightarrow{P} \infty,$

(iv) $|\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n| \xrightarrow{P} \infty \;\Leftrightarrow\; |\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n| \wedge 1 \xrightarrow{P} 1.$

Proof: (i) If $\xi A_n \wedge 1 \xrightarrow{P} 1$, then $e^{-\nu A_n} = P\{\xi A_n = 0\} \to 0$, and so $\nu A_n \to \infty$, which yields $\xi A_n \xrightarrow{P} \infty$, by Theorem 3.26 (iii). This proves the assertion for $d = 1$. Now suppose that $\xi_1 \cdots \xi_d A_n \wedge 1 \xrightarrow{P} 1$ for some $d \geq 2$. Since ξ_1 is Z_+-valued, this is equivalent to $\xi_1(\xi_2 \cdots \xi_d A_n \wedge 1) \wedge 1 \xrightarrow{P} 1$. Taking sub-sequences to reduce to the a.s. case, conditioning on $\xi_2, \dots, \xi_d$, and using the result for $d = 1$, we obtain $\xi_1 \cdots \xi_d A_n \geq \xi_1(\xi_2 \cdots \xi_d A_n \wedge 1) \xrightarrow{P} \infty$.

(ii) If $\xi f_n^2 \xrightarrow{P} \infty$, then $\nu(\psi \circ f_n^2) \to \infty$ by Theorem 3.26 (iii), which implies $\nu\{\psi \circ (f_n^2 \wedge 1)\} \to \infty$, since $\psi(x \wedge 1) \geq \psi(1)\psi(x)$. Hence, $\xi(f_n^2 \wedge 1) \xrightarrow{P} \infty$ by the same theorem, which proves the assertion for $d = 1$. Now assume that $\xi_1 \cdots \xi_d f_n^2 \xrightarrow{P} \infty$ for some $d \geq 2$. Since $\xi_2 \cdots \xi_d$ is Z_+-valued,

$$\begin{aligned} \xi_2 \cdots \xi_d f_n^2 \wedge 1 &= \xi_2 \cdots \xi_d (f_n^2 \wedge 1) \wedge 1 \\ &\leq \xi_2 \cdots \xi_d (f_n^2 \wedge 1). \end{aligned}$$

Conditioning as before, and using the result for $d = 1$, we obtain

$$\xi_1 \cdots \xi_d (f_n^2 \wedge 1) \geq \xi_1 \Big(\xi_2 \cdots \xi_d f_n^2 \wedge 1 \Big) \xrightarrow{P} \infty.$$

(ii′) Use (ii) and Theorem 10.19.

(iii) Combine (ii) with Theorem 10.19 (i).

(iv) Turning to a.s. convergent sub-sequences, and conditioning on $\tilde{\xi}_1, \dots, \tilde{\xi}_d$, we may reduce to the case of $d = 1$. Assuming $|\tilde{\xi} f_n| \wedge 1 \xrightarrow{P} 1$, we need to show that $|\tilde{\xi} f_n| \xrightarrow{P} \infty$, which holds by Theorems 3.26 (iii) and 10.19 (i), iff $\nu(f_n^2 \wedge 1) \to \infty$. If this fails, we have $\nu(f_n^2 \wedge 1) \leq c < \infty$ along a sub-sequence $N' \subset \mathsf{N}$. Writing $f_{n,h} = f_n 1\{|f_n| \leq h\}$ for $h > 0$, we obtain

$$\nu\{f_{n,h} \neq f_n\} = \nu\{|f_n| > h\} \leq c h^{-1} < \infty, \quad h \in (0,1),\; n \in N',$$

and so the condition $|\tilde{\xi} f_n| \wedge 1 \xrightarrow{P} 1$ can be strengthened to

$$|\tilde{\xi} f_{n,h}| \wedge 1 \xrightarrow{P} 1 \text{ along } N', \quad h > 0.$$

We may then choose some constants $h_n \to 0$, such that for every $n \in N'$, the functions $g_n = f_{n,h_n}$ satisfy

$$|g_n| \leq h_n \to 0, \qquad \sup_n \nu g_n^2 \leq c, \qquad |\tilde{\xi} g_n| \wedge 1 \xrightarrow{P} 1. \tag{12}$$

By FMP 15.15, the first two conditions yield $\tilde{\xi} g_n \xrightarrow{d} \zeta$ along a further sub-sequence N'', where ζ is $N(0, \sigma^2)$ for some constant $\sigma^2 \leq c$. This contradicts the third condition in (12), and the claim follows. □

The counterpart of (iv) fails for positive integrals $\xi_1 \cdots \xi_d f_n$. For a counterexample, take $S = \mathsf{R}_+$ and $\nu = \lambda$, and choose $f_n = n^{-d} 1_{[0,n]^d}$. Then $\xi_1 \cdots \xi_d f_n \to 1$ a.s., by the law of large numbers, and so $\xi_1 \cdots \xi_d f_n \wedge 1 \xrightarrow{P} 1$, whereas $\xi_1 \cdots \xi_d f_n \stackrel{P}{\not\to} \infty$.

We conclude with a strengthened version of Theorem 10.11. For any random variables α, α_n and event $A \subset \Omega$, the convergence $\alpha_n \xrightarrow{P} \alpha$ on A means that $E(|\,\alpha_n - \alpha| \wedge 1; A) \to 0$. The definition for $\alpha = \infty$ is similar. The symmetric integrals $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ can now be defined as before, on suitable subsets.

Corollary 10.26 *(convergence dichotomy) Let $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ and f be such as in Theorem 10.11, and consider some functions $f_n \to f$ with bounded supports satisfying $|f_n| \le |f|$. Then*

(i) $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} \tilde{\xi}_1 \cdots \tilde{\xi}_d f$ *on* $\{\xi_1 \cdots \xi_d f^2 < \infty\}$,

(ii) $|\,\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n| \xrightarrow{P} \infty$ *on* $\{\xi_1 \cdots \xi_d f^2 = \infty\}$.

We conjecture that (ii) remains true for the integrals $\tilde{\xi}^d f$.

Proof: (i) This is clear from the definition of $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$.

(ii) By the sign-sequence representation in Theorem 10.11 and Fubini's theorem, we may reduce to the case of non-random $\xi_1, \ldots, \xi_d$. The assertion then follows from the equivalence (i) $\Leftrightarrow$ (ii) in Theorem 10.19. □

10.4 Lévy and Related Integrals

We may now apply the previous theory to the existence and basic properties of multiple integrals, with respect to the components of a positive or symmetric Lévy process $X = (X_1, \ldots, X_d)$ in R^d of pure jump type. Here X is said to be *positive* if $X_1, \ldots, X_d \ge 0$, and *symmetric* if $-X \stackrel{d}{=} X$, and we use the shorthand notation

$$X_1 \cdots X_d f = \int \cdots \int f(t_1, \ldots, t_d)\, dX_1(t_1) \cdots dX_d(t_d), \qquad (13)$$

for suitable functions f on R_+^d.

Writing $\Delta X_t = X_t - X_{t-}$, we introduce the associated *jump point process* $\eta = \sum_t \delta_{t, \Delta X_t}$ on $\mathsf{R}_+ \times (\mathsf{R}^d \setminus \{0\})$. Here η is Poisson by Theorem 3.19, and the obvious stationarity yields $E\eta = \lambda \otimes \nu$, for some measure ν on $\mathsf{R}^d \setminus \{0\}$, known as the *Lévy measure* of X. In the positive case, ν and η are restricted to $\mathsf{R}_+^d \setminus \{0\}$ and $\mathsf{R}_+^{d+1} \setminus \{0\}$, respectively, and we may recover X from η through the formula

$$X_t = \int_0^{t+} \!\! \int x\, \eta(dt\, dx), \quad t \ge 0, \qquad (14)$$

where the right-hand side is a pathwise Lebesgue-type integral, which converges by Lemma 3.26 (i) iff $\int(|x|\wedge 1)\,\nu(dx) < \infty$. The individual components $X_1, \ldots, X_d$ may be recovered by similar formulas involving the coordinate projections $\eta_1, \ldots, \eta_d$ of η, which satisfy similar integrability conditions.

For any measurable function $f \geq 0$ on R_+^d, we may express the associated multiple integral (13) in terms of η or $\eta_1, \ldots, \eta_d$, as in

$$X_1 \cdots X_d f = \eta^d(L'f) = \eta_1 \cdots \eta_d(Lf), \tag{15}$$

where the operators L and L' are given by

$$\begin{aligned} Lf(t_1, \ldots, t_d;\, x_1, \ldots, x_d) &= x_1 \cdots x_d\, f(t_1, \ldots, t_d), \\ L'f\big(t_1, \ldots, t_d;\, x_{ij},\, i, j \leq d\big) &= x_{11} \cdots x_{dd}\, f(t_1, \ldots, t_d). \end{aligned}$$

For convenience, we shall henceforth write L for both operators, the interpretation being obvious from the context. By the Poisson representation (15), the basic integrability and convergence criteria for Lévy integrals $X_1 \cdots X_d f$ follow immediately from the corresponding results for Poisson processes.

The theory for symmetric Lévy processes X is much more sophisticated, since the integrals in (13)−(15) may no longer exist, in the elementary, pathwise sense. In all three cases, we then need to replace η by a suitable symmetric Poisson process ζ, which allows us to apply the previous theory of symmetric Poisson integrals. We consider two different constructions of such a process ζ.

For the first approach, we fix a *half-space* S of $\mathsf{R}^d \setminus \{0\}$, defined as a measurable subset of R^d, such that $\pm S$ are disjoint with union $\mathsf{R}^d \setminus \{0\}$. We may then define a signed random measure ζ on $\mathsf{R}_+ \times S$ by

$$\zeta B = \eta B - \eta(-B), \quad B \in \mathcal{B}(\mathsf{R}_+ \times S),$$

where $-B = \{(t, -x);\, (t, x) \in B\}$. The symmetry of X yields a corresponding symmetry of ν, which ensures that ζ will be a symmetric Poisson process on $\mathsf{R}_+ \times S$ with intensity $2\lambda \otimes \nu$. We may now define a process X' in R^d as in (14), though with η replaced by ζ, and with the integral interpreted in the sense of Theorem 10.11. By Theorem 10.11 and Lemma 3.26, the integral exists for $t > 0$ iff $\int(|x|^2 \wedge 1)\,\nu(dx) < \infty$, in which case we see from martingale theory that X' has a right-continuous version with left-hand limits, defining a symmetric, purely discontinuous Lévy process. Since the jumps of X and X' agree, we have in fact $X = X'$ a.s., which means that X has a representation (14), though with η replaced by ζ. The multiple integral (13) can now be defined by (15) with η replaced by ζ.

For an alternative construction, we introduce a symmetrization $\tilde{\eta}$ of η on the entire domain $\mathsf{R}_+ \times (\mathsf{R}^d \setminus \{0\})$, and define a process $\tilde{X}$ as in (14), though with η replaced by $\tilde{\eta}$. Then $\tilde{X}$ is again a symmetric, purely discontinuous Lévy process, and the symmetry of ν implies that even $\tilde{X}$ has Lévy measure $\lambda \otimes \nu$. Thus, $\tilde{X} \overset{d}{=} X$, and so a transfer argument yields a Poisson process

$\zeta \stackrel{d}{=} \eta$ with $E\zeta = \lambda \otimes \nu$, along with a symmetrized version $\tilde{\zeta}$, satisfying (14) with η replaced by $\tilde{\zeta}$. We may now define the multiple integral (13) as in (15), though with η replaced by $\tilde{\zeta}$. This approach avoids the artificial construction in terms of an arbitrary half-space S. On the other hand, it has the disadvantage that the underlying point processes ζ and $\tilde{\zeta}$ are no longer measurably determined by X or η.

To justify our construction of the integral, we start from an elementary random set function. Thus, for any disjoint intervals $I_j = (s_j, t_j]$ with $0 \leq s_j < t_j < \infty$, $j = 1, \ldots, d$, we introduce the associated random variables

$$X_1 \cdots X_d(I_1 \times \cdots \times I_d) = \prod_{i \leq d} \{X_i(t_i) - X_i(s_i)\}. \tag{16}$$

Our aim is to extend the definition of $X_1 \cdots X_d$ to more general integrands f, subject to suitable linearity and continuity constraints.

For the *domain* of $X_1 \cdots X_d$, we choose a class $\mathcal{D}$ of measurable functions f on R_+^d, which is said to be

- *linear* if $\quad f, g \in \mathcal{D},\ a, b \in \mathsf{R} \quad \Rightarrow \quad af + bg \in \mathcal{D}$,
- *solid* if $\quad |f| \leq |g|,\ g \in \mathcal{D} \quad \Rightarrow \quad f \in \mathcal{D}$.

Given such a class $\mathcal{D}$, we say that a process Ξ on $\mathcal{D}$ is

- *linear* if

$$f, g \in \mathcal{D}, \ \ a, b \in \mathsf{R} \quad \Rightarrow \quad \Xi(af + bg) = a\,\Xi f + b\,\Xi g \text{ a.s.},$$

- *continuous* if

$$f_n \to f \text{ in } \mathcal{D}, \ \ |f_n| \leq g \in \mathcal{D} \quad \Rightarrow \quad \Xi f_n \xrightarrow{P} \xi f.$$

We may now state the basic integrability criterion, for multiple integrals with respect to symmetric, purely discontinuous Lévy processes X.

Theorem 10.27 *(symmetric Lévy integral) Let $X = (X_1, \ldots, X_d)$ be a symmetric, purely discontinuous Lévy process in R^d, represented as in* (14) *in terms of a symmetric Poisson process ζ. Then the set function* (16) *extends a.s. uniquely to a linear, continuous process $X_1 \cdots X_d f$, on the class $\mathcal{D}_X$ of non-diagonal, measurable functions f on R_+^d with $\eta^d(Lf)^2 < \infty$ a.s., and*

$$X_1 \cdots X_d f = \zeta^d(Lf) = \zeta_1 \cdots \zeta_d(Lf) \ \ \text{a.s.} \tag{17}$$

This class $\mathcal{D}_X$ is the largest domain with the stated properties.

Proof: First we show that $\mathcal{D}_X$ contains all indicator functions of bounded, non-diagonal Borel sets. By Corollary 10.17, we may then assume that $X_1, \ldots, X_d$ are independent, in which case the claim follows from Lemma 10.16. When X is a symmetric, compound Poisson process, it is obvious that the integral in (17) satisfies (16). The general result then follows by a simple truncation argument, based on the continuity of the integrals in (14) and (17).

The linearity and solidity of $\mathcal{D}_X$ are clear from the condition $\eta^d(Lf)^2 < \infty$ a.s., and the integral $\zeta^d(Lf)$ exists on $\mathcal{D}_X$ by Theorem 10.11. The linearity and continuity of $\zeta^d(Lf)$ follow from the corresponding properties of ζ^d, in Theorems 10.11 (i) and 10.18 (i), together with the obvious linearity and continuity of L. The maximality of the domain $\mathcal{D}_X$ is clear from Theorem 10.11 (iii).

To prove the asserted uniqueness, consider any linear and continuous extensions Ξ and Ξ' of the set function (16), defined on some linear and solid domains $\mathcal{D}$ and $\mathcal{D}'$, both containing the indicator functions of all bounded, non-diagonal rectangles. By a monotone-class argument based on the linearity and continuity of Ξ and Ξ', we conclude that $\Xi B = \Xi' B$ a.s., for every Borel set B contained in such a rectangle. By linearity and continuity, we get $\Xi f = \Xi' f$ a.s. for every indicator function f in $\mathcal{D} \cap \mathcal{D}'$, and the equality for general $f \in \mathcal{D} \cap \mathcal{D}'$ then follows by an obvious approximation. □

We list some basic properties of multiple integrals with respect to positive or symmetric Lévy processes, all consequences of the corresponding statements for multiple Poisson integrals given in previous sections.

Corollary 10.28 *(decoupling) Given a positive or symmetric, purely discontinuous Lévy process* $X = (X_1, \ldots, X_d)$*, let* $X_1', \ldots, X_d'$ *be independent with* $X_k' \overset{d}{=} X_k$ *for all* k*. Then for any tetrahedral, positive or real, measurable functions* $f, f_1, f_2, \ldots$ *on* R_+^d*,*

(i) $X_1 \cdots X_d f$ *and* $X_1' \cdots X_d' f$ *exist simultaneously,*

(ii) $X_1 \cdots X_d f_n \xrightarrow{P} X_1 \cdots X_d f \;\Leftrightarrow\; X_1' \cdots X_d' f_n \xrightarrow{P} X_1' \cdots X_d' f$*,*

(iii) $(X_1 \cdots X_d f_n)$ *and* $(X_1' \cdots X_d' f_n)$ *are simultaneously tight,*

(iv) $|X_1 \cdots X_d f_n| \xrightarrow{P} \infty \;\Rightarrow\; |X_1' \cdots X_d' f_n| \xrightarrow{P} \infty$*.*

Proof: Use Corollary 10.3 (i), (iii)−(iv), Corollary 10.17 (i)−(iii), and Theorem 10.19. □

To state the next result, let $X = (X_1, \ldots, X_d)$ be a symmetric, purely discontinuous Lévy process in R^d with associated jump point process η, and introduce the *quadratic variation* processes

$$[X_j]_t = \int_0^{t+} \!\! \int x_j^2 \, \eta(dt\, dx), \quad t \geq 0, \;\; j \leq d,$$

which exist by Lemma 3.26, since $\int (|x|^2 \wedge 1) \nu(dx) < \infty$. Writing $[X_1 \cdots X_d]$ for the associated tensor product with induced product measure $\bigotimes_k [X_k]$, we note that

$$\begin{aligned} [X_1 \cdots X_d] f^2 &= \eta^d (Lf)^2 \\ &= \eta_1 \cdots \eta_d (Lf)^2 \;\text{ a.s.} \end{aligned}$$

Defining $\tilde{X}$ as in (14), but with η replaced by $\tilde{\eta}$, we have

$$\begin{aligned}\tilde{X}_1 \cdots \tilde{X}_d f &= \tilde{\eta}^d(Lf)\\ &= \tilde{\eta}_1 \cdots \tilde{\eta}_d(Lf) \text{ a.s.,}\end{aligned}$$

which shows that $\tilde{X}_1 \cdots \tilde{X}_d \stackrel{d}{=} X_1 \cdots X_d$. We may now state some basic relations between the properties of $X_1 \cdots X_d$ and $[X_1 \cdots X_d]$.

Corollary 10.29 *(quadratic variation) Given a symmetric, purely discontinuous Lévy process $X = (X_1, \ldots, X_d)$, define $[X_1 \cdots X_d] = \bigotimes_k [X_k]$. Then for any tetrahedral, measurable functions $f, f_1, f_2, \ldots$ on R_+^d,*

(i) $X_1 \cdots X_d f$ *exists* $\Leftrightarrow$ $[X_1 \cdots X_d] f^2 < \infty$ *a.s.,*

(ii) $X_1 \cdots X_d f_n \xrightarrow{P} 0 \;\Leftrightarrow\; [X_1 \cdots X_d] f_n^2 \xrightarrow{P} 0$,

(iii) $(X_1 \cdots X_d f_n)$ *and* $([X_1 \cdots X_d] f_n^2)$ *are simultaneously tight,*

(iv) $|X_1 \cdots X_d f_n| \xrightarrow{P} \infty \;\Rightarrow\; [X_1 \cdots X_d] f_n^2 \xrightarrow{P} \infty$, *with equivalence when* $X_1, \ldots, X_d$ *are independent.*

Proof: (i) Use Theorem 10.11 (i) and (iii).

(ii)−(iii): Use Theorem 10.15 (ii) and (iii).

(iv) Use Theorem 10.19. □

We turn to some dominated convergence and related properties.

Corollary 10.30 *(convergence and closure) Consider a positive or symmetric, purely discontinuous Lévy process $X = (X_1, \ldots, X_d)$, and some tetrahedral, positive or real, measurable functions $f, f_1, f_2, \ldots$ on R_+^d. Then*

(i) $X_1 \cdots X_d f_n \xrightarrow{P} X_1 \cdots X_d f \;\Rightarrow\; f_n \to f$ *in* λ^d_{loc}, *with equivalence when* $|f_n| \le g \in \mathcal{D}_X$,

(ii) $X_1 \cdots X_d f_n \xrightarrow{P} X_1 \cdots X_d f \;\Leftrightarrow\; X_1 \cdots X_d |f_n - f| \xrightarrow{P} 0$,

(iii) *when X is symmetric,* $X_1 \cdots X_d f_n \xrightarrow{P} \gamma \;\Rightarrow\; \gamma = X_1 \cdots X_d f$ *for some* $f \in \mathcal{D}_X$.

Proof: (i) Use Theorems 10.2 (i) and 10.18, along with the equivalence $f_n \to f$ a.e. $\lambda^d \;\Leftrightarrow\; Lf_n \to Lf$ a.e. $\lambda^d \otimes \nu^d$.

(ii) Use Theorems 10.2 (ii) and 10.15 (ii), along with the equality $(Lf)^2 = (L|f|)^2$.

(iii) Use Theorem 10.18 (ii), and the fact that $Lf_n \to g$ a.e. $\lambda^d \otimes \nu^d$ implies $g = Lf$ a.e., for some measurable function f on R_+^d. □

We conclude with a Fubini-type theorem for multiple integrals, based on independent Lévy processes.

Corollary 10.31 *(recursion) Consider a positive or symmetric, purely discontinuous Lévy process $X = (X_1, \ldots, X_d)$, and a positive or real, measurable function f on R_+^d. Then the following relation holds, whenever either side exists:*

$$X_1 \cdots X_d f = X_1(X_2(\cdots(X_d f)\cdots)) \ \textit{a.s.}$$

Proof: For positive X, this holds by the classical Fubini theorem. In the symmetric case, use Theorem 10.16. □

10.5 Multiple Series and Integrals

Here we consider multiple random series of the form

$$\sum_{k \in \mathsf{N}^d} f_k(\vartheta_{k_1}, \ldots, \vartheta_{k_d}), \tag{18}$$

where $\vartheta_1, \vartheta_2, \ldots$ is a sequence of i.i.d. $U(0,1)$ random variables, and $f_k = f_{k_1,\ldots,k_d}$, $k \in \mathsf{N}^d$, is an array of measurable functions on $[0,1]^d$. Our aim is to show how criteria for the existence, convergence to 0, and tightness of a sequence of such sums can be obtained from the corresponding criteria for suitable multiple Poisson integrals.

Letting $I_k = I_{k_1} \times \cdots \times I_{k_d}$ with $I_{k_i} = [k_i, k_i + 1)$, $k \in \mathsf{N}^d$, we may write the sum (18) as a multiple stochastic integral $\eta^d f$, where

$$f(t) = \sum_{k \in \mathsf{N}^d} 1\{t \in I_k\}\, f_k(t - k), \quad t \in \mathsf{R}_+^d, \tag{19}$$

and η denotes the point process on R_+, given by

$$\eta = \sum_{k \in \mathsf{N}} \delta_{\vartheta_k + k}. \tag{20}$$

As a special case, we may consider multiple integrals of the form $\eta_1 \cdots \eta_d f$, where $\eta_1, \ldots, \eta_d$ are independent copies of η.

We proceed to compare the multiple series $\eta_1 \cdots \eta_d f$ and $\eta^d f$ with the corresponding multiple Poisson integrals $\xi_1 \cdots \xi_d f$ and $\xi^d f$. We may also compare the symmetric series $\tilde\eta_1 \cdots \tilde\eta_d f$ and $\tilde\eta^d f$ with the correspondingly symmetrized Poisson integrals $\tilde\xi_1 \cdots \tilde\xi_d f$ and $\tilde\xi^d f$, respectively, where the symmetric point processes $\tilde\eta, \tilde\eta_1, \ldots, \tilde\eta_d$ are defined as in Section 8.2. The random multi-linear forms of Theorem 10.12 constitute a simple special case.

Theorem 10.32 *(multiple series and integrals) Define $\eta, \eta_1, \ldots, \eta_d$ as in (20), in terms of some i.i.d. $U(0,1)$ random variables ϑ_{kj}, and consider any tetrahedral, measurable functions $f_1, f_2, \ldots$ on R_+^d. Then the following conditions are mutually equivalent, and also equivalent to the corresponding conditions for Poisson integrals:*

(i) $\eta_1 \cdots \eta_d f_n^2 \xrightarrow{P} 0$,

(ii) $\eta^d f_n^2 \xrightarrow{P} 0$,

(iii) $\tilde\eta_1 \cdots \tilde\eta_d f_n \xrightarrow{P} 0$,

(iv) $\tilde\eta^d f_n \xrightarrow{P} 0$.

As an immediately consequence, we obtain criteria for existence and tightness of the various sums and integrals. This is because a sequence of random variables $\alpha_1, \alpha_2, \ldots$ is tight, iff $c_n \alpha_n \xrightarrow{P} 0$ for every sequence of constants $c_n \to 0$, and a single random variable α is a.s. finite, iff $c_n \alpha \xrightarrow{P} 0$ for any constants $c_n \to 0$.

Proof: Let (i′)−(iv′) denote the corresponding conditions for Poisson integrals, and note that (i′) ⇔ (ii′) ⇔ (iii′) ⇔ (iv′), by Corollary 10.3 (ii) and Theorem 10.15.

(i) ⇔ (ii): Lemma 10.4 remains true for the η-integrals, by a similar proof, based on Corollary 9.64 in place of Theorem 9.57. The assertion now follows as in Corollary 10.3 (ii).

(i) ⇔ (iii) and (ii) ⇔ (iv): Proceed as in the proof of Theorem 10.15.

(i) ⇔ (i′): Let $f_n \geq 0$. Using Lemma 3.26 (ii), along with Theorem 4.22 (i) for a singleton S, we get

$$\xi f_n \xrightarrow{P} 0 \quad \Leftrightarrow \quad \lambda(f_n \wedge 1) \to 0 \quad \Leftrightarrow \quad \eta f_n \xrightarrow{P} 0,$$

which proves the assertion for $d = 1$. For general $d \in \mathsf{N}$, we may choose $\xi_1, \ldots, \xi_d$ and $\eta_1, \ldots, \eta_d$ to be independent. In that case we claim that, for any $k \in \{1, \ldots, d\}$,

$$\xi_1 \cdots \xi_k \eta_{k+1} \cdots \eta_d f_n \xrightarrow{P} 0 \quad \Leftrightarrow \quad \xi_1 \cdots \xi_{k-1} \eta_k \cdots \eta_d f_n \xrightarrow{P} 0.$$

The asserted equivalence will then follow by induction in finitely many steps.

For the implication to the right, we may assume that $\xi_1 \cdots \xi_k\, \eta_{k+1} \cdots \eta_d f_n \to 0$ a.s., by a sub-sequence argument based on FMP 4.2. The same convergence then holds conditionally, given $\xi_1, \ldots, \xi_{k-1}$ and $\eta_{k+1}, \ldots, \eta_d$. By independence, the conditioning does not affect the distributions of ξ_k and η_k, and so the result for $d = 1$ yields $\xi_1 \cdots \xi_{k-1} \eta_k \cdots \eta_d f_n \xrightarrow{P} 0$, under the same conditioning. The corresponding unconditional convergence now follows by dominated convergence. A similar argument yields the implication to the left. □

Reversing the approach in Section 10.2, we may next express convergence and related criteria for random multi-linear forms in terms of the corresponding criteria for multiple compound Poisson integrals. Then replace the random sign sequences $\sigma, \sigma_1, \ldots, \sigma_d$ in Theorem 10.12 by more general sequences $\gamma = (\gamma_k)$ and $\gamma_j = (\gamma_{jk})$, respectively, and write as before

$$\gamma_1 \cdots \gamma_d\, A = \sum_{k \in \mathsf{N}^d} a_k (\gamma_{1,k_1} \cdots \gamma_{d,k_d}),$$

$$\gamma^d A = \sum_{k \in \mathsf{N}^d} a_k(\gamma_{k_1} \cdots \gamma_{k_d}), \tag{21}$$

for any arrays $A = (a_k)$ of real numbers indexed by N^d, such that the relevant sums exist. We assume that the random vectors $\gamma^k = (\gamma_{1,k}, \ldots, \gamma_{d,k})$ are i.i.d. and symmetrically distributed, in the sense that $-\gamma^k \overset{d}{=} \gamma^k$. Of special interest are the cases where the components $\gamma_{1,k}, \ldots, \gamma_{d,k}$ are either independent or equal.

Writing ν for the common distribution of the vectors $(\gamma_{1,k}, \ldots, \gamma_{d,k})$, we may introduce a symmetric, compound Poisson process $X = (X_1, \ldots, X_d)$ in R^d, with Lévy measure ν restricted to $\mathsf{R}^d \setminus \{0\}$, and consider the associated multiple integrals $X_1 \cdots X_d f_A$, where

$$f_A(t_1, \ldots, t_d) = \sum_{k \in \mathsf{N}^d} a_k 1\{t \in I_k\}, \quad t = (t_1, \ldots, t_d) \in \mathsf{R}_+^d.$$

When $\gamma_{1,k} = \cdots = \gamma_{d,k}$, then even $X_1 = \cdots = X_d$ a.s., and we may write the multiple integral as $X^d f_A$. Existence of the sums $\gamma_1 \cdots \gamma_d\, A$ and integrals $X_1 \cdots X_d f_A$ is defined as in Theorem 10.11.

Theorem 10.33 *(multi-linear forms and compound Poisson integrals) Let $(\gamma_{1,k}, \ldots, \gamma_{d,k})$, $k \in \mathsf{N}$, be i.i.d., positive or symmetric random vectors with distribution ν, and let $X = (X_1, \ldots, X_d)$ be a compound Poisson process in R^d with Lévy measure ν. Then for any positive or real, tetrahedral array A, the sum $\gamma_1 \cdots \gamma_d\, A$ and integral $X_1 \cdots X_d f_A$ exist simultaneously, in which case they are related by*

$$\gamma_1 \cdots \gamma_d\, A_n \xrightarrow{P} 0 \quad \Leftrightarrow \quad X_1 \cdots X_d f_{A_n} \xrightarrow{P} 0.$$

Proof: It is enough to prove the last assertion when both sides exist, since the existence assertion then follows from the special case of arrays with finitely many non-zero entries. By Corollary 10.29, we have $X_1 \cdots X_d f_n \xrightarrow{P} 0$ iff $[X_1 \cdots X_d] f_n^2 \xrightarrow{P} 0$, where $[X_1 \cdots X_d] = [X_1] \cdots [X_d]$ denotes the quadratic variation process of X. A similar argument based on Theorem 10.32 shows that $\gamma_1 \cdots \gamma_d\, A_n \xrightarrow{P} 0$ iff $[\gamma_1 \cdots \gamma_d]\, A_n^2 \xrightarrow{P} 0$, where $[\gamma_1 \cdots \gamma_d] = [\gamma_1] \cdots [\gamma_d]$ with $[\gamma_j]_k = \gamma_{j,k}^2$, and A_n^2 denotes the array with entries $a_{n,k}^2$. Since $f_{A_n^2} = f_{A_n}^2$, and since the Lévy measure of $([X_1], \ldots, [X_d])$ agrees with the distribution of $([\gamma_1]_k, \ldots, [\gamma_d]_k)$, it suffices to consider the case of non-negative variables $\gamma_{j,k}$ and entries a_k, so that X becomes a compound Poisson process in R_+^d.

By a coding argument (FMP 3.19), we may write $\gamma_{j,k} = h_j(\vartheta_k)$ a.s., for some measurable functions $h_1, \ldots, h_d$ and i.i.d. $U(0,1)$ random variables $\vartheta_1, \vartheta_2, \ldots$. Inserting this into (21) yields the representation

$$\begin{aligned}\gamma_1 \cdots \gamma_d\, A &= \sum_{k \in \mathsf{N}^d} a_k \Big\{ h_1(\vartheta_{k_1}) \cdots h_d(\vartheta_{k_d}) \Big\} \\ &= \sum_{k \in \mathsf{N}^d} f_k(\vartheta_{k_1}, \ldots, \vartheta_{k_d}) = \eta^d f,\end{aligned}$$

where $f_k = a_k(h_1 \otimes \cdots \otimes h_d)$, and the function f on R_+^d and point process η on R_+ are given by (19) and (20).

To compare with the corresponding multiple Poisson integral $\xi^d f$, we may assume, by Theorem 3.4, that

$$\xi = \sum_{k\in\mathsf{N}} \sum_{j\le\kappa_k} \delta_{\vartheta_{jk}+k},$$

for some i.i.d. Poisson random variables $\kappa_1, \kappa_2, \dots$ with mean 1, and an independent array (ϑ_{jk}) of i.i.d. $U(0,1)$ random variables. Writing $\gamma_{ijk} = h_i(\vartheta_{jk})$ and $\kappa_k = (\kappa_{k_1}, \dots, \kappa_{k_d})$, we obtain

$$\begin{aligned}
\xi^d f &= \sum_{k\in\mathsf{N}^d} \sum_{j\le\kappa_k} f_k(\vartheta_{j_1,k_1}, \dots, \vartheta_{j_d,k_d}) \\
&= \sum_{k\in\mathsf{N}^d} \sum_{j\le\kappa_k} a_k\Big\{h_1(\vartheta_{j_1,k_1}) \cdots h_d(\vartheta_{j_d,k_d})\Big\} \\
&= \sum_{k\in\mathsf{N}^d} \sum_{j\le\kappa_k} a_k\left(\gamma_{1,j_1,k_1} \cdots \gamma_{d,j_d,k_d}\right) \\
&= X_1 \cdots X_d f_A,
\end{aligned}$$

where

$$X_i(t) = \sum_{k\in\mathsf{N}} \sum_{j\le\kappa_k} \gamma_{ijk} 1\{\vartheta_{jk} + k \le t\}, \quad t \ge 0, \ i \le d.$$

Using Theorem 3.4 again, we see that $X = (X_1, \dots, X_d)$ is a compound Poisson process with Lévy measure ν.

Now let $A_1, A_2, \dots$ be arrays with positive entries, and denote the associated functions f in (19) by $f_1, f_2, \dots$. Then Theorem 10.32 yields $\eta^d f_n \xrightarrow{P} 0$ iff $\xi^d f_n \xrightarrow{P} 0$, and the asserted equivalence follows by the previous calculations. $\square$

10.6 Tangential Reduction and Decoupling

Here we use the tangential methods of Chapter 9 to extend the existence and convergence criteria of Sections 10.1–2 to multiple integrals $\xi_1 \cdots \xi_d f$ with respect to arbitrary marked point processes. Our main conclusion is that the previous results remain valid, if we replace the intensities $\mu_k = E\xi_k$ in the independence case by certain *sequential compensators* $\eta_1, \dots, \eta_d$, as defined below. This essentially solves the existence and convergence problems for arbitrary non-decreasing processes.

Given some T-marked point processes $\xi_1, \dots, \xi_d$ on $(0, \infty)$, adapted to a filtration $\mathcal{F}$, we take η_1 to be the $\mathcal{F}$-compensator of ξ_1. Continuing recursively, suppose that $\eta_1, \dots, \eta_{k-1}$ have already been defined for some $k < d$. Then let η_k be the compensator of ξ_k with respect to the extended filtration

$$\mathcal{F}_t^k = \mathcal{F}_t \vee \sigma(\eta_1, \dots, \eta_{k-1}), \quad t \ge 0, \ k = 1, \dots, d.$$

We further introduce the *sequentially tangential* processes $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$, where each $\tilde{\xi}_k$ is $\mathcal{F}^k$-tangential to ξ_k, with conditionally independent increments and intensity η_k, and the $\tilde{\xi}_k$ are conditionally independent, given $\eta_1, \ldots, \eta_d$.

We may now state our basic existence and convergence criteria.

Theorem 10.34 *(integrability and convergence) Let $\xi_1, \ldots, \xi_d$ be $\mathcal{F}$-adapted, T-marked point processes on $(0, \infty)$, with associated sequentially tangential processes $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$, and similarly for $\xi_1^n, \ldots, \xi_d^n$ and $\tilde{\xi}_1^n, \ldots, \tilde{\xi}_d^n$. Then for any tetrahedral functions $f, f_1, f_2, \ldots \geq 0$ on $(\mathsf{R}_+ \times T)^d$, we have*

(i) $\xi_1 \cdots \xi_d f < \infty$ *a.s.* $\Leftrightarrow$ $\tilde{\xi}_1 \cdots \tilde{\xi}_d f < \infty$ *a.s.,*

(ii) $\xi_1^n \cdots \xi_d^n f_n \xrightarrow{P} 0 \;\Leftrightarrow\; \tilde{\xi}_1^n \cdots \tilde{\xi}_d^n f_n \xrightarrow{P} 0.$

When the processes $\xi_1, \ldots, \xi_d$ are mutually independent with independent increments and intensities $\mu_1, \ldots, \mu_d$, we introduce the class

$$\mathcal{C}_d(\mu_1, \ldots, \mu_d) = \Big\{ f \geq 0;\, \xi_1 \cdots \xi_d f < \infty \text{ a.s.} \Big\},$$

which is essentially given by Proposition 10.5 above. Assertion (i) is then equivalent to

$$\xi_1 \cdots \xi_d f < \infty \;\text{ a.s.} \quad \Leftrightarrow \quad f \in \mathcal{C}_d(\eta_1, \ldots, \eta_d) \;\text{ a.s.}$$

Similarly, we may restate part (ii) in terms of the conditions in Theorem 10.9.

Both assertions follow easily from the following comparison lemma.

Lemma 10.35 *(sequential comparison) Let $\xi_1, \ldots, \xi_d$ and $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ be such as in Theorem 10.34. Then for any tetrahedral function $f \geq 0$ on $(\mathsf{R}_+ \times T)^d$ and increasing function $\varphi \geq 0$ of moderate growth, we have*

$$E\varphi(\xi_1 \cdots \xi_d f) \asymp E\varphi(\tilde{\xi}_1 \cdots \tilde{\xi}_d f),$$

where the domination constants depend only on φ and d.

Proof: The statement for $d = 1$ follows from Theorem 9.57. Now assume the assertion to be true in dimension $d - 1$. Proceeding to d dimensions, we see from Lemma 9.26 that the processes $\tilde{\xi}_2, \ldots, \tilde{\xi}_d$ are sequentially tangential with respect to the filtration $\mathcal{F}^2$ and probability measure $P(\cdot \,|\, \eta_1)$, and so the induction hypothesis yields a.s.

$$E\Big\{ \varphi(\xi_2 \cdots \xi_d\, g) \,\Big|\, \eta_1 \Big\} \asymp E\Big\{ \varphi(\tilde{\xi}_2 \cdots \tilde{\xi}_d\, g) \,\Big|\, \eta_1 \Big\}, \tag{22}$$

simultaneously for all measurable functions $g \geq 0$ on $(\mathsf{R}_+ \times T)^{d-1}$. Since f is tetrahedral, the partial integral $\xi_2 \cdots \xi_d f$ is predictable in the remaining argument, and so the associated integrals with respect to ξ_1 and $\tilde{\xi}_1$ are again tangential. Using Theorem 9.57, the disintegration theorem (FMP 6.4), the conditional independence

$$\tilde{\xi}_1 \perp\!\!\!\perp_{\eta_1} (\xi_2, \ldots, \xi_d, \tilde{\xi}_2, \ldots, \tilde{\xi}_d),$$

and formula (22) with g replaced by the partial integral μf, we get

$$\begin{aligned} E\,\varphi(\xi_1 \cdots \xi_d f) &\asymp E\varphi(\tilde{\xi}_1\, \xi_2 \cdots \xi_d f) \\ &= EE\Big\{\varphi(\mu\, \xi_2 \cdots \xi_d f) \,\Big|\, \eta_1, \tilde{\xi}_1 \Big\}_{\mu=\tilde{\xi}_1} \\ &= EE\Big\{\varphi(\mu\, \xi_2 \cdots \xi_d f) \,\Big|\, \eta_1 \Big\}_{\mu=\tilde{\xi}_1} \\ &\asymp EE\Big\{\varphi(\mu\, \tilde{\xi}_2 \cdots \tilde{\xi}_d f) \,\Big|\, \eta_1 \Big\}_{\mu=\tilde{\xi}_1} \\ &= EE\Big\{\varphi(\mu\, \tilde{\xi}_2 \cdots \tilde{\xi}_d f) \,\Big|\, \eta_1, \tilde{\xi}_1 \Big\}_{\mu=\tilde{\xi}_1} \\ &= E\,\varphi(\tilde{\xi}_1\, \tilde{\xi}_2 \cdots \tilde{\xi}_d f). \end{aligned}$$

In both approximation steps, we note that the domination constants depend only on φ and d. The induction is then complete. □

Proof of Theorem 10.34: Taking $\varphi(x) = x \wedge 1$, we see from Lemma 10.35 that, under the stated conditions,

$$E\Big(\xi_n^1 \cdots \xi_n^d f_n \wedge 1\Big) \asymp E\Big(\tilde{\xi}_n^1 \cdots \tilde{\xi}_n^d f_n \wedge 1\Big),$$

which yields the stated equivalence. The assertions now follow, since $\xi_1 \cdots \xi_d f < \infty$ a.s. iff $\xi_1 \cdots \xi_d(cf) \xrightarrow{P} 0$ as $c \to 0$, and similarly for $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$. □

The previous results simplify when the processes $\xi_1, \ldots, \xi_d$ are independent. Though the independence may not carry over to the associated compensators $\eta_1, \ldots, \eta_d$, we do have a simple criterion in terms of independent copies $\tilde{\eta}_1, \ldots, \tilde{\eta}_d$ of the η_k.

Corollary 10.36 *(decoupling) Let $\xi_1, \ldots, \xi_d$ be independent, adapted, T-marked point processes on $(0, \infty)$, with compensators $\eta_1, \ldots, \eta_d$ and their independent copies $\tilde{\eta}_1, \ldots, \tilde{\eta}_d$, and correspondingly for some processes indexed by n. Then for any functions $f, f_1, f_2, \ldots \geq 0$ on $(\mathsf{R}_+ \times T)^d$, we have*

(i) $P\Big\{\xi_1 \cdots \xi_d f < \infty\Big\} = P\Big\{f \in \mathcal{C}_d(\tilde{\eta}_1, \ldots, \tilde{\eta}_d)\Big\}$,

(ii) $\xi_1^n \cdots \xi_d^n f_n \xrightarrow{P} 0 \quad \Leftrightarrow \quad \tilde{\xi}_1^n \cdots \tilde{\xi}_d^n f_n \xrightarrow{P} 0.$

Proof: (i) Choose some independent copies $\xi_1', \ldots, \xi_d'$ of $\xi_1, \ldots, \xi_d$ and $\tilde{\eta}_1, \ldots, \tilde{\eta}_d$ of $\eta_1, \ldots, \eta_d$, and let $\tilde{\xi}_1', \ldots, \tilde{\xi}_d'$ be conditionally independent of $\xi_1, \ldots, \xi_d$ with independent increments and intensities $\tilde{\eta}_1, \ldots, \tilde{\eta}_d$. Then Theorem 9.60 (i) yields

$$\begin{aligned} P\{\xi_1 \cdots \xi_d f < \infty\} &= P\{\xi_1 \xi_2' \cdots \xi_d' f < \infty\} \\ &= P\{\tilde{\xi}_1 \xi_2' \cdots \xi_d' f < \infty\} \\ &= P\{\tilde{\xi}_1' \xi_2 \cdots \xi_d f < \infty\}. \end{aligned}$$

Since $\tilde{\xi}_1' \perp\!\!\!\perp (\xi_2, \ldots, \xi_d)$, we may continue recursively in d steps to get

$$P\{\xi_1 \cdots \xi_d f < \infty\} = P\{\tilde{\xi}_1' \cdots \tilde{\xi}_d' f < \infty\},$$

which is equivalent to the asserted relation.

(ii) Proceeding as before, we get for any non-decreasing function $\varphi \geq 0$ of moderate growth

$$E\,\varphi(\xi_1^n \cdots \xi_d^n f_n) \asymp E\,\varphi(\tilde{\xi}_1^n \cdots \tilde{\xi}_d^n f_n),$$

where the domination constants depend only on φ and d. Now let $n \to \infty$. □

The signed case is more difficult, since existence is then defined by a Cauchy criterion for the final values, whereas the tangential comparison is stated in terms of the maximum values. We then need the tangential processes $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$ to satisfy the maximum condition in Corollary 3.32. A further complication is that possible martingale properties of $\xi_1, \ldots, \xi_d$ may not be preserved by the sequentially tangential processes $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$. For simplicity, we consider only the following one-sided result:

Theorem 10.37 *(existence and convergence) Consider some signed, marked point processes $\xi_1, \ldots, \xi_d$, with sequentially tangential versions $\tilde{\xi}_1, \ldots, \tilde{\xi}_d$, satisfying the stated maximum property, and fix any measurable functions f and $f_1, f_2, \ldots$. Then*

(i) $\tilde{\xi}_1 \cdots \tilde{\xi}_d f$ *exists* $\Rightarrow$ $\xi_1 \cdots \xi_d f$ *exists,*

(ii) $\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n \xrightarrow{P} 0 \;\Rightarrow\; \xi_1 \cdots \xi_d f_n \xrightarrow{P} 0.$

Here the integrals on the left of (ii) are understood to exist.

Proof: First we prove (ii) for functions of bounded supports. Putting $\varphi(x) = |x| \wedge 1$, we may write our claim as

$$E\,\varphi(\tilde{\xi}_1 \cdots \tilde{\xi}_d f_n) \to 0 \quad \Rightarrow \quad E\,\varphi(\xi_1 \cdots \xi_d f_n) \to 0,$$

which will be proved by induction. For $d = 1$, the maximum property yields

$$E\,\varphi(\tilde{\xi} f_n) \to 0 \quad \Rightarrow \quad E\,\varphi\Big\{(f_n \cdot \tilde{\xi})^*\Big\} \to 0,$$

which by Theorem 9.59 implies

$$\begin{aligned} E\,\varphi(\xi f_n) \;&\leq\; E\,\varphi \circ (f_n \cdot \xi)^* \\ &\lesssim\; E\,\varphi \circ (f_n \cdot \tilde{\xi})^* \to 0. \end{aligned}$$

Assuming the statement in dimension $d-1$, we turn to the d-dimensional case. Using Lemma 9.26, we get as before

$$E\Big\{\varphi(\tilde{\xi}_2 \cdots \tilde{\xi}_d g_n)\,\Big|\,\eta_1\Big\} \to 0 \quad \Rightarrow \quad E\Big\{\varphi(\xi_2 \cdots \xi_d g_n)\,\Big|\,\eta_1\Big\} \to 0, \qquad (23)$$

outside a fixed P-null set. Now assume $E\,\varphi(\tilde\xi_1\cdots\tilde\xi_d f_n)\to 0$, so that

$$EE\Big\{\varphi(\mu\,\tilde\xi_2\cdots\tilde\xi_d f_n)\,\Big|\,\eta_1\Big\}_{\mu=\tilde\xi_1}\to 0,$$

by the disintegration theorem. Then for any sub-sequence $N'\subset\mathsf{N}$, we have

$$E\Big\{\varphi(\mu\,\tilde\xi_2\cdots\tilde\xi_d f_n)\,\Big|\,\eta_1\Big\}_{\mu=\tilde\xi_1}\to 0\ \text{ a.s.}$$

along a further sub-sequence $N''\subset N'$, which implies

$$E\Big\{\varphi(\mu\,\xi_2\cdots\xi_d f_n)\,\Big|\,\eta_1\Big\}_{\mu=\tilde\xi_1}\to 0\ \text{ a.s.,}$$

by (23). By dominated convergence and the disintegration theorem, we get $E\,\varphi(\tilde\xi_1\,\xi_2\cdots\xi_d f_n)\to 0$. Hence, by the maximum property and Theorem 9.57, we have along N''

$$\begin{aligned} E\,\varphi(\xi_1\cdots\xi_d f_n) &\le E\,\varphi\circ\Big\{(\xi_2\cdots\xi_d f_n)\cdot\xi_1\Big\}^* \\ &\asymp E\,\varphi\circ\Big\{(\xi_2\cdots\xi_d f_n)\cdot\tilde\xi_1\Big\}^*\to 0. \end{aligned}$$

Since N' was arbitrary, this completes the induction.

We now proceed to the general case:

(i) Suppose that $\tilde\xi_1\cdots\tilde\xi_d f$ exists, and consider any functions $f_1,f_2,\dots$ with bounded supports and $|f_n|\le|f|$, such that $f_n\to f$. Then $\tilde\xi_1\cdots\tilde\xi_d f_n \xrightarrow{P} \tilde\xi_1\cdots\tilde\xi_d f$, and so the sequence of integrals $\tilde\xi_1\cdots\tilde\xi_d f_n$ is Cauchy in probability. By the special case of (ii), the same property holds for the integrals $\xi_1\cdots\xi_d f_n$, which then converge to some limiting random variable. Considering alternating sequences of the form $f_1,f_1',f_2,f_2',\dots$, we see that the limit is a.s. independent of the choice of sequence (f_n), and hence defines a version of the integral $\xi_1\cdots\xi_d f$.

(ii) Suppose that $\tilde\xi_1\cdots\tilde\xi_d f_n\xrightarrow{P}0$. For every n, we may then choose some functions f_{nk} with bounded supports and $|f_{nk}|\le|f_n|$, such that $f_{nk}\to f_n$ as $k\to\infty$. By (i) and linearity, we get

$$\tilde\xi_1\cdots\tilde\xi_d(f_n-f_{nk})\xrightarrow{P}0,\qquad \xi_1\cdots\xi_d(f_n-f_{nk})\xrightarrow{P}0.$$

Writing $\psi(x)=|x|\wedge 1$, we may choose $k=k_n$ so large that the functions $f_n'=f_{n,k_n}$ satisfy

$$\psi\Big\{\tilde\xi_1\cdots\tilde\xi_d(f_n-f_n')\Big\}\vee\psi\Big\{\xi_1\cdots\xi_d(f_n-f_n')\Big\}<n^{-1}. \tag{24}$$

Then

$$\tilde\xi_1\cdots\tilde\xi_d f_n'=\tilde\xi_1\cdots\tilde\xi_d f_n-\tilde\xi_1\cdots\tilde\xi_d(f_n-f_n')\xrightarrow{P}0,$$

and so by (24) and the special case of (ii),

$$\xi_1\cdots\xi_d f_n=\xi_1\cdots\xi_d f_n'+\xi_1\cdots\xi_d(f_n-f_n')\xrightarrow{P}0. \qquad \Box$$

Chapter 11

Line and Flat Processes

Random measure theory has important applications to stochastic geometry. This is because some typical geometrical objects, such as lines or spheres, may be represented as points in suitable parameter spaces, so that random collections of such objects can be regarded as point processes on those spaces. The associated Papangelou conditional intensities then become more general random measures, and similarly for the directing random measures of Cox processes.

In this chapter, we study the special case of line and flat processes in Euclidean spaces R^d, where a *flat* is simply a translate of a linear subspace. We write F_k^d for the class of all k-dimensional flats in R^d, and Φ_k^d for the subclass of flats through the origin. The two spaces may be parametrized in a natural way as smooth manifolds of dimensions $(k+1)(d-k)$ and $k(d-k)$, respectively. The manifold structure induces a topology with associated Borel σ-field, along with a local structure, and it makes sense to consider point processes and more general random measures on F_k^d and Φ_k^d. Note in particular that a random measure η on F_k^d gives a.s. finite mass to the set of flats intersecting any bounded Borel set in R^d.

Simple point processes ξ on F_k^d are called *flat processes*, and when $k=1$ they are also called *line processes*. They clearly represent discrete random collections of lines or flats in R^d. We say that ξ is *stationary*, if its distribution is invariant[1] under arbitrary translations in R^d. Similarly, a random measure η on F_k^d is said to be *a.s. invariant*, if the measures $\eta(\omega)$ are invariant under such shifts for P-almost every $\omega \in \Omega$. We may also consider stationarity under the larger group of rigid motions, which has the advantage of making the action on F_k^d transitive. However, most results in the latter case are equivalent to the corresponding statements for the group of translations, which leads to significant simplifications.

The simplest case, though still general enough to exhibit some important features, is that of lines in the plane. Already in Theorem 11.1, we derive a remarkable factorization of the second order moment measure of a stationary random measure η on F_1^2. Assuming η to be ergodic with locally finite intensity $E\eta$, we have $\mathrm{Cov}(\eta A, \eta B) = 0$ for any disjoint sets A and B. In

[1]The distinction between stationarity and invariance is crucial here. Note that ξ is stationary iff $\mathcal{L}(\xi)$ is invariant. Invariance of ξ itself is a much stronger requirement.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_11

particular, a stationary line process in the plane, with locally finite intensity and a.s. no pairs of parallel lines, has the same first and second moments as a Cox process directed by an a.s. invariant random measure.

Stationary line processes of the mentioned Poisson and Cox types are easy to construct. However, excluding the occurrence of parallel lines, it is not so obvious how to construct stationary line processes with more general distributions. In fact, it was thought for some time that no such processes might exist, at least subject to some mild moment restriction. In Theorem 11.6 we construct a broad class of stationary line processes, which have no parallel lines or multiple intersections, and yet are not Cox distributed. As an immediate corollary, we obtain similar counter-examples for the case of stationarity under rigid motions.

More interesting are the results in the positive direction, where the proposed Cox property is established under additional regularity conditions. Since it is usually hard to prove directly that a given line or flat process ξ is Cox, a basic strategy is to prove instead that the associated Papangelou random measure η of Corollary 8.3 is a.s. invariant under translations. The required Cox property of ξ itself is then a consequence of Theorem 8.19. This reduces the original problem to that of proving, under suitable regularity conditions, that a stationary random measure η on the space F_k^d of k-flats in R^d is a.s. invariant under arbitrary shifts. Much of the remainder of the chapter is devoted to the latter invariance problem.

The basic assumptions on η are essentially of two types, referred to below as *spanning* and *absolute-continuity* conditions. To explain the former, we introduce the projection operator π, mapping any flat $x \in F_k^d$ into its parallel subspace $\pi x \in \Phi_k^d$. The basic spanning condition in Theorem 11.10 requires that, for $E\eta^2$-almost every pair of flats $x, y \in F_k^d$, the direction spaces πx and πy span the entire space R^d. This condition has the obvious limitation that it can only be fulfilled when $k \geq d/2$.

Conditions of absolute continuity apply more generally. Assuming η to be stationary and such that[2] $\eta \circ \pi^{-1} \ll \mu$ a.s. for some fixed measure μ on Φ_k^d, we show that η is a.s. invariant under suitable conditions on μ. The simplest case is that of Theorem 11.12, where the desired invariance is proved when μ equals λ_k—the unique probability measure on Φ_k^d invariant under arbitrary rotations. Using Corollary 7.36, we may reduce the discussion to the corresponding problem for stationary and product-measurable processes on F_k^d, where the required invariance follows by an argument based on the mean continuity of measurable functions.

The general absolute-continuity criteria, considered in the last section, are much deeper and based on the dual notions of inner and outer degeneracies, defined as follows. Given a σ-finite measure μ on Φ_k^d, we say that a flat $v \in \Phi_m^d$ with $m \geq k$ is an *outer degeneracy* of μ, if the set of flats $x \in F_k^d$ contained in v has positive μ-measure. If instead $m \leq k$, we say that $v \in \Phi_m^d$

[2]Recall that the measure $\mu \circ f^{-1}$ is defined by $(\mu \circ f^{-1})g = \mu(g \circ f)$, where $\mu f = \int f d\mu$.

is an *inner degeneracy* of μ, whenever the set of flats $x \in F_k^d$ containing v has positive μ-measure. Excluding the trivial cases of $m = d$ or 0, we say that μ has *no outer/inner degeneracies*, if no such flats exist of dimension $< d$ or > 0, respectively.

In Corollary 11.28 we show that a stationary random measure η on F_k^d is a.s. invariant, whenever $\eta \circ \pi^{-1} \ll \mu$ a.s. for some bounded measure μ on Φ_k^d with no outer degeneracies. This condition is best possible, in the sense that, whenever μ has an outer degeneracy, there exists a random measure η on F_k^d with $\eta \circ \pi^{-1} \ll \mu$ a.s., which is stationary but not a.s. invariant. Our proof is based on the unique degeneracy decomposition in Lemma 11.22, a reduction to the case $k = 1$ of lines, and an induction on the dimension d of the underlying space.

In Section 11.5, we explore the relationship between line processes and non-interacting particle systems. A connection arises from the fact that a particle moving at constant speed in R^d describes a straight line in the associated space-time diagram. Thus, a random collection of such particles in R^d gives rise to a line process in R^{d+1}, which is stationary under arbitrary translations, iff the underlying particle system is stationary in both space and time. Assuming stationarity in the space variables only, we may look for conditions ensuring that the system will approach a steady-state distribution, corresponding to stationarity even in time. Geometrically, this amounts to studying the asymptotic behavior of line processes that are stationary in all but one variables, under shifts in the remaining direction.

Here a classical result is the Breiman–Stone theorem, where the particle positions at time 0 form a stationary point process with a.s. finite sample intensity, and the associated velocities are independently chosen from an absolutely continuous distribution μ. As $t \to \infty$, the process ξ_t of positions at time t will then converge toward a Cox process with invariant directing random measure. Unfortunately, the independence assumption is rather artificial, since the independence will be destroyed at all times $t > 0$, unless the process is steady-state Cox to begin with. In Theorem 11.20, we consider a more realistic setting of conditional independence. We also discuss some general conditions for asymptotic invariance, incidentally leading, in Theorem 11.16, to some extensions of the invariance criteria for line processes considered earlier.

11.1 Stationary Line Processes in the Plane

Here we consider random measures ξ on the space of lines in R^2, equipped with the topology of local convergence and the generated Borel σ-field. We require ξ to give finite mass to the set of lines passing through an arbitrary bounded set. The space of lines may be parametrized in different ways as a two-dimensional manifold. If nothing else is said, stationarity and invariance of ξ are defined with respect to translations. By a *line process* in R^2 we mean a simple point process on the space of lines in R^2.

Let $\mathcal{I}_\xi$ denote the σ-field of ξ-measurable events, invariant under translations of R^2, and write D for the set of pairs of parallel lines in R^2. Applying Lemma 2.7 to $\xi \circ \pi^{-1}$, we see that $E\xi^2 D = 0$, iff a.s. every set of parallel lines has ξ-measure 0. Furthermore, when ξ is a line process, $E\xi^{(2)} D = 0$ iff ξ has a.s. no pairs of parallel lines. We say that a line process ξ has no *multiple intersections*, if at most two lines of ξ pass through each point, and that ξ is *non-degenerate*, if it has no parallel lines or multiple intersections.

Much of the theory relies on the following remarkable moment identity, which has no counterpart for general random measures.

Theorem 11.1 *(second moment measure, Davidson, OK) Let ξ be a stationary random measure on the space of lines in R^2, such that $\eta = E(\xi \,|\, \mathcal{I}_\xi)$ is a.s. locally finite*[3]*. Then $E(\xi^2 | \mathcal{I}_\xi) = \eta^2$ a.s. on D^c.*

Proof: Let $f \geq 0$ be a measurable function on the set of lines in R^2, and put $f_p = f(p, \cdot)$, where p denotes direction. Writing $I_a = [-a/2, a/2]$, we get for any $p \neq q$

$$\begin{aligned}
&\iint_{I_a^2} f_p(sp + tq)\, f_q(sp + tq)\, ds\, dt \\
&\qquad = \int_{I_a} f_p(tq)\, dt \int_{I_a} f_q(sp)\, ds \\
&\qquad = a^{-2} \iint_{I_a^2} f_p(sp + tq)\, ds\, dt \iint_{I_a^2} f_q(sp + tq)\, ds\, dt. \qquad (1)
\end{aligned}$$

Write $R_{p,q} = \{sp + tq;\, s, t \in I_1\}$. For any $p_0 \neq q_0$ and $\varepsilon > 0$, let A be a non-diagonal neighborhood of (p_0, q_0), such that

$$(1 + \varepsilon)^{-1} R_{p_0,q_0} \subset R_{p,q} \subset (1 + \varepsilon) R_{p_0,q_0}, \quad (p, q) \in A.$$

For fixed $\omega \in \Omega$, consider a disintegration $\xi = \int \nu(dp)\, \xi_p$, where ν is a measure on the set of directions, and ξ_p is restricted to lines with direction p. Writing θ for possibly different quantities of the form $1 + O(\varepsilon)$, we get by (1) for measurable $B, C \subset A \times \mathsf{R}^2$

$$\begin{aligned}
&\iint_{I_a^2} \xi(B + sp_0 + tq_0)\, \xi(C + sp_0 + tq_0)\, ds\, dt \\
&\quad = \iint \nu^2(dp\, dq) \iint_{I_a^2} \xi_p(B + sp_0 + tq_0)\, \xi_q(C + sp_0 + tq_0)\, ds\, dt \\
&\quad \leq \theta \iint \nu^2(dp\, dq) \iint_{\theta I_a^2} \xi_p(B + sp + tq)\, \xi_q(C + sp + tq)\, ds\, dt \\
&\quad = a^{-2}\, \theta \iint \nu^2(dp\, dq) \iint_{\theta I_a^2} \xi_p(B + sp + tq)\, ds\, dt \\
&\qquad\qquad\qquad\qquad \times \iint_{\theta I_a^2} \xi_q(C + sp + tq)\, ds\, dt
\end{aligned}$$

[3]For the existence of conditional intensities such as $E(\xi \,|\, \mathcal{F})$, see Lemma 2.10.

$$\begin{aligned}
&\le a^{-2}\,\theta \iint \nu^2(dp\,dq) \iint_{\theta I_a^2} \xi_p(B+sp_0+tq_0)\,ds\,dt \\
&\qquad\qquad\qquad\qquad\times \iint_{\theta I_a^2} \xi_q(C+sp_0+tq_0)\,ds\,dt \\
&= a^{-2}\,\theta \iint_{\theta I_a^2} \xi(B+sp_0+tq_0)\,ds\,dt \iint_{\theta I_a^2} \xi(C+sp_0+tq_0)\,ds\,dt,
\end{aligned}$$

and similarly for the corresponding lower bound. Letting $a \to \infty$, and applying the pointwise ergodic theorem in R^2 three times, we get a.s.

$$\begin{aligned}
\theta^{-1}\eta^2(B\times C) &\le E\Big\{\xi^2(B\times C)\,\Big|\,\mathcal{I}_\xi\Big\} \\
&\le \theta\,\eta^2(B\times C).
\end{aligned}$$

The rectangles $B\times C$ form a dissecting semiring in $D^c\times\mathsf{R}^2$, and so $\theta^{-1}\eta^2 \le E(\xi^2\,|\,\mathcal{I}_\xi) \le \theta\,\eta^2$ a.s. on D^c, by Lemma 2.1. As $\varepsilon \to 0$, we get $E(\xi^2\,|\,\mathcal{I}_\xi) = \eta^2$ a.s. on D^c. □

Under stronger moment conditions, we can also give a simple non-computational proof:

Proof for locally finite $E\xi^2$ (Krickeberg): Let μ be a σ-finite measure on D^c, invariant under joint translations. By Theorem 7.3 it allows a disintegration on D^c into jointly invariant measures $\mu_{p,q}$ on $(\pi^{-1}p)\times(\pi^{-1}q)$ with directions $p \ne q$. Letting S, S' and T, T' be arbitrary translations in directions p and q, we get for any lines $x \in \pi^{-1}p$ and $y \in \pi^{-1}q$

$$\begin{aligned}
(STx, S'T'y) &= (Tx, S'y) = (S'Tx, S'Ty) \\
&= S'T(x,y),
\end{aligned}$$

which shows that every separate translation is also a joint one. Thus, every $\mu_{p,q}$ is even separately invariant, which remains true for the mixture μ on D^c. In particular, the jointly invariant random measure $E(\xi^2|\,\mathcal{I}_\xi)$ is also a.s. separately invariant on D^c. Averaging separately in each component, and using the L^2-ergodic theorem in R^2 three times, we get

$$\begin{aligned}
E(\xi^2\,|\,\mathcal{I}_\xi) &= E\Big(\Big\{E(\xi\,|\,\mathcal{I}_\xi)\Big\}^2\Big|\;\mathcal{I}_\xi\Big) \\
&= \Big\{E(\xi\,|\,\mathcal{I}_\xi)\Big\}^2 = \eta^2,
\end{aligned}$$

a.s. on every measurable product set in D^c. The assertion now follows by Lemma 2.1. □

The last theorem yields a remarkable invariance property.

Corollary 11.2 *(stationarity and invariance) Let η be a stationary random measure on the space of lines in R^2, such that $E\eta^2 D = 0$, and $\zeta = E(\eta\,|\,\mathcal{I}_\eta)$ is a.s. locally finite. Then η is a.s. invariant.*

We conjecture that the moment condition can be dropped.

Proof: Since $\eta^2 D = 0$ a.s., Theorem 11.1 yields[4]

$$\begin{aligned} E(\eta^2 | \mathcal{I}_\eta) &= E\Big(1_{D^c}\eta^2\Big|\,\mathcal{I}_\eta\Big) \\ &= 1_{D^c} E(\eta^2 | \mathcal{I}_\eta) \\ &= 1_{D^c}\zeta^2 \le \zeta^2, \end{aligned}$$

and so $\mathrm{Var}(\eta B \,|\, \mathcal{I}_\eta) = 0$ a.s. for any $B \in \hat{\mathcal{B}}^2$, which implies $E(\eta B - \zeta B)^2 = 0$, and finally $\eta B = \zeta B$ a.s. Since B was arbitrary, Lemma 2.1 yields $\eta = \zeta$ a.s. Finally, ζ is a.s. invariant, by the stationarity of η and the invariance of $\mathcal{I}_\eta$. $\Box$

We list some interesting properties of stationary line processes:

Corollary 11.3 *(stationary line processes) Let ξ be a stationary line process in R^2, such that $\xi^{(2)} D = 0$ a.s., and $\eta = E(\xi \,|\, \mathcal{I}_\xi)$ is a.s. locally finite. Then*

(i) *η is a.s. invariant with $\eta^2 D = 0$ a.s., and $E(\xi^{(2)} \,|\, \mathcal{I}_\xi) = \eta^2$ a.s.,*
(ii) *$E\xi^{(2)} = (E\xi)^2$ with $E\xi$ locally finite, iff η is a.s. non-random,*
(iii) *if ξ is stationary under rotations, then so is η.*

Proof: (i) The a.s. invariance of η follows from the stationarity of ξ and the definition of $\mathcal{I}_\xi$. Since $\xi^{(2)} D = 0$ a.s., ξ has a.s. no pairs of parallel lines, and so the ergodic theorem shows that η a.s. gives zero mass to every set of parallel lines, which implies $\eta^2 D = 0$ a.s. Now Theorem 11.1 yields $E(\xi^{(2)} \,|\, \mathcal{I}_\xi) = \eta^2$ a.s. on D^c, which extends to L^2, since $\xi^{(2)} D = \eta^2 D = 0$ a.s.

(ii) If η is a.s. non-random, then $E\xi = EE(\xi \,|\, \mathcal{I}_\xi) = E\eta = \eta$ is locally finite, and (i) yields

$$\begin{aligned} E\xi^{(2)} &= EE\Big(\xi^{(2)}\Big|\,\mathcal{I}_\xi\Big) \\ &= E\eta^2 = \eta^2 \\ &= (E\eta)^2 = (E\xi)^2. \end{aligned}$$

Conversely, if $E\xi^{(2)} = (E\xi)^2$, then $E\eta^2 = (E\eta)^2$ by the same calculation. When $E\xi = E\eta$ is locally finite, this yields $\mathrm{Var}(\eta B) = 0$ for every $B \in \hat{\mathcal{B}}^2$, and so $\eta B = E\eta B$ a.s. for all B, which implies $\eta = E\eta$ a.s. by Lemma 2.1.

(iii) This may be seen most easily from the ergodic theorem. $\Box$

The most obvious examples of stationary line processes are the Cox processes directed by invariant random measures.

Corollary 11.4 *(stationary Cox line processes) Let ξ be a stationary Cox line process directed by a random measure η, such that $\eta^2 D = 0$ a.s. and $E(\eta \,|\, \mathcal{I}_\eta)$ is a.s. locally finite. Then*

[4]Recall that $1_B\mu$ denotes the restriction of μ to B, given by $(1_B\mu)f = \int_B f d\mu$.

(i) *η is a.s. invariant with $\sigma(\eta) = \mathcal{I}_\xi$, and $E(\xi^{(n)}|\,\mathcal{I}_\xi) = \eta^n$ for all $n \in \mathsf{N}$,*
(ii) *ξ is a.s. non-degenerate,*
(iii) *ξ and η are simultaneously stationary under rotations.*

Proof: (i) Since η is again stationary, by Lemma 3.3, it is a.s. invariant by Corollary 11.2. Proceeding as in the proof of Theorem 5.30, we obtain $\sigma(\eta) = \mathcal{I}_\xi$ a.s. Hence, Theorem 6.27 (i) yields $E(\xi^{(n)}|\,\mathcal{I}_\xi) = E(\xi^{(n)}|\,\eta) = \eta^n$, a.s. for all $n \in \mathsf{N}$.

(ii) Since $E\xi^{(2)}D = E\eta^2 D = 0$ by (i), ξ has a.s. no parallel lines. To see that ξ has a.s. no multiple intersections either, it is enough to consider its conditional distribution, given η. Then for any disjoint Borel sets of directions A and B, the associated restrictions $1_A\xi$ and $1_B\xi$ are stationary and independent. Hence, by Fubini's theorem, the lines of $1_B\xi$ a.s. fail to hit the countably many intersection points of $1_A\xi$.

(iii) This is again obvious by Lemma 3.3. □

Combining Corollary 11.2 with results from Chapter 8, we get conditions for a stationary line process to be Cox. Similar results can be obtained for all invariance criteria in subsequent sections.

Corollary 11.5 *(Cox criterion) Let ξ be a stationary line process in R^2 satisfying (Σ), and such that the Papangelou kernel η of ξ is locally integrable and satisfies $\eta^2 D = 0$ a.s. Then η is a.s. invariant, and ξ is a Cox process directed by η.*

Proof: The stationarity of ξ carries over to η. The latter is then a.s. invariant by Corollary 11.2, and so the Cox property of ξ follows by Theorem 8.19. □

11.2 A non-Cox Counterexample

Given a stationary line process ξ in R^2, such that $\xi^{(2)}D = 0$ a.s. and $\eta = E(\xi\,|\,\mathcal{I}_\xi)$ is a.s. locally finite, let $\tilde{\xi}$ be a Cox line process directed by η, and conclude from the preceding corollaries that $E\tilde{\xi} = E\xi$ and $E(\tilde{\xi})^2 = E\xi^2$. This suggests that $\xi \overset{d}{=} \tilde{\xi}$, so that any line process ξ with the stated properties would be a Cox process. This is not so, and we proceed to construct a broad class of counter-examples.

Mutual singularity of measures is denoted by $\perp$. For random measures, the relation is measurable by Lemma 1.28.

Theorem 11.6 *(stationary line processes that are not Cox) Let η be an a.s. invariant random measure on the space of lines in R^2, such that $\eta^2 D = 0$ a.s. and $P\{\eta \perp \lambda^2\} < 1$. Then there exists a stationary, non-degenerate, non-Cox line process ξ in R^2 with $E(\xi|\,\mathcal{I}_\xi) = \eta$ a.s. If η is even stationary under rotations, we may choose ξ to have the same property.*

For the proof, it is convenient to replace the usual cylindrical parametrization of lines in R^2 by the following *linear* representation. Then fix a directed line l_0 through the origin. For lines l that are not parallel to l_0, let ϑ be the angle between l and l_0, measured counter-clockwise from l_0 to l, and let y be the distance along l_0 from the origin to the point of intersection. Writing $x = \cot\vartheta$, we may identify l with the pair $(x, y) \in \mathsf{R}^2$. By elementary geometry, a translation of l in a finite direction r, measured as before, corresponds to a *vertical shear* $S_{r,t}$ in the parameter space of the form

$$S_{r,t}(x, y) = \Big\{x,\ y + t(r - x)\Big\}, \quad x, y, r, t \in \mathsf{R}.$$

Furthermore, translations in the direction of l_0 correspond to vertical shifts of R^2. Such vertical shears and shifts clearly form a subgroup of affine transformations on R^2. For any random measure η on the space of lines, we may choose l_0 such that $E\eta$ gives zero mass to the set of lines parallel to l_0.

Our construction is based on a certain lattice-type point process in the parameter plane, which is stationary under the mentioned shifts and shears. By a *lattice point process* on R^2 we mean the image of Z^2, or the associated counting measure, under a random, non-singular, affine transformation.

Lemma 11.7 *(shear-stationary lattice process) There exists a lattice point process ξ on R^2, which is stationary under translations and vertical shears, and satisfies*

$$E\xi = \lambda^2, \qquad E\xi^{(2)} = \lambda^4, \qquad E\xi^{(4)} \perp \lambda^8.$$

Proof: Applying a uniform shift in $[0, 1]^2$ to the counting measure on Z^2, we obtain a stationary lattice point process ζ on R^2. Next, we apply to ζ an independent, uniformly distributed rotation about the origin, to form a rotationally stationary lattice point process η. Finally, we choose $\vartheta_n \perp\!\!\!\perp \eta$ for each $n \in \mathsf{N}$ to be uniformly distributed on $[n, 2n]$, and form a lattice point process ξ_n by applying the shear $S_{0,\vartheta_n} = S_{\vartheta_n}$ to η. By the invariance of λ^2 under shifts, rotations, and shears, we have $E\xi_n = E\eta = E\zeta = \lambda^2$. In particular, the sequence (ξ_n) is vaguely tight by Theorem 4.10, and so $\xi_n \xrightarrow{vd} \xi$ along a sub-sequence, where the limit ξ is stationary under translations and vertical shears.

To determine $E\xi$, let $\hat\zeta$, $\hat\eta$, $\hat\xi_n$ denote the Palm versions of ζ, η, ξ_n, respectively, obtained by shifting one point of the process to 0, and put $B_r = \{x \in \mathsf{R}^2;\ |x| < r\}$. Approximating each atom of $\hat\zeta$ by Lebesgue measure on the surrounding unit square, we get $\hat\zeta B_r \le \lambda^2 B_{r+c}$ with $c = 2^{-1/2}$, and so for any $r \ge 1$,

$$\begin{aligned} E(\hat\eta - \delta_0)B_r &\le \lambda^2 B_{r+c} - 1 \\ &= \pi(r + c)^2 - 1 \\ &\le 3\pi r^2 = 3\lambda^2 B_r, \end{aligned}$$

which extends to all $r \ge 0$, since $\hat\eta B_r = 1$ for $r < 1$. By rotational symmetry, we obtain $E(\hat\eta - \delta_0)C \le 3\lambda^2 C$ for any bounded, convex set $C \subset \mathsf{R}^2$ containing the origin. Since the shears $S_t = S_{0,t}$ preserve convexity and area, we get

$$\begin{aligned} E(\hat{\xi}_n - \delta_0)C &= n^{-1}\int_n^{2n} E(\hat{\eta} - \delta_0)S_t C\,dt \\ &\leq 3\,n^{-1}\int_n^{2n} \lambda^2 S_t C\,dt = 3\lambda^2 C. \end{aligned} \tag{2}$$

Since $x \in C$ implies $0 \in C - x$, a moment disintegration yields for any convex set C

$$\begin{aligned} E(\xi_n C)^2 &\leq \lambda^2 C + 3\int_C \lambda^2(dx)\,\lambda^2(C-x) \\ &= \lambda^2 C\,(1 + 3\lambda^2 C) < \infty. \end{aligned}$$

The variables $\xi_n C$ are then uniformly integrable in n for fixed C, and so $\lambda^2 f = E\xi_n f \xrightarrow{v} E\xi f$, for every continuous function with bounded support $f \geq 0$ on R^2, which implies $E\xi = \lambda^2$.

To determine $E\xi^{(2)}$, we introduce the sets

$$B_{p,q,r,s} = \Big\{(x,y) \in \mathsf{R}^2;\ x \in [p,q],\ y \in [rx, sx]\Big\}, \quad 0 < p < q,\ r < s,$$

and note that

$$S_t B_{p,q,r,s} = B_{p,q,r+t,s+t}, \quad t \in \mathsf{R}.$$

Since $E\hat{\eta} - \delta_0$ can be obtained from λ^2 by shifting mass by a distance $\leq c$ in the radial direction, we conclude that for $r + t > 0$,

$$\begin{aligned} \lambda^2 B_{p+\varepsilon,q-\varepsilon,r,s} &\leq E\,\hat{\eta} S_t B_{p,q,r,s} \\ &\leq \lambda^2 B_{p-\varepsilon,q+\varepsilon,r,s}, \end{aligned}$$

where $\varepsilon = c\{1+(r+t)^2\}^{-1/2}$. Hence, $E\,\hat{\eta} S_t B \to \lambda^2 B$ for every set $B = B_{p,q,r,s}$, and so by dominated convergence $E\hat{\xi}_n B \to \lambda^2 B$. Assuming $E\hat{\xi}_n \xrightarrow{v} m_2$ along a sub-sequence, we obtain $m_2 B = \lambda^2 B$ for sets B as above, and so a monotone-class argument yields $m_2 = \lambda^2$ on $(0,\infty) \times \mathsf{R}$. This extends by symmetry to $\mathsf{R}' \times \mathsf{R}$ with $\mathsf{R}' = \mathsf{R} \setminus \{0\}$, and since $m_2 = 0$ on $\{0\} \times \mathsf{R}'$ by (2), we obtain $m_2 = \delta_0 + \lambda^2$. Since the limit is independent of sub-sequence, we conclude that $E\hat{\xi}_n \xrightarrow{v} \delta_0 + \lambda^2$ along the original sequence.

Now consider any continuous function $f \geq 0$ on R^4 with bounded support. By moment disintegration and dominated convergence,

$$\begin{aligned} E\,\xi_n^2 f &= \int \lambda^2(dx) \int E\,\hat{\xi}_n(dy)\, f(x, x+y) \\ &\to \int \lambda^2(dx) \int (\delta_0 + \lambda^2)(dy)\, f(x, x+y) \\ &= (\lambda^2)_D\, f + \lambda^4 f. \end{aligned}$$

Since $\xi_n^2 \xrightarrow{vd} \xi^2$, we obtain $E\,\xi^2 f \leq (\lambda^2)_D\, f + \lambda^4 f$ by FMP 4.11, and so $E\xi^2 \leq (\lambda^2)_D + \lambda^4$ and $E\xi^{(2)} \leq \lambda^4$. In particular, $E\,\hat{\xi}(\{0\} \times \mathsf{R}) = 0$. The process ξ is clearly stationary under translations and vertical shifts, and its

Palm version $\hat{\xi}$ can be chosen to be stationary under the shears S_t. Hence, the measure $E\hat{\xi}$ is invariant under each S_t. Changing variables and using Theorem 7.1, we get $E\hat{\xi} - \delta_0 = \mu \otimes \lambda$ for some $\mu \in \mathcal{M}_{\mathsf{R}}$.

For measurable functions $f \geq 0$, a moment disintegration yields

$$\lambda^4 f \geq E\,\xi^{(2)} f = \int \lambda^2(dx) \int (\mu \otimes \lambda)(dy)\, f(x, x+y).$$

Writing $f(x,y) = g(x, y-x)$, and using the invariance of λ and Fubini's theorem, we obtain $\lambda^4 \geq \lambda^3 \otimes \mu$. Applying this to any measurable rectangle gives $\mu \leq \lambda$.

Next, define $B_j = [0,a] \times (j-1, j)$, $j \in \mathsf{N}$, for a fixed $a > 0$. Then for any $i \neq j$,

$$\begin{aligned} E(\xi B_i)(\xi B_j) &= \int_{B_i} \lambda^2(dx)\, (\mu \otimes \lambda)(B_j - x) \\ &= \int_0^a \mu(-x, a-x)\, dx \\ &= \int_{-a}^a \mu(dx)\, |a - x|, \end{aligned}$$

and so

$$\begin{aligned} \mathrm{Cov}(\xi B_i,\, \xi B_j) &= E\,\xi(B_i)\xi(B_j) - E\xi(B_i)\, E\xi(B_j) \\ &= -\int_{-a}^a (\lambda - \mu)(dx)\, |a - x| \leq 0, \end{aligned}$$

since $\lambda \geq \mu$. Letting $n \to \infty$ in the relation

$$\begin{aligned} 0 &\leq n^{-2}\, \mathrm{Var} \sum_{i \leq n} \xi B_i \\ &= n^{-1}\, \mathrm{Var}(\xi B_i) + (1 - n^{-1})\, \mathrm{Cov}(\xi B_i,\, \xi B_j), \end{aligned}$$

we obtain $\mathrm{Cov}(\xi B_i,\, \xi B_j) \geq 0$, and so $\mu = \lambda$ on $(-a, a)$. Since $a > 0$ was arbitrary, the equality extends to all of R, and we get $E\xi^{(2)} = \lambda^4$.

Since $E\xi^{(2)} = (E\xi)^2$, Theorem 11.3 yields $E(\xi \,|\, \mathcal{I}_\xi) = E\xi = \lambda^2$, and so the ergodic theorem yields $\xi \neq 0$ a.s. By Lemma 5.2 (i), ξ is then a lattice process with the desired symmetry properties. Since $\mu\{0\} = \lambda\{0\} = 0$, ξ has a.s. at most one point on every vertical line.

To see that $E\xi^{(4)} \perp \lambda^8$, let $x_1, x_2, x_3, x_4 \in \mathsf{R}^2$ be distinct. Since any triple of colinear points x_1, x_2, x_3 is five-dimensional and hence generates a λ^6-singular set in $(\mathsf{R}^2)^4$, it remains to consider the case where the three points generate a two-dimensional lattice L in R^2. Then any lattice L' containing x_4 must also contain L as a sub-lattice, which determines a countable collection of possible lattices L'. Since each lattice is countable, there are countably many possible positions for x_4. Now consider a disintegration $E\xi^{(4)} = E\xi^{(3)} \otimes \mu$, in terms of a kernel $\mu\colon (\mathsf{R}^2)^3 \to \mathsf{R}^2$, which exists since $E\xi^{(3)}$ is a supporting measure of $E\xi^{(4)}$. Then for almost every (x_1, x_2, x_3) outside a λ^6-singular set, the measure μ_{x_1,x_2,x_3} has countable support and is therefore λ^2-singular. Since $\nu \otimes \mu$ is singular when either ν is singular or μ is singular a.e. ν, the asserted singularity of $E\xi^{(4)}$ follows. □

Line processes generated by stationary lattice point processes have typically multiple intersections, which can be dissolved by a simple randomization.

Lemma 11.8 *(dissolving multiplicities) Let ξ be a stationary line process in R^2 with $\xi^{(2)}D = 0$ a.s., such that $\eta = E(\xi|\mathcal{I}_\xi)$ is a.s. locally finite. Form a new line process ζ by applying some independent, perpendicular shifts to the lines of ξ, according to a common symmetric and diffuse distribution. Then ζ is stationary and non-degenerate with $E(\zeta|\mathcal{I}_\zeta) = \eta$ a.s. If ξ is even stationary under rotations, then so is ζ.*

Proof: The stationarity of ξ is preserved by a uniform randomization, from which we can form ζ by an invariant construction. This yields the stationarity of ζ, and the a.s. relation $E(\xi|\mathcal{I}_\xi) = E(\zeta|\mathcal{I}_\zeta)$ follows by the ergodic theorem. Since the directions of lines are preserved by shifts, ζ has a.s. no parallel lines either. To see that ζ has a.s. no multiple intersections, we may condition on ξ and proceed as in the proof of Corollary 11.3 (iii). The last assertion is clear from the rotational invariance of the construction. □

Proof of Theorem 11.6: Since η is a.s. invariant under vertical shifts in the parameter space, Theorem 7.1 gives $\eta = \rho \otimes \lambda$ a.s., for some random measure ρ on R. Next, Lemma 1.28 yields a measurable decomposition $\rho = \rho_a + \rho_s$ with $\rho_a \ll \lambda$ and $\rho_s \perp \lambda$ a.s., along with a product-measurable process $Y \geq 0$ on R with $\rho_a = Y \cdot \lambda$ a.s. Since $\rho \ll \lambda$ iff $\rho \otimes \lambda \ll \lambda^2$, we have $P\{\rho \perp \lambda\} = P\{\eta \perp \lambda^2\} < 1$, and so $E\lambda Y > 0$. Then for small enough $p \in (0, 1]$, the random set $A_p = \{Y \geq p\}$ satisfies $E\lambda A_p > 0$. Putting $\rho' = p1_{A_p} \cdot \lambda$ and $\rho'' = \rho - \rho'$, we get $P\{\rho' \neq 0\} > 0$ and $\rho'' \geq 0$.

Now define $\xi = \xi' + \xi''$, where $\xi' \perp\!\!\!\perp_{\rho',\rho''} \xi''$, and ξ' is a p-thinning of the lattice process in Lemma 11.7, restricted to $A \times \mathsf{R}$, while ξ'' is a Cox process directed by $\rho'' \otimes \lambda$. Conditioning on ρ' and ρ'', we see that ξ is again stationary under vertical shears and has at most one point on every vertical line. The ergodic theorem yields a.s.

$$\begin{aligned} E(\xi\,|\,\mathcal{I}_\xi) &= E(\xi'\,|\,\mathcal{I}_{\xi'}) + E(\xi''\,|\,\mathcal{I}_{\xi''}) \\ &= p1_A\lambda \otimes \lambda + \rho'' \otimes \lambda \\ &= \rho \otimes \lambda = \eta. \end{aligned}$$

If ξ is Cox, its directing random measure equals η by Corollary 11.3. Combining this with the last relation in Lemma 11.7, we get on $(A_p \times \mathsf{R})^4$ the a.s. relations

$$\begin{aligned} \lambda^8 \perp E(\xi'^{(4)}|\,\eta) &\leq E(\xi^{(4)}|\,\eta) \\ &= \eta^4 \ll \lambda^8, \end{aligned} \tag{3}$$

which are contradictory since $E\xi'^{(4)} \neq 0$. Hence, ξ is not a Cox process. Since ξ has a.s. at most one point on every vertical line, the corresponding line process has a.s. no parallel lines. Finally, we may randomize, as in Lemma 11.8, to eliminate any multiple intersections. The Cox property fails for the resulting processes when the shifts are small enough, since the class of Cox processes is closed under convergence in distribution, by Lemma 4.17.

Now suppose that η is even stationary under rotations. Then construct ξ as above, apart from the final randomization, and form ξ_ϑ by rotating ξ about the origin by an independent, uniformly distributed angle ϑ, so that ξ_ϑ becomes stationary under such special rotations. Conditioning on ϑ, we see that ξ_ϑ is also stationary under translations, and so, by iteration, it is stationary under arbitrary rotations. Since the relations in (3) are preserved by rotations, we see as before that ξ_ϑ is not Cox. Finally, we may randomize, as in Lemma 4.17, to eliminate any multiple intersections.

The ergodic theorem yields $E(\xi_\vartheta | \mathcal{I}_{\xi_\vartheta}) = \eta_\vartheta$ a.s., where η_ϑ is formed from η by rotation about the origin by the same random angle ϑ. The rotational stationarity of η yields $\eta_\vartheta \stackrel{d}{=} \eta$, and so the transfer theorem yields a line process ζ with $(\zeta, \eta) \stackrel{d}{=} (\xi_\vartheta, \eta_\vartheta)$. The process ζ is again non-Cox, non-degenerate, and stationary under rotations with $E(\zeta | \mathcal{I}_\zeta) = \eta$ a.s. $\Box$

11.3 Spanning Criteria for Invariance

A *k-flat* in R^d is defined as a translate of a k-dimensional linear subspace of R^d, for arbitrary $k \le d$. We denote the space of k-flats in R^d by $F_k = F_k^d$, so that F_1^d is the space of lines and F_{d-1}^d the space of hyperplanes in R^d. For completeness, we may even allow $k = 0$ or d, so that $F_0^d = \mathsf{R}^d$ and $F_d^d = \{\mathsf{R}^d\}$. The *direction* of a flat $x \in F_k^d$ is defined as the shifted flat πx containing the origin, and we write $\Phi_k = \Phi_k^d$ for the set of such subspaces. Let $F^d = \bigcup_k F_k^d$ and $\Phi^d = \bigcup_k \Phi_k^d$, and put $F' = F^d \setminus F_d^d$. For fixed $u \in \Phi^d$, we define $\Phi_k^{(u)} = \{x \in \Phi_k^d;\, x \subset u\}$, and similarly for $\Phi^{(u)}$, $F_k^{(u)}$, and $F^{(u)}$. Write $\mathrm{span}\{x, y, \dots\}$ for the linear subspace of R^d spanned by $x, y, \dots$.

The spaces Φ_k^d and F_k^d may be parametrized as manifolds of dimensions $k(d-k)$ and $(k+1)(d-k)$, respectively. As such, they are endowed with the obvious topologies of local convergence, with associated Borel σ-fields, which make them into Borel spaces. The space Φ_k^d is clearly a compact sub-manifold of F_k^d. Both spaces may also be regarded as homogeneous spaces, under the action of the groups of rotations about the origin and general rotations, respectively, where the latter includes the translations by definition. However, F_k^d fails to be transitive under the group of translations, for which we have instead the orbit decomposition $F_k^d = \bigcup\{\pi^{-1}x;\, x \in \Phi_k^d\}$.

Flats with the same direction are said to be *parallel*. More generally, given any $x \in F_h^d$ and $u \in F_k^d$ with $h \le k$, we say that x is *parallel to* u if $\pi x \subset \pi u$. For any $u \in \Phi^d$ and $x \in F^d$, we define $\pi_u x = u \cap \pi x$. Writing $\sigma_u x$ for the set of points in u closest to x, we note that $\sigma_u x \in F^{(u)}$ and $\pi_u = \pi \circ \sigma_u$. For fixed u, we may parametrize the flats $x \in F^d$ by the pairs $(\sigma_u x, \kappa_u x)$, where $\kappa_u x$ records the orthogonal complement of $\pi_u x$ in x, along with the perpendicular, directed distance from σ_u to x. This allows us to regard flats in R^d as marked flats in u of a possibly lower dimension.

Random measures on F^d are assumed to be a.s. locally finite, in the sense of giving finite mass to the set of flats intersecting an arbitrary bounded Borel set in R^d. Simple point processes on F^d are called *flat processes*. Stationarity or invariance are defined with respect to the group of translations of the underlying space R^d, unless otherwise specified. For any $u \in \Phi^d$, we may also define *u-stationarity* and *u-invariance* as stationarity and invariance with respect to the subgroup of translations in directions contained in u only.

We say that a measure μ on Φ^d has no *inner degeneracies*, if it gives zero mass to every set of flats containing a fixed line, and no *outer degeneracies*, if no mass is assigned to any set of flats contained in a fixed hyper-plane. In the former case, we do allow mass ≤ 1 when μ is a point measure. For measures μ on F^d, the same conditions are required for the projections $\mu \circ \pi^{-1}$.

First we show how a partial invariance property can sometimes be extended to a larger space of directions.

Lemma 11.9 *(extended invariance) For fixed $u, v \in \Phi$, let $\mu \in \mathcal{M}_F$ be u-invariant with*

$$\operatorname{span}\{\pi x, u\} \supset v, \quad x \in F \ \text{a.e. } \mu.$$

Then μ is even $\{u, v\}$-invariant.

Proof: Applying Theorem 7.5 to a suitable parametrization, we get $\mu = \nu \otimes \rho$, for some measure $\nu \sim \mu \circ \pi^{-1}$ and u-invariant kernel $\rho \colon \Phi \to F$, such that each measure ρ_p is supported by the set $\pi^{-1}\{p\}$. Since ρ_p is trivially p-invariant, it is even (u, p)-invariant. Since also $v \subset \operatorname{span}\{u, p\}$ for $p \in \Phi$ a.e. ν, we conclude that ρ_p is a.e. v-invariant, hence even (u, v)-invariant. The latter property carries over to the mixture μ. $\square$

We may now extend Corollary 11.2 to arbitrary dimensions. Recall that $\pi_u x = \pi x \cap u$.

Theorem 11.10 *(spanning criterion) For fixed $u, v \in \Phi$, let η be a u-stationary random measure on F_k, such that $E(\eta \mid \mathcal{I}_u)$ is a.s. locally finite and*

(i) $\operatorname{span}\{\pi x, u\} \supset v, \quad x \in F_k$ *a.e.* $E\eta$,

(ii) $\operatorname{span}\{\pi_u x, \pi_u y\} = u, \quad (x, y) \in F_k^2$ *a.e.* $E\eta^2$.

Then η is a.s. v-invariant.

Writing $h = \dim u$, we note that typically

$$\dim(\pi_u x) = (h + k - d)_+.$$

But then (ii) requires $2(h + k - d) \geq h$, or

$$2(d - k) \leq h \leq d,$$

which is possible only when $k \geq d/2$.

Proof: Since $E\eta B = 0$ implies $\eta B = 0$ a.s., it suffices, by (i) and Lemma 11.9, to prove that η is a.s. u-invariant. Next, we may identify the flats $x \in F_k$ with appropriate pairs $(\sigma_u x, \kappa_u x)$, and recall that $\pi_u x = \pi \circ \sigma_u x$. We may then regard η as a stationary random measure, on the space $F^{(u)}$ of flats in u, equipped with marks in a suitable Borel space K. Since the a.s. invariance of η is equivalent to a.s. invariance of $\eta(\cdot \times B)$, for every $B \in \hat{\mathcal{K}}$, we may omit the marks, and consider random measures on $F^{(u)}$ only. This reduces the proof to the case of $u = \mathsf{R}^d$.

For any measurable function $f \geq 0$ on F, consider the restrictions $f_p = f(p, \cdot)$ to flats of direction $p \in \Phi$. If $\operatorname{span}\{p, q\} = \mathsf{R}^d$, we may choose some linearly independent vectors $p_1, \ldots, p_h \in p$ and $q_1, \ldots, q_k \in q$ with $h + k = d$. Proceeding as in (1), we have

$$\begin{aligned}
&\iint_{I_a^d} f_p(sp + tq)\, f_q(sp + tq)\, ds\, dt \\
&\qquad = a^{-d} \iint_{I_a^d} f_p(sp + tq)\, ds\, dt \iint_{I_a^d} f_q(sp + tq)\, ds\, dt,
\end{aligned}$$

with integrations over $s \in I_a^h$ and $t \in I_a^k$. For fixed $\omega \in \Omega$, we get by disintegration

$$\begin{aligned}
&\iint_{I_a^d} \eta(B + sp_0 + tq_0)\, \eta(C + sp_0 + tq_0)\, ds\, dt \\
&\qquad \leq a^{-d}\, \theta \iint_{\theta I_a^d} \eta(B + sp_0 + tq_0)\, ds\, dt \iint_{\theta I_a^d} \eta(C + sp_0 + tq_0)\, ds\, dt,
\end{aligned}$$

along with a similar estimate in the reverse direction. Here $B, C \subset A \times \mathsf{R}^d$, where A denotes an arbitrary ε-neighborhood of (p_0, q_0) in Φ^2, such that $\operatorname{span}\{p, q\} = \mathsf{R}^d$ for all $(p, q) \in A$. Using the ergodic theorem, and letting $\varepsilon \to 0$, we get as before $E(\eta^2 | \mathcal{I}_\eta) = \{E(\eta \,|\, \mathcal{I}_\eta)\}^2$, a.s. on the set of pairs $(x, y) \in F^2$ with $\operatorname{span}\{\pi x, \pi y\} = \mathsf{R}^d$. This extends to all of F^2, since the complementary set has a.s. η^2-measure 0. Since $E(\eta \,|\, \mathcal{I}_\eta)$ is a.s. locally finite, we may take differences, and see as before that $\eta = E(\eta \,|\, \mathcal{I}_\eta)$ a.s., which implies the asserted invariance. $\square$

Similar arguments apply to some higher order moment measures. To illustrate the idea, we consider the case of hyper-plane processes ξ on R^d.

Corollary 11.11 *(moments of hyper-plane processes) Let ξ be a stationary hyper-plane process in R^d, such that a.s. $E(\xi | \mathcal{I}_\xi)$ is locally finite, and ξ has no inner degeneracies. Then*

$$E\Big(\xi^{(m)} | \mathcal{I}_\xi\Big) = \Big\{E(\xi | \mathcal{I}_\xi)\Big\}^m \ \text{ a.s. }, \quad 1 \leq m \leq d.$$

In particular, there exists a Cox process with the same moment measures of orders $\leq d$.

Proof: The argument for Theorem 11.10 yields a.s.

$$E\Big(\xi^{(m+1)}\Big|\,\mathcal{I}_\xi\Big) = E\Big(\xi^{(m)}\Big|\,\mathcal{I}_\xi\Big) \otimes E(\xi|\,\mathcal{I}_\xi), \quad 1 \le m < d,$$

provided that even $E(\xi^{(m)}|\,\mathcal{I}_\xi)$ is a.s. locally finite. Here the non-degeneracy condition ensures the existence of an appropriate set of linearly independent vectors. The asserted formula now follows by induction in finitely many steps. By Theorem 6.27, a Cox process directed by $\eta = E(\xi|\,\mathcal{I}_\xi)$ has the same conditional moments. □

11.4 Invariance under Absolute Continuity

Here we consider the case of arbitrary dimensions $1 \le k < d$, where d is fixed. In Theorem 11.10, we saw that, under a suitable spanning condition, a stationary random measure on F_k is a.s. invariant. The same implication will now be established, under a suitable condition of absolute continuity. Let λ_k denote the unique probability measure on Φ_k that is invariant under arbitrary rotations, which exists by Theorem 7.3. By *u-stationarity* we mean stationarity under arbitrary translations in u, and similarly for the notion of *u-invariance.*

Theorem 11.12 *(invariance criterion) For any $k \in [1, d)$ and $u \in \Phi_{d-k+1}$, let η be a u-stationary random measure on F_k with $\eta \circ \pi^{-1} \ll \lambda_k$. Then η is a.s. invariant.*

Some more general invariance criteria are given in Theorems 11.16 and 11.27 below. The stated theorem will follow from the following version for stationary processes.

Lemma 11.13 *(invariant processes) For any $k \in [1, d)$ and $u \in \Phi^d_{d-k+1}$, let X be a u-stationary, product-measurable process on F^d_k. Then*

$$X_a = X_b \;\; a.s., \quad a, b \in \pi^{-1}t, \;\; t \in \Phi_k \;\; a.e. \;\; \lambda_k.$$

Proof: For every $t = t_0 \in \Phi_k$, we may choose some ortho-normal vectors $x_1, \dots, x_{d-k} \in \mathsf{R}^d$, such that

$$\operatorname{span}\{x_1, \dots, x_{d-k}, t\} = \mathsf{R}^d, \qquad x_1, \dots, x_{d-k} \not\perp t. \tag{4}$$

For fixed $x_1, \dots, x_{d-k}$, the set G of flats $t \in \Phi_k$ satisfying (4) is clearly open with full λ_k-measure, containing in particular a compact neighborhood C of t_0. The compact space Φ_k is covered by finitely many such sets C, with possibly different $x_1, \dots, x_{d-k}$, and it is enough to consider the restriction of X to any one of the sets $\pi^{-1}C$. Equivalently, we may take X to be supported by $\pi^{-1}C$ for a fixed C.

Fixing $x_1, \dots, x_{d-k}$ accordingly, and letting $a, b \in \pi^{-1}t$ with $t \in C$, we have $a = b + x$ for some linear combination x of $x_1, \dots, x_{d-k}$, and by iteration we may assume that $x = rx_i$ for some $r \in \mathsf{R}$ and $i \le d-k$. By symmetry, we may take $i = 1$, and we may also choose $b = t$, so that $a = t + rx_1$.

We may now parametrize the flats $t \in G$ as follows, relative to the vector x_1. Writing $\hat{x}_1$ for the projection of x_1 onto t, we put $p = |\hat{x}_1| \in (0,1)$. By continuity of the map $t \mapsto p$, we note that p is bounded away from 0 and 1, on the compact set C. Next, we introduce the unique unit vector $y \in \operatorname{span}\{x_1, \hat{x}_1\}$, satisfying $y \perp x_1$ and $y \cdot \hat{x}_1 > 0$. Finally, we take $s \in \Phi_{k-1}$ to be the orthogonal complement of $\hat{x}_1$ in t. The map $t \mapsto (p, y, s)$ is $1-1$ and measurable on G, and by symmetry we note that p, y, and s are independent under λ_k.

The image $\hat{\lambda}_k = \lambda_k \circ p^{-1}$ is absolutely continuous on $[0,1]$ with a positive density, and so for any $\varepsilon \in (0, \varepsilon_0]$ with $\varepsilon_0 > 0$ fixed, there exists a continuous function $g_\varepsilon : C \to \Phi_k$, mapping $t = (p, y, s)$ into some $t' = (p', y, s)$ with $p' > p$ and $\hat{\lambda}_k(p, p') = \varepsilon$. By independence[5],

$$1_C \lambda_k \circ g_\varepsilon^{-1} = 1_{g_\varepsilon C} \lambda_k, \quad \varepsilon \in (0, \varepsilon_0). \tag{5}$$

Write p_0 for the maximum of p on C, and define $B_0 = \{(p, y, s);\, p \le p_0\}$. By a monotone-class argument, any function $f \in L^2(\mu)$ supported by B_0 can be approximated in L^2 by some continuous functions f_n with the same support (cf. FMP 1.35). Combining with (5), we get as $n \to \infty$ for fixed $\varepsilon \in (0, \varepsilon_0)$

$$\begin{aligned} \left\| f \circ g_\varepsilon - f_n \circ g_\varepsilon \right\|_2^2 &= (\mu \circ g_\varepsilon^{-1})(f - f_n)^2 \\ &= \|f - f_n\|_2^2 \to 0. \end{aligned}$$

Furthermore, we get by continuity and dominated convergence

$$\lim_{\varepsilon \to 0} \|f_n - f_n \circ g_\varepsilon\|_2 \to 0, \quad n \in \mathsf{N}.$$

Using Minkowski's inequality, and letting $\varepsilon \to 0$ and then $n \to \infty$, we obtain

$$\|f - f \circ g_\varepsilon\|_2 \le \|f_n - f_n \circ g_\varepsilon\|_2 + 2\|f - f_n\|_2 \to 0. \tag{6}$$

Truncating, if necessary, we may assume that X is bounded. Since it is also product-measurable, it is measurable on M_k for fixed $\omega \in \Omega$ (cf. FMP 1.26), and so by (6),

$$\lim_{\varepsilon \to 0} \int \mu(dt) \left\{ X(t) - X \circ g_\varepsilon(t) \right\}^2 = 0.$$

Hence, by Fubini's theorem and dominated convergence,

$$\begin{aligned} &\int \mu(dt)\, E\left\{ X(t) - X \circ g_\varepsilon(t) \right\}^2 \\ &\qquad = E \int \mu(dt) \left\{ X(t) - X \circ g_\varepsilon(t) \right\}^2 \to 0, \end{aligned}$$

[5] Recall that $1_B \mu = 1_B \cdot \mu$, where the measure $f \cdot \mu$ is given by $(f \cdot \mu) g = \mu(fg)$.

and so we may choose some $\varepsilon_n \to 0$ satisfying

$$\left\|X(t) - X \circ g_{\varepsilon_n}(t)\right\|_2 \to 0, \quad t \in C \text{ a.e. } \mu, \tag{7}$$

say for $t \in C'$, where the displayed norm is in $L^2(P)$.

For any $t \in C$ and $r \in \mathsf{R}$, the flats $g_\varepsilon(t)$ and $a = t + rx_1$ are non-parallel and lie in the $(k+1)$-dimensional span of x_1, y, and s, and so their intersection is non-empty. Thus, the flat $t_n = g_{\varepsilon_n}(t)$ intersects both t and a, and the flats of intersection are parallel to s.

Now choose recursively some unit vectors $s_1, \dots, s_{k-1}$, each uniformly distributed in the orthogonal complement of the preceding ones. The generated subspace s' has the same distribution as s under μ, and s' is further a.s. linearly independent of u, which implies

$$\operatorname{span}\{u, s\} = \mathsf{R}^d, \quad t \in C' \text{ a.e. } \mu,$$

say for $t \in C''$. For such a t, every flat in $\pi^{-1}s$ equals $s + x$ for some $x \in u$, and so by (7) and the u-stationarity of X,

$$\begin{aligned}\left\|X(a) - X(t_n)\right\|_2 &= \left\|X(t) - X(t_n)\right\|_2 \\ &= \left\|X(t) - X \circ g_{\varepsilon_n}(t)\right\|_2 \to 0.\end{aligned}$$

Hence, by Minkowski's inequality,

$$\left\|X(a) - X(t)\right\|_2 \le \left\|X(a) - X(t_n)\right\|_2 + \left\|X(t_n) - X(t)\right\|_2 \to 0,$$

and so $X(a) = X(t)$ a.s., as long as $t \in C''$. $\square$

Proof of Theorem 11.12: Any flat $x \in F_k$ can be written uniquely as $x = \pi(x) + r(x)$, where $\pi(x) \in \Phi_k$ and $r(x) \in \pi(x)^\perp$—the orthogonal complement of $\pi(x)$ in R^d. We may then think of the pair $(\pi(x), r(x))$ as a parametrization of x.

Now consider a random flat $\varphi \in \Phi_{d-k}$ with distribution λ_{d-k}. Since φ can be generated by some ortho-normal vectors $\alpha_1, \dots, \alpha_{d-k}$, each of which is uniformly distributed in the orthogonal complement of the preceding ones, we have a.s.

$$\begin{aligned}\dim \operatorname{span}\{u^\perp, \varphi\} &= \dim u^\perp + \dim \varphi \\ &= (k-1) + (d-k) = d-1,\end{aligned}$$

and so the intersection $u^\perp \cap \varphi$ has a.s. dimension 0. If a vector $y \in \varphi$ is not in the φ-projection of u, then $x \perp y$ for every $x \in u$, which means that $y \in u^\perp$. Assuming $y \in u^\perp \cap \varphi = \{0\}$, we obtain $y = 0$, which shows that the φ-projection of u agrees a.s. with φ itself. In particular, any measure $\nu \ll \lambda^{d-k+1}$ on u has a φ-projection ν' satisfying $\nu' \ll \lambda^{d-k}$ a.s. on φ.

Let ν_ε denote the uniform distribution on the ε-ball in u around 0. For any $r \in \Phi_k$, let ν_ε^r be the orthogonal projection of ν_ε onto the subspace

$r^\perp \in \Phi_{d-k}$, so that $\nu_\varepsilon^r \ll \lambda^{d-k}$ on Φ_{d-k}, for $r \in \Phi_k$ a.e. λ_k. Since $\eta \circ \pi^{-1} \ll \mu$ a.s., η has a disintegration $\int \lambda_k(dr)\, \eta_r$ a.s., and so

$$\eta * \nu_\varepsilon = \int \lambda_k(dr)(\eta_r * \nu_\varepsilon^r) \text{ a.s.,} \quad \varepsilon > 0,$$

where the convolutions on the left are in R^d, while those on the right are in $\pi^{-1}r$ for each $r \in \Phi_k$. The u-stationarity of η carries over to $\eta * \nu_\varepsilon$, whereas the absolute continuity of ν_ε^r carries over to $\eta_r * \nu_\varepsilon^r$, and then also to the mixture $\eta * \nu_\varepsilon$.

For every $\varepsilon > 0$, Corollary 7.36 yields a u-stationary, product-measurable process $X^\varepsilon \geq 0$ on F_k, such that $\eta * \nu_\varepsilon = X^\varepsilon \cdot (\lambda_k \otimes \lambda^{d-k})$, in a suitable parameter space $\Phi_k \times \mathsf{R}^{d-k}$. By Lemma 11.13,

$$X_a^\varepsilon = X_b^\varepsilon \text{ a.s.,} \quad a, b \in \pi^{-1}r, \;\; r \in \Phi_k \text{ a.e. } \lambda_k.$$

Since $\eta \circ \pi^{-1} \ll \lambda_k$, we may extend this to all $r \in \Phi_k$, by redefining X to be 0 on the exceptional λ_k-set, so that $X_a = X_b$ a.s. whenever $\pi a = \pi b$.

Fix any bounded, measurable sets $I, J \in L_k$ with $J = I + h$, for some translation $h \in \mathsf{R}^{d-k}$ in the parameter space. Then

$$\begin{aligned} E\Big| (\eta * \nu_\varepsilon)I - (\eta * \nu_\varepsilon)J \Big| &= E\left| \int_I (X_s - X_{s+h})\, ds \right| \\ &\leq E \int_I |X_s - X_{s+h}|\, ds \\ &= \int_I E|X_s - X_{s+h}|\, ds = 0, \end{aligned}$$

where all integrations are with respect to $\lambda_k \otimes \lambda^{d-k}$, and so $(\eta * \nu_\varepsilon)I = (\eta * \nu_\varepsilon)J$ a.s. Since I was arbitrary, Lemma 2.1 shows that $\eta * \nu_\varepsilon$ is a.s. invariant under any fixed shift h, and since $\eta * \nu_\varepsilon \xrightarrow{v} \eta$ as $\varepsilon \to 0$, even η itself is a.s. h-invariant for fixed h. The a.s. invariance then holds simultaneously for any dense sequence of shifts $h_1, h_2, \ldots$, and since $\eta * \delta_h$ is vaguely continuous in h, it extends to arbitrary h, for all $\omega \in \Omega$ outside a fixed P-null set. □

11.5 Non-Interactive Particle Systems

A particle moving with constant velocity through space generates a straight line in a space-time diagram. In this way, any random collection of non-interacting particles in R^d is equivalent to a line process in R^{d+1}, and either description illuminates the other.

In a similar way, any random measure η on the $2d$-dimensional space of positions and velocities may be identified with a random measure on the space of lines in R^{d+1}. Assuming stationarity in the initial space variables, we may study the asymptotic behavior of the generated line process in the space-time diagram. We show that a suitable local invariance of the velocity distribution implies asymptotic invariance in the space variables. Though the particle description yields a useful parametrization, we prefer to state all results in terms of line processes.

Theorem 11.14 *(asymptotic L^p-invariance) Fix a closed subgroup $G \subset \mathbb{R}^d$ spanning $u \in \Phi^d_{d-1}$, a subspace $v \subset u$, a unit vector $y \perp u$, and a constant $p \geq 1$. Let η be a G-stationary random measure on F^d_1, such that $1_B \eta \circ \pi^{-1}$ is locally v-invariant in L^p for bounded $B \in \mathcal{B}^{d-1}$. Then $\theta_{ty}\eta$ is asymptotically v-invariant in L^p as $t \to \infty$.*

Proof: First let $G = u$ and $\eta(\pi^{-1}u) = 0$ a.s. Fix a probability measure $\nu = g \cdot \lambda^{d-1}$ on u, and write for any $I, B \in \mathcal{B}^{d-1}$

$$\begin{aligned}(\nu * \eta)(I \times B) &= \int_I ds \int g(s-x)\, \eta(dx \times B) \\ &= \int_I \eta_s B\, ds. \qquad (8)\end{aligned}$$

Shifting a measure on F^d_1 by ty corresponds to a mapping of the associated parameter pair (s, r) into $(s + tr, r)$. Fixing any $h \in v$, and writing $\varepsilon = 1/t$, we get by (8), Minkowski's inequality, and the u-stationarity of η

$$\begin{aligned}&\Big\| \Big\{ (\delta_{ty} - \delta_{ty+h}) * \nu * \eta \Big\} (I \times B) \Big\|_p \\ &\quad = \Big\| \int \Big\{ 1_B \eta_s(\varepsilon I - \varepsilon s) - 1_B \eta_s(\varepsilon I - \varepsilon h - \varepsilon s) \Big\} ds \Big\|_p \\ &\quad \leq \int \Big\| 1_B \eta_s(\varepsilon I - \varepsilon s) - 1_B \eta_s(\varepsilon I - \varepsilon h - \varepsilon s) \Big\|_p ds \\ &\quad = \int \Big\| 1_B \eta_0(\varepsilon I - \varepsilon s) - 1_B \eta_0(\varepsilon I - \varepsilon h - \varepsilon s) \Big\|_p ds.\end{aligned}$$

Since $1_B \eta_0$ is locally v-invariant in L^p, by Corollary 5.48, the right-hand side tends to 0 as $t \to \infty$, at least for rectangular I. A similar argument shows that $\|(\delta_{ty} * \nu * \eta)(I \times B)\|_p$ remains bounded as $t \to \infty$. Hence, $\delta_{-ty} * \nu * \eta$ is asymptotically v-invariant in L^p, and so is $\theta_{ty}\eta = \delta_{ty} * \eta$ by Lemma 5.22, since ν was arbitrary.

For a closed subgroup G spanning u, the quotient group u/G is compact, and hence admits a normalized Haar measure ν. Introducing a random vector $\gamma \perp\!\!\!\perp \eta$ with distribution ν, we note that $\theta_\gamma \eta$ is u-stationary. Indeed, we may assume that $G = \mathbb{Z}^m \times \mathbb{R}^n$ with $m + n = d - 1$, and choose $\gamma = (\gamma_1, \ldots, \gamma_m, 0 \ldots, 0)$, where the γ_j are i.i.d. $U(0,1)$, in which case our claim follows from the invariance of $U(0,1)$, under shifts modulo 1. By Corollary 5.48, the local L^p-invariance of η carries over to $\theta_\gamma \eta$. Using the result for $G = u$, we conclude that $\theta_{\gamma + ty}\eta$ is asymptotically v-invariant in L^p, and so by Lemma 5.21, the same property holds for the random measures $\theta_{-ty}\eta$.

Now assume instead that $\eta(\pi^{-1}u^c) = 0$ a.s. For any bounded set $B \in \mathcal{B}^1$, we may apply the previous version to the projection $\eta_B = 1_{u \times B}\eta \circ \pi^{-1}$, which is then asymptotically v-invariant in L^p, under u-shifts perpendicular to v.

By the u-stationarity of η, we conclude that η_B itself is a.s. v-invariant. Applying this to a measure-determining sequence of sets B, we see that the stated invariance holds a.s. for η. For general η, we may apply the previous results to the restrictions of η to the u-invariant subsets $\pi^{-1}u$ and $\pi^{-1}u^c$, both of which satisfy the hypotheses of the theorem. $\square$

Under slightly stronger conditions, we can prove the corresponding convergence. Write $\bar\eta_u$ for the sample intensity of η, under averages in u. Recall from Theorem 5.47, with succeeding remark, that a random measure ζ on R^d is *strongly locally L^p-invariant*, if it is locally L^p-invariant in the weak sense, and satisfies $\zeta \ll \mu$ a.s. for some fixed measure μ. This condition is equivalent to local L^p-invariance when $p > 1$, but for $p = 1$ it is strictly stronger.

Theorem 11.15 *(convergence to invariant limit) Fix a closed subgroup $G \subset \mathsf{R}^d$ spanning $u \in \Phi_{d-1}$, a unit vector $y \perp u$, and a constant $p \geq 1$. Let η be a G-stationary random measure on F_1^d, such that $1_B \eta \circ \pi^{-1}$ is strongly locally u-invariant in L^p, for bounded $B \in \mathcal{B}^d$. Then*

$$\theta_{ty}\eta \xrightarrow{v} \bar\eta_u \ \textit{ in } L^p \ \textit{ as } t \to \infty.$$

Proof: As before, we may reduce to the case where $G = u$. By Theorems 2.12 and 5.47, we have $\eta \circ \pi^{-1} \ll E\eta \circ \pi^{-1}$ a.s. Replacing η by $\eta * \nu$ for some $\nu \ll \lambda^{d-1}$, if necessary, we may assume that in fact $\eta \ll E\eta$ a.s. By Corollary 7.36, we may then choose a stationary and product-measurable density Y of η, so that $\eta = Y \cdot E\eta$ a.s. By Theorem 2.13 and Corollary 5.48, we may assume that $\|Y\|_p < \infty$. We may also write $E\eta = \lambda^{d-1} \otimes \mu$, where μ is a locally invariant measure on Φ_1.

For any $h > 0$, let ν_h denote the uniform distribution on the cube $[0, h]^{d-1}$. For any bounded sets $B, C \in \mathcal{B}^{d-1}$, we get by Fubini's theorem, Minkowski's inequality, and the stationarity of Y

$$\begin{aligned}
&\Big\| \Big\{ \delta_{ty} * \nu_h * \eta - \bar\eta_u \Big\} (B \times C) \Big\|_p \\
&\quad = \Big\| \int_C \mu(dr) \int_{B+rt} \Big\{ (\nu_h * Y)(x, r) - \bar Y(r) \Big\}\, dx \Big\|_p \\
&\quad = \Big\| \int_C \mu(dr) \int_B \Big\{ (\nu_h * Y)(x + rt, r) - \bar Y(r) \Big\}\, dx \Big\|_p \\
&\quad \leq \int_C \mu(dr) \int_B \Big\| (\nu_h * Y)(x + rt, r) - \bar Y(r) \Big\|_p\, dx \\
&\quad = \int_C \mu(dr) \int_B \Big\| (\nu_h * Y)(x, r) - \bar Y(r) \Big\|_p\, dx,
\end{aligned}$$

which tends to 0 as $h \to \infty$, by the multivariate L^p-ergodic Theorem A2.4 and dominated convergence, since the integrand is bounded by $2\|Y(x, r)\|_p$. Noting that the right-hand side is independent of t, we get

$$\lim_{h \to \infty} \sup_{t \geq 0} \Big\| (\delta_{ty} * \nu_h * \eta - \bar\eta_u) A \Big\|_p = 0, \quad A = B \times C \in \hat{\mathcal{B}}^{2(d-1)}.$$

Next, we see from Theorem 11.14 that $\delta_{ty} * \eta$ is asymptotically u-invariant in L^p as $t \to \infty$. Arguing as in Lemma 5.20, we conclude that

$$\lim_{t\to\infty} \left\| (\delta_{ty} * (\nu_{nh} - \nu_h) * \eta) A \right\|_p = 0, \quad h > 0,\ n \in \mathsf{N}.$$

By Minkowski's inequality, we may combine the previous L^p-norms into

$$\begin{aligned} &\left\| (\delta_{ty} * \nu_h * \eta - \bar{\eta}_u) A \right\|_p \\ &\qquad \le \left\| (\delta_{ty} * (\nu_{nh} - \nu_h) * \eta) A \right\|_p + \left\| (\delta_{ty} * \nu_{nh} * \eta - \bar{\eta}_u) A \right\|_p, \end{aligned}$$

which tends to 0 as $t \to \infty$, and then $n \to \infty$. Since h was arbitrary, it follows that $\delta_{ty} * \eta \to \bar{\eta}_u$, vaguely in L^p, as asserted. □

The previous results yield a joint extension of the invariance criteria in Theorems 11.10 and 11.12.

Theorem 11.16 *(stationarity and invariance) Fix a closed subgroup G spanning R^d, and a subspace v. Let η be a G-stationary random measure on F_1^d, such that $(1_B \eta) \circ \pi^{-1}$ is locally v-invariant in L^1, for bounded $B \in \mathcal{B}^d$. Then η is a.s. v-invariant, and it is even a.s. R^d-invariant, whenever*

$$\operatorname{span}\{v, \pi x, \pi y\} = \mathsf{R}^d, \quad (x, y) \in (F_1^d)^2 \ \textit{a.e. } E\eta^2. \tag{9}$$

Proof: As before, we may reduce to the case of $G = \mathsf{R}^d$. The first assertion is then an easy consequence of Theorem 11.14 and its proof (especially the last paragraph). It remains to prove the second assertion for $\dim(v) < d$. Since (9) can only hold when $\dim(v) \ge d - 2$, we may assume that either $v \in \Phi_{d-1}$ or $v \in \Phi_{d-2}$. In the former case, (9) yields

$$\operatorname{span}\{v, \pi x\} = \mathsf{R}^d, \quad x \in M_1^d \ \text{a.e. } E\eta,$$

which implies the asserted R^d-invariance of η, by Lemma 11.9. It remains to take $v \in \Phi_{d-2}$.

Then let (x, r) be the phase representation in F_1^d, based on a flat $u \in \Phi_{d-1}$ with $u \supset v$. Note that any such u is permissible, since $\pi x \notin u$ for $x \in F_1^d$ a.e. $E\eta$, by (9). Write $x = (x', x'')$ and $r = (r', r'')$, where x'' and r'' are the orthogonal projections of x and r onto v. By the first assertion, η is a.s. invariant in $x'' \in \mathsf{R}^{d-2}$, so that $\eta = \zeta \otimes \lambda^{d-2}$ a.s., for some random measure ζ on R^d, corresponding to the remaining coordinates (x', r). We need to prove that ζ is a.s. invariant under shifts in $x' \in \mathsf{R}$, which will imply the required a.s. invariance of η in $x \in \mathsf{R}^{d-1}$.

Then regard the pairs (x', r') as a parametrization of F_1^2, with marks $x'' \in \mathsf{R}^{d-2}$ attached to the lines, so that ζ becomes a random measure on the space of marked lines in R^2. To see that ζ is stationary, write ty for a perpendicular shift in R^2, consider any measurable function $f \geq 0$ on R^d, and put $B = [0,1]^{d-2}$. Using Fubini's theorem and the stationarity of η, we get

$$\begin{aligned}
(\delta_{ty} * \zeta)f &= \int f(x' - r't, r)\,\zeta(dx'\,dr) \\
&= \int f(x' - r't, r)\,\zeta(dx'\,dr) \int 1_B(x'' - r''t)\,dx'' \\
&= \iint f(x' - r't, r)\,1_B(x'' - r''t)\,\eta(dx\,dr) \\
&= (\delta_{ty} * \eta)(f \otimes 1_B) \overset{d}{=} \eta(f \otimes 1_B) \\
&= \zeta f \cdot \lambda^{d-2} B = \zeta f,
\end{aligned}$$

as desired.

By Corollary 11.2, it remains to prove that

$$\operatorname{span}\{\pi x_1, \pi x_2\} = \mathsf{R}^2, \quad (x_1, x_2) \in (\mathsf{R}^2 \times \mathsf{R}^{d-2})^2 \ \text{ a.e. } E\zeta^2,$$

which holds iff the r'-projection of ζ is a.s. diffuse. If it is not, then with positive probability, we may choose an $\alpha \in \mathsf{R}$ satisfying $\zeta\pi^{-1}\{\alpha\} > 0$, which implies

$$\eta^2\Big\{(x_1, r_1, x_2, r_2);\ r_1' = r_2' = \alpha\Big\} > 0.$$

Now the relations $r_1' = r_2' = \alpha$ yield

$$\begin{aligned}
\operatorname{span}\Big\{v, \pi(x_1, r_1), \pi(x_2, r_2)\Big\} &= \operatorname{span}\Big\{v, (x_1, 1), (x_2, 1)\Big\} \\
&= \operatorname{span}\Big\{v, (\alpha, x_1'', 1), (\alpha, x_2'', 1)\Big\} \\
&= \operatorname{span}\Big\{v, (\alpha, 0, 1)\Big\} \neq \mathsf{R}^d,
\end{aligned}$$

and so

$$E\,\eta^2\Big\{(x_1, x_2);\ \operatorname{span}\{v, \pi x_1, \pi x_2\} \neq \mathsf{R}^d\Big\} > 0,$$

contradicting (9). □

Turning to the asymptotic behavior of line processes with associated particle systems, we begin some classical results.

Proposition 11.17 *(independent velocities, Breiman, Stone) Let ξ be a stationary point process ξ on R^d with $\bar{\xi} < \infty$ a.s., and form ξ_t for $t > 0$, by attaching independent velocities to the points of ξ, with a common distribution μ. Writing ζ for a Cox process directed by $\bar{\xi}\lambda^d$, we have as $t \to \infty$*

(i) $\xi_t \xrightarrow{vd} \zeta$, *when μ is locally invariant,*

(ii) $\xi_t \overset{uld}{\to} \zeta$, *when $\mu \ll \lambda^d$.*

In general, the ξ_t are vaguely tight, and if μ is diffuse, then all limiting processes are Cox.

Proof: Choosing a γ with distribution μ, let μ_t be the distribution of $t\gamma$, and note that the μ_t are weakly asymptotically invariant, iff μ is locally invariant. From Lemma 5.44, we see that μ_t is strictly asymptotically invariant iff $\mu \ll \lambda^d$. Claims (i) and (ii) now follow by Theorem 5.27.

For general μ, all processes ξ_t have clearly the same sample intensity $\bar{\xi}$, and so the stated tightness follows by Theorem 4.10 and Chebyshev's inequality. If μ is diffuse, a simple compactness argument shows that the μ_t are dissipative as $t \to \infty$, and so the distributional limits are Cox by Theorem 4.40. □

The previous independence assumption may be too restrictive for most purposes, for the following reason.

Proposition 11.18 *(independence dichotomy) Let ξ be a stationary point process on R^d, and form ξ_t for $t > 0$, by attaching independent velocities to the points of ξ, with a common non-lattice distribution μ. Then exactly one of these cases will occur:*

(i) *The independence fails for ξ_t, at every time $t > 0$.*

(ii) *The ξ_t form a space-time stationary Cox line process in R^{d+1}, with a.s. invariant directing random measure, and the independence is preserved at all times $t \geq 0$.*

Proof: Suppose that (i) fails, so that the independence holds at some time $t > 0$. Then Theorem 5.29 shows that ξ is Cox and directed by $\eta = \rho\, e_{tr} \cdot \lambda^d$, for some random variable $\rho \geq 0$ and constant $r \in \mathsf{R}^d$, where $e_r(s) = e^{rs}$. Since ξ is stationary, we have $r = 0$, so in fact $\eta = \rho\,\lambda^d$. By Theorem 3.2, the process of positions and velocities is again Cox and directed by $\rho\,\lambda^d \otimes \mu$, which is clearly invariant under shifts in both space and time. Hence, the generated line process satisfies (ii). □

To avoid imposing any restrictive independence assumptions on ξ, we may apply the previous invariance criteria to the associated Papangelou kernel η, which leads to some more general conditions for Cox convergence.

Theorem 11.19 *(Cox convergence) Fix a plane $u \in \Phi^d_{d-1}$, and a unit vector $y \perp u$. Let ξ be a u-stationary line process in R^d satisfying (Σ), and such that the associated Papangelou kernel η satisfies $\theta_{ty}\eta \xrightarrow{v} \bar{\eta}_u$ in L^1, as $t \to \infty$. Then $\theta_{ty}\xi \xrightarrow{vd} \zeta$, where ζ is a Cox process directed by $\bar{\eta}_u$.*

Proof: By Theorem 8.20 (ii), it is enough to prove that $\bar{\eta}_u$ is $\mathcal{R}_\xi$-measurable, which clearly implies measurability on $\mathcal{R}(\theta_{ty}\xi)$, for every $t > 0$. Thus, we need to show that $f(\bar{\eta}_u)$ is $\mathcal{R}_\xi$-measurable, for any bounded, measurable function $f \geq 0$. Here $f(\bar{\eta}_u) = g(\xi)$ a.s., for some bounded, measurable function $g \geq 0$. Using the stationarity of (ξ, η) and invariance of $\bar{\eta}_u$, we get

$$f(\bar{\eta}_u) = f \circ (\overline{\theta_h\eta})_u = g(\theta_h\xi) \text{ a.s.,} \quad h \in u,$$

and so the multi-variate ergodic Theorem A2.4 yields

$$f(\bar{\eta}_u) = \overline{g(\theta_h \xi)} = E\Big\{ g(\xi) \,|\, \mathcal{I}_\xi \Big\} \text{ a.s.}$$

It remains to note that $\mathcal{I}_\xi \subset \mathcal{R}_\xi$, by FMP 3.16. □

Conditions involving the Papangelou kernel η are often hard to translate into requirements on the underlying point process ξ. Here we consider a special case where such a translation is possible. The result may be regarded as an extension of Proposition 11.17 to the case of conditional independence.

Theorem 11.20 *(conditionally independent velocities) Let ξ be a stationary point process on R^d, with marks $\mu \in \hat{\mathcal{M}}_d$ satisfying $\mu \ll \lambda^d$ and $\|\mu\| = 1$, where $E(\xi|\,\mathcal{I}_\xi) = \lambda^d \otimes \beta$ for some random measure β on R^d. To the points (s_i, μ_i) of ξ, attach some conditionally independent velocities γ_i with distributions μ_i. Then the generated line processes ζ_t in R^{d+1} satisfy $\zeta_t \xrightarrow{vd} \tilde{\zeta}$ as $t \to \infty$, where $\tilde{\zeta}$ is a Cox process directed by $\lambda^d \otimes \beta$.*

Proof: We may take ξ to be simple, since we may otherwise consider a uniform randomization $\tilde{\xi}$ of ξ on $\mathsf{R}^d \times [0,1] \times \hat{\mathcal{M}}_d$. We may further assume that ξ satisfies condition (Σ), since we may otherwise consider a p-thinning of ξ for some fixed $p \in (0,1)$, and use Lemma 4.17. Finally, we may take $\|\beta\|$ to be a.s. bounded, since we may otherwise consider the restriction of ξ to the invariant set $\{\|\beta\| \le r\}$, for any $r > 0$.

Now form a point process $\tilde{\xi}$, by attaching to each point (s_i, μ_i) of ξ a conditionally independent mark γ_i with distribution μ_i, and note that ζ_0 is equivalent to the projection of $\tilde{\xi}$ onto R^{2d}. Then ξ is a ν-randomization of ξ, where $\nu(s, \mu, \cdot) = \mu$. Writing η for the Papangelou kernel of ξ, we see from Theorem 8.9 that $\tilde{\xi}$ has Papangelou kernel $\eta \otimes \nu$.

The stated absolute continuity implies that, for any bounded, measurable set B,

$$(1_B \eta \otimes \nu) \circ \pi^{-1} \ll \lambda^d \text{ a.s.}$$

By the boundedness of $\|\beta\|$ and Theorem 5.47, the random measures on the left are in fact strongly locally L^1-invariant. Hence, Theorem 11.15 yields $\eta_t \otimes \nu \xrightarrow{v} \bar{\eta} \otimes \nu$ in L^1 as $t \to \infty$, and the asserted Cox convergence $\zeta_t \xrightarrow{vd} \tilde{\zeta}$ follows by Theorem 11.19. □

11.6 Degeneracies of Flat Processes

Degeneracies of measures on F_k^d and Φ_k^d, mentioned briefly in previous sections, become important in the context of more general invariance criteria. Here our key results are the degeneracy decompositions in Lemma 11.22, and the dimension reduction in Lemma 11.23, through the intersection with a fixed hyper-plane. Both results will be useful in the next section. In Theorem 11.24, we also prove a $0-\infty$ law for stationary random measures on F_k^d, similar to the elementary laws of this type in Lemma 5.2.

For any $u \in F^d$, put $\sigma_u x = u \cap x$, and write $\rho_u x$ for the orthogonal projection of x onto u. For $u \in \Phi_{d-1}^d$ and $x \in \Phi_k^d$, let $\tilde{x}$ and $(\tilde{x})_u$ denote the orthogonal complements of x in R^d and u, respectively. The operators σ_u and ρ_u are *dual*, in the following sense:

Lemma 11.21 *(intersection and projection duality) For any $k < d$, we have*

$$(\widetilde{\sigma_u x})_u = \rho_u \tilde{x}, \quad u \in \Phi_{d-1}^d, \ x \in \Phi_k^d.$$

Proof: The claim is obvious when $x \subset u$, and also when $x = \operatorname{span}\{v, y\}$, with $v = u \cap x$ and $y = u^\perp$. It remains to take $x = \operatorname{span}\{v, y\}$ with $v = u \cap x$, and $y \notin u$ with $y \not\perp u$. Then consider the contributions of each side to $\operatorname{span}\{y, u^\perp\}$ and its orthogonal complement. Since both sides are closed linear subspaces of R^d, the general result follows, as we take linear combinations. □

We turn to the degeneracies of measures μ on F_k^d or Φ_k^d. Given a flat $v \in F_m^d$, let $\hat{v}$ and $\check{v}$ denote the sets of flats contained in v, or containing v, respectively. When $m \geq k$, we say that $v \in \Phi_m^d$ is an *outer degeneracy flat* of μ, if $\mu\hat{v} > 0$. Similarly, v is said to be an *inner degeneracy flat* of μ, when $m \leq k$ and $\mu\check{v} > 0$. We say that μ has *no outer degeneracies* if $\mu\hat{v} = 0$ for all $v \in \Phi_{d-1}^d$, and *no inner degeneracies* if $\mu\check{v} = 0$ for all $v \in \Phi_1^d$.

For any $\mu \in \mathcal{M}(\Phi_k^d)$, let $\tilde{\mu} \in \mathcal{M}(\Phi_{d-k}^d)$ be the measure induced by the mapping $x \mapsto \tilde{x}$, and note that μ and $\tilde{\mu}$ are again *dual*, in the sense that $\mu x > 0$ iff $\tilde{\mu}\tilde{x} > 0$. Thus, every outer/inner degeneracy of μ corresponds uniquely to an inner/outer degeneracy of $\tilde{\mu}$. We show how μ can be decomposed uniquely, according to its inner or outer degeneracies.

Lemma 11.22 *(degeneracy decompositions) For fixed $k < d$, every measure $\mu \in \mathcal{M}(\Phi_k^d)$ can be written, uniquely up to the order of terms, in the forms*

$$\mu = \sum_{m=k}^{d} \sum_{i \geq 1} \mu_{mi} = \sum_{n=0}^{k} \sum_{j \geq 1} \nu_{nj},$$

where the μ_{mi} and ν_{nj} are non-zero measures on $\hat{\alpha}_{mi}$ and $\check{\beta}_{nj}$, for some distinct flats $\alpha_{mi} \in \Phi_m^d$ and $\beta_{nj} \in \Phi_n^d$, and the μ_{mi} and ν_{nj} have no outer or inner degeneracies of dimension $< m$ or $> n$, respectively.

We refer to the flats α_{mi} and β_{nj} above as the *minimum outer* and *maximum inner degeneracy flats* of μ.

Proof: Let m be the smallest dimension $\geq k$ of all outer degeneracy flats of μ. For any distinct outer degeneracy flats α_1 and α_2 of dimension m, the intersection $\alpha = \alpha_1 \cap \alpha_2$ satisfies $\mu\hat{\alpha} = 0$, since α would otherwise be an outer degeneracy flat of dimension $< m$. Hence, the outer degeneracy flats of dimension m are a.e. disjoint, and there can be at most countably many of them, say $\alpha_{mi} \in \Phi_m^d$ with $i \geq 1$. The associated restrictions μ_{mi} of μ to $\hat{\alpha}_{mi}$ are clearly mutually singular, and have no outer degeneracies of dimension $< m$. Furthermore, the remainder $\mu' = \mu - \sum_i \mu_{mi}$ has no outer degeneracies of dimension $\leq m$.

Continuing recursively in m, we ultimately end up with a single measure μ_{d1}, with no outer degeneracies of dimension $< d$. The stated uniqueness is obvious from the construction. This completes the proof of the outer degeneracy decomposition. The inner degeneracy decomposition can be established by a similar argument. Alternatively, it follows by duality from the outer version. □

The following result can be used to reduce the dimensions of a proposed statement, by replacing the original flats in R^d by their intersections with a fixed flat $u \in F_{d-1}^d$. Here the phrase "for almost every u" refers to the homogeneous measure λ_{d-1} on F_{d-1}^d or Φ_{d-1}^d.

Lemma 11.23 *(reduction of dimension) For fixed $k, m \in \{1, \ldots, d\}$, let $\mu \in \mathcal{M}(F_k^d)$ with $\mu\hat{v}$ or $\mu\check{v} = 0$ for all $v \in F_m^d$. Writing $\mu_u = \mu \circ \sigma_u^{-1}$, we have for $u \in F_{d-1}^d$ a.e. λ_{d-1}*

(i) *$\sigma_u x \in F_{k-1}^{(u)}$, for $x \in F_k^d$ a.e. μ,*

(ii) *$\mu_u\hat{v}$ or $\mu_u\check{v} = 0$, for all $v \in F_{m-1}^{(u)}$.*

This remains true for the subspaces Φ_h^d, when $k, m \geq 2$.

Proof: (i) Fix any $x \in F_k^d$. If $k < d$, then $\tilde{u} \not\perp x$ for almost every $u \in F_{d-1}^d$, and so $\dim(x \cap u) = \dim(\pi x \cap \pi u) = k - 1$, which means that $\sigma_u x \in F_{k-1}^{(u)}$. This is also trivially true when $k = d$. The assertion now follows by Fubini's theorem.

(ii) First we show that the F-version for dimensions (d, k, m) follows from the Φ-version for $(d+1, k+1, m+1)$. Then embed R^d as a flat $w \in F_d^{d+1} \setminus \Phi_d^{d+1}$ in R^{d+1}, and embed F^d correspondingly into $F^{(w)} \subset F^{d+1}$. The linear span of a flat in $F^{(w)}$ is a linear subspace of R^{d+1}, not contained in πw. Conversely, any such subspace is the span of a flat in $F^{(w)}$. Hence, the spanning operation defines a bi-measurable $1-1$ correspondence between F^d and $\Phi^{d+1} \setminus \Phi^{(\pi w)}$, with inverse σ_w. Since both hypothesis and assertion hold simultaneously for a measure $\mu \in \mathcal{M}(F_k^d)$ and its spanning image in $\mathcal{M}(\Phi_{k+1}^{d+1})$, we may henceforth take $\mu \in \mathcal{M}(\Phi_k^d)$ with $\mu\hat{v}$ or $\mu\check{v} = 0$ for all $v \in \Phi_m^d$, and prove (ii) for $v \in \Phi_{m-1}^{(u)}$ and almost every $u \in \Phi_{d-1}^d$.

When $m \leq k$, let $v_0, v_1, \ldots$ be the maximum inner degeneracy flats of μ, and let $\mu_0, \mu_1, \ldots$ be the corresponding components of μ. Then, by hypothesis, $\dim v_j < m$ for all j, and we may assume that $\dim v_j \geq 1$ iff $j \geq 1$. Noting that

$$\dim \rho_u v = \dim v \;\Leftrightarrow\; \tilde{u} \notin v, \quad u \in \Phi_{d-1}^d,\ v \in \Phi^d,$$

we have

$$\dim \rho_u \tilde{v}_j = \dim \tilde{v}_j > d - m, \quad j \geq 1, \; u \in \Phi^d_{d-1} \text{ a.e. } \lambda_{d-1}.$$

Since $\tilde{\mu}_j$ is non-degenerate on $\tilde{v}_j$, by definition, it follows that $\tilde{\mu}_j \circ \rho_u^{-1}$ is non-degenerate on $\rho_u \tilde{v}_j$ for all such u and j. Letting $v \in \Phi^{(u)}_{m-1}$ be arbitrary, and noting that $\dim(\tilde{v})_u = d - m$, we get by Lemma 11.21

$$\begin{aligned}\left(\mu_j \circ \sigma_u^{-1}\right)\check{v} &= \left(\mu_j \circ \sigma_u^{-1}\right)^{\tilde{}}_u (\tilde{v})\hat{}_u \\ &= \left(\tilde{\mu}_j \circ \rho_u^{-1}\right)(\tilde{v})\hat{}_u = 0, \quad j \geq 1,\end{aligned}$$

which remains true for $j = 0$, since $\mu_0 \circ \sigma_u^{-1}$ is non-degenerate on u.

The case $m = d$ being trivial, it remains to take $k \leq m < d$. Assuming $\|\mu\| = 1$, we may consider a random flat ξ with distribution μ. Next, choose a random line ζ in ξ, at a distance 1 from 0, but otherwise uniformly distributed, and let γ be the point on ζ closest to 0. The non-degeneracy condition on μ yields

$$\begin{aligned}P\{\zeta \subset v\} &\leq P\{\gamma \in v\} \\ &= EP(\gamma \in v \mid \xi) \\ &= P\{\xi \subset v\} = 0, \quad v \in \Phi^d_m.\end{aligned} \tag{10}$$

Next, consider an independent random flat η in Φ^d_{d-1} with distribution λ_{d-1}, and note that $\alpha = \zeta \cap \eta$ is a.s. a singleton. Indeed, assuming η to be spanned by some independent, uniformly distributed random lines $\chi_1, \ldots, \chi_{d-1}$ in Φ^d_1, we note that $\pi\zeta, \chi_1, \ldots, \chi_{d-1}$ are a.s. linearly independent. Then $\xi \not\subset \eta$ a.s., and so $\dim(\xi \cap \eta) = k - 1$ a.s.

Independently of η, we now choose some i.i.d. copies $(\xi_1, \zeta_1), (\xi_2, \zeta_2), \ldots$ of (ξ, ζ), and put $\alpha_j = \zeta_j \cap \eta$. We claim that

$$H_n\colon \;\; \overline{\dim}\,\{\alpha_1, \ldots, \alpha_n\} = n \text{ a.s.}, \quad n = 1, \ldots, m,$$

where $\overline{\dim}$ is short for dim span. Anticipating H_m, and using Fubini's theorem, we get for $u \in \Phi^d_{d-1}$ a.e. λ_{d-1}

$$\overline{\dim}\,\{\sigma_u \xi_1, \ldots, \sigma_u \xi_m\} \geq \overline{\dim}\,\{\sigma_u \zeta_1, \ldots, \sigma_u \zeta_m\} = m \text{ a.s.},$$

which implies $\sigma_u \xi \not\subset v$ a.s. for every $v \in \Phi^{(u)}_{m-1}$, proving the assertion in (ii).

Statement H_1 is clear, since $\alpha_1 \neq 0$. Proceeding by induction, we assume H_{n-1} to be true for some n, so that $\alpha_1, \ldots, \alpha_n$ are a.s. singletons satisfying

$$\overline{\dim}\,\{\alpha_1, \ldots, \alpha_{n-1}\} = \overline{\dim}\,\{\alpha_1, \ldots, \alpha_{n-2}, \alpha_n\} = n - 1. \tag{11}$$

Then Fubini's theorem shows that, for almost all $\eta, \zeta_1, \ldots, \zeta_{n-1}$, the line ζ_n satisfies (11) a.s. Fixing $\eta, \zeta_1, \ldots, \zeta_{n-1}$ accordingly, let β denote the orthogonal complement in η of $\text{span}\{\alpha_1, \ldots, \alpha_{n-1}\}$. Since $\zeta_{n-1} \not\subset \eta$, almost every point $\alpha'_{n-1} \in \zeta_{n-1} \setminus \{\alpha_{n-1}\}$ is such that $\eta' = \text{span}\{\alpha_1, \ldots, \alpha_{n-2}, \alpha'_{n-1}, \beta\}$ has dimension $d-1$, and intersects $\zeta_1,\dots,\zeta_{n-1}$ uniquely at the points $\alpha_1,\dots,\alpha_{n-2},\alpha'_{n-1}$, as well as ζ_n at a unique point α'_n. By the choice of α'_{n-1}, we get from (11)

$$\overline{\dim}\,\{\alpha_1,\dots,\alpha_{n-1},\alpha'_{n-1}\} = n \le m, \tag{12}$$

and so by (10)

$$\zeta_n \not\subset \mathrm{span}\{\alpha_1,\dots,\alpha_{n-1},\alpha'_{n-1}\} \ \text{ a.s.} \tag{13}$$

Fix ζ_n accordingly, and such that ζ_n intersects η and η' uniquely at α_n and α'_n, respectively.

Assuming $\zeta_1,\dots,\zeta_n$ to have linearly dependent intersections with both η and η', we get by (12)

$$\begin{aligned} \alpha_n &\in \mathrm{span}\{\alpha_1,\dots,\alpha_{n-1}\},\\ \alpha'_n &\in \mathrm{span}\{\alpha_1,\dots,\alpha_{n-2},\alpha'_{n-1}\}. \end{aligned} \tag{14}$$

If $\alpha_n = \alpha'_n$, we see from (12) and (14) that

$$\begin{aligned} \alpha_n &\in \mathrm{span}\{\alpha_1,\dots,\alpha_{n-1}\} \cap \mathrm{span}\{\alpha_1,\dots,\alpha_{n-2},\alpha'_{n-1}\}\\ &= \mathrm{span}\{\alpha_1,\dots,\alpha_{n-2}\}, \end{aligned}$$

which contradicts (11). If instead $\alpha_n \ne \alpha'_n$, then by (14),

$$\begin{aligned} \zeta_n &\subset \mathrm{span}\{\alpha_n,\alpha'_n\}\\ &\subset \mathrm{span}\{\alpha_1,\dots,\alpha_{n-1},\alpha'_{n-1}\}, \end{aligned}$$

contrary to (13). From those contradictions and Fubini's theorem, we conclude that, for almost every choice of $\zeta_1,\dots,\zeta_n$, there exists a $u \in \Phi^d_{d-1}$, such that the intersections $\sigma_u\zeta_1,\dots,\sigma_u\zeta_n$ are unique and linearly independent.

Fixing any lines $\zeta_1,\dots,\zeta_n$ with the latter property, we may choose $a_i \in \mathsf{R}^d$ and $b_i \in \mathsf{R}^d \setminus \{0\}$, such that $\zeta_i \equiv \{a_i + tb_i;\, t \in \mathsf{R}\}$. If $r \in \mathsf{R}^d \setminus \{0\}$ satisfies $rb_i \ne 0$ for all i, then the flat $u = \{x \in \mathsf{R}^d;\, rx = 0\}$ intersects $\zeta_1,\dots,\zeta_n$ uniquely at the points

$$x_i = a_i - b_i\,\frac{ra_i}{rb_i}, \quad i = 1,\dots,n.$$

The intersections are then unique and linearly independent, iff the $d \times n$ matrix

$$(rb_i)\Big\{a_i(rb_i) - b_i(ra_i)\Big\}, \quad i = 1,\dots,n,$$

has full rank n, so that at least one of its $n \times n$ sub-determinants is non-zero. Now this was seen to occur for at least one r. Since a polynomial in $r_1,\dots,r_d$ is a.e. non-zero, unless it vanishes identically, the intersections are indeed unique and linearly independent, for almost every u. By Fubini's theorem, this proves H_n, and hence completes the induction. □

We proceed with a $0-\infty$ law for the inner and outer degeneracies of a stationary random measure on F^d_k.

Theorem 11.24 *($0-\infty$ law) Let η be a stationary random measure on F_k^d with locally finite intensity $E\eta$, such that a.s. $(\eta \circ \pi^{-1})v = 0$ for all $v \in \Phi_k^d$. Then a.s.*

$$(\eta \circ \pi^{-1})\hat{u} \in \{0, \infty\}, \quad u \in \Phi_m^d, \; m \in \{k, \dots, d\},$$

and similarly for $(\eta \circ \pi^{-1})\check{u}$ when $m \in \{1, \dots, k\}$.

Proof: We consider only the case $m \geq k$, the argument for $m \leq k$ being similar. Choose a disjoint partition of bounded Borel sets $B_1, B_2, \dots$ in F_k^d, and introduce on Φ_k^d the random measure

$$\zeta_1 = \sum_{n \geq 1} \frac{(1_{B_n}\eta) \circ \pi^{-1}}{1 + \eta B_n} 2^{-n}.$$

Next, we choose a measurable map $g_1 \colon \Phi_k^d \times \Phi_1^k \to \Phi_1^d$, such that $g_1(v, \cdot)$ is an isomorphism of Φ_1^k onto $\Phi_1^{(v)}$, for every $v \in \Phi_k^d$. Given an invariant measure λ_1 on Φ_1^k, we may form the random measure $\zeta_2 = (\zeta_1 \otimes \lambda_1) \circ g_1^{-1}$ on Φ_1^d, so that $(\zeta_2)^m$ becomes a random measure on $(\Phi_1^d)^m$.

For any $\omega \in \Omega$, consider the smallest integer $m \geq k$, such that $(\eta \circ \pi^{-1})\hat{u} > 0$ for some $u \in \Phi_m^d$. Since $x_1, \dots, x_m \in \Phi_1^d$ are linearly independent a.e. $(\zeta_2)^m$, they span a subspace $g_2(x_1, \dots, x_m)$ in Φ_m^d where the function g_2 is measurable. Consider the random measure $\zeta_3 = (\zeta_2)^m \circ g_2^{-1}$ on Φ_m^d, and note that $\zeta_3\{u\} > 0$ iff $(\eta \circ \pi^{-1})\hat{u} > 0$. Since the integer m and the associated sequence of degeneracy flats $u \in \Phi_m^d$ are measurable, by Lemma 1.6, we may introduce the smallest integer $m \geq k$, such that with positive probability, $(\eta \circ \pi^{-1})\hat{u} > 0$ for some $u \in \Phi_m^d$. Writing $\alpha_1, \alpha_2, \dots$ for the atom sites of ζ_3, measurably enumerated as in Lemma 1.6, we may introduce the restrictions $\eta_1, \eta_2, \dots$ of η to the inverse sets $\beta_n = \{x \in F_k^d;\, \pi x \subset \alpha_n\}$. By Lemma 11.22, the η_n are mutually singular, and a.s. unique up to a random order. Denote the bounded components η_n by η'_n, and put $\eta' = \sum_n \eta'_n$.

By Lemma 11.23 and Fubini's theorem, we may fix a flat $v \in F_{d-k+1}^d$, such that a.s. $\sigma_v u \equiv u \cap v \in F_1^{(v)}$ for $u \in F_k^d$ a.e. η, and $\eta \circ \sigma_v^{-1} \circ \pi^{-1}\{x\} = 0$ for all $x \in \Phi_1^{(v)}$. Since σ_v is measurable, we may introduce some random measures $\chi_i = \eta'_i \circ \sigma_v^{-1}$ on $F_1^{(v)}$, so that the $(\chi_i)^2$ become random measures on $(F_1^{(v)})^2$. For any $x_1, x_2 \in F_1^{(v)}$ with $\pi x_1 \neq \pi x_2$, there exists a unique point $g_4(x_1, x_2)$, at equal minimum distance from x_1 and x_2, and since g_4 is measurable, we may introduce the random measures $\rho_i = (\chi_i)^2 \circ g_4^{-1}$ on v. Finally, we put $\gamma_i = g_5(\rho_i)$, where for bounded measures $\mu \neq 0$ on v, we choose $g_5(\mu)$ to be a measurable and translation-preserving "center" in v. (The precise definition of g_5 is not important; it could be based on the medians in $d - k + 1$ directions.)

For any $x \in F_1^{(v)}$ and $y \in v$, let $g_6(x, y)$ be the point in x closest to y, and put $g_7(x, y) = \{\pi x, g_6(x, y)\}$, where both functions are measurable. Since for every n, the product $\chi_n \otimes \delta_{\gamma_n}$ is a random measure on $F_1^{(v)} \times v$, we may form a random measure $\zeta = \sum_n (\chi_n \otimes \delta_{\gamma_n}) \circ g_7^{-1}$ on $\Phi_1^{(v)} \times v$. Putting $\chi = \sum_n \chi_n$, we note that $\eta' = \chi \circ g_8^{-1}$, where $g_8(x,y) = x + y$ for $x \in \Phi_1^{(v)}$ and $y \in v$. Note that χ is v-stationary, since our construction is coordinate-free, and that $E\eta'$ is locally finite, since the σ_u-inverses of bounded sets are bounded, and $\eta' \le \eta$.

For fixed $w \in \Phi_{d-k}^{(v)}$, we may divide v into congruent "slices" S_j, $j \in \mathsf{Z}$, parallel to w. Writing κ_j for the restriction of χ to $\Phi_1^{(v)} \times S_j$, we note that the κ_j are w-stationary and equally distributed, apart from shifts. Those properties carry over to the random measures $\kappa_j \circ g_8^{-1}$, whose intensities then agree, by Lemma 11.9. Hence, $E\eta' B = \sum_j E\kappa_j \circ g_8^{-1} B$ equals 0 or ∞, for bounded Borel sets B. Since the latter possibility is excluded, we have in fact $\eta' = 0$ a.s., and so $\eta'_n = 0$ a.s. for all n. Thus, $\|\eta_n\| \in \{0, \infty\}$ a.s., which proves the assertion for degeneracy flats of dimension $\le m$. Since $\eta' = \eta - \sum_n \eta_n$ is stationary, and a.s. $(\eta' \circ \pi^{-1})\hat{u} = 0$ for all $u \in \Phi_m^d$, we may continue recursively to complete the proof. □

11.7 General Criteria of Absolute Continuity

Invariance criteria for random measures η on F_k^d, based on absolute-continuity conditions of the form $\eta \circ \pi^{-1} \ll \mu$ a.s., have already been considered in a previous section. Here we return to the subject, now exploring results of the mentioned type under minimal hypotheses on the measure μ. Our key result is Lemma 11.25, where for $k = 1$, we prove that a.s. invariance holds in a direction φ_p, depending on the direction p of the lines, under the sole condition that μ be diffuse. As a consequence, we show in Theorem 11.26 that, when ξ is a flat process with Papangelou random measure η, the latter condition alone ensures that ξ be Cox and directed by η. We also prove in Corollary 11.28 that the a.s. invariance of η remains true in all directions, provided that μ has no outer degeneracies, a condition also shown to be best possible.

For our initial technical result, we use the parametrization of the lines $l \in F_1^{d+1}$ as points $(q,p) \in (\mathsf{R}^d)^2$, where q is the intersection of l with a fixed subspace $u \in \Phi_d^{d+1}$, and p is the slope of l, measured as the rate of change of p, as u is moved in the perpendicular direction. We can always choose u, such that the random measure η below gives a.s. no charge to lines parallel to u. Say that a measure η_p is *$\bar{q}$-invariant*, if it is tq-invariant for every $t \in \mathsf{R}$.

Lemma 11.25 *(invariant disintegration) Let η be a stationary random measure on F_1^{d+1}, such that $\eta \circ \pi^{-1} \ll \mu$ a.s. for some diffuse measure $\mu \in \mathcal{M}_{\mathsf{R}^d}$. Then we may choose a measurable a.s. disintegration $\eta = \mu \otimes (\eta_p)$, and a measurable function φ on R^d, such that η_p is stationary and a.s. $\bar{\varphi}_p$-invariant, for every $p \in \mathsf{R}^d$.*

Proof: By Theorem 7.35, we may choose a strongly stationary disintegration kernel (η_p) of η with respect to μ, where η_p is a random measure on

R^d, for every $p \in \mathsf{R}^d$. Given a measure-determining sequence of continuous functions $f_1, f_2, \ldots \geq 0$ on R^d with bounded supports, we may introduce the processes $Y_k(q,p) = (\eta_p * \delta_q) f_k$. Next, we fix a dissecting system (I_{nj}) in R^d, write $I_n^{(p)}$ for the unique set I_{nj} containing p, and define $\mu_n^{(p)} = I_n^{(p)} \cdot \mu / \mu I_n^{(p)}$, where $B \cdot \mu = 1_B \mu$, and $0/0$ is interpreted as 0. Letting $n \to \infty$ for fixed $k \in \mathsf{N}$ and $\omega \in \Omega$, we get as in the proof of Lemma 1.28

$$\int \Big| Y_k(0,p) - Y_k(0,x) \Big| \, \mu_n^{(p)}(dx) \to 0, \quad p \in \mathsf{R}^d \text{ a.e. } \mu.$$

Hence, by Fubini's theorem and dominated convergence,

$$\int \mu_n^{(p)}(dx) \sum_{k \geq 1} 2^{-k} E\Big\{ |Y_k(0,p) - Y_k(0,x)| \wedge 1 \Big\} \to 0, \quad p \in \mathsf{R}^d \text{ a.e. } \mu.$$

Now fix a non-exceptional point $p \in \operatorname{supp} \mu$. Thinking of the last condition as the L^1-convergence $\|\xi_n\|_1 \to 0$ of some random variables $\xi_n \geq 0$, we get $\xi_n \to 0$ a.s. along a sub-sequence $N' \subset \mathsf{N}$. Since μ is diffuse, we may then choose some non-zero points $h_n \to 0$ in R^d, such that

$$\sum_{k \geq 1} 2^{-k} E\Big\{ |Y_k(0,p) - Y_k(0,p+h_n)| \wedge 1 \Big\} \to 0,$$

which implies

$$Y_k(0, p+h_n) \xrightarrow{P} Y_k(0,p), \quad n \to \infty \text{ along } N', \;\; k \in \mathsf{N}. \tag{15}$$

A compactness argument yields $h_n/|h_n| \to q$ along a further sub-sequence $N'' \subset N'$, for some $q \in \mathsf{R}^d$ with $|q| = 1$. Fixing $k \in \mathsf{N}$ and $t > 0$, and writing $r_n = t/|h_n|$, we get

$$\begin{aligned}
\Big| Y_k(tq,p) - Y_k(0,p) \Big| &\leq \Big| Y_k(tq,p) - Y_k(r_n h_n, p) \Big| \\
&\quad + \Big| Y_k(r_n h_n, p) - Y_k(r_n h_n, p+h_n) \Big| \\
&\quad + \Big| Y_k(r_n h_n, p+h_n) - Y_k(0,p) \Big| \\
&= \alpha_n + \beta_n + \gamma_n.
\end{aligned}$$

Here $\alpha_n \to 0$ along N'', since $Y_k(\cdot, p)$ is continuous, and $r_n h_n = t h_n/|h_n| \to tq$. Next, we see from (15) and the space-time stationarity of Y_k that

$$\begin{aligned}
\beta_n &\overset{d}{=} \Big| Y_k(0,p) - Y_k(0,p+h_n) \Big| \xrightarrow{P} 0, \\
\gamma_n &\overset{d}{=} \Big| Y_k\Big\{ r_n h_n - r_n(p+h_n),\, p+h_n \Big\} - Y_k(-r_n p,\, p) \Big| \\
&\overset{d}{=} \Big| Y_k(0, p+h_n) - Y_k(0,p) \Big| \xrightarrow{P} 0.
\end{aligned}$$

By combination, we obtain $Y_k(tq,p) = Y_k(0,p)$ a.s., where the exceptional P-null set can be chosen by continuity to be independent of t. Since the f_k are measure determining, it follows that η_p is a.s. q-invariant.

Writing S for the unit sphere in R^d, we define

$$f(q,p) = \int_0^1 dt \sum_{k\geq 1} 2^{-k}\, E\Big\{ \Big| Y_k(tq,p) - Y_k(0,p) \Big| \wedge 1 \Big\}, \quad p \in \mathsf{R}^d,\ q \in S,$$

and note that η_p is a.s. $\bar{q}$-invariant iff $f(q,p) = 0$. Since f is product measurable by Fubini's theorem, the set $A = \{(q,p);\ f(q,p) = 0\}$ is measurable in $S \times \mathsf{R}^d$. The previous argument shows that $\mu(\pi A)^c = 0$, where πA denotes the projection $\{p \in \mathsf{R}^d;\ \bigcup_q \{f(q,p) = 0\}\}$. Hence, the general section Theorem A1.1 (ii) yields a measurable function $\varphi\colon \mathsf{R}^d \to S$, such that $(\varphi_p, p) \in A$ for $p \in \mathsf{R}^d$ a.e. μ, which means that η_p is a.s. $\bar{\varphi}_p$-invariant for every such p. Redefining $\eta_p = 0$ on the exceptional μ-null set, we can make this hold identically in p. $\square$

The last result has some remarkable consequences. First we show that, for a sufficiently regular flat process ξ in R^d, a simple absolute-continuity condition on the associated Papangelou kernel η guarantees that ξ will be Cox distributed.

Theorem 11.26 *(Cox criterion) Let ξ be a stationary simple point process on F_k^d satisfying (Σ), and such that the associated Papangelou kernel η satisfies $\eta \circ \pi^{-1} \ll \mu$ a.s., for some diffuse measure $\mu \in \mathcal{M}(\Phi_k^d)$. Then ξ is a Cox process directed by η.*

Proof: By an obvious approximation argument based on Lemma 11.23, we may take $k = 1$, hence assuming ξ to be a point process on F_1^{d+1}. Let φ and η_p be such as in Lemma 11.25, and choose $u \in \Phi_{d-1}^d$, such that $\varphi_p \not\subset u$ for $p \in \mathsf{R}^d$ a.e. μ. Fix a unit vector $y \perp u$ in R^d, write $a_p = (y + u) \cap \bar{\varphi}_p$, and introduce on R^d some linear transformations f_p with inverses f_p^{-1}, given by

$$\begin{aligned} f_p(q) &= q + (qy)(y - a_p), \\ f_p^{-1}(q) &= q - (qy)(y - a_p), \quad p, q \in \mathsf{R}^d. \end{aligned} \tag{16}$$

Both functions are clearly measurable in the pair (p,q), which remains true for the mutual inverses

$$\begin{aligned} h(q,p) &= \{f_p(q), p\}, \\ h^{-1}(q,p) &= \Big\{f_p^{-1}(q), p\Big\}, \quad p, q \in \mathsf{R}^d. \end{aligned} \tag{17}$$

The disintegration $\eta = \mu \otimes (\eta_p)$ gives

$$\eta \circ h^{-1} = \mu \otimes \Big(\eta_p \circ f_p^{-1}\Big). \tag{18}$$

Introducing the shifts $\theta_r q = q + r$ in R^d, we further obtain, for any $t \in \mathsf{R}$ and $q, p \in \mathsf{R}^d$,

$$\begin{aligned} f_p \circ \theta_{ta_p}(q) &= q + ta_p + \{y(q + ta_p)\}\,(y - a_p) \\ &= q + (yq)(y - a_p) + ty \\ &= \theta_{ty} \circ f_p(q), \end{aligned} \tag{19}$$

and so the $\bar{\varphi}_p$-invariance of η_p yields

$$\begin{aligned} \eta_p \circ f_p^{-1} \circ \theta_{ty}^{-1} &= \eta_p \circ \theta_{ta_p}^{-1} \circ f_p^{-1} \\ &= \eta_p \circ f_p^{-1}, \qquad t \in \mathsf{R},\ \ p \in \mathsf{R}^d \ \text{ a.e. } \mu, \end{aligned} \tag{20}$$

outside a fixed P-null set. By (18) it follows that $\eta \circ h^{-1}$ is a.s. $\bar{y}$-invariant. Since $\xi \circ h^{-1}$ has Papangelou kernel $\eta \circ h^{-1}$, Theorem 8.19 shows that the process $\xi \circ h^{-1}$ is Cox and directed by $\eta \circ h^{-1}$. Applying the inverse mapping (17) to both $\xi \circ h^{-1}$ and $\eta \circ h^{-1}$, we see from Theorem 3.2 that ξ itself is Cox and directed by η. □

The directing random measure η of the last result may fail to be invariant. Here we give a stronger condition, ensuring the desired a.s. invariance.

Theorem 11.27 *(partial invariance) For any $\mu \in \mathcal{M}(\Phi_k^d)$ and $y \in \Phi_1^d$, these conditions are equivalent:*

(i) *y lies in every flat $v \in \Phi_{d-1}^d$ with $\mu\hat{v} > 0$,*

(ii) *any stationary random measure η on F_k^d with $\eta \circ \pi^{-1} \ll \mu$ a.s. is a.s. y-invariant.*

Proof, (i) ⇒ (ii): Assume (i). The outer degeneracy decomposition of μ in Lemma 11.22 yields a corresponding decomposition of η, and it is enough to prove that each component is a.s. y-invariant. We may then take μ to be supported by the set of flats in some subspace $v \in \Phi_m^d$ with $y \subset v$, such that v contains no proper degeneracy space. It suffices to prove the a.s. invariance of the v-projection $\eta(\cdot \times B)$ of η onto $F_k^{(v)}$, for any bounded Borel set B in the $(d-m)$-dimensional subspace $v^\perp$. Equivalently, we may take $v = \mathsf{R}^d$, so that μ has no outer degeneracies of order $< d$, and prove that η is then a.s. invariant in all directions.

Applying Lemma 11.23 recursively $k-1$ times, we obtain a flat $u \in \Phi_{d-k+1}^d$, such that $\sigma_u x \in \Phi_1^{(u)}$ for $x \in \Phi_k^d$ a.e. μ, and $(\mu \circ \sigma_u^{-1})\hat{v} = 0$ for all $v \in \Phi_{d-k}^{(u)}$. Since $\sigma_u x \in \Phi_1^{(u)}$ implies $\sigma_u v \in F_1^{(u)}$ for every $v \in F_k^d$ with $\pi v = x$, we have a.s. $\sigma_u v \in F_1^{(u)}$ a.e. η. Thus, the function $v \mapsto (\sigma_u v, \pi v)$ maps η into a random measure η' on $F_1^{(u)} \times \Phi_k^d$, whereas σ_u maps μ into a measure $\mu' = \mu \circ \sigma_u^{-1}$ on $\Phi_1^{(u)}$, with $\mu'\hat{v} = 0$ for all $v \in \Phi_{d-k}^{(u)}$. Moreover, η' is u-stationary, and $\eta'(B \times (\cdot) \times \Phi_k^d) \ll \mu'$ a.s., for any measurable set $B \subset F_1^{(u)}/\Phi_1^{(u)}$. If the claim is true for $k=1$, then the projection $\eta'(\cdot \times A)$ onto $F_1^{(u)}$ is a.s. invariant, for any measurable set $A \subset \Phi_k^d$, and so η' itself is a.s. invariant, which extends by Lemma 11.9 to the a.s. invariance of η. This reduces the proof to the case of $k = 1$.

For $d = 2$, we may choose a stationary disintegration $\eta = \mu \otimes (\eta_p)$ and a measurable function φ on Φ_1^2, as in Lemma 11.25, so that η_p is a.s. $\bar{\varphi}_p$-invariant for every $p \in \Phi_1^2$. Since $\bar{\varphi}_p \neq p$, Lemma 11.9 shows that the η_p are a.s. invariant under arbitrary translations, which then remains true for η itself.

Proceeding by induction on d, assume the assertion to be true for some fixed $d \geq 2$, and let μ and η fulfill the hypotheses in dimension $d+1$. Fixing a reference plane $u \in \Phi_d^{d+1}$, and ignoring the null set of lines parallel to u, we may identify F_1^{d+1} with $\mathsf{R}^{2d} = (\mathsf{R}^d)^2$, so that $\eta \circ \pi^{-1} \ll \mu$ a.s. on R^d. Writing $(\mathsf{R}^d)^2 = (\mathsf{R}^{d-1})^2 \times \mathsf{R}^2$, we may fix any bounded Borel set $B \subset \mathsf{R}^2$, and regard the associated projection $\eta_B = \eta(\cdot \times B)$ onto $(\mathsf{R}^{d-1})^2$ as representing a stationary random measure on F_1^d. Here $\eta_B \circ \pi^{-1} \ll \mu'$ a.s., where μ' denotes the projection $\mu(\cdot \times \mathsf{R})$ onto R^{d-1}, which again has no outer degeneracies. Applying the induction hypothesis, we conclude that η_B is a.s. invariant. Since B was arbitrary, η is then a.s. invariant under shifts in the first $d-1$ variables. Permuting the coordinates, we may extend the a.s. invariance to arbitrary directions. This completes the induction, and the assertion in (ii) follows.

(ii) $\Rightarrow$ (i): Let $\mu \in \mathcal{M}(\Phi_k^d)$ and $y \in \Phi_1^d$ be arbitrary. If (i) fails, there exists a flat $v \in \Phi_{d-1}^d$ with $\mu\hat{v} > 0$, such that $y \not\subset v$. Choose $z \perp v$ with $|z| = 1$, let ϑ be $U(0, \pi)$, and consider on F_k^d the stationary process

$$Y(x) = \sin^2\{(\rho_z x)z + \vartheta\} \cdot 1\{\pi x \subset v\}, \quad x \in F_k^d,$$

where $\rho_z x$ denotes the projection of x onto the z-axis, which is unique when $\pi x \subset v$. Writing λ^{d-k} for Lebesgue measure on $F_k^d/\Phi_k^d \sim \mathsf{R}^{d-k}$, we see that the random measure $\eta = Y \cdot (\lambda^{d-k} \otimes \mu)$ is stationary but a.s. not y-invariant. Thus, even (ii) is violated. □

The last result leads in particular to the following remarkable extension of Theorem 11.12.

Corollary 11.28 *(invariance criterion) For any $\mu \in \mathcal{M}(\Phi_k^d)$, these conditions are equivalent:*

(i) *μ has no outer degeneracies,*

(ii) *any stationary random measure η on F_k^d with $\eta \circ \pi^{-1} \ll \mu$ a.s. is a.s. invariant.*

Proof, (i) $\Rightarrow$ (ii): Suppose that μ has no outer degeneracies, and let η be a stationary random measure on F_k^d with $\eta \circ \pi^{-1} \ll \mu$ a.s. Then condition (i) of Theorem 11.27 is vacuously satisfied for any $y \in \Phi_1^d$, and so η is a.s. y-invariant for every such y. Applying this to some linearly independent directions $y_1, \ldots, y_d \in \Phi_1^d$, we conclude that η is a.s. invariant under arbitrary translations.

(ii) $\Rightarrow$ (i): Suppose that every stationary random measure η on F_k^d with $\eta \circ \pi^{-1} \ll \mu$ a.s. is a.s. invariant. If $v \in \Phi_{d-1}^d$ with $\mu\hat{v} > 0$, then Theorem 11.27 shows that $y \subset v$ for every $y \in \Phi_1^d$, although this is clearly false for $y = v^\perp$. The contradiction shows that $\mu\hat{v} = 0$ for all $v \in \Phi_{d-1}^d$, which means that μ has no outer degeneracies. □

Chapter 12

Regeneration and Local Time

Regenerative sets and processes, along with the associated notions of local time and occupation measures, form an important application area of random measure theory. Here our criterion is not primarily whether the central notions and results can be expressed in terms of random measures and point processes, which is true for practically every area of probability theory, but rather whether the powerful theory of the previous chapters, including results for Poisson-type processes and Palm distributions, contributes in any significant way to their understanding. In this chapter, we will highlight some aspects of regenerative sets and processes, where general ideas from random measure theory play a basic and often striking role.

The classical renewal process, regarded as a point process ξ on R, provides the simplest non-trivial example illustrating the notion of Palm measures, since its distribution agrees with the Palm measure of the stationary version $\tilde{\xi}$ of ξ, whenever such a version exists. Thinking of ξ as the occupation measure of a random walk in R, it becomes important to explore the asymptotic properties of ξ and its intensity $E\xi$, which leads to the classical *renewal theorem*. The result shows in essence that, as $t \to \infty$, the shifted processes $\theta_t \xi$ tend in distribution to the stationary version $\tilde{\xi}$, whereas their intensities $\theta_t E\xi$ tend to a multiple of Lebesgue measure λ.

The elementary renewal property generalizes to the *regenerative property* of a process X at a state a, defined as the strong Markov property of X at any optional time τ, taking values in the random set $\Xi = \{t \geq 0;\, X_t = a\}$. Familiar examples are the level sets of a Brownian motion X. Excluding some elementary cases, including the classical renewal process, we show that the *regenerative set* Ξ is nowhere dense with a perfect closure $\bar{\Xi}$, typically with Lebesgue measure $\lambda\Xi = 0$. It is naturally described by a random measure ξ with support $\bar{\Xi}$, or by the associated non-decreasing, continuous process $L_t = \xi[0,t]$, both referred to as the *local time* of Ξ. Indeed, the entire excursion structure of the underlying process X is determined by a Poisson process η on the local time scale, with intensity $E\eta = \lambda \otimes \nu$, where ν is a σ-finite measure on the space D_0 of excursion paths, known as the *Itô excursion law.* We may also describe Ξ as the range of a subordinator T, and ξ as the image of Lebesgue measure λ under mapping by the paths of T.

Many approaches to the local time are known, valid under different conditions on the underlying process X. Most classical constructions use ap-

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_12

proximation from the excursions. A more elegant and powerful approach is based on *Tanaka's formula* from stochastic calculus, which yields the local time L_t^x of a continuous semi-martingale X, at an arbitrary level $x \in \mathsf{R}$. Under mild regularity conditions, the process L has a jointly continuous version on $\mathsf{R}_+ \times R$, admitting an interpretation as *occupation density*, defined as a density of the occupation random measure of X up to time t. The space-time local time L_t^x also leads to a generalized version of Itô's formula.

Itô's celebrated excursion theorem shows that, for any regenerative set Ξ with associated local time ξ, the excursions of X are conditionally independent, with distributions given by the kernel (ν_r), obtained by disintegration of ν with respect to excursion lengths r. A stronger result is Theorem 12.28, where for fixed times $t_1 < \cdots < t_n$, we introduce the straddling excursion intervals $I_k = [\sigma_k, \tau_k]$, along with the intermediate intervals $J_k = [\tau_{k-1}, \sigma_k]$. Here the restrictions of X to the intervals I_k and J_k are conditionally independent, with distributions given by the excursion kernel ν, respectively by the Palm kernel Q with respect to the local time random measure ξ. A similar factorization holds for the joint distribution of all interval lengths $\gamma_k = \tau_k - \sigma_k$ and $\beta_k = \sigma_k - \tau_{k-1}$.

Some more delicate local properties of X are obtainable, under special regularity conditions on the Palm kernel Q. Since specific versions of Q may be constructed by duality from the intensity measure $E\xi$, we need the latter to be absolutely continuous with a nice density p. When Ξ has positive Lebesgue measure, we may choose[1] $\xi = 1_\Xi \lambda$, in which case a continuous density always exists. In the more subtle case of $\lambda\Xi = 0$ a.s., the desired absolute continuity still holds, under an exceedingly weak regularity condition. By imposing slightly stronger conditions, we can give a fairly precise description of the continuity set for the associated density p. In particular, the desired continuity holds on the entire interval $(0, \infty)$, for the level sets of sufficiently regular diffusion processes.

When $\lambda\Xi \neq 0$, we can simply define the Palm distributions Q_t by

$$Q_t = \mathcal{L}(X \mid t \in \Xi\,), \quad t \geq 0.$$

If instead $\lambda\Xi = 0$, then under the mentioned regularity assumptions, we can choose

$$Q_t = p_t^{-1} \int_0^\infty p_{s,t}\, Q_{s,t}\, ds, \quad t > 0 \text{ with } p_t < \infty,$$

where $p_{s,t}$ is the density of the distribution $\mu_s = \mathcal{L}(T_s)$, and $Q_{s,t}$ denotes the associated set of conditional distributions $\mathcal{L}(X \mid T_s = t)$. Both cases are also covered by the formula

$$Q_t^- \{X^{\tau_r} \in \cdot\} = p_t^{-1} E\{\, p_{t-\tau_r};\, X^{\tau_r} \in \cdot\}, \quad t > 0 \text{ with } p_t < \infty,$$

from which much more information can be extracted. Using those versions of the Palm measures Q_t, we prove in Theorem 12.44 some strong continuity

[1] Recall that $1_B\mu$ denotes the restriction of μ to B, given by $(1_B\mu)f = \int_B f d\mu$.

and asymptotic factorization properties, of the form

$$\left\| Q_t \circ (X, \tilde{X}_t)^{-1} - Q_0 \otimes \tilde{Q}_t \right\|_{u,v} \to 0,$$

under each of the four conditions $u \to 0$, $v \to 0$, $u+v \to 0$, and $t-u-v \to \infty$, where the "tilde" denotes time reversal on $[0,t]$, and $\|\cdot\|_{u,v}$ denotes total variation for the restrictions to $[0,u] \cup [0,v]$.

In the final Section 12.8, we establish some local hitting and conditioning properties, similar to, but much deeper than those for general point processes in Sections 5.1 and 6.5. It should be noted that the stated results rely profoundly on the special nature of regenerative sets and processes, and that no such properties can be expected for general random measures. In the next chapter, we will see how similar results can also be proved for suitable super-processes.

For the mentioned local results, we consider any small neighborhoods $I_1, \ldots, I_n$ of some fixed times $t_1 < \cdots < t_n$, whose differences Δt_j are continuity points of the renewal density p. Then in Theorems 12.45 and 12.49, we show that as $|I_k| \to 0$ for each k, the hitting events $\xi I_k > 0$ are asymptotically independent, and the contributions of ξ to those intervals essentially agree with those for the stationary version $\tilde{\xi}$ of ξ. Indeed, in Theorem 12.49, similar asymptotic properties are established for the underlying regenerative process X, under some more general forms of hitting conditions. In Theorem 12.50 and its Corollary 12.53, we consider instead the conditional distribution of X outside $I_1, \ldots, I_n$, which is asymptotically given by the multi-variate Palm distribution $Q_{t_1,\ldots,t_n}$. Furthermore, by the factorization properties of the latter, the conditional contributions of X to the intermediate intervals J_k are asymptotically independent.

12.1 Renewal and Occupation Measures

Given a random walk X with $X_0 = 0$, based on a distribution μ on $(0, \infty)$, we define the associated *renewal process* ξ as the point process $\sum_{n \ge 0} \delta_{X_n}$. More generally, we may form a *delayed renewal process* η, by letting X_0 be independent of $X - X_0$ with an arbitrary *delay distribution* ν on R_+. We say that η is a *stationary version* of ξ, if it is stationary on R_+, with the same underlying spacing distribution μ.

Theorem 12.1 *(stationary renewal process) Let ξ be a renewal process, based on a distribution μ with mean m. Then ξ has a stationary version η on R_+ iff $m < \infty$, in which case $E\eta = m^{-1}\lambda$, and the delay distribution ν of η is given uniquely by*

$$\nu[0,t] = m^{-1} \int_0^t \mu(s, \infty)\, ds, \quad t \ge 0.$$

Furthermore, $\nu = \mu$ iff ξ is stationary Poisson on $(0, \infty)$.

Proof: For a delayed renewal process η based on μ and ν, Fubini's theorem yields

$$\begin{aligned} E\eta &= E\sum_{n\geq 0}\delta_{X_n} = \sum_{n\geq 0}\mathcal{L}(X_n) \\ &= \sum_{n\geq 0}\nu * \mu^{*n} \\ &= \nu + \mu * \sum_{n\geq 0}\nu * \mu^{*n} \\ &= \nu + \mu * E\eta, \end{aligned}$$

and so $\nu = (\delta_0 - \mu) * E\eta$. If η is stationary, then $E\eta$ is invariant, and so $E\eta = c\lambda$ for some constant $c > 0$. Thus, $\nu = c(\delta_0 - \mu) * \lambda$, which yields the displayed formula, with m^{-1} replaced by c. As $t \to \infty$, we get $1 = cm$, which implies $m < \infty$ and $c = m^{-1}$.

Conversely, suppose that $m < \infty$, and choose ν as stated. Then

$$\begin{aligned} E\eta &= \nu * \sum_{n\geq 0}\mu^{*n} \\ &= m^{-1}(\delta_0 - \mu) * \lambda * \sum_{n\geq 0}\mu^{*n} \\ &= m^{-1}\lambda * \left(\textstyle\sum_{n\geq 0}\mu^{*n} - \sum_{n\geq 1}\mu^{*n}\right) = m^{-1}\lambda, \end{aligned}$$

which shows that $E\eta$ is invariant. By the strong Markov property, the shifted random measure $\theta_t\eta$ is again a renewal process based on μ, say with delay distribution ν_t. Since $E\theta_t\eta = \theta_t E\eta$, we get as before

$$\begin{aligned} \nu_t &= (\delta_0 - \mu) * (\theta_t E\eta) \\ &= (\delta_0 - \mu) * E\eta = \nu, \end{aligned}$$

which implies $\theta_t\eta \overset{d}{=} \eta$, showing that η is stationary.

If ξ is a stationary Poisson process on $(0,\infty)$ with rate $c > 0$, then by Theorem 3.12 (i), it is a delayed renewal process based on the exponential distribution $\mu = \nu$ with mean c^{-1}. Conversely, let ξ be a stationary renewal process, based on a common distribution $\nu = \mu$ with mean $m < \infty$. Then the tail function $f(t) = \mu(t,\infty)$ satisfies the differential equation $mf' + f = 0$ with initial condition $f(0) = 1$, which gives $f(t) = e^{-t/m}$. This shows that $\mu = \nu$ is the exponential distribution with mean m. Using Theorem 3.12 again, we conclude that ξ is stationary Poisson on R_+ with rate m^{-1}. $\square$

We may form a *two-sided renewal process* $\tilde{\xi}$, by extending the ordinary renewal process ξ on R_+ to a simple point process on R with points at $\cdots < \tau_{-1} < \tau_0 < \tau_1 < \tau_2 < \cdots$, where $\tau_0 = 0$, and the differences $\tau_n - \tau_{n-1}$ are i.i.d. with distribution μ. It may be generated in the obvious way by two independent random walks $X \overset{d}{=} Y$, based on the common distribution μ. Since $\tilde{\xi}$ is trivially cycle-stationary, we may use Theorem 5.4 to construct a stationary point process η on R with Palm distribution $\mathcal{L}(\tilde{\xi})$.

Corollary 12.2 *(Palm inversion) Let ξ be a two-sided renewal process, based on a distribution μ with mean $m < \infty$, and form a point process η on R, by inserting an interval (σ, τ) at the origin with $\tau = \vartheta\gamma$ and $\sigma = \tau - \gamma$, where ξ, γ, and ϑ are independent, and ϑ is $U(0,1)$, while γ has distribution $m^{-1}\, t\, \mu(dt)$. Then η is stationary with Palm distribution $\mathcal{L}(\xi)$.*

Proof: By scaling, we may assume that $m = 1$, in which case the statement follows immediately from Theorem 5.4, where part (ii) gives the desired description of η in terms of ξ. □

A random walk X in R^d is said to be *recurrent* if $\liminf_{n\to\infty} |X_n| = 0$ a.s., and *transient* if $|X_n| \to \infty$ a.s. In terms of the *occupation measure* $\xi = \sum_n \delta_{X_n}$, transience means that ξ is a.s. locally finite, and recurrence that $\xi B_0^\varepsilon = \infty$ a.s. for every $\varepsilon > 0$. We define the *recurrence set* R of X as the set of points $x \in \mathsf{R}^d$ with $\xi B_x^\varepsilon = \infty$ a.s. for every $\varepsilon > 0$.

Theorem 12.3 *(recurrence dichotomy) Let X be a random walk in R^d, with occupation measure ξ and recurrence set R. Then exactly one of these cases occurs:*

(i) *X is recurrent, and $R = \operatorname{supp} E\xi$, which is then a closed subgroup of R^d,*

(ii) *X is transient, and $E\xi$ is locally finite.*

Proof: The event $|X_n| \to \infty$ has probability 0 or 1, by Kolmogorov's $0-1$ law. First assume that $|X_n| \to \infty$ a.s. For any $m, n \in \mathsf{N}$ and $r > 0$, the Markov property at time m yields

$$\begin{aligned} P\Big\{ |X_m| < r,\ \inf_{k\ge n} |X_{m+k}| \ge r \Big\} &\\ \ge P\Big\{ |X_m| < r,\ \inf_{k\ge n} |X_{m+k} - X_m| \ge 2r \Big\} &\\ = P\{|X_m| < r\}\, P\Big\{ \inf_{k\ge n} |X_k| \ge 2r \Big\}. & \end{aligned}$$

Here the event on the left can occur for at most n different values of m, and so

$$P\Big\{ \inf_{k\ge n} |X_k| \ge 2r \Big\} \sum_{m\ge 1} P\{|X_m| < r\} \le n, \quad n \in \mathsf{N}.$$

Since $|X_k| \to \infty$ a.s., the probability on the left is positive for large enough n, and we get

$$\begin{aligned} E\,\xi B_0^r &= E \sum_{m\ge 1} 1\{|X_m| < r\} \\ &= \sum_{m\ge 1} P\{|X_m| < r\} < \infty, \quad r > 0, \end{aligned}$$

which shows that $E\xi$ is locally finite.

Next suppose that $|X_n| \not\to \infty$ a.s., and note that $P\{|X_n| < r \text{ i.o.}\} > 0$ for some $r > 0$. Covering B_0^r by finitely many balls $G_1, \dots, G_n$ of radius $\varepsilon/2$, we conclude that $P\{X_n \in G_k \text{ i.o.}\} > 0$ for at least one k. By the Hewitt-Savage $0-1$ law, the latter probability is in fact 1. Thus, the optional time $\tau = \inf\{n;\, X_n \in G_k\}$ is a.s. finite, and so the strong Markov property at τ yields

$$\begin{aligned} 1 &= P\Big\{X_n \in G_k \text{ i.o.}\Big\} \\ &\le P\Big\{|X_{\tau+n} - X_\tau| < \varepsilon \text{ i.o.}\Big\} \\ &= P\Big\{|X_n| < \varepsilon \text{ i.o.}\Big\} \\ &= P\{\xi B_0^\varepsilon = \infty\}, \end{aligned}$$

which shows that $0 \in R$, and X is recurrent.

Putting $A = \operatorname{supp} E\xi$, we note that trivially $R \subset A$. For the converse, fix any $x \in A$ and $\varepsilon > 0$. Using the strong Markov property at $\sigma = \inf\{n;\, |X_n - x| < \varepsilon/2\}$, along with the recurrence of X, we get

$$\begin{aligned} P\{\xi B_x^\varepsilon = \infty\} &= P\Big\{|X_n - x| < \varepsilon \text{ i.o.}\Big\} \\ &\ge P\Big\{\sigma < \infty,\ |X_{\sigma+n} - X_\sigma| < \varepsilon/2 \text{ i.o.}\Big\} \\ &= P\{\sigma < \infty\}\, P\Big\{|X_n| < \varepsilon/2 \text{ i.o.}\Big\} > 0. \end{aligned}$$

By the Hewitt-Savage $0-1$ law, the probability on the left then equals 1, which means that $x \in R$. Thus, $A \subset R \subset A$, and so in this case $R = A$.

The set A is clearly a closed additive semigroup in R^d. To see that in this case it is even a group, it remains to show that $x \in A$ implies $-x \in A$. Defining σ as before, and using the strong Markov property, along with the fact that $x \in A \subset R$, we get

$$\begin{aligned} P\{\xi B_{-x}^\varepsilon = \infty\} &= P\Big\{|X_n + x| < \varepsilon \text{ i.o.}\Big\} \\ &= P\Big\{|X_{\sigma+n} - X_\sigma + x| < \varepsilon \text{ i.o.}\Big\} \\ &\ge P\Big\{|X_\sigma - x| < \varepsilon,\ |X_n| < \varepsilon/2 \text{ i.o.}\Big\} = 1, \end{aligned}$$

which shows that $-x \in R = A$. □

In particular, we obtain the following sufficient condition for recurrence in dimension $d = 1$:

Corollary 12.4 *(recurrence criterion) A random walk X in R is recurrent when $n^{-1}X_n \xrightarrow{P} 0$, which holds in particular when EX_1 exists and equals 0.*

Proof: For any $\varepsilon > 0$ and $r \ge 1$, we may cover $B_0^{r\varepsilon}$ by some intervals $I_1, \dots, I_m$ of length ε, where $m \lesssim r$. Using the strong Markov property at the optional times $\tau_k = \inf\{n;\, X_n \in I_k\}$, we get

$$\begin{aligned} E\,\xi B_0^{r\varepsilon} &\le \sum_{k\le m} E\,\xi I_k = \sum_{k\le m}\sum_{n\ge 0} P\{X_n \in I_k\} \\ &\le \sum_{k\le m}\sum_{n\ge 0} P\Big\{\tau_k < \infty;\, |X_{\tau_k+n} - X_{\tau_k}| \le \varepsilon\Big\} \\ &= \sum_{k\le m}\sum_{n\ge 0} P\{\tau_k < \infty\}\, P\Big\{|X_{\tau_k+n} - X_{\tau_k}| \le \varepsilon\Big\} \\ &\lesssim r \sum_{n\ge 0} P\{|X_n| \le \varepsilon\} = rE\,\xi B_0^{\varepsilon}, \end{aligned}$$

and so

$$E\,\xi B_0^{\varepsilon} \gtrsim r^{-1} E\,\xi B_0^{r\varepsilon} = \int_0^\infty P\Big\{|X_{[rt]}| \le r\varepsilon\Big\}\, dt.$$

If $n^{-1}X_n \xrightarrow{P} 0$, then the integrand on the right tends to 1 as $r \to \infty$, and so $E\,\xi B_0^{\varepsilon} = \infty$ by Fatou's lemma, and X is recurrent by Theorem 12.3. The last assertion is clear from the law of large numbers. $\square$

We now consider the occupation measure of a random walk in R, based on a distribution with positive mean. Though the renewal property fails in general, we still have stationarity on R_+.

Theorem 12.5 *(stationary occupation measure) Let ξ be the occupation measure of a random walk in R, based on some distributions μ and ν, where μ has mean $m \in (0,\infty)$, and ν is defined as in Theorem 12.1, in terms of the ladder height distribution $\tilde\mu$ and its mean $\tilde m$. Then ξ is stationary on R_+ with intensity m^{-1}.*

Our proof depends on the following result, involving the strictly and weakly *ascending ladder times* τ_k and σ_k, defined by $\tau_0 = \sigma_0 = 0$, and then recursively

$$\begin{aligned} \tau_n &= \inf\Big\{k > \tau_{n-1};\, X_k > X_{\tau_{n-1}}\Big\}, \\ \sigma_n &= \inf\Big\{k > \sigma_{n-1};\, X_k \ge X_{\sigma_{n-1}}\Big\}, \quad n \in \mathsf{N}. \end{aligned}$$

The definitions of the strictly and weakly *descending ladder times* τ_k^- and σ_k^- are similar.

Lemma 12.6 *(ladder times and heights) Let X be a random walk in R with $X_0 = 0$ and $EX_1 \in (0,\infty]$. Then the first strictly ascending ladder time τ_1 and height X_{τ_1} satisfy $E\tau_1 < \infty$ and $EX_{\tau_1} = E\tau_1\, EX_1$.*

Proof: Since the increments ΔX_n are exchangeable, we have

$$(X_1,\dots,X_n) \stackrel{d}{=} \Big(X_n - X_{n-1},\dots,X_n - X_0\Big), \quad n \in \mathsf{N},$$

and so

$$\begin{aligned} P\{\tau_1 > n\} &= P\{X_1, \dots, X_n \le 0\} \\ &= P\Big\{X_0, \dots, X_{n-1} \ge X_n\Big\} \\ &= \sum_{k \ge 0} P\{\sigma_k^- = n\}. \end{aligned}$$

Summing over $n \ge 0$ gives

$$E\tau_1 = E\sum_{k \ge 0} 1\{\sigma_k^- < \infty\} = E\kappa.$$

Since $X_n \to \infty$ a.s., by the law of large numbers, we have $\kappa < \infty$ a.s. By the strong Markov property, it follows that κ is geometrically distributed, and so $E\tau_1 = E\kappa < \infty$.

Next we note that the differences ΔX_n, $\Delta\tau_n$, and ΔX_{τ_n} are all i.i.d. in $(0, \infty)$. The last relation then follows, as we apply the strong law of large numbers to each ratio in the equation

$$\frac{X_{\tau_n}}{n} = \frac{X_{\tau_n}}{\tau_n}\,\frac{\tau_n}{n}, \quad n \in \mathsf{N}.$$

□

Proof of Theorem 12.5: Since $m > 0$, Lemma 12.6 shows that the ascending ladder times τ_n and heights $\kappa_n = X_{\tau_n}$ have finite means, and so by Theorem 12.1 the renewal process $\zeta = \sum_n \delta_{\kappa_n}$ is stationary for the stated choice of initial distribution ν. Fixing $t \ge 0$, and writing $\sigma_t = \inf\{n;\, X_n \ge t\}$, we conclude that $X_{\sigma_t} - t$ has distribution ν. By the strong Markov property at σ_t, the sequence $X_{\sigma_t + n} - t$, $n \in \mathsf{Z}_+$, has then the same distribution as X. Since also $X_k < t$ for all $k < \sigma_t$, we get $\theta_t\xi \overset{d}{=} \xi$ on R_+, which proves the asserted stationarity.

To identify the intensity, let η_n denote the occupation measure of the sequence $X_k - \kappa_n$, $\tau_n \le k < \tau_{n+1}$, and note that $\kappa_n \perp\!\!\!\perp \eta_n \overset{d}{=} \eta_0$ for $n \in \mathsf{Z}_+$, by the strong Markov property. Using Fubini's theorem, we get

$$\begin{aligned} E\xi &= E\sum_{n \ge 0} \eta_n * \delta_{\kappa_n} \\ &= \sum_{n \ge 0} E\eta_n * E\delta_{\kappa_n} \\ &= E\eta_0 * E\sum_{n \ge 0} \delta_{\kappa_n} = E\eta_0 * E\zeta. \end{aligned}$$

Noting that $m\,E\zeta = \lambda$ by Theorem 12.1, that $E\eta_0\mathsf{R} = E\eta_0\mathsf{R}_- = E\tau_1$, and that $\tilde{m} = m\,E\tau_1$ by Lemma 12.6, we get on R_+

$$E\xi = \frac{E\tau_1}{\tilde{m}}\,\lambda = \frac{\tilde{m}}{m \cdot \tilde{m}}\,\lambda = m^{-1}\,\lambda,$$

which shows that ξ has intensity m^{-1}. □

For a random walk X, based on a distribution μ with possibly infinite mean $m \neq 0$, we proceed to examine the asymptotic behavior of the associated occupation measure ξ and its intensity $E\xi$. Say that μ is *non-arithmetic*, if the additive subgroup generated by μ is dense in R. The results in the arithmetic case are similar and more elementary. When μ is restricted to $(0, \infty)$, the result for $E\xi$ reduces to the classical *renewal theorem.*

Theorem 12.7 *(extended renewal theorem, Blackwell, Feller & Orey) Let ξ be the occupation measure of a random walk X in R, based on some distributions μ and ν, where μ is non-arithmetic with mean $m \in [-\infty, \infty] \setminus \{0\}$. When $m \in (0, \infty)$, let $\tilde{\xi}$ denote the stationary version in Theorem 12.5, and put $\tilde{\xi} = 0$ otherwise. Then as $t \to \infty$,*

(i) $\theta_t \xi \xrightarrow{vd} \tilde{\xi}$,

(ii) $\theta_t E\xi \xrightarrow{v} E\tilde{\xi} = (m^{-1} \vee 0)\, \lambda$.

Our proof relies on two additional lemmas. When $m < \infty$, the crucial step is to prove convergence of the ladder height distributions of the sequences $X - t$. This is most easily accomplished by a coupling argument.

Lemma 12.8 *(delay distributions) For X as in Theorem 12.7 with $m \in (0, \infty)$, let ν_t be the distribution of the first ladder height ≥ 0 of the sequence $X - t$. Then $\nu_t \xrightarrow{w} \tilde{\nu}$ as $t \to \infty$.*

Proof: Let α and α' be independent random variables with distributions ν and $\tilde{\nu}$. Choose some i.i.d. sequences $(\xi_k) \perp\!\!\!\perp (\vartheta_k)$, independent of α and α', with $\mathcal{L}(\xi_k) = \mu$ and $P\{\vartheta_k = \pm 1\} = \frac{1}{2}$. Then

$$M_n = \alpha' - \alpha - \sum_{k \leq n} \vartheta_k \, \xi_k, \quad n \in \mathsf{Z}_+,$$

is a random walk, based on a non-arithmetic distribution with mean 0. By Theorem 12.3 and Corollary 12.4, the range $\{M_n\}$ is then dense in R, and so for any $\varepsilon > 0$, the optional time $\tau = \inf\{n;\, M_n \in [0, \varepsilon]\}$ is a.s. finite.

Next define $\vartheta'_k = \pm \vartheta_k$, with sign '+' iff $k > \tau$, and note that $\{\alpha', (\xi_k, \vartheta_k)\} \overset{d}{=} \{\alpha', (\xi_k, \vartheta'_k)\}$. Let $\kappa_1, \kappa_2, \ldots$ and $\kappa'_1, \kappa'_2, \ldots$ be the values of k with $\vartheta_k = 1$ or $\vartheta'_k = 1$, respectively. Writing

$$X_n = \alpha + \sum_{i \leq n} \xi_{\kappa_i}, \quad X'_n = \alpha' + \sum_{i \leq n} \xi_{\kappa'_i}, \qquad n \in \mathsf{Z}_+,$$

we note that the sequences X and X' are again random walks based on μ, with initial distributions ν and $\tilde{\nu}$, respectively. Putting $\sigma_\pm = \sum_{k \leq \tau} (\pm \vartheta_k \vee 0)$, we further note that

$$X'_{\sigma_- + n} - X_{\sigma_+ + n} = M_\tau \in [0, \varepsilon], \quad n \in \mathsf{Z}_+.$$

Putting $\beta = X^*_{\sigma_+} \vee X'^*_{\sigma_-}$, and considering the first entry of X and X' into the interval $[t, \infty)$, we obtain for any $r \geq \varepsilon$

$$\begin{aligned}\tilde{\nu}[\varepsilon, r] - P\{\beta \geq t\} &\leq \nu_t[0, r] \\ &\leq \tilde{\nu}[0, r + \varepsilon] + P\{\beta \geq t\}.\end{aligned}$$

Letting $t \to \infty$ and then $\varepsilon \to 0$, and noting that $\tilde{\nu}\{0\} = 0$ by stationarity, we get $\nu_t[0, r] \to \tilde{\nu}[0, r]$, which implies $\nu_t \xrightarrow{w} \tilde{\nu}$. □

To deduce (ii) from (i) in the main theorem, we need the following technical result, which will also play a crucial role when $m = \infty$.

Lemma 12.9 *(uniform integrability) Let ξ be the occupation measure of a transient random walk X in R^d, with arbitrary initial distribution. Then for any bounded $B \in \mathcal{B}^d$, the random variables $\xi(B + x)$, $x \in \mathsf{R}^d$, are uniformly integrable.*

Proof: Fix any $x \in \mathsf{R}^d$, and put $\tau = \inf\{n;\ X_n \in B + x\}$. Letting η be the occupation measure of an independent random walk, starting at 0, we get by the strong Markov property

$$\begin{aligned}\xi(B + x) &\stackrel{d}{=} \eta(B + x - X_\tau)\, 1\{\tau < \infty\} \\ &\leq \eta(B - B).\end{aligned}$$

It remains to note that $E\eta(B - B) < \infty$ by Theorem 12.3, since X is transient. □

Proof of Theorem 12.7 ($m < \infty$): By Lemma 12.9, it is enough to prove (i). If $m < 0$, then $X_n \to -\infty$ a.s., by the law of large numbers, and so $\theta_t \xi = 0$ for sufficiently large t, and (i) follows. If instead $m \in (0, \infty)$, then $\nu_t \xrightarrow{w} \tilde{\nu}$ by Lemma 12.8, and we may choose some random variables α_t and α with distributions ν_t and $\tilde{\nu}$, respectively, such that $\alpha_t \to \alpha$ a.s. We may also introduce the occupation measure ξ_0 of an independent random walk starting at 0.

Now consider any $f \in \hat{C}_{\mathsf{R}_+}$, and extend f to R by putting $f(x) = 0$ for $x < 0$. Note that $\xi_0\{-\alpha\} = 0$ a.s., since $\tilde{\nu} \ll \lambda$. By the strong Markov property and dominated convergence, we have[2]

$$\begin{aligned}(\theta_t \xi) f &\stackrel{d}{=} \int f(\alpha_t + x)\, \xi_0(dx) \\ &\to \int f(\alpha + x)\, \xi_0(dx) \stackrel{d}{=} \tilde{\xi} f.\end{aligned}$$

and (i) follows by Theorem 4.11.

[2]Recall that $\mu f = \int f\, d\mu$.

($m = \infty$): Here it is clearly enough to prove (ii). By the strong Markov property, we have $E\xi = \nu * E\chi * E\zeta$, where χ denotes the occupation measure of the ladder height sequence of $X - X_0$, and ζ is the occupation measure of $X - X_0$, prior to the first ladder time. Since $E\zeta\mathsf{R} = E\zeta\mathsf{R}_- < \infty$, by Lemma 12.6, it suffices by dominated convergence to show that $\theta_t E\chi \xrightarrow{v} 0$. Since the ladder heights again have infinite mean, by Lemma 12.6, we may assume instead that $\nu = \delta_0$, and let μ be an arbitrary distribution on $(0,\infty)$ with mean $m = \infty$.

Writing $I = [0,1]$, we see from Lemma 12.9 that $E\xi(I+t)$ is bounded. Define $b = \limsup_t E\xi(I+t)$, and choose some $t_k \to \infty$ with $E\xi(I+t_k) \to b$. Subtracting the bounded measures μ^{*j} with $j < n$, we get $(\mu^{*n} * E\xi)(I+t_k) \to b$ for all $n \in \mathsf{Z}_+$. Using the reverse Fatou lemma, we get for any $B \in \mathcal{B}_{\mathsf{R}_+}$

$$\begin{aligned}
&\liminf_{k\to\infty} E\xi(I - B + t_k)\,\mu^{*n}B \\
&\qquad \geq \liminf_{k\to\infty} \int_B E\xi(I - x + t_k)\,\mu^{*n}(dx) \\
&\qquad = b - \limsup_{k\to\infty} \int_{B^c} E\xi(I - x + t_k)\,\mu^{*n}(dx) \\
&\qquad \geq b - \int_{B^c} \limsup_{k\to\infty} E\xi(I - x + t_k)\,\mu^{*n}(dx) \\
&\qquad \geq b - \int_{B^c} b\,\mu^{*m}(dx) = b\,\mu^{*n}B.
\end{aligned}$$

Since n was arbitrary, we conclude that the lim inf on the left is $\geq b$ for any B with $E\xi B > 0$.

Now fix any $h > 0$ with $\mu(0,h] > 0$, and write $J = [0,a]$ with $a = h+1$. Noting that $E\xi[r, r+h] > 0$ for all $r \geq 0$, we get

$$\liminf_{k\to\infty} E\xi(J + t_k - r) \geq b, \quad r \geq a.$$

We also see from the identity $\delta_0 = (\delta_0 - \mu) * E\xi$ that

$$\begin{aligned}
1 &= \int_0^{t_k} \mu(t_k - x, \infty)\,E\xi(dx) \\
&\geq \sum_{n\geq 1} \mu(na,\infty)\,E\xi(J + t_k - na).
\end{aligned}$$

As $k \to \infty$, we get $1 \geq b\sum_{n\geq 1}\mu(na,\infty)$, by Fatou's lemma. Here the sum diverges since $m = \infty$, and so $b = 0$, which implies $\theta_t E\xi \xrightarrow{v} 0$. □

We turn to the special case where $E\xi \ll \lambda$ on $(0,\infty)$, with a uniformly continuous density. The result will be needed in the proof of Theorem 12.36 below.

Corollary 12.10 *(renewal density) Let ξ be a renewal process, based on a distribution μ with mean $m \in (0,\infty]$, possessing a continuous density p on R_+ with limit $p_\infty = 0$. Then $E\xi$ has a uniformly continuous density $q = \sum_{n\geq 1} p^{*n}$ on $(0,\infty)$ with limit $q_\infty = m^{-1}$.*

Proof: First we consider the convolutions $\mu * \rho$, where μ has has a continuous density p with bounded support in $[0,a]$, and ρ satisfies $m(r) \equiv \sup_x \rho[x, x+r] < \infty$, for every $r \geq 0$. Writing $\delta_p(h) = \sup_x |p_{x+h} - p_x|$, we get as $h \to 0$

$$\begin{aligned} \Big| (p * \rho)_{x+h} - (p * \rho)_x \Big| &\leq \int \Big| p_{x+h-t} - p_{x-t} \Big| \rho(dt) \\ &\leq \delta_p(h)\, m(a+h) \to 0. \end{aligned}$$

If instead μ is bounded, and ρ has a uniformly continuous density q, then as $h \to 0$,

$$\begin{aligned} \Big| (q * \mu)_{x+h} - (q * \mu)_x \Big| &\leq \int \Big| q_{x+h-t} - q_{x-t} \Big| \mu(dt) \\ &\leq \delta_q(h)\, \|\mu\| \to 0. \end{aligned}$$

Thus, in both cases, $\mu * \rho$ has a uniformly continuous density. We also note that, if μ and ρ have densities p and q, then this continuous density agrees with $p * q$.

Returning to the renewal context, we may write $\mu = \frac{1}{2}(\alpha + \beta)$, where α and β are probability measures with uniformly continuous densities p' and p'', and p' has bounded support. Letting $\xi = \sum_{n \geq 1} \delta_{\sigma_n}$ with $\sigma_n = \gamma_1 + \cdots + \gamma_n$, we may choose a Bernoulli sequence $\vartheta_1, \vartheta_2, \ldots$, such that the pairs (γ_n, ϑ_n) are i.i.d., and $\mathcal{L}(\gamma_k \,|\, \vartheta_k = 1) \equiv \alpha$. Introducing the optional time $\tau = \inf\{n \geq 1;\, \vartheta_n = 1\}$, and writing

$$\rho_n = \beta + \beta^{*2} + \cdots + \beta^{*(n-1)} * (\delta_0 + \alpha), \quad n \geq 1,$$

we get on $(0, \infty)$

$$\begin{aligned} \rho \equiv E\xi &= \sum_{n \geq 1} E(\, \xi;\, \tau = n\,) \\ &= \sum_{n \geq 1} P\{\tau = n\}\, E(\, \xi \,|\, \tau = n\,) \\ &= \sum_{n \geq 1} 2^{-n} \Big\{ \rho_n + \beta^{*(n-1)} * \alpha * \rho \Big\} \\ &= \sum_{n \geq 1} 2^{-n} \rho_n + \alpha * \rho * \sum_{n \geq 1} 2^{-n} \beta^{*(n-1)} \\ &= \nu' + \alpha * \rho * \nu'', \end{aligned}$$

where $\|\nu'\| = \sum_{n \geq 1} n\, 2^{-n} < \infty$ and $\|\nu''\| = \sum_{n \geq 1} 2^{-n} = 1$. Hence,

$$\begin{aligned} \rho = \sum_{n \geq 1} \mu^{*n} &= \mu * \sum_{n \geq 0} \mu^{*n} \\ &= \mu * (\delta_0 + \rho) \\ &= \mu * \Big(\delta_0 + \nu' + \alpha * \rho * \nu'' \Big) \\ &= \mu * (\delta_0 + \nu') + (\alpha * \rho) * (\nu'' * \mu). \end{aligned}$$

Applying the two special cases above, we see that the density q of ρ is uniformly continuous. The limiting value of q then follows from Theorem 12.7. $\square$

12.2 Excursion Local Time and Point Process

Consider an rcll process X in a Polish space S, adapted to a right-continuous and complete filtration $\mathcal{F} = (\mathcal{F}_t)$. It is said to be *regenerative* at a point $a \in S$, if it satisfies the strong Markov property at a, so that for any optional time τ,

$$\mathcal{L}(\theta_\tau X | \mathcal{F}_\tau) = P_a \text{ a.s. on the set } \{\tau < \infty, X_\tau = a\},$$

for some probability measure P_a, on the path space $D(\mathsf{R}_+, S)$ of X. In particular, a strong Markov process in S is regenerative at every point.

The associated random set $\Xi = \{t \geq 0; X_t = a\}$ is regenerative in the same sense. From the right continuity of X, we see that $\Xi \ni t_n \downarrow t$ implies $t \in \Xi$, which means that the points of $\overline{\Xi} \setminus \Xi$ are isolated from the right. In particular, the first entry times $\tau_r = \inf\{t \geq r; t \in \Xi\}$ lie in Ξ, a.s. on $\{\tau_r < \infty\}$. Since $(\overline{\Xi})^c$ is open and hence a countable union of disjoint intervals (u, v), it also follows that Ξ^c is a countable union of intervals of the form (u, v) or $[u, v)$. If we are only interested in regeneration at a fixed point a, we may further assume that $X_0 = a$, or $0 \in \Xi$ a.s. The mentioned properties of Ξ will be assumed below, even when no reference is made to an underlying process X.

First, we may classify of regenerative sets into five different categories. Recall that a set $B \subset \mathsf{R}_+$ is said to be *nowhere dense* if $(\overline{B})^c = \emptyset$, and *perfect* if it is closed and has no isolated points.

Theorem 12.11 *(regenerative sets) Every regenerative set Ξ satisfies exactly one of these conditions:*

(i) *Ξ is a.s. locally finite,*

(ii) *$\Xi = \mathsf{R}_+$ a.s.,*

(iii) *Ξ is a.s. a locally finite union of disjoint intervals, whose lengths are i.i.d. exponentially distributed,*

(iv) *Ξ is a.s. nowhere dense with no isolated points, and $\overline{\Xi} = \operatorname{supp}(1_\Xi \lambda)$,*

(v) *Ξ is a.s. nowhere dense with no isolated points, and $\lambda\Xi = 0$.*

Case (iii) *is excluded when Ξ is closed.*

In case (i), Ξ is a generalized renewal process based on a distribution μ on $(0, \infty]$. In cases (ii)−(iv), Ξ is known as a *regenerative phenomenon*[3], whose distribution is determined by the so-called *p-function* $p(t) = P\{t \in \Xi\}$. Our proof of Theorem 12.11 is based on the following dichotomies.

Lemma 12.12 *(local dichotomies) For any regenerative set Ξ,*

(i) *either $(\overline{\Xi})^o = \emptyset$ a.s., or $\overline{\Xi^o} = \overline{\Xi}$ a.s.,*

(ii) *either all points of Ξ are isolated a.s., or a.s. none of them is,*

(iii) *either $\lambda\Xi = 0$ a.s., or* $\operatorname{supp}(1_\Xi \lambda) = \overline{\Xi}$ *a.s.*

[3]in the sense of Kingman

Proof: We may take $\mathcal{F}$ to be the right-continuous and complete filtration induced by Ξ, which allows us to use a canonical notation. By the regenerative property, we have for any optional time τ

$$\begin{aligned} P\{\tau = 0\} &= E\Big\{P(\tau = 0 \,|\, \mathcal{F}_0);\ \tau = 0\Big\} \\ &= (P\{\tau = 0\})^2, \end{aligned}$$

and so $P\{\tau = 0\} = 0$ or 1. If σ is another optional time, then $\tau' = \sigma + \tau \circ \theta_\sigma$ is again optional, and so for any $h \geq 0$,

$$\begin{aligned} P\{\,\tau' - h \leq \sigma \in \Xi\,\} &= P\{\,\tau \circ \theta_\sigma \leq h,\ \sigma \in \Xi\,\} \\ &= P\{\tau \leq h\}\, P\{\sigma \in \Xi\}, \end{aligned}$$

which gives $\mathcal{L}(\tau' - \sigma \,|\, \sigma \in \Xi) = \mathcal{L}(\tau)$. In particular, $\tau = 0$ a.s. implies $\tau' = \sigma$, a.s. on $\{\sigma \in \Xi\}$.

(i) Apply the previous observations to the optional times $\tau = \inf \Xi^c$ and $\sigma = \tau_r$. If $\tau > 0$ a.s., then $\tau \circ \theta_{\tau_r} > 0$ a.s. on $\{\tau_r < \infty\}$, and so $\tau_r \in \overline{\Xi^o}$ a.s. on the same set. Since the set $\{\tau_r;\ r \in \mathsf{Q}_+\}$ is dense in $\overline{\Xi}$, it follows that $\overline{\Xi} = \overline{\Xi^o}$ a.s. If instead $\tau = 0$ a.s., then $\tau \circ \theta_{\tau_r} = 0$ a.s. on $\{\tau_r < \infty\}$, and so $\tau_r \in \overline{\Xi^c}$ a.s. on the same set. Hence, $\overline{\Xi} \subset \overline{\Xi^c}$ a.s., and therefore $\overline{\Xi^c} = \mathsf{R}_+$ a.s. It remains to note that $\overline{\Xi^c} = (\overline{\Xi})^c$, since Ξ^c is a disjoint union of intervals (u, v) or $[u, v)$.

(ii) Define $\tau = \inf(\Xi \setminus \{0\})$. If $\tau = 0$ a.s., then $\tau \circ \theta_{\tau_r} = 0$ a.s. on $\{\tau_r < \infty\}$. Since every isolated point of Ξ is of the form τ_r for some $r \in \mathsf{Q}_+$, it follows that Ξ has no isolated points. If instead $\tau > 0$ a.s., we may define recursively some optional times σ_n by $\sigma_{n+1} = \sigma_n + \tau \circ \theta_{\sigma_n}$, starting from $\sigma_1 = \tau$. Then the σ_n form a renewal process based on $\mathcal{L}(\tau)$, and so $\sigma_n \to \infty$ a.s., by the law of large numbers. Thus, $\Xi = \{\sigma_n < \infty;\ n \in \mathsf{Z}_+\}$ a.s. with $\sigma_0 = 0$, and a.s. all points of Ξ are isolated.

(iii) Take[4] $\tau = \inf\{t > 0;\ (1_\Xi \cdot \lambda)_t > 0\}$. If $\tau = 0$, then $\tau \circ \theta_{\tau_r} = 0$ a.s. on $\{\tau_t < \infty\}$, and so $\tau_r \in \operatorname{supp}(1_\Xi \lambda)$ a.s. on the same set. Hence, $\overline{\Xi} \subset \operatorname{supp}(1_\Xi \lambda)$ a.s., and the two sets agree a.s. If instead $\tau > 0$ a.s., then $\tau = \tau + \tau \circ \theta_\tau > \tau$ a.s. on $\{\tau < \infty\}$, which implies $\tau = \infty$ a.s., showing that $\lambda \Xi = 0$ a.s. $\Box$

Proof of Theorem 12.11: First, we may use Lemma 12.12 (ii) to eliminate case (i). Here Ξ is clearly a renewal process, which is a.s. locally finite. Next, use parts (i) and (iii) of the same lemma to separate cases (iv) and (v). Excluding (i) and (iv)$-$(v), we see that Ξ is a.s. a locally finite union of intervals of i.i.d. lengths. Writing $\gamma = \inf \Xi^c$, and using the regenerative property at τ_s, we get

$$\begin{aligned} P\{\gamma > s + t\} &= P\{\,\gamma \circ \theta_s > t,\ \gamma > s\,\} \\ &= P\{\gamma > s\}\, P\{\gamma > t\}, \quad s, t \geq 0. \end{aligned}$$

[4]For distinction, we write $1_\Xi \cdot \lambda$ for the process $\int_0^t 1\{s \in \Xi\}\, ds$ and $1_\Xi \lambda$ for the associated measure.

Thus, the monotone function $f_t = P\{\gamma > t\}$ satisfies the Cauchy equation $f_{s+t} = f_s f_t$ with initial condition $f_0 = 1$, and so $f_t \equiv e^{-ct}$ for some constant $c \geq 0$. Here $c = 0$ corresponds to case (ii), and $c > 0$ to case (iii). If Ξ is closed, then $\gamma \in \Xi$ on $\{\gamma < \infty\}$, and the regenerative property at γ yields $\gamma = \gamma + \gamma \circ \theta_\gamma > \gamma$ a.s. on $\{\gamma < \infty\}$, which implies $\gamma = \infty$ a.s., corresponding to case (ii). □

The complement $(\overline{\Xi})^c$ is a countable union of disjoint intervals (σ, τ), each supporting an *excursion* of X away from a. The shifted excursions $Y_{\sigma,\tau} = \theta_\sigma X^\tau$ are rcll processes in their own right. Let D_0 denote the space of possible excursion paths, write $l(x)$ for the length of excursion $x \in D_0$, and put $D_h = \{x \in D_0;\, l(x) > h\}$. Write κ_h for the number of excursions of X longer than h. We need some basic facts:

Lemma 12.13 *(long excursions) Fix any $h > 0$, and allow even $h = 0$, when the recurrence time is positive. Then*

(i) *we have the only possibilities $\kappa_h = 0$ a.s., $\kappa_h = \infty$ a.s., or $\kappa_h \in \mathsf{N}$ is geometrically distributed with mean $m_h \in [1, \infty)$,*

(ii) *the processes Y_j^h are i.i.d. in D_h, and $l(Y_{\kappa_h}^h) = \infty$ a.s. when $\kappa_h \in (0, \infty)$ a.s.*

Proof: Let κ_t^h be the number of excursions in D_h, completed by time $t \in (0, \infty]$, and note that $\kappa_{\tau_t}^h > 0$ when $\tau_t = \infty$. Writing $p_h = P\{\kappa_h > 0\}$, we get for any $t > 0$

$$\begin{aligned} p_h &= P\{\kappa_{\tau_t}^h > 0\} + P\Big\{\kappa_{\tau_t}^h = 0,\, \kappa_h \circ \theta_{\tau_t} > 0\Big\} \\ &= P\{\kappa_{\tau_t}^h > 0\} + P\{\kappa_{\tau_t}^h = 0\}\, p_h. \end{aligned}$$

As $t \to \infty$, we obtain $p_h = p_h + (1 - p_h)\, p_h$, which implies $p_h = 0$ or 1.

Assuming $p_h = 1$, let $\sigma = \sigma_1$ be the right endpoint of the first D_h-excursion, and define recursively $\sigma_{n+1} = \sigma_n + \sigma_1 \circ \theta_{\sigma_n}$, starting with $\sigma_0 = 0$. Writing $q = P\{\sigma < \infty\}$, and using the regenerative property at each σ_n, we obtain $P\{\kappa_h > n\} = q^n$. Thus, $q = 1$ yields $\kappa_h = \infty$ a.s., whereas for $q < 1$, we see that κ_h is geometrically distributed with mean $m_h = (1-q)^{-1}$, proving (i). Even (ii) is clear by regeneration at each σ_n. □

Write $\hat{h} = \inf\{h > 0;\, \kappa_h = 0 \text{ a.s.}\}$, and for any $h < \hat{h}$, let ν_h denote the common distribution of all excursions in D_h. We show how the measures ν_h can be combined into a single σ-finite measure ν.

Lemma 12.14 *(excursion measure, Itô) There exists a σ-finite measure ν on D_0, such that $\nu D_h \in (0, \infty)$ and $\nu_h = \nu(\cdot \mid D_h)$ for all $h \in (0, \hat{h})$. The measure ν is unique up to a normalization, and it is bounded iff the recurrence time is a.s. positive.*

Proof: Fix any $h \le k$ in $(0, \hat{h})$, and let $Y_1^h, Y_2^h, \dots$ be such as in Lemma 12.13. Then the first excursion in D_k is the first process Y_j^h that also belongs to D_k, and since the Y_j^h are i.i.d. ν_h, we have

$$\nu_k = \nu_h(\,\cdot\,|\,D_k), \quad 0 < h \le k < \hat{h}. \tag{1}$$

Now fix any $k \in (0, \hat{h})$, and define $\tilde{\nu}_h = \nu_h/\nu_h D_k$ for every $h \in (0, k]$. Then (1) yields $\tilde{\nu}_{h'} = \tilde{\nu}_h(\cdot \cap D_{h'})$ for any $h \le h' \le k$, and so $\tilde{\nu}$ increases as $h \to 0$ toward a σ-finite measure ν on D_0 with $\nu(\cdot \cap D_h) = \tilde{\nu}_h$ for all $h \le k$. For any $h \in (0, \hat{h})$, we get

$$\begin{aligned} \nu(\cdot\,|\,D_h) &= \tilde{\nu}_{h\wedge k}(\cdot\,|\,D_h) \\ &= \nu_{h\wedge k}(\cdot\,|\,D_h) = \nu_h. \end{aligned}$$

If ν' is another measure with the stated properties, then

$$\frac{\nu(\cdot \cap D_h)}{\nu D_k} = \frac{\nu_h}{\nu_h D_k} = \frac{\nu'(\cdot \cap D_h)}{\nu' D_k}, \quad h \le k < \hat{h}.$$

As $h \to 0$ for fixed k, we get $\nu = r\nu'$ with $r = \nu D_k/\nu' D_k$.

If the recurrence time is positive, then (1) remains true for $h = 0$, and we may take $\nu = \nu_0$. Otherwise, let $h \le k$ in $(0, \hat{h})$, and let $\kappa_{h,k}$ be the number of D_h-excursions, up to the first completed excursion in D_k. For fixed k, we have $\kappa_{h,k} \to \infty$ a.s. as $h \to 0$, since $\overline{\Xi}$ is perfect and nowhere dense. Noting that $\kappa_{h,k}$ is geometrically distributed with mean

$$E\kappa_{h,k} = \frac{1}{\nu_h D_k} = \frac{1}{\nu(D_k|\,D_h)} = \frac{\nu D_h}{\nu D_k},$$

we get $\nu D_h \to \infty$, which shows that ν is unbounded. □

We turn to the fundamental theorem of excursion theory, describing the excursion structure of X, in terms of a Poisson process η on $\mathsf{R}_+ \times D_0$, and a diffuse random measure ξ on R_+. The measure ξ, with associated cumulative process $L_t = \xi[0, t]$, is unique up to a normalization, and will be referred to as the *excursion local time* of X at a. Similarly, η is referred to as the *excursion point process* of X at a.

Theorem 12.15 *(excursion local time and point process, Lévy, Itô) Let the process X be regenerative at a, and such that the set $\Xi = \{t;\, X_t = a\}$ has a.s. perfect closure. Then there exist a non-decreasing, continuous, adapted process L on R_+ with support $\overline{\Xi}$ a.s., a Poisson process η on $\mathsf{R}_+ \times D_0$ with intensity of the form $\lambda \otimes \nu$, and a constant $c \ge 0$, such that $1_\Xi \cdot \lambda = c\,L$ a.s., and the excursions of X with associated L-values are given by the restriction of η to $[0, L_\infty]$. Here the product $\nu \cdot L$ is a.s. unique.*

For reference, we note a useful independence property:

Corollary 12.16 *(conditional independence) For any regenerative process X with local time ξ, the excursions are conditionally independent given ξ, with distributions given by the kernel ν.*

Partial proof of Theorem 12.15: When $E\gamma = m > 0$, we may define $\nu = \nu_0/m$, and introduce a Poisson process η on $\mathsf{R}_+ \times D_0$ with intensity $\lambda \otimes \nu$, say with points $(\sigma_j, \tilde{Y}_j)$, $j \in \mathsf{N}$. Putting $\sigma_0 = 0$, we see from Theorem 3.12 (i) that the differences $\tilde{\gamma}_j = \sigma_j - \sigma_{j-1}$ are i.i.d. exponentially distributed with mean m. By Theorem 3.2 (i), the processes $\tilde{Y}_j$ are further independent of the σ_j and i.i.d. ν_0. Writing $\tilde{\kappa} = \inf\{j;\, l(\tilde{Y}_j) = \infty\}$, we see from Lemmas 12.11 and 12.13 that

$$\{\, \gamma_j, Y_j;\, j \le \kappa \,\} \stackrel{d}{=} \{\, \tilde{\gamma}_j, \tilde{Y}_j;\, j \le \tilde{\kappa} \,\}, \tag{2}$$

where the quantities on the left are the holding times and subsequent excursions of X. By the transfer theorem (FMP 6.10), we may redefine η to make (2) hold a.s. The stated conditions are then fulfilled with $L = 1_\Xi \cdot \lambda$.

When $E\gamma = 0$, we may again choose η to be Poisson $\lambda \otimes \nu$, now with ν defined as in Lemma 12.14. For any $h \in (0, \hat{h})$, we may enumerate the points of η in $\mathsf{R}_+ \times D_h$ as $(\sigma_j^h, \tilde{Y}_j^h)$, $j \in \mathsf{N}$, and define $\tilde{\kappa}_h = \inf\{j;\, l(Y_j^h) = \infty\}$. The processes $\tilde{Y}_j^h$ are i.i.d. ν_h, and by Lemma 12.13 we have

$$\{\, \gamma_j, Y_j;\, j \le \kappa \,\} \stackrel{d}{=} \{\, \tilde{\gamma}_j, \tilde{Y}_j;\, j \le \tilde{\kappa} \,\}, \quad h \in (0, \hat{h}). \tag{3}$$

Since the excursions in D_h form nested subarrays, the entire collections have the same finite-dimensional distributions, and so the transfer theorem allows us to redefine η, to make (3) hold a.s.

Let τ_j^h be the right endpoint of the j-th excursion in D_h, and define

$$L_t = \inf\left\{\sigma_j^h;\, hj > 0,\, \tau_j^h \ge t\right\}, \quad t \ge 0.$$

It is easy to verify that, for any $t \ge 0$ and $h, j > 0$,

$$L_t < \sigma_j^h \quad \Rightarrow \quad t \le \tau_j^h \quad \Rightarrow \quad L_t \le \sigma_j^h. \tag{4}$$

To see that L is continuous, we may assume that (3) holds identically. Since ν is unbounded, we may further assume that the set $\{\sigma_j^h;\, h, j > 0\}$ is dense in the interval $[0, L_\infty]$. If $\Delta L_t > 0$, there exist some $i, j, h > 0$ with $L_{t-} < \sigma_i^h < \sigma_j^h < L_{t+}$. Then (4) yields $t - \varepsilon \le \tau_i^h < \tau_j^h \le t + \varepsilon$ for every $\varepsilon > 0$, which is impossible. Thus, $\Delta L_t = 0$ for all t.

To prove that $\overline{\Xi} \subset \operatorname{supp} L$ a.s., we may assume that $\overline{\Xi}_\omega$ is perfect and nowhere dense, for every $\omega \in \Omega$. If $t \in \overline{\Xi}$, then for every $\varepsilon > 0$, there exist some $i, j, h > 0$ with $t - \varepsilon < \tau_j^h < \tau_j^h < t + \varepsilon$. By (4), we get $L_{t-\varepsilon} \le \sigma_i^h < \sigma_j^h \le L^{t+\varepsilon}$, and so $L_{t-\varepsilon} < L_{t+\varepsilon}$ for all $\varepsilon > 0$, which implies $t \in \operatorname{supp} L$. □

In the perfect case, it remains to establish the a.s. relation $1_\Xi \cdot \lambda = c\, L$ for some constant $c \ge 0$, and to show that L is unique and adapted. The former claim is a consequence of Theorem 12.18 below, whereas the latter statement follows immediately from the next theorem.

Most standard constructions of L are special cases of the following result, where $\eta_t A$ denotes the number of excursions in a set $A \in D_0$ completed by time $t \geq 0$, so that η is an adapted, measure-valued process on D_0. The result may be compared with the global construction in Corollary 12.24.

Theorem 12.17 *(local approximation) Let $A_1, A_2, \ldots \in D_0$ with finite measures $\nu A_n \to \infty$. Then*

$$\sup_{t \leq u} \left| \frac{\eta_t A_n}{\nu A_n} - L_t \right| \xrightarrow{P} 0, \quad u \geq 0.$$

The convergence holds a.s. when the A_n are increasing.

In particular, $\eta_t D_h / \nu D_h \to L_t$ a.s. as $h \to 0$ for fixed t, which shows that L is a.s. determined by the regenerative set Ξ.

Proof: Let ξ be the excursion point process of Theorem 12.15, and put $\xi_s = \xi([0, s] \times \cdot)$. If the A_n are increasing, then for any unit rate Poisson process N on R_+, we have $(\xi_s A_n) \overset{d}{=} (N_{s\,\nu A_n})$ for every $s \geq 0$. Since $t^{-1} N_t \to 1$ a.s., by the law of large numbers, we get

$$\frac{\xi_s A_n}{\nu A_n} \to s \ \text{ a.s.}, \quad s \geq 0.$$

Using the monotonicity of both sides, we may strengthen this to

$$\sup_{s \leq r} \left| \frac{\xi_s A_n}{\nu A_n} - s \right| \to 0 \ \text{ a.s.}, \quad r \geq 0.$$

The desired convergence now follows, by the continuity of L, and the fact that $\xi_{L_t -} \leq \eta_t \leq \xi_{L_t}$ for all $t \geq 0$.

Dropping the monotonicity condition, but assuming for simplicity that $\nu A_n \uparrow \infty$, we may choose a non-decreasing sequence $A'_n \in D_0$, such that $\nu A_n = \nu A'_n$ for all n. Then clearly $(\xi_s A_n) \overset{d}{=} (\xi_s A'_n)$ for fixed n, and so

$$\sup_{s \leq r} \left| \frac{\xi_s A_n}{\nu A_n} - s \right| \overset{d}{=} \sup_{s \leq r} \left| \frac{\xi_s A'_n}{\nu A'_n} - s \right| \to 0 \ \text{ a.s.}, \quad r \geq 0,$$

which implies the convergence $\xrightarrow{P} 0$ on the left. The asserted convergence now follows as before. $\square$

The local time L may be described in terms of its right-continuous inverse

$$T_s = \inf\left\{ t \geq 0;\, L_t > s \right\}, \quad s \geq 0,$$

which is shown below to be a *generalized subordinator*, defined as a non-decreasing Lévy process with possibly infinite jumps. By Theorem 3.19, T

may be represented in terms of a stationary Poisson process on the product space $\mathsf{R}_+ \times [0, \infty]$, related to the excursion point process η. Here we introduce the subset $\Xi' \subset \Xi$, obtained by omission of all points of Ξ that are isolated from the right.

Theorem 12.18 *(inverse local time) For L, ξ, η, ν, and c as in Theorem 12.15, the right-continuous inverse $T = L^{-1}$ is a generalized subordinator, with characteristics $(c,\, \nu \circ l^{-1})$ and a.s. range Ξ', satisfying $\xi = \lambda \circ T^{-1}$ and*

$$T_s = c\, s + \int_0^{s+} \!\! \int l(u)\, \eta(dr\, du), \quad s \geq 0. \tag{5}$$

Proof: We may first discard the P-null set where L fails to be continuous with support $\overline{\Xi}$. If $T_s < \infty$ for some $s \geq 0$, then $T_s \in \operatorname{supp} L = \overline{\Xi}$, by the definition of T, and since L is continuous, we get $T_s \notin \overline{\Xi} \setminus \Xi'$. Thus, $T(\mathsf{R}_+) \subset \Xi' \cup \{\infty\}$ a.s. Conversely, let $t \in \Xi'$. Then for any $\varepsilon > 0$, we have $L_{t+\varepsilon} > L_t$, and so $t \leq T \circ L_t \leq t + \varepsilon$. As $\varepsilon \to 0$, we get $T \circ L_t = t$. Thus, $\Xi' \subset T(\mathsf{R}_+)$ a.s.

The times T_s are optional, by the right continuity of $\mathcal{F}$. Furthermore, Theorem 12.17 shows that, as long as $T_s < \infty$, the process $\theta_s T - T_s$ is obtainable from $\theta_{T_s} X$ by a measurable map, independent of s. Using the regenerative property at each T_s, we conclude that T is a generalized subordinator (cf. FMP 15.11), hence admitting a representation as in Theorem 3.19. Since the jump sizes of T agree with the interval lengths in $(\overline{\Xi})^c$, we obtain (5) for some constant $c \geq 0$.

The double integral in (5) equals $\int x(\xi_s \circ l^{-1})(dx)$, which shows that T has Lévy measure $E(\xi_1 \circ l^{-1}) = \nu \circ l^{-1}$. Substituting $s = L_t$ into (5), we get a.s. for any $t \in \Xi'$

$$\begin{aligned} t = T \circ L_t &= c\, L_t + \int_0^{L_t+} \!\! \int l(u)\, \xi(dr\, du) \\ &= c\, L_t + (\Xi^c \cdot \lambda)_t, \end{aligned}$$

and so by subtraction $c\, L_t = (1_\Xi \cdot \lambda)_t$ a.s., which extends by continuity to arbitrary $t \geq 0$. Finally, noting that a.s. $T^{-1}[0,t] = [0, L_t]$ or $[0, L_t)$ for all $t \geq 0$, since T is a.s. strictly increasing, we have a.s.

$$\xi[0,t] = L_t = \lambda \circ T^{-1}[0,t], \quad t \geq 0,$$

which extends by FMP 3.3 to $\xi = \lambda \circ T^{-1}$ a.s. □

We conclude with a version of Corollary 12.2, for two-sided regenerative processes $\tilde{X}$. For a construction, we may first extend the excursion point process η of Theorem 12.15 to a stationary point process $\tilde{\eta}$ on $\mathsf{R} \times D_0$. If $\lambda\Xi = 0$ a.s., the extended process $\tilde{X}$ may be constructed directly from $\tilde{\eta}$. If instead $\lambda\Xi \neq 0$, we need to adjust the time scale accordingly, by taking the linear drift term in (5) into account. The same construction also yields an extension of the local time ξ to a random measure $\tilde{\xi}$ on R. Note the the latter is time-reversible, whereas $\tilde{X}$ is clearly not in general (since we are not reversing the excursions).

Under a moment condition, we may construct a stationary version Y of $\tilde{X}$, whose Palm distribution with respect to the associated local time random measure equals $\mathcal{L}(\tilde{X})$.

Theorem 12.19 *(Palm inversion) Let $\tilde{X}$ be a two-sided regenerative process with characteristics (c, ν), where $m = c + \nu l < \infty$. Form a process Y on R, by inserting an excursion Z on $[\sigma, \tau]$ at the origin, where $\tau = \vartheta l_Z$ and $\sigma = \tau - l_Z$, with $\tilde{X}$, Z, ϑ independent and such that $\mathcal{L}(Z) = m^{-1}(c\,\delta_0 + l \cdot \nu)$ and ϑ is $U(0, 1)$. Then Y is stationary with Palm distribution $\mathcal{L}(\tilde{X})$.*

Proof: First consider the local time random measure ξ only. By Theorem 12.18, the inverse $T = L^{-1}$ is a subordinator with characteristics $(c, \nu \circ l^{-1})$, and we introduce the associated random measure with independent increments ζ, along with its stationary extension to R. By Lemma 6.16, the associated Palm distribution at 0 may be written as $\mathcal{L}(\zeta^0) = \mathcal{L}(\zeta + \beta\,\delta_0)$, for a random variable $\beta \perp\!\!\!\perp \zeta$ with distribution $m^{-1}(c\,\delta_0 + l \cdot \nu)$. By Theorem 5.6 (ii), the corresponding spacing measure $\tilde{\zeta}$ is obtained from ξ by insertion of an interval $[\sigma, \tau)$ across the origin, where $\tau = \vartheta\, l_Z$ and $\sigma = \tau - l_Z$, and part (iii) of the same theorem shows that $\tilde{\xi}$ is again stationary. Then by Theorem 5.6 (iv), the spacing measure of $\tilde{\xi}$ equals $\mathcal{L}(\zeta)$, and so, by part (i) of the same theorem, its Palm distribution agrees with the two-sided version of $\mathcal{L}(\xi)$.

For a general regenerative process X, we introduce the disintegration[5] $\nu = (\nu \circ l^{-1}) \otimes (\nu_r)$ with respect to excursion length. By Corollary 12.16, we have $\mathcal{L}(X \,|\, \xi) = \tilde{\nu}(\xi, \cdot)$ a.s., where $\tilde{\nu}$ is the kernel, which to every diffuse measure μ on R assigns a process X, with independent excursions on the connected components of $(\operatorname{supp} \mu)^c$, each distributed according to the appropriate measure ν_r. By Lemma 6.1, we conclude that $\mathcal{L}(\xi, X \,\|\, \xi) = \mathcal{L}(\xi \,\|\, \xi) \otimes \tilde{\nu}$ a.e. $E\xi$. It remains to note that the processes $\tilde{X}$ and Y enjoy the same property of conditional independence, again with excursions distributed according to the measures ν_r. □

12.3 Semi-Martingale Local Time

Given a real-valued, continuous semi-martingale X, we define the *local time* of X at 0 as the process

$$L_t^0 = |X_t| - |X_0| - \int_0^t \operatorname{sgn}(X_s-)\, dX_s, \quad t \geq 0.$$

More generally, we define the local time L_t^x at an arbitrary point $x \in \mathsf{R}$ as the local time at 0 of the process $X_t - x$.

[5] Recall that the kernel composition $\mu \otimes \nu$ is given by $(\mu \otimes \nu) f = \int \mu(ds) \int f(s, t)\, \nu_s(dt)$.

Theorem 12.20 *(local time at 0, Tanaka) Let X be a real, continuous semi-martingale, with local time (L_t^0) at 0. Then a.s.*

(i) *L^0 is non-decreasing, continuous, and supported by the set $\Xi_0 = \{t \geq 0;\ X_t = 0\}$,*

(ii) $L_t^0 = \Big\{ -|X_0| - \inf_{s \leq t} \int_0^s \mathrm{sgn}(X_r-)\, dX_r \Big\} \vee 0, \quad t \geq 0.$

Our proof is based on Itô's formula, together with the following elementary but clever observation.

Lemma 12.21 *(supporting function, Skorohod) For any continuous function f on R_+ with $f_0 \geq 0$, there exists a non-decreasing, continuous function g on R_+, satisfying the conditions*

$$g_0 = 0, \qquad h \equiv f + g \geq 0, \qquad \int 1\{h_t > 0\}\, dg_t = 0.$$

It is given uniquely by

$$g_t = -\inf_{s \leq t} f_s \wedge 0 = \sup_{s \leq t}(-f_s) \vee 0, \quad t \geq 0.$$

Proof: The given function has clearly the required properties. To prove the uniqueness, suppose that both g and g' have the stated properties, and put $h = f + g$ and $h' = f + g'$. If $g_t < g'_t$ for some $t > 0$, define $s = \sup\{r < t;\ g_r = g'_r\}$, and note that $h' \geq h' - h = g' - g > 0$ on $(s, t]$. Hence, $g'_s = g'_t$, and so $0 < g'_t - g_t \leq g'_s - g_s = 0$, a contradiction. □

Proof of Theorem 12.20: (i) For any $h > 0$, choose a convex function $f_h \in C^2$ with $f_h(x) = -x$ for $x \leq 0$, and $f_h(x) = x - h$ for $x \geq h$. By Itô's formula (FMP 17.18), we get, a.s. for any $t \geq 0$,

$$\begin{aligned} Y_t^h &\equiv f_h(X_t) - f_h(X_0) - \int_0^t f'_h(X_s)\, dX_s \\ &= \tfrac{1}{2} \int_0^t f''_h(X_s)\, d[X]_s. \end{aligned}$$

As $h \to 0$, we have $f_h(x) \to |x|$ and $f'_h \to \mathrm{sgn}(x-)$. By dominated convergence for Itô and Stieltjes integrals (FMP 17.13), we get $(Y^h - L^0)_t^* \xrightarrow{P} 0$ for every $t > 0$. Now (i) follows, from the facts that the processes Y^h are non-decreasing and satisfy

$$\int_0^\infty 1\big\{X_s \notin [0, h]\big\}\, dY_s^h = 0 \ \text{ a.s.}, \quad h > 0.$$

(ii) This is immediate from Lemma 12.21. □

For any continuous local martingale M, we may choose a jointly continuous version of the associated local time L_t^x. More generally, we have the following regularization theorem:

Theorem 12.22 *(spatial dependence, Trotter, Yor) For any real, continuous semi-martingale X with canonical decomposition $M+A$, the associated local time has a version L_t^x, $x\in \mathsf{R}$, $t\geq 0$, which is rcll in x, uniformly for bounded t, and satisfies*

$$L_t^x - L_t^{x-} = 2\int_0^t 1\{X_s = x\}\,dA_s, \quad x\in\mathsf{R},\ t\geq 0.$$

Proof: By the definition of L, we have for any $x\in\mathsf{R}$ and $t\geq 0$

$$\begin{aligned} L_t^x &= |X_t - x| - |X_0 - x| \\ &\qquad - \int_0^t \operatorname{sgn}(X_s - x-)\,dM_s - \int_0^t \operatorname{sgn}(X_s - x-)\,dA_s. \end{aligned}$$

By dominated convergence, the last term has the stated continuity properties, with discontinuities given by the displayed formula. Since the first two terms are trivially jointly continuous in x and t, it remains to show that the first integral term, henceforth denoted by I_t^x, has a jointly continuous version.

By localization, we may then assume that the processes $X - X_0$, $[M]^{1/2}$, and $\int |dA|$ are all bounded by some constant $c<\infty$. Fixing any $p>2$, and using a BDG inequality (FMP 17.7), we get[6] for any $x<y$

$$\begin{aligned} E(I^x - I^y)_t^{*p} &\leq 2^p\, E\Big\{1_{(x,y]}(X)\cdot M\Big\}_t^{*p} \\ &\lesssim E\Big\{1_{(x,y]}(X)\cdot [M]\Big\}_t^{p/2}. \end{aligned}$$

To estimate the integral on the right, put $y-x=h$, and choose a function $f\in C^2$ with $f''\geq 2\cdot 1_{(x,y]}$ and $|f'|\leq 2h$. Then by Itô's formula,

$$\begin{aligned} 1_{(x,y]}(X)\cdot[M] &\leq \tfrac{1}{2} f''(X)\cdot[X] \\ &= f(X) - f(X_0) - f'(X)\cdot X \\ &\leq 4\,ch + |f'(X)\cdot M|, \end{aligned}$$

and another application of the same BDG inequality yields

$$\begin{aligned} E\Big\{f'(X)\cdot M\Big\}_t^{*p/2} &\lesssim E\Big(\{f'(X)\}^2\cdot[M]\Big)_t^{p/4} \\ &\leq (2\,ch)^{p/2}. \end{aligned}$$

Combining the last three estimates, we get $E(I^x - I^y)_t^{*p} \lesssim (ch)^{p/2}$, and since $p/2>1$, the desired continuity follows by the Kolmogorov–Chentsov criterion (FMP 3.23). $\square$

From this point on, we choose L to be the regularized version of the local time. Given a continuous semi-martingale X, we write ξ_t for the associated *occupation measures*, defined as below with respect to the quadratic variation process $[X]$. We show that $\xi_t \ll \lambda$ a.s. for every $t\geq 0$, and that the local time

[6]Recall that $X\cdot Y$ denotes the process $\int_0^t X_s\,dY_s$.

L_t^x provides a family of associated densities. This also leads to a simultaneous extension of the Itô and Tanaka formulas.

For any convex function f on R, the left derivative $f'(x-)$ exists and is non-decreasing and left-continuous, and we may introduce an associated measure μ_f on R, satisfying

$$\mu_f[x,y) = f'(y-) - f'(x-), \quad x \le y \text{ in } \mathsf{R}.$$

More generally, if f is the difference between two convex functions, then μ_f becomes a locally finite, signed measure.

Theorem 12.23 *(occupation density, substitution rule, Meyer, Wang) Let X be a real, continuous semi-martingale with local time L. Then outside a fixed P-null set, we have*

(i) *for any measurable function $f \ge 0$ on R,*

$$\xi_t f \equiv \int_0^t f(X_s)\, d[X]_s = \int_{-\infty}^{\infty} f(x)\, L_t^x\, dx, \quad t \ge 0,$$

(ii) *for any difference f between two convex functions,*

$$f(X_t) - f(X_0) = \int_0^t f'(X_s-)\, dX_s + \tfrac{1}{2}\int_{-\infty}^{\infty} L_t^x\, \mu_f(dx), \quad t \ge 0.$$

Replacing L by its left-continuous version, we get a corresponding formula involving the integral $f'(X+)\cdot X$.

Proof: (ii) When $f(x) = |x-a|$, the formula reduces to the definition of L_t^a. Since the equation is also trivially true for functions $f(x) = ax+b$, it extends by linearity to the case where μ_f has finite support. By linearity and a suitable truncation, it remains to prove the formula when μ_f is positive with bounded support, and $f(-\infty) = f'(-\infty) = 0$. Then for every $n \in \mathsf{N}$, we introduce the functions

$$g_n(x) = f'\big(2^{-n}[2^n x]-\big), \quad f_n(x) = \int_{-\infty}^{x} g_n(r)\, dr, \qquad x \in \mathsf{R},$$

and note that (ii) holds for every f_n. As $n \to \infty$, we get $f_n'(x-) = g_n(x-) \uparrow f'(x-)$, and so by dominated convergence (FMP 17.13),

$$f_n'(X-)\cdot X \xrightarrow{P} f'(X-)\cdot X.$$

Since also $f_n \to f$ by monotone convergence, it remains to show that

$$\int L_t^x\, \mu_{f_n}(dx) \to \int L_t^x\, \mu_f(dx).$$

Then let h be any bounded, right-continuous function on R, and note that $\mu_{f_n} h = \mu_f h_n$ with $h_n(x) = h(2^{-n}[2^n x + 1])$. Since $h_n \to h$ as $n \to \infty$, we get $\mu_f h_n \to \mu_f h$, by dominated convergence.

(i) Comparing (ii) with Itô's formula, we obtain the asserted formula for fixed $t \ge 0$ and $f \in C$. Since both sides define random measures on R, a simple approximation from Lemma 2.1 yields $\xi_t = L_t \cdot \lambda$, a.s. for fixed $t \ge 0$. Finally, by the continuity of both sides, the exceptional null set can be chosen to be independent of t. $\Box$

The last two theorems yield a spatial construction of the local time L, which may be compared with the temporal construction in Theorem 12.17.

Corollary 12.24 *(spatial approximation) Let X be a real, continuous semi-martingale with local time L and occupation measures ξ_t. Then outside a fixed P-null set, we have*

$$L_t^x = \lim_{h \downarrow 0} h^{-1} \xi_t[x, +h), \quad t \ge 0, \ x \in \mathsf{R}.$$

Finally, we show that the excursion and semi-martingale local times agree up to a normalization, whenever both approaches apply.

Theorem 12.25 *(reconciliation) Let X be a real, continuous semi-martingale with local time L, such that X is regenerative at some point $a \in \mathsf{R}$ with $P\{L_\infty^a > 0\} > 0$. Then the set $\Xi = \{t;\, X_t = a\}$ is a.s. perfect and nowhere dense, and the process L_t^a, $t \ge 0$, is a version of the excursion local time at a.*

Proof: The regenerative set $\Xi = \{t;\, X_t = a\}$ is closed, since X is continuous. If Ξ were a countable union of closed intervals, then L^a would vanish a.s., contrary to our hypothesis, and so by Lemma 12.11 it is perfect and nowhere dense. Let L be a version of the excursion local time at a, and put $T = L^{-1}$. Define $Y_s = L^a \circ T_s$ for $s < L_\infty$, and put $Y_s = \infty$ otherwise. The continuity of L^a yields $Y_{s\pm} = L^a \circ T_{s\pm}$ for every $s < L_\infty$. If $\Delta T_s > 0$, then $L^a \circ T_{s-} = L^a \circ T_s$, since (T_{s-}, T_s) is an excursion interval of X, and L^a is continuous with support in Ξ. Thus, Y is continuous on $[0, L_\infty)$.

By Theorem 12.17 and Corollary 12.24, the processes $\theta_s Y - Y_s$ are obtainable from $\theta_{T_s} X$ by a common measurable map for all $s < L_\infty$. By the regenerative property at a, Y is then a generalized subordinator, and so, by Theorem 3.19 and the continuity of Y, we have $Y_s \equiv c\, s$, a.s. on $[0, L_\infty)$, for some constant $c \ge 0$. For $t \in \Xi'$, we have a.s. $T \circ L_t = t$, and therefore

$$\begin{aligned} L_t^a &= L^a \circ (T \circ L_t) \\ &= (L^a \circ T) \circ L_t = c\, L_t, \end{aligned}$$

which extends to R_+, since the extreme members are continuous with support in Ξ. $\Box$

12.4 Moment and Palm Measures

The local time random measure ξ of Theorem 12.15 has a locally finite intensity $E\xi$, by Theorem 12.3, which ensures the existence of the associated Palm distributions Q_t, $t \ge 0$. The following evolution formula plays a basic role for our continued study of regenerative processes, along with their moment and Palm measures. As before, we write ν for the excursion measure and $l(x)$ for the length of excursion x. Let $[\sigma_t, \tau_t]$ or $[\sigma_t^-, \sigma_t^+]$ denote the excursion interval straddling t, where $\sigma_t = \tau_t = t$ when $t \in \bar{\Xi}$. Write X^t for the stopped process $X_s^t = X_{s\wedge t}$, and define $X^{s,t} = \theta_s X^t$, where $(\theta_t X)_s = X_{s+t}$, as usual.

Lemma 12.26 *(time evolution) Let X be a regenerative process with local time ξ, excursion measure ν, and Palm distributions Q_s. Then for any time $t > 0$ and measurable function $f \ge 0$,*

$$Ef\Big(X^{\sigma_t}, X^{\sigma_t,\tau_t}\Big) = \int_0^t E\xi(ds) \int Q_s(dx) \int_{l(y)>t-s} \nu(dy)\, f(x^s, y)$$
$$+ P\{t \in \Xi\} \int Q_t(dx)\, f(x^t, 0).$$

Proof: Let η denote the Poisson process of excursions, and note that η has compensator $\nu \otimes \lambda$. Summing over the excursions of X, and noting that the processes T_{r-} and $X^{T_{r-}}$ are predictable on the local time scale, we get by Theorem 9.21, followed by an elementary substitution

$$\begin{aligned}
&E\Big\{f(X^{\sigma_t}, X^{\sigma_t,\tau_t});\ \sigma_t < \tau_t\Big\} \\
&\qquad = E\iint f(X^{T_{r-}}, y)\, 1\Big\{0 < t - T_{r-} < l(y)\Big\}\, \eta(dy\,dr) \\
&\qquad = E\iint f(X^{T_{r-}}, y)\, 1\Big\{0 < t - T_{r-} < l(y)\Big\}\, \nu(dy)\, dr \\
&\qquad = E\int_0^1 \xi(ds) \int_{l(y)>t-s} f(X^s, y)\, \nu(dy).
\end{aligned}$$

Here the inner integral on the right can be written as $g(X, s)$, for a jointly measurable function $g \ge 0$, and so by Palm disintegration,

$$E\int_0^1 g(X,s)\, \xi(ds) = \int_0^1 E\xi(ds) \int g(x,s)\, Q_s(dx).$$

Using the right-continuity of X, and arguing as before, we get

$$\begin{aligned}
P\Big\{t \in \bar{\Xi} \setminus \Xi\Big\} &= P\{\sigma_t = t < \tau_{t+}\} \\
&= E\iint 1\{0 = t - T_{r-} < l(y)\}\, \eta(dy\,dr) \\
&= E\iint 1\Big\{0 = t - T_{r-} < l(y)\Big\}\, \nu(dy)\, dr \\
&= E\xi\{t\}\, \nu D = 0.
\end{aligned}$$

Hence, $\{\sigma_t = \tau_t\} = \{t \in \Xi\}$ a.s., and so

$$\begin{aligned} E\Big\{ f(X^{\sigma_t}, X^{\sigma_t,\tau_t});\ \sigma_t = \tau_t \Big\} &= P\{t \in \Xi\}\, E\Big\{ f(X^t, 0) \,\Big|\, t \in \Xi \Big\} \\ &= P\{t \in \Xi\} \int f(x^t, 0)\, Q_t(dx). \end{aligned}$$

The assertion now follows by addition of the two expressions. □

For a first application, we consider some distributional properties of the excursion endpoints σ_t and τ_t.

Corollary 12.27 *(endpoint distributions) For any $t > 0$, the distributions of σ_t and τ_t are absolutely continuous with respect to $E\xi + \delta_t$, and*

$$1_{[0,t]}\xi \perp\!\!\!\perp_{\sigma_t} \tau_t, \quad t > 0.$$

Proof: Writing $\hat\nu = \nu \circ l^{-1}$, we see from Theorem 12.26 that, for any measurable set $B \subset (0,t) \times (t,\infty)$,

$$P\Big\{(\sigma_t, \tau_t) \in B\Big\} = \int_0^t E\xi(ds) \int_{t-s}^\infty 1_B(s, s+r)\, \hat\nu(dr).$$

Hence, $\mathcal{L}(\sigma_t) \ll E\xi$ and $\mathcal{L}(\tau_t) \le E\xi * \hat\nu$, and it remains to show that $E\xi * \hat\nu \ll E\xi$. We may then replace $\hat\nu$ by $\rho_u = 1_{(u,\infty)} \cdot \hat\nu$, for arbitrary $u > 0$. Since $E\xi = \int \mu_s ds$ with $\mu_s = \mathcal{L}(T_s)$, it suffices to show that $\mu_s * \rho_u \ll \mu_s$, for all $s, u > 0$ with $\rho_u \ne 0$.

Then let $\gamma_1, \gamma_2, \ldots$ denote the jumps of T exceeding u, and let κ be the number of such jumps up to times s. Writing T' for the remaining component of T, we note that

$$\mu_s = \mathcal{L}\Big(T'_s + \textstyle\sum_{k\le\kappa} \gamma_k\Big), \qquad \mu_s * \hat\rho_u = \mathcal{L}\Big(T'_s + \textstyle\sum_{k\le\kappa+1} \gamma_k\Big),$$

where $\hat\rho_u = \rho_u/\hat\nu(u,\infty)$. Since κ is independent of T' and all γ_k, it suffices to show that $\mathcal{L}(\kappa+1) \ll \mathcal{L}(\kappa)$, which is obvious since κ is Poisson distributed with mean $s\,\hat\nu(u,\infty) > 0$. The last assertion may be proved by similar arguments. □

We turn to a fundamental independence property, for the intervals $[\sigma_t, \tau_t]$, along with the associated excursions and intermediate processes. This also provides a basic connection to the Palm distributions. For any times $t_1 < \cdots < t_n$, we introduce the set

$$\Delta(t_1, \ldots, t_n) = \Big\{(b,c) \in \mathsf{R}_+^{2n};\ \textstyle\sum_{i\le k}(b_i + c_i) < t_k + c_k,\ k \le n\Big\},$$

needed to exclude the possibility that two or more excursion intervals $[\sigma_{t_i}, \tau_{t_i}]$ coincide. The *excursion kernel* (ν_r) is defined by the disintegration $\nu = \int \nu_r\, \hat\nu(dr)$, where $\hat\nu = \nu \circ l^{-1}$, and ν_r is restricted to excursions of length r. By Q_t^- we denote the distribution of X^t under Q_t, often written as Q_t, when there is no risk for confusion.

Theorem 12.28 *(excursion decomposition) Let X be a regenerative process with local time ξ, Palm distributions Q_s, and excursion kernel (ν_r), and fix any $t_1 < \cdots < t_n$. Let $[\sigma_j, \tau_j]$ be the excursion interval containing t_j, and put $\beta_j = \sigma_j - \tau_{j-1}$ and $\gamma_j = \tau_j - \sigma_j$, where $\tau_0 = 0$ and $\sigma_{n+1} = \infty$. Then*

(i) *the pairs (β_j, γ_j) have joint distribution $(E\xi \otimes \hat{\nu})^n$ on $\Delta(t_1, \ldots, t_n)$,*

(ii) *the restrictions of (X, ξ) to the intervals $[\tau_{j-1}, \sigma_j)$ and $[\sigma_j, \tau_j)$, shifted back to time 0, are conditionally independent with distributions $Q^-_{\beta_j}$ and ν_{γ_j}, given all the β_i and γ_i.*

Proof: (i) This holds for $n = 1$, by Theorem 12.26. Proceeding by induction, assume the statement to be true for up to $n-1$ points. In case of n points, consider any Borel sets $A \subset \mathsf{R}^2_+$ and $B \subset \mathsf{R}^{2n-2}_+$ with $A \times B \subset \Delta(t_1, \ldots, t_n)$. Here clearly $A \subset \Delta(t_1)$ and $B \subset \Delta(t_2 - r, \ldots, t_n - r)$, for any $r = b + c$ with $(b, c) \in A$. Using the regenerative property at τ_1 and the induction hypothesis, we get

$$\begin{aligned} P\Big\{(\beta_1, \gamma_1, \ldots, \beta_n, \gamma_n) \in A \times B\Big\} \\ &= P\Big\{(\beta_1, \gamma_1) \in A\Big\}\, P\Big\{(\beta_2, \gamma_2, \ldots, \beta_n, \gamma_n) \in B\Big\} \\ &= (E\xi \otimes \hat{\nu})^n (A \times B). \end{aligned}$$

Since $\Delta(t_1, \ldots, t_n)$ is open, its Borel σ-field is generated by all measurable rectangles $A \times B$ as above, and the general result follows by a monotone-class argument.

(ii) Fix any $t > 0$, and put $\gamma_t = \tau_t - \sigma_t$. Using Theorem 12.26, we get for any measurable function $f \geq 0$ on $\mathsf{R}^2_+ \times D^2$

$$\begin{aligned} Ef\Big(\sigma_t, \gamma_t, X^{\sigma_t}, X^{\sigma_t, \gamma_t}\Big) \\ &= \int_0^t E\xi(ds) \int_{t-s}^\infty \hat{\nu}(dr) \iint f(s, r, x, y)\, Q^-_s(dx)\, \nu_r(dy) \\ &\quad + \int f(t, 0, x, 0)\, Q^-_t(dx) \\ &= E \iint f(\sigma_t, \gamma_t, x, y)\, Q^-_{\sigma_t}(dx)\, \nu_{\gamma_t}(dy), \end{aligned}$$

which shows that X^{σ_t} and X^{σ_t, τ_t} are conditionally independent, given (σ_t, τ_t), with distributions Q_{σ_t} and ν_{γ_t}. Furthermore, by the regenerative property at τ_t, the process $\theta_{\tau_t} X$ is independent of X^{τ_t} on the set $\{\tau_t < \infty\}$, with distribution Q_∞. It follows easily that the processes X^{σ_t}, X^{σ_t, τ_t}, and $\theta_{\tau_t} X$ are conditionally independent, given (σ_t, τ_t), with distributions Q_{σ_t}, ν_{γ_t}, and Q_∞, which proves the assertion for $n = 1$.

Proceeding by induction, assume the stated property for up to $n-1$ times t_j, and turn to the case of n times $t_1 < \cdots < t_n$. Given $\sigma_1, \ldots, \sigma_{n-1}$ and $\tau_1, \ldots, \tau_{n-1}$ with $\sigma_1 < \cdots < \sigma_{n-1}$, the restrictions of (X, ξ) to the intervals $[\tau_{j-1}, \sigma_j)$ and $[\sigma_j, \tau_j)$ with $j < n$, shifted back to the origin, are then conditionally independent with distributions Q_{β_j} and ν_{γ_j}. Furthermore, by the univariate result and the regenerative property at τ_{n-1}, the restrictions to $[\tau_{n-1},\sigma_n)$, $[\sigma_n,\tau_n)$, and $\tau_n,\infty)$ are conditionally independent with distributions Q_{β_n}, ν_{γ_n}, and Q_∞, given β_n and γ_n, on the set where $\sigma_{n-1}<\sigma_n$. Since the two sets of times and processes are mutually independent, by the regenerative property at τ_{n-1}, we may combine the two statements, by means of Lemma A1.4 below, into the desired property for all n times. □

We proceed with some basic factorizations of multi-variate moment and Palm measures.

Theorem 12.29 *(moment and Palm factorization) Let X be a regenerative process with local time ξ and Palm distributions Q_t. Then the n-th order moment and Palm measures are given, for any $t_1<\cdots<t_n$, by*

(i) $E\,\xi^n(dt_1\cdots dt_n)=\prod_{j\le n} E\,\xi(dt_j-t_{j-1})$,

(ii) $Q_{t_1,\dots,t_n}\circ\left(X^{t_1},X^{t_1,t_2},\dots,X^{t_n,\infty}\right)^{-1}=\bigotimes_{j\le n+1} Q^-_{\Delta t_j}$.

Proof: To prove the a.s. uniqueness in (ii), let $Q_{t_1,\dots,t_n}$ and $Q'_{t_1,\dots,t_n}$ be given by (ii), for two different, univariate versions Q_t and Q'_t. Noting that

$$\left\|Q_{t_1,\dots,t_n}-Q'_{t_1,\dots,t_n}\right\|\le\sum_{k\le n}\left\|Q_{t_k-t_{k-1}}-Q'_{t_k-t_{k-1}}\right\|,$$

we get by a substitution and Fubini's theorem

$$\begin{aligned}\int\!\cdots\!\int_{\Delta_n}&\left\|Q_{t_1,\dots,t_n}-Q'_{t_1,\dots,t_n}\right\|E\,\xi^n(dt_1\cdots dt_n)\\&\le\sum_{k\le n}\int_0^\infty\!\!\cdots\!\int_0^\infty\left\|Q_{s_k}-Q'_{s_k}\right\|E\,\xi(ds_1)\cdots E\xi(ds_n)=0.\end{aligned}$$

If T has drift $a>0$, then (ii) holds identically, with $Q_{t_1,\dots,t_n}$ defined by elementary conditioning, as in Lemma 6.4. Note that this definition makes sense for arbitrary $t_1,\dots,t_n$, by Proposition 12.33 below.

For $n=1$, consider any bounded interval $I\subset[u,v]$, and let $A\in\mathcal{F}_u$ and $B\in\mathcal{F}_\infty$ be arbitrary. Writing $R_t=\mathcal{L}(\theta_tX)$, and using the regenerative property at T_s, we get

$$P\left(\theta_vX\in B\,\middle|\,\mathcal{F}_{T_s}\right)=R_{v-T_s}B\ \text{ a.s. on }\ \{T_s\le v\}.$$

Using the definitions of Palm distributions and local time, Fubini's theorem, and the substitution rule for general integrals, we conclude that

$$\begin{aligned}\int_I Q_t\left(A\cap\theta_v^{-1}B\right)E\xi(dt)&=E\left\{\xi I;\ X\in A\cap\theta_v^{-1}B\right\}\\&=E\left\{\int_0^\infty 1_I(T_s)\,ds;\ X\in A\cap\theta_v^{-1}B\right\}\end{aligned}$$

$$\begin{aligned}
&= \int_0^\infty P\Big\{T_s \in I,\, X \in A \cap \theta_v^{-1} B\Big\}\, ds \\
&= \int_0^\infty E\Big\{R_{v-T_s}(B);\, T_s \in I,\, X \in A\Big\}\, ds \\
&= E\Big\{\int_0^\infty R_{v-T_s}(B)\, 1_I(T_s)\, ds;\, X \in A\Big\} \\
&= E\Big\{\int_I R_{v-t}(B)\, \xi(dt);\, X \in A\Big\} \\
&= \int_I R_{v-t}(B)\, E\Big\{\xi(dt);\, X \in A\Big\} \\
&= \int_I Q_t(A)\, R_{v-t}(B)\, E\xi(dt).
\end{aligned}$$

Since I was arbitrary, we obtain

$$Q_t\Big(A \cap \theta_v^{-1} B\Big) = Q_t(A)\, R_{v-t}(B), \quad t \in [u, v] \text{ a.e. } E\xi.$$

By a monotone-class argument, we may choose the exceptional null set N to be independent of A and B, and we may also take N to be common to all intervals $[u, v]$ with rational endpoints. Fixing any $t > 0$ in N^c, and letting $u \uparrow t$ and $v \downarrow t$ along Q, we get the required relation

$$Q_t\Big(A \cap \theta_t^{-1} B\Big) = Q_t(A)\, Q_\infty(B), \quad A \in \mathcal{F}_t,\ B \in \mathcal{F}_\infty.$$

Now suppose that (i) and (ii) hold in dimensions $\leq n$. Proceeding by induction, we get for any intervals $I_1 \leq \cdots \leq I_{n+1}$

$$\begin{aligned}
E \prod_{k \leq n+1} \xi I_k &= \int_{I_1} \cdots \int_{I_n} E\, \xi(I_{n+1} - t_n)\, E \prod_{k \leq n} \xi(dt_k) \\
&= \int_{I_1} \cdots \int_{I_n} E\, \xi(I_{n+1} - t_n) \prod_{k \leq n} E\, \xi(dt_k - t_{k-1}) \\
&= \int_{I_1} \cdots \int_{I_{n+1}} \prod_{k \leq n+1} E\, \xi(dt_k - t_{k-1}),
\end{aligned}$$

which extends (i) to dimension $n + 1$.

For $I_1, \ldots, I_{n+1}$ as before, we assume $I_n = [u, v]$ and $I_{n+1} = [u', v']$, and let $A \in \mathcal{F}_u$, $B \in \mathcal{F}_{u'-v}$, and $C \in \mathcal{F}_\infty$ be arbitrary. Using (ii) for indices n and 1, we obtain

$$\begin{aligned}
&E\Big\{\prod_{k \leq n+1} \xi I_k;\, X \in A \cap \theta_v^{-1} B \cap \theta_{v'}^{-1} C\Big\} \\
&= \int_{I_1} \cdots \int_{I_n} Q_{t_1, \ldots, t_n}(A)\, E \prod_{k \leq n} \xi(dt_k)\, E\Big\{\xi(I_{n+1} - t_n);\, X \in \theta_{v-t_n}^{-1} B \cap \theta_{v'-t_n}^{-1} C\Big\} \\
&= \int_{I_1} \cdots \int_{I_n} Q_{t_1, \ldots, t_n}(A)\, E \prod_{k \leq n} \xi(dt_k) \int_{I_{n+1}-t_n} Q_t\Big(\theta_{v-t_n}^{-1} B\Big)\, R_{v'-t_n-t}(C)\, E\xi(dt) \\
&= \int_{I_1} \cdots \int_{I_n} Q_{t_1, \ldots, t_n}(A)\, Q_{t_{n+1}-t_n}\Big(\theta_{v-t_n}^{-1} B\Big)\, R_{v'-t_{n+1}}(C) \\
&\qquad \times E \prod_{k \leq n} \xi(dt_k)\, E\, \xi\Big(dt_{n+1} - t_n\Big).
\end{aligned}$$

Using (i) in dimension $n+1$, we get for $(t_1,\dots,t_{n+1}) \in I_1 \times \cdots \times I_{n+1}$ a.e. $E\xi^{n+1}$

$$\begin{aligned} &Q_{t_1,\dots,t_{n+1}}\Big(A \cap \theta_v^{-1} B \cap \theta_{v'}^{-1} C\Big) \\ &\qquad = Q_{t_1,\dots,t_n}(A)\, Q_{t_{n+1}-t_n}\Big(\theta_{v-t_n}^{-1} B\Big)\, R_{v'-t_{n+1}}(C), \end{aligned}$$

which extends as before to

$$\begin{aligned} &Q_{t_1,\dots,t_{n+1}}\Big(A \cap \theta_{t_n}^{-1} B \cap \theta_{t_{n+1}}^{-1} C\Big) \\ &\qquad = Q_{t_1,\dots,t_n}(A)\, Q_{t_{n+1}-t_n}(B)\, Q_\infty(C), \end{aligned}$$

for any $A \in \mathcal{F}_{t_n}$, $B \in \mathcal{F}_{t_{n+1}-t_n}$, and $C \in \mathcal{F}_\infty$. By (ii), we obtain the same relation for $n+1$. $\Box$

In particular, we show how some structural properties carry over to suitable conditional distributions. Say that ξ is *inversely exchangeable*, if the generating process T is exchangeable in the ordinary sense.

Corollary 12.30 *(symmetry and independence) Let X be a regenerative process with local time $\xi = \lambda \circ T^{-1}$. Then for any fixed $s > 0$ we have, a.s. under $\mathcal{L}(X, \xi \,|\, T_s)$,*

(i) *ξ is inversely exchangeable on $[0, T_s]$,*

(ii) *the excursions of X are conditionally independent, given ξ, with distributions given by the kernel (ν_r).*

Proof: (i) The exchangeability of T on $[0, s]$ may be expressed by countably many conditions, and hence remains valid under conditioning on T_s. This implies the inverse exchangeability of ξ on $[0, T_s]$.

(ii) The stated property holds for the unconditional distribution of (X, ξ), by Corollary 12.16. Since T is measurably determined by ξ, the conditional version then follows by Lemma 6.7 (ii). $\Box$

Similar structural properties hold for the Palm distributions Q_t.

Proposition 12.31 *(Palm distributions) Let X be a regenerative process with local time ξ and associated Palm distributions Q_t. Then for $t > 0$ a.e. $E\xi$, we have under Q_t*

(i) *ξ is inversely exchangeable on $[0, t]$,*

(ii) *the excursions of X are conditionally independent, given ξ, with distributions given by the kernel (ν_r).*

Proof: Let T be the generator of ξ, put $\tilde{T}_s - (s, T_s)$, and define $\eta = \lambda \circ \tilde{T}^{-1}$. Writing $Q_{s,t}$ for the Palm distributions of (X, ξ) with respect to η, and letting

A_t be the set of distributions with the required properties on $[0,t]$, we get by Lemmas 6.6 and Corollary 12.30

$$Q_{s,T_s} = \mathcal{L}(X,\xi \,|\, T_s) \in A_{T_s} \text{ a.s.}, \quad s \geq 0 \text{ a.e. } \lambda.$$

By the definition of η and Fubini's theorem, we get formally

$$\begin{aligned} E\eta\Big\{(s,t) \in \mathsf{R}_+^2;\, Q_{s,t} \notin A_t\Big\} &= E\int_0^\infty 1\Big\{Q_{s,T_s} \notin A_{T_s}\Big\}\, ds \\ &= \int_0^\infty P\Big\{Q_{s,T_s} \notin A_{T_s}\Big\}\, ds = 0, \end{aligned}$$

which means that $Q_{s,t} \in A_t$ a.e. $E\eta$. To justify the use of Fubini's theorem, we may introduce the appropriate scaling operators S_t on $D \times \mathcal{M}$, and note that

$$\Big\{(s,t);\, Q_{s,t} \in A_t\Big\} = \Big\{(s,t);\, Q_{s,t} \circ S_t^{-1} \in A_1\Big\},$$

which is Borel measurable by Lemmas 1.15 (ii) and 6.7 (i). Changing to versions of the Palm distributions $Q_{s,t}$ satisfying $Q_{s,t} \in A_t$ identically, we conclude from Lemma 6.5 that $Q_t \in A_t$ a.e. $E\xi$. $\square$

12.5 Density Existence and Continuity

For a more detailed analysis of regenerative processes X, along with their local time random measures ξ and Palm distributions Q_t, we need to impose some weak regularity conditions, to ensure that the intensity measure $E\xi$ will have a nice density p. Such conditions will be stated in terms of the local characteristics (a,ν) or one-dimensional distributions $\mu_s = \mathcal{L}(T_s)$ of the generating subordinator T.

First we note that the measures μ_s have continuous densities for all $s > 0$, whenever the associated characteristic functions $\hat{\mu}_s$ are integrable. The latter condition will be shown in Theorem 12.36 to be exceedingly weak, in requiring little more than the unboundedness of the Lévy measure ν.

Lemma 12.32 *(densities of μ_s) Let $\xi = \lambda \circ T^{-1}$ with $\hat{\mu}_s \in L^1$ for all $s > 0$. Then $\mu_s \ll \lambda$ for $s > 0$, with some jointly continuous densities $p_{s,t}$ on $(0,\infty] \times [0,\infty]$, satisfying $p_{s,0} = p_{s,\infty} = 0$ for all $s > 0$, $p_{\infty,t} = 0$ for all $t \geq 0$, and $p_{s,t} > 0$ for all $t > sa$.*

Proof: By Fourier inversion, the μ_s have densities

$$p_{s,t} = (2\pi)^{-1} \int e^{-itu}\, \hat{\mu}_{s,u}\, du, \quad s > 0,\ t \geq 0.$$

Here the joint continuity on $(0,\infty) \times \mathsf{R}$ is clear, by dominated convergence, from the estimates, for $s \geq \varepsilon > 0$ and $h > 0$,

$$\begin{aligned} |\,p_{s,t} - p_{s,t+h}| &\leq (2\pi)^{-1} \int \Big|(e^{ihu} - 1)\, \hat{\mu}_{\varepsilon,u}\Big|\, du, \\ |\,p_{s,t} - p_{s+h,t}| &\leq (2\pi)^{-1} \int \Big|(\hat{\mu}_{h,u} - 1)\, \hat{\mu}_{\varepsilon,u}\Big|\, du, \end{aligned} \tag{6}$$

where $\hat{\mu}_{h,u} = Ee^{iuT_h} \to 1$ as $h \to 0$, by dominated convergence, since $T_h \to 0$ a.s. In particular, $p_{s,t} = 0$ by continuity, for $t \le 0$ and $s > 0$. Since $|\hat{\mu}_{s,u}| = |\hat{\mu}_{1,u}|^s \to 0$ for $u \ne 0$, we also have $\sup_t p_{s,t} \to 0$ as $s \to 0$. Finally, $p_{s,t} \to 0$ as $t \to \infty$, by the Riemann–Lebesgue lemma, and by (6) the convergence is uniform, for s belonging to any compact set in $(0, \infty)$.

To prove the last assertion, we may assume that $a = 0$. First we show that $\operatorname{supp} \mu_s = \mathsf{R}_+$ for all $s > 0$. Noting that $0 \in \operatorname{supp} \nu$, since $\|\nu\| = \infty$, and that $\operatorname{supp}(\mu + \mu') = \operatorname{supp} \mu + \operatorname{supp} \mu'$, we may assume that ν is bounded with $0 \in \operatorname{supp} \nu$, in which case $\operatorname{supp} \mu_s = \bigcup_{n \ge 0} \operatorname{supp} \nu^{*n} = \mathsf{R}_+$. Now fix any $s, t > 0$, and put $u = s/2$, so that $p_s = p_u * p_u$. Since p_u is continuous with support R_+, there exists an interval $I \subset (0, t)$, such that $p_u \ge \varepsilon > 0$ on I. Then

$$\begin{aligned} p_{s,t} &= \int_0^t p_u(t-r)\, p_u(r)\, dr \\ &\ge \varepsilon \int_I p_u(t-r)\, dr \\ &= \varepsilon\, \mu_u(t - I) > 0, \end{aligned}$$

where the support property is used again in the last step. □

Integrating the densities $p_{s,t}$ of the previous lemma with respect to s, we obtain a density p_t of the renewal measure $E\xi$, which will be shown in Theorem 12.36 to have nice continuity properties, under some further regularity conditions on ν. A continuous density p_t also exists, when the generating subordinator T has a drift $a > 0$, so that $a\,\xi = 1_\Xi \lambda$ a.s. We say that the local time random measure ξ, or the underlying regenerative process X, is *regular* if either of these conditions is fulfilled.

Theorem 12.33 *(renewal density) Let $\xi = \lambda \circ T^{-1}$, where either $\hat{\mu}_s \in L^1$ for all $s > 0$, or $a > 0$. Then $E\xi = p \cdot \lambda$, for some measurable function $p > 0$ on $(0, \infty)$, given as follows:*

(i) *If $\hat{\mu}_s \in L^1$ for all $s > 0$, we may choose $p_t = \int_0^\infty p_{s,t}\, ds$, with $p_{s,t}$ as in Lemma 12.32.*

(ii) *If $a > 0$, we may choose $p_t = a^{-1} P\{t \in \Xi\}$, which is absolutely continuous on $[0, \infty]$, with $p_0 = a^{-1}$ and $p_\infty = (a + \nu l)^{-1}$.*

The two versions of p_t agree when both conditions are fulfilled.

Proof: (ii) We may assume that $a = 1$, so that $\xi = 1_\Xi \lambda$ a.s. Then $E\xi \le \lambda$ has a density $p \le 1$. To determine the associated Laplace transform $\hat{p}$, we introduce the tail function $m(t) = \nu(t, \infty)$, with Laplace transform

$$\begin{aligned} \hat{m}(u) &= \int_0^\infty e^{-ut} dt \int_t^\infty \nu(ds) \\ &= \int_0^\infty \nu(ds) \int_0^s e^{-ut} dt \\ &= u^{-1} \int_0^\infty (1 - e^{-us})\, \nu(ds). \end{aligned} \tag{7}$$

Since $\int (1 \wedge s)\, \nu(ds) < \infty$, and

$$\begin{aligned} u^{-1}(1 - e^{-us}) &\le u^{-1}(1 \wedge us) \\ &= u^{-1} \wedge s, \quad u, s > 0, \end{aligned}$$

we have $\hat{m}(u) \to 0$ as $u \to \infty$, by dominated convergence, and so $\hat{m}(u) < 1$ for u greater than some $u_0 \ge 0$. For such a u, we get by Fubini's theorem, the representation $\xi = \lambda \circ T^{-1}$, the Lévy–Khinchin formula, relation (7), and the summation of a geometric series

$$\begin{aligned} \hat{p}(u) &= \int_0^\infty e^{-ut} E\xi(dt) = E \int_0^\infty e^{-ut} \xi(dt) \\ &= E \int_0^\infty e^{-uT_s} ds = \int_0^\infty E e^{-uT_s} ds \\ &= \int_0^\infty \exp\Big\{ - us - s \int_0^\infty (1 - e^{-ux})\, \nu(dx) \Big\}\, ds \\ &= \Big\{ u + \int_0^\infty (1 - e^{-ux})\, \nu(dx) \Big\}^{-1} \\ &= \big\{ u + u\, \hat{m}(u) \big\}^{-1} \\ &= u^{-1} \Big\{ 1 + \textstyle\sum_{n \ge 1} (-1)^n\, \hat{m}^n(u) \Big\}. \end{aligned}$$

By a formal Laplace inversion, the density p has then a version

$$p(t) = 1 + \int_0^t ds \sum_{n \ge 1} (-1)^n m^{*n}(s), \tag{8}$$

where m^{*n} denotes the n-th convolution power of m. To justify the formal calculations, we note that for any $u > u_0$,

$$\begin{aligned} \int_0^t ds \sum_{n \ge 1} m^{*n}(s) &\le e^{ut} \int_0^t e^{-us} ds \sum_{n \ge 1} m^{*n}(s) \\ &\le e^{ut} \sum_{n \ge 1} \hat{m}^n(u) \\ &= e^{ut} \frac{\hat{m}(u)}{1 - \hat{m}(u)} < \infty, \end{aligned}$$

which shows that the sum in (8) is absolutely convergent, for $s \ge 0$ a.e. λ. Thus, p is an absolutely continuous density of $E\xi$.

From (8), we see that $p(0) = 1$. To see that the version $\tilde{p}(t) = P\{t \in \Xi\}$ agrees with $p(t)$, we may use the regenerative property of Ξ, along with the fact that $p = \tilde{p}$ a.e. λ, to write for any $t, h \ge 0$

$$\begin{aligned} &h^{-1} \int_0^h p(t+s)\, ds \\ &\qquad = \frac{\tilde{p}(t)}{h} \int_0^h p(s)\, ds + h^{-1} \int_0^h P\big\{ t \notin \Xi,\, t + s \in \Xi \big\}\, ds. \end{aligned}$$

Letting $h \to 0$, and using the continuity of p, and the fact that $t \in \Xi^c$ implies $[t, t+\varepsilon) \subset \Xi^c$ for small enough $\varepsilon > 0$, we get $p(t) = \tilde{p}(t)$.

By the regenerative property of Ξ, we get for any $s, t \geq 0$

$$\begin{aligned} p(s)\, p(t) &\leq p(s+t) \\ &\leq p(s)\, p(t) + 1 - p(s), \end{aligned}$$

and so, by iteration on the left, we have $p(t) > 0$ for all $t \geq 0$. Rearranging the inequalities into

$$\begin{aligned} -p(t)\,(1-p(s)) &\leq p(s+t) - p(t) \\ &\leq (1-p(s))\,(1-p(t)), \end{aligned}$$

we obtain

$$\big| p(s+t) - p(t) \big| \leq 1 - p(s),$$

which shows that p is uniformly continuous on R_+. Applying the renewal Theorem 12.7 to the discrete skeletons $p(nh)$, $n \in \mathsf{N}$, for arbitrary $h > 0$, we see that the limit $p(\infty)$ exists, and applying the law of large numbers to the process T yields $p(\infty) = (1 + \nu l)^{-1}$.

(i) By Fubini's theorem and the definitions of ξ, $p_{s,t}$, and p_t, we have for any $B \in \mathcal{B}_{\mathsf{R}_+}$

$$\begin{aligned} E\xi B &= E \int_0^\infty 1_B(T_s)\, ds = \int_0^\infty P\{T_s \in B\}\, ds \\ &= \int_0^\infty ds \int_B p_{s,t}\, dt \\ &= \int_B dt \int_0^\infty p_{s,t}\, ds = \int_B p_t\, dt. \end{aligned}$$

To prove the last assertion, suppose that $\hat{\mu}_s \in L^1$ for all $s > 0$. Then for any $\varepsilon > 0$, the function

$$\int_\varepsilon^\infty p_{s,t}\, ds = \int_\varepsilon^{t/a} p_{s,t}\, ds \vee 0, \quad t \geq 0,$$

is continuous on R_+, since $p_{s,t}$ is jointly continuous on $[\varepsilon, \infty) \times \mathsf{R}_+$, by Lemma 12.32. Next let $B \in \mathcal{B}_{(0,\infty)}$, with $\inf B \geq h > 0$, and use Fubini's theorem, Theorems 12.15 and 12.18, and the definitions of $p_{s,t}$ and Ξ to obtain

$$\begin{aligned} \int_B dt \int_0^\varepsilon p_{s,t}\, ds &= \int_0^\varepsilon ds \int_B p_{s,t}\, dt \\ &= \int_0^\varepsilon P\{T_s \in B\}\, ds = E \int_0^\varepsilon 1_B(T_s)\, ds \\ &\leq E\Big\{ \int_0^\infty 1_B(T_s)\, ds\, ;\, T_s \geq h \Big\} \\ &= E(\,\xi B;\, T_\varepsilon \geq h\,) \\ &\leq a^{-1}\, \lambda(B)\, P\{T_s \geq h\}. \end{aligned}$$

Here the inner integral on the left is lower semi-continuous in t, by Fatou's lemma, and so

$$\int_0^\varepsilon p_{s,t}\,ds \le a^{-1} P\{T_\varepsilon \ge h\}, \quad t \ge h,$$

which tends to 0, as $\varepsilon \to 0$ for fixed $h > 0$. Thus, the density in (i) is continuous on $(0,\infty)$, and hence agrees for $t > 0$ with the continuous density $p_t = a^{-1} P\{t \in \Xi\}$ in (ii). □

When the drift coefficient a is positive, we may use the previous result to prove a continuity property for the underlying regenerative set Ξ.

Corollary 12.34 *(renewal set) If $\xi = \lambda \circ T^{-1}$ with $a > 0$, then the process $1\{t \in \Xi\}$ is uniformly continuous in probability.*

Proof: Write $p_t = P\{t \in \Xi\}$, and let $t < t+h$ be arbitrary. By the regenerative property at t,

$$\begin{aligned} P\Big\{t \in \Xi,\, t+h \notin \Xi\Big\} &= P\{t \in \Xi\}\, P\Big(t+h \notin \Xi \,\Big|\, t \in \Xi\Big) \\ &= p_t(1-p_h) \le 1 - p_h, \end{aligned}$$

which tends to 0 as $h \to 0$, by the continuity of p. On the other hand, we see from Theorem 12.28, the regenerative property at τ_t, the disintegration theorem (FMP 6.4), and Fubini's theorem that

$$\begin{aligned} P\Big\{t \notin \Xi,\, t+h \in \Xi\Big\} &= \int_0^t p_s\,ds \int_{t-s}^{t-s+h} p_{t+h-s-r}\,\hat\nu(dr) \\ &\le \int_0^\infty ds \int_s^{s+h} \hat\nu(dr) \\ &= \int_0^\infty (r \wedge h)\,\hat\nu(dr), \end{aligned}$$

which tends to 0 as $h \to 0$, by dominated convergence. □

We turn to conditions on the Lévy measure ν of T, ensuring the integrability in Lemma 12.32. Under stronger regularity conditions, we may nearly describe the set where the renewal density p_t is continuous. This set will play a basic role in subsequent sections, and it becomes crucial for the local independence and conditioning properties in the last section.

Writing ν for the Lévy measure of the generating process T, and putting

$$\nu_2(u) = \int_0^u x^2\,\nu(dx), \quad u > 0,$$

we introduce the tail characteristics

$$\begin{aligned} m_r &= \limsup_{u\to 0}\, u^{-2}\,|\log u|^r\,\nu_2(u), \quad r \in \mathsf{R}, \\ c &= \sup\Big\{r \ge 0;\, \lim_{u\to 0} u^{r-2}\,\nu_2(u) = \infty\Big\}, \\ c' &= \sup\Big\{r \ge 0;\, \lim_{u\to 0} u^r \int (1-e^{-x/u})\,\nu(dx) = \infty\Big\}. \end{aligned}$$

Lemma 12.35 *(regularity indices) Let ν be the Lévy measure of T, and define m_r, c, and c' as above. Then*

(i) *$m_r = 0$ for some $r > 1$ implies $\|\nu\| < \infty$, and $m_0 > 0$ implies $\|\nu\| = \infty$,*

(ii) *$m_{-1} = \infty$ implies $\hat\mu_s \in L^1$ for all $s > 0$,*

(iii) *$0 \le c \le c' \le 1$, and $m_{-1} = \infty$ when $c > 0$.*

Proof: (i) If ν is bounded, then as $u \to 0$,

$$\begin{aligned} u^{-2}\, \nu_2(u) &= \int_0^u \left(\frac{x}{u}\right)^2 \nu(dx) \\ &\le \nu(0, u] \to 0, \end{aligned}$$

which shows that $m_0 = 0$. Now suppose instead that $m_r = 0$, for some $r > 1$. Integrating by parts, and letting $u \to 0$, we get

$$\begin{aligned} \nu(0, u] &= \int_0^u \nu(dx) = \int_0^u x^{-2}\, d\nu_2(x) \\ &= u^{-2}\, \nu_2(u) + 2 \int_0^u x^{-3}\, \nu_2(x)\, dx \\ &\lesssim |\log u|^{-r} + 2 \int_0^u \frac{dx}{x\, |\log x|^r} \to 0, \end{aligned}$$

which shows that ν is bounded.

(ii) If $m_{-1} = \infty$, then for any $s > 0$ and large enough $t > 0$,

$$\begin{aligned} -\log |\hat\mu_s(t)| &= s \int (1 - \cos tx)\, \nu(dx) \\ &\ge \frac{s}{3}\, t^2 \nu_2(t^{-1}) \ge 2\, |\log t| \\ &= -\log t^{-2}, \end{aligned}$$

and so $|\hat\mu_s(t)| \lesssim t^{-2}$ as $|t| \to \infty$, which implies $\hat\mu_s \in L^1$.

(iii) For $u > 0$, we have

$$\begin{aligned} u^{-2} \nu_2(u) &= \int_0^u \left(\frac{x}{u}\right)^2 \nu(dx) \\ &\le \int_0^\infty \left(\frac{x}{u} \wedge 1\right) \nu(dx) \\ &\lesssim \int_0^\infty (1 - e^{-x/u})\, \nu(dx), \end{aligned}$$

which implies $c \le c'$. Furthermore, we get by dominated convergence, as $u \to 0$

$$u \int_0^\infty (1 - e^{-x/u})\, \nu(dx) \le \int_0^\infty (x \wedge u)\, \nu(dx) \to 0,$$

which shows that $c' \le 1$. If $c > 0$, then $u^{r-2} \nu_2(u) \to \infty$ for some $r > 0$, and $m_{-1} = \infty$ follows, since $|\log u| \lesssim u^{-r}$ as $u \to 0$. □

We show how c and c' give fairly precise information about the continuity set of p. To this end, define $d = [c^{-1}] - 1$ when $c > 0$, and $d' = [(c')^{-1}-] - 1$ when $c' > 0$. Put $S_0 = \{0\}$, let S_1 be the set of points $t \in [0, \infty]$, such that ν has no bounded density in any neighborhood of t, and define recursively $S_n = S_{n-1} + S_1$, for $n \in \mathsf{N}$. Similarly, put $S'_0 = \{0\}$, let $S'_1 = \{0\} \cup \{t > 0;\, \nu\{t\} > 0\}$, and define $S'_n = S'_{n-1} + S'_1$ for $n \in \mathsf{N}$.

Theorem 12.36 *(continuity set) Let $\xi = \lambda \circ T^{-1}$ with $a = 0$. Then*

(i) *for $c > 0$, the density p of $E\xi$ is continuous on $(0, \infty] \setminus S_d$,*
(ii) *for $c' > 0$, every density of $E\xi$ is discontinuous on $S'_{d'}$.*

Combining Theorems 12.33 and 12.36, we see that if $a > 0$ or $m_{-1} = \infty$, then $E\xi \ll \lambda$ with a density $p > 0$, whereas if $a > 0$ or $c > \frac{1}{2}$, or if $c > 0$ and ν has a locally bounded density on $(0, \infty)$, then $E\xi$ has a continuous density on $(0, \infty)$. For the local time of Brownian motion at a fixed point, the generating subordinator T is $\frac{1}{2}$-stable, which implies $c = c' = \frac{1}{2}$, and so in this case $d = 1$, while $d' = 0$. However, the density p is still continuous on $(0, \infty]$, since ν has the continuous density $\asymp x^{-3/2}$. The result for Brownian motion carries over to any sufficiently regular diffusion process.

Our proof of Theorem 12.36 is based on the following technical result.

Lemma 12.37 *(uniform continuity) When $m_{-1} = \infty$, the density $p'_t = \int_1^\infty p_{s,t}\, ds$ of the measure $\mu' = \int_1^\infty \mu_s\, ds$ is continuous on $[0, \infty]$.*

Proof: Writing $p^s_t = p_{s,t}$, we see from the semi-group property $\mu_s * \mu_t = \mu_{s+t}$ that $p^s * \mu_t = p^{s+t}$ a.e. λ, and since both sides are continuous, the relation $p^s * p^t = p^{s+t}$ holds identically. Writing $\bar{p} = \int_0^1 p^s\, ds$, we get by Fubini's theorem

$$\begin{aligned} p' = \int_1^\infty p^s\, ds &= \sum_{n \geq 1} \int_n^{n+1} p^s\, ds \\ &= \sum_{n \geq 1} p^n * \int_0^1 p^s\, ds \\ &= \bar{p} * \sum_{n \geq 1} p^n. \end{aligned}$$

Here the sum on the right is continuous on $[0, \infty]$, by Corollary 12.10, and so the same thing is true for the convolution with $\bar{p}$, by the proof of the same result. $\square$

Proof of Theorem 12.36: (i) By Lemma 12.37 and a suitable scaling, the function $\int_\varepsilon^\infty p^s\, ds$ is continuous on $[0, \infty]$, for every $\varepsilon > 0$. Invoking Lemma 1.34, it is then enough to prove that the measure $\bar{\mu}_\varepsilon = \int_0^\varepsilon \mu_s\, ds$ has a density $\bar{p}^\varepsilon$ that tends to 0 as $\varepsilon \to 0$, uniformly on compacts in $(0, \infty] \setminus S_d$.

From this point on, the proof is analytic and extremely technical, and so we indicate only the main steps. First let ν be supported by an interval $[0, b]$.

By repeated differentiation in the expression for the Laplace transform $\tilde{\mu}^t$, followed by an analytic continuation and some elementary estimates, we get for any $u \in \mathsf{R}$ and $n \in \mathsf{N}$ the bound

$$\left| \int_0^\infty x^n \, e^{iux} \, \mu^t(dx) \right| \leq 2tm \, b^{n-1} e^{-t\psi(u)}, \quad t \leq b \, e^{-n}/m, \tag{9}$$

where $\psi(u) = u^2 \, \nu_2(u^{-1})/3$. Now choose

$$n = n(t) = [\, \log(b/mt) \,], \quad t \leq t_0 \equiv b/me,$$

put $h = b \, e^d$, and define the measure ρ on R_+ by

$$\rho f = \int_0^{t_0} dt \int_0^\infty (x/h)^{n(t)} \, f(x) \, \mu^t(dx), \quad f \geq 0.$$

Using (9) and some elementary estimates, we get

$$\begin{aligned} |\hat{\rho}(u)| &\leq \frac{2m}{b} \int_0^{t_0} t \, (b/h)^{n(t)} \, e^{-t\psi(u)} \, dt \\ &\leq \frac{2m}{b} \left(\frac{em}{b} \right)^d \left\{ \frac{t_0^{d+2}}{d+2} \wedge \frac{(d+1)!}{\{\psi(u)\}^{d+2}} \right\}. \end{aligned}$$

Since $d + 2 = [c^{-1}] + 1 > c^{-1}$, the right-hand side is integrable, and so ρ has a uniformly continuous density. Thus, $\bar{\mu}_\varepsilon$ has a continuous density on $[h, \infty)$ that tends to 0 at ∞.

Returning to a general ν, write $\nu = \nu' + \nu''$, where ν' is supported by $(0, b]$ and $\|\nu''\| = a < \infty$, and let $\mu'^t, \psi', \ldots$ denote the quantities $\mu^t, \psi, \ldots$ based on ν'. Note that ν and ν' have the same index c. The decomposition of ν induces a division of the generating subordinator T into independent components, where the jumps $> b$ occur at local times given by a Poisson process with rate a. Thus, the n-th such time has probability density $a^n \, t^{n-1} \, e^{-at}/(n-1)!$ on R_+, which gives

$$\int_0^\varepsilon \mu^t dt \leq \sum_{n \geq 0} \frac{1}{n!} \int_0^\varepsilon t^n \mu'^t dt * \nu''^n.$$

Here it is enough to consider the sum over $n \leq d$, since the characteristic function of the sum over $n > d$ is bounded by

$$\begin{aligned} \sum_{n>d} \frac{a^n}{n!} \int_0^1 t^n |\hat{\mu}'|^t dt &\leq \sum_{n>d} \frac{a^n}{n!} \int_0^1 t^{d+1} e^{-t\psi'} dt \\ &\leq e^a \left\{ \frac{1}{d+2} \wedge \frac{(d+1)!}{(\psi')^{d+2}} \right\}, \end{aligned}$$

which is integrable and hence corresponds to a uniformly continuous density.

In the remaining finite sum, we may write $\nu'' = \nu''' + (\nu'' - \nu''')$, where the second term equals $\nu'' \wedge r\lambda$, for some $r > 0$. This yields for every n the estimate

$$\frac{1}{n!} \int_0^\varepsilon t^n \mu'^t dt * \nu''^n \leq \frac{1}{n!} \int_0^\varepsilon t^n \mu'^t dt * \nu'''^n + \frac{n \, a^{n-1} \, r \, \varepsilon^{n+1} \, \lambda}{(n+1)!}. \tag{10}$$

Summing over n, we get in the last term a measure with density $\leq r\varepsilon\, e^{a\varepsilon}/a$, which tends to zero as $\varepsilon \to 0$.

It remains to consider the contributions, for $n \leq d$, of the first term in (10). Here the support is contained in $[0, h] + (A_r)_d$, where A_r denotes the support of $\nu - \nu \wedge r\lambda$, and $(A_r)_d$ is defined by iteration in d steps, as before. Since $(A_r)_d$ is closed and $h > 0$ is arbitrary, we conclude that p is continuous outside $(A_r)_d$. Since even $r > 0$ was arbitrary, the last statement remains true for the intersection $\bigcap_r (A_r)_d = S_d$.

(ii) Here we note that the Laplace transforms $\tilde{\mu}^t$ of the measures μ^t satisfy

$$\int_0^\infty t^n\, \tilde{\mu}^t(u)\, dt = n! \left\{ \int_0^\infty (1 - e^{-ux})\, \nu(dx) \right\}^{-n-1},$$

for all n. Since

$$\liminf_{u \to 0} u^{-1} \left\{ \int_0^\infty (1 - e^{-ux})\, \nu(dx) \right\}^{n+1} = 0, \quad n \leq d',$$

by the definition of c', it follows that the measure $\int_0^\infty t^n\, \mu^t\, dt$ has no bounded density near the origin, when $n \leq d'$. Arguing as in part (i), we conclude that, for any $b > 0$ and $n \leq d'$, any density p of $E\xi$ must be unbounded to the right of every atom of $(1_{[b,\infty)}\nu)^n$, which shows that p is indeed discontinuous on $S'_{d'}$. □

12.6 Conditioning via Duality

To obtain some good versions of the Palm distributions, with respect to the local time random measure ξ, we might try to use the powerful duality theory of Section 6.7. Unfortunately, a direct application would require the intensity measure $E\xi$ to have a continuous and nice density p, which may not be true, even under the regularity conditions of Lemma 12.32. Instead we will proceed in two steps, first constructing some regular conditional distributions with respect to the variables T_s, and then forming the associated Palm distributions, by a suitable mixing.

In this section, we will focus on the somewhat technical first step of the construction, postponing our study of the resulting Palm distributions with respect to ξ until the next section. We begin with an important averaging property, which will play a basic role in the sequel.

Lemma 12.38 *(fundamental identity) If $\hat{\mu}_s \in L^1$ for all $s > 0$, then the density p in Theorem 12.33* (i) *satisfies*

$$Ep(t - \tau_y) = p(t), \quad y < t.$$

Proof: The stated relation being trivial for $y \leq 0$, we may assume that $y > 0$. Noting that $X_{L_y} = \tau_y$, and using the strong Markov property at L_y, and the disintegration theorem (FMP 6.4), we get for any $B \in \mathcal{B}_{(y,\infty)}$ and $t \geq 0$

$$\begin{aligned}
\mu^t B &= P\{X_t \in B\} \\
&= P\{X_t \in B,\, L_y \leq t\} \\
&= P\Big\{(\theta_{L_y} X)_{t-L_y} \in B - \tau_y;\, L_y \leq t\Big\} \\
&= E\Big\{\mu^{t-L_y}(B - \tau_y);\, L_y \leq t\Big\}.
\end{aligned}$$

Taking densities of both sides, we obtain for almost every $x > y$

$$p^t(x) = E\Big\{p^{t-L_y}(x - \tau_y);\, L_y \leq t\Big\}. \tag{11}$$

To extend this to an identity, we may use Lemma 12.26 and its corollary, Fubini's theorem, and the uniform boundedness of $p^s(u)$ for $s \geq t - \varepsilon$ to get

$$\begin{aligned}
&E\Big\{p^{t-L_y}(x - \tau_y);\, L_y \in (t-\varepsilon, t]\Big\} \\
&\quad = \int_{t-\varepsilon}^t ds \int_0^y p^s(u)\, du \int_{y-u}^{x-u} p^{t-s}(x-u-v)\, \nu(dv) \\
&\quad \lesssim \int_0^\varepsilon ds \int_{-\infty}^y du \int_{y-u}^\infty (x-u-v)\, \nu(dv) \\
&\quad = \int_0^\varepsilon ds \int_0^\infty \nu(dv) \int_{y-v}^y p^s(x-u-v)\, du \\
&\quad = \int_0^\varepsilon ds \int_0^\infty \nu(dv) \int_0^v p^s(x-y-r)\, dr \\
&\quad = \int_0^\varepsilon ds \int_0^{x-y} p^s(x-y-r)\, \nu(r,\infty)\, dr \\
&\quad = \int_0^\varepsilon ds \int_0^{x-y} \nu(x-y-r,\, \infty)\, p^s(r)\, dr \\
&\quad = \int_0^{x-y} \nu(x-y-r,\, \infty)\, E\xi_\varepsilon(dr) \\
&\quad = P\{X_\varepsilon > x - y\}.
\end{aligned}$$

Here the right-hand side tends to 0 as $\varepsilon \to 0$, uniformly for $x > y$ bounded away from y, and so

$$E\Big\{p^{t-L_y}(x-\tau_y);\, L_y \leq t - \varepsilon\Big\} \to E\Big\{p^{t-L_y}(x-\tau_y);\, L_y \leq t\Big\},$$

uniformly on compacts in (y,∞). Since the densities p^s, $s \geq \varepsilon$, are uniformly equi-continuous, the left-hand side is continuous in x, and the continuity extends to the limit, by the uniformity of the convergence. Thus, both sides of (11) are continuous, and the relation holds identically. Integrating in t, and using Fubini's theorem, we get

$$\begin{aligned}
p(x) &= \int_0^\infty p^t(x)\, dt \\
&= \int_0^\infty E\Big\{p^{t-L_y}(x-\tau_y);\, L_y \leq t\Big\}\, dt \\
&= E\int_{L_y}^\infty p^{t-L_y}(x-\tau_y)\, dt \\
&= E\, p(x - \tau_y).
\end{aligned}$$

□

We also need the following somewhat technical continuity property.

Lemma 12.39 *(kernel continuity) Let X be a regular, regenerative process with local time ξ. Then the probability kernel*

$$P_t = \mathcal{L}\left\{\theta_{\sigma_t}(X,\xi),\, t-\sigma_t\right\}, \quad t \ge 0,$$

is continuous in total variation on $(0,\infty)$, and even on $[0,\infty)$ when $a>0$.

Proof: If $a>0$, then for any $t,h \ge 0$,

$$\begin{aligned}\left\|P\circ\theta_t^{-1} - P\circ\theta_{t+h}^{-1}\right\| &\le \left\|P - P\circ\theta_h^{-1}\right\| \\ &= 1 - P\{h\in\Xi\} = 1-p_h,\end{aligned}$$

which tends to 0 as $h\to 0$, by the continuity of p. If instead $\hat\mu_s \in L^1$ for all $s>0$, then for any $t>\varepsilon>0$ and $s,h>0$,

$$\begin{aligned}\left\|P\circ\theta_t^{-1} - P\circ\theta_{t+h}^{-1}\right\| &\le 2\,P\{T_s>t\} + \left\|\mathcal{L}(T_s) - \mathcal{L}(T_s-h)\right\| \\ &\le 2\,P\{T_s>\varepsilon\} + \int_0^\infty \left|p_{s,x} - p_{s,x+h}\right| dx,\end{aligned}$$

which tends to 0 as $h\to 0$ and then $s\to 0$, by the right continuity of T, the continuity of $p_{s,x}$, and extended dominated convergence (FMP 1.21). This shows that the mapping $t\mapsto P\circ\theta_t^{-1}$ is continuous in total variation on R_+ or $(0,\infty)$, respectively.

To extend this result to the asserted continuity of the measures P_t, we note that, for any $t>\varepsilon>0$ and $h>0$,

$$\|P_t - P_{t+h}\| \le 2\,P\{\sigma_\varepsilon < r\} + \left\|P\circ\theta_t^{-1} - P\circ\theta_{t+h}^{-1}\right\|,$$

which tends to 0 as $h\to 0$ and then $r\to 0$, since $\sigma_\varepsilon>0$ a.s. Finally, we see as before that $\|P_h - P_0\| \le 1-p_h \to 0$ when $a>0$. $\Box$

We may now construct versions of the conditional distributions $Q_{s,t} = \mathcal{L}(X\,|\,T_s)_t$, with nice continuity properties. Those will in turn be used, in the next section, to form good versions of the Palm distributions $Q_t = \mathcal{L}(X\,\|\,\xi)_t$. Write $W_a = \{(s,t);\, sa<t\}$.

Theorem 12.40 *(conditioning kernel) Let X be a regenerative process with local time $\xi = \lambda\circ T^{-1}$, where $\hat\mu_s \in L^1$ for all $s>0$. Choose $p_{s,t}$ as in Lemma 12.32, and put $Y=(X,\xi)$. Then*

(i) *for fixed $(s,t)\in W_a$, the relations*

$$Q_{s,t}\circ(Y^{\tau_r},\theta_t Y)^{-1} = p_{s,t}^{-1}\,E\left\{p_{s-\xi_r,\,t-\tau_r};\, Y^{\tau_t}\in\cdot\right\}\otimes\mathcal{L}(Y)$$

are consistent in $r\in(0,t)$, and determine uniquely a version of the probability kernel $(Q_{s,t})\colon W_a\to D$, with $\xi\{t\}=0$ a.s. $Q_{s,t}$,

(ii) *the kernel* $(Q_{s,t})$ *in* (i) *satisfies*

$$\mathcal{L}(X, \xi \,|\, T_s) = Q_{s,T_s} \ \ a.s., \quad s > 0,$$

(iii) *the mapping* $(s,t) \mapsto Q_{s,t}$ *is continuous at every* $(s,t) \in W_a$*, both weakly for the uniform topology on* D*, and in total variation on* $\mathcal{F}_{\tau_u} \vee \mathcal{G}_{\sigma_v}$*, whenever* $t \in (u,v)$*,*

(iv) *under* $Q_{s,t}$ *for fixed* $(s,t) \in W_a$*, the local time* ξ *is inversely exchangeable on* $[0,t]$*, with* $\xi_t = s$ *a.s., and the excursions of* X *are conditionally independent, given* ξ*, with distributions given by the kernel* (ν_r)*.*

Proof, (i)−(ii): Using the regenerative property at τ_r and the disintegration theorem (FMP 6.4), we get for any $s, r > 0$ and $B \in \mathcal{B}_{(r,\infty)}$

$$\begin{aligned} P(\,T_s \in B \,|\, \mathcal{F}_{\tau_r}) &= \mu_{s-\xi_r}(B - \tau_r) \\ &= 1_B(\tau_r)\, 1\{\xi_r = s\} + \int_B p(s - \xi_r,\, t - \tau_r)\, dt, \end{aligned} \tag{12}$$

where $\mu_s = 0$ for $s < 0$. Now $\xi_r \neq s$ a.s. for all $s, r > 0$, since $\mu_s \ll \lambda$ and $\Delta T_s = 0$ a.s. Hence, the first term on the right vanishes a.s., and the integrand $M^r_{s,t} = p(s - \xi_r,\, t - \tau_r)$ in the second term on the right is a.s. jointly continuous in (s,t). Since also $EM^r_{s,t} = p_{s,t}$, as in the proof of Lemma 12.38, we conclude that $M^r_{s,t}$ is L^1-continuous in (s,t). By Lemma 2.14 and Theorem 6.42 we note that, for any $(s,t) \in W_a$, the process $M^r_{s,t}$ is a martingale in $r \in (0,t)$, with respect to the filtration $\mathcal{F}_{\tau_r}$, which ensures the asserted consistency of the defining relations. The measurability of $Q_{s,t}$ follows by Fubini's theorem from that for the functions p, ξ_r, and τ_r. Property (ii) follows from (12) by Theorem 6.41, and the total variation claim in (iii) follows from the L^1-continuity of $M^r_{s,t}$, by Theorem 6.42 and Lemma 12.39.

(iv) To show that $\xi_t = s$ a.s. $Q_{s,t}$, for any $(s,t) \in W_a$, let $r \in (0,t)$ be arbitrary, and conclude from (i) and the relation $EM^r_{s,t} = p_{s,t}$ that

$$\begin{aligned} Q_{s,t}\{\xi_r < s\} &= p^{-1}_{s,t}\, E\Big\{p(s - \xi_r,\, t - \tau_r);\, \xi_r < s\Big\} \\ &= p^{-1}_{s,t}\, E\, p(s - \xi_r,\, t - \tau_r) = 1. \end{aligned} \tag{13}$$

Letting $r \to t$, and recalling that $\xi\{t\} = 0$ a.s. $Q_{s,t}$, we obtain

$$Q_{s,t}\{\xi_t \le s\} = Q_{s,t}\Big\{\xi(0,t) \le s\Big\} = 1, \quad (s,t) \in W_a. \tag{14}$$

By (ii) and Corollary 12.30, we may next choose some $t_n \uparrow t$, such that the properties in (iv) hold for the measures Q_{s,t_n}. In particular, we have for any $h \in (0,t)$

$$\begin{aligned} Q_{s,t_n}\xi(0, t-h) &\ge Q_{s,t_n}\xi(0, t_n - h) \\ &= s - Q_{s,t_n}\xi(0,h), \end{aligned}$$

and so, by the total variation part of (iii),

$$Q_{s,t}\xi(0,t-h) \geq s - Q_{s,t}\xi(0,h).$$

As $h \to 0$, we get by monotone and dominated convergence $Q_{s,t}\xi(0,t) \geq s$, which along with (14) yields the desired relation $Q_{s,t}\{\xi_t = s\} = 1$ or $\xi_t = s$ a.s. $Q_{s,t}$. Combining with (13) gives $\xi(r,t) > 0$ a.s. $Q_{s,t}$, for every $r \in (0,t)$, which implies

$$\sigma_{t-} = \sup_{r<t} \sigma_r = t \;\text{ a.s. } Q_{s,t}, \quad (s,t) \in W_a.$$

To prove the inverse exchangeability of ξ, we first need to extend (12) to the form

$$Q_{s,t}(A) = p_{s,t}^{-1}\, E\Big\{p(s-r,\, t-T_r);\, A\Big\}, \quad A \in \mathcal{F}_{T_r}, \;\; r < s. \tag{15}$$

For a justification, we note that for fixed $(s,t) \in W_a$, the martingale $M_{s,t}^r$ is continuous in $r < t$, apart from possible discontinuities at T_s and at the left endpoints of the excursion intervals. Since the latter points are a.s. avoided by the times T_r with $r < s$, by the right-continuity of T, we get by optional sampling

$$E\Big(M_{s,t}^{T_r \wedge v}\Big|\, \mathcal{F}_u\Big) = M_{s,t}^{T_r \wedge u} \;\text{ a.s.,} \quad 0 < u < v < t, \;\; r < s. \tag{16}$$

Now recall from Lemma 12.32 that $p_{s,t}$ is uniformly bounded on the sets $[\varepsilon, \infty) \times \mathsf{R}$ with $\varepsilon > 0$. We may then let $v \to t$ in (16), and conclude by dominated convergence that, for any $r < s$ and $u < t$,

$$E\Big\{p(s-r,\, t-T_r)\,\Big|\, \mathcal{F}_u\Big\} = p(s-\xi_u,\, t-\tau_u) \;\text{ a.s. on } \{T_r \geq u\}.$$

which yields the required consistency of (i) and (15).

(iii) We already noted the continuity in total variation. To prove the asserted weak continuity, fix any $(s,t) \in W_a$, let $r_n < t < t_n$ with $r_n \uparrow t$ and $t_n \downarrow t$, put $I_n = [\sigma_{r_n}, \tau_{t_n})$, and write X^n and ξ^n for the restrictions of X and ξ to the complements I_n^c. Define $L_r = \xi[0,r]$ and $L_r^n = \xi^n[0,r]$. For any bounded, measurable function f on D, the continuity in total variation yields

$$\lim_{u\to t} Q_{s,t}f(X^n, L^n) = Q_{s,t}f(X^n, L^n), \quad n \in \mathsf{N}.$$

By Lemma 4.21, it then suffices to prove that

$$\lim_{n\to\infty} \limsup_{u\to t} Q_{s,u} \sup_{r>0} \Big\{\, \rho(X_r^n, X_r) + |L_r - L_r^n| \wedge 1 \Big\} = 0,$$

for any bounded metrization ρ of S. Hence, we need to verify that

$$\lim_{n\to\infty} \limsup_{u\to t} Q_{s,u}\Big\{ \sup_{r\in I_n} \rho(X_r, 0) + \xi I_n \wedge 1\Big\} = 0. \tag{17}$$

Beginning with the contribution to $[u,\infty)$, we need to show that

$$\lim_{\varepsilon\to 0} E\Big\{ \sup_{r\le\tau_\varepsilon} \rho(X_r,0) + L_{\tau_\varepsilon} \wedge 1\Big\} = 0.$$

Since X and L are both continuous at 0, with $X_0 = L_0 = 0$, the last relation will follow, if we can prove that $\tau_\varepsilon \to 0$ a.s. as $\varepsilon \to 0$. But this is clear from the fact that the generating subordinator T is continuous and strictly increasing at 0.

Turning to the contribution to (17) from the interval $[0,u)$, we see from (ii) that the distribution $Q_{s,u} \circ (X,\xi)^{-1}$ is invariant under a reversal on $[0,u]$, followed by reversals of the individual excursions (cf. Lemma 12.43 below). Since the latter reversals do not affect the maxima of $\rho(X,0)$, we get under $Q_{s,u}$, for arbitrary $b \in (0,u)$,

$$\sup_{r\in[\sigma_b,u)} \rho(X_r,0) \overset{d}{=} \sup_{r\le\tau_{u-b}} \rho(X_r,0), \qquad \xi[b,u) \overset{d}{=} L_{u-b}.$$

Hence, it only remains to show that

$$\lim_{\varepsilon\to 0} \limsup_{u\to t} Q_{s,u}\Big\{ \sup_{r\le\tau_\varepsilon} \rho(X_r,0) + L_\varepsilon \wedge 1\Big\} = 0.$$

Assuming $\rho \le 1$, and using the total variation part of (i), we note that the lim sup on the left is bounded, for any $h > \varepsilon$, by

$$Q_{s,t}\Big\{ \sup_{r<h} \rho(X_r,0) + L_\varepsilon \wedge 1\Big\} + Q_{s,t}\{\tau_\varepsilon > h\}.$$

Since X_r, L_r, and τ_r tend a.s. to 0 as $r \to 0$, we may use dominated convergence to see from (i) that the last expression tends to 0 as $\varepsilon \to 0$ and then $h \to 0$, which completes the proof of (iii). □

12.7 Regular Palm Distributions

Here we use results from the previous sections to construct good versions of the Palm distributions Q_t of X with respect to ξ. Basic properties of X involving those distributions will then be studied throughout the remainder of the chapter.

Our construction requires either one of the regularity properties in Theorem 12.33, namely the integrability of the characteristic functions $\hat{\mu}_s$, or the positivity of the drift coefficient a. Elementary constructions are available in both cases, but there is also a general approach, similar to the one used in Theorem 12.40, to construct versions of the conditional distributions $Q_{s,t}$. Both constructions are useful whenever they apply.

Theorem 12.41 *(Palm distributions) Let X be a regular, regenerative process with local time $\xi = \lambda \circ T^{-1}$, define $(p_{s,t})$, (p_t), and $(Q_{s,t})$ as in Lemma 12.32 and Theorems 12.33 and 12.40, and put $Y = (X,\xi)$. Then*

(i) *for fixed $t > 0$ with $p_t < \infty$, the relations*

$$Q_t\Big\{(Y^{\tau_r}, \theta_t Y) \in \cdot\Big\} = p_t^{-1} E\Big\{p_{t-\tau_r};\, Y^{\tau_r} \in \cdot\Big\} \otimes \mathcal{L}(Y)$$

are consistent in $r \in (0,t)$, and determine uniquely some versions of the Palm distributions $Q_t = \mathcal{L}(X, \xi \,\|\, \xi)_t$, with $\xi\{t\} = 0$ a.s. Q_t,

(ii) *when $\hat\mu_s \in L^1$ for all $s > 0$, the measures Q_t in* (i) *are also given by*

$$Q_t = p_t^{-1} \int_0^\infty p_{s,t}\, Q_{s,t}\, ds, \quad t > 0 \text{ with } p_t < \infty,$$

(iii) *when $a > 0$, the measures Q_t in* (i) *are also given by*

$$Q_t = P\Big\{\,(X, \xi) \in \cdot \,\Big|\, t \in \Xi\Big\}, \quad t \ge 0.$$

(iv) *If $\hat\mu_s \in L^1$ for all $s > 0$, then for every continuity point $t \ge 0$ of p, the kernel Q_s is continuous at t, both weakly for the uniform topology on D, and in total variation on $\mathcal{F}_{\tau_u} \vee \mathcal{G}_{\sigma_v}$, whenever $t \in (u,v)$. The induced kernel $Q_s \circ \xi_s^{-1}$ is further continuous in total variation at t. If instead $a > 0$, then Q_t is continuous in total variation on D.*

(v) *Under Q_t for fixed $t > 0$ with $p_t < \infty$, the local time ξ is inversely exchangeable on $[0,t]$, and the excursions of X are conditionally independent, given ξ, with distributions given by the kernel (ν_r).*

Proof, (i)$-$(iii): When $\hat\mu_s \in L^1$ for all $s > 0$, Lemmas 6.5 and 6.6 show that the measures in (ii) are versions of the Palm distributions Q_t. To prove (i) in this case, we note that the relation $Ep(s - \xi_r,\, t - \tau_r) = p_{s,t}$ remains valid for all $s, t \in \mathsf{R}$. By Fubini's theorem and the definitions of $Q_{s,t}$ and p_t, we have for any $r \in (0,t)$

$$\begin{aligned} p_t\, Q_t \circ (\theta_0^{\tau_r}, \theta_t)^{-1} &= \int_0^\infty E\Big\{p(s - \xi_r,\, t - \tau_r);\, (X,\xi)^{\tau_r} \in \cdot\Big\}\, ds \otimes Q_0 \\ &= E\Big\{\int_0^\infty p(s - \xi_r,\, t - \tau_r)\, ds;\, (X,\xi)^{\tau_r} \in \cdot\,\Big\} \otimes Q_0 \\ &= E\Big\{p(t - \tau_r);\, (X,\xi)^{\tau_r} \in \cdot\Big\} \otimes Q_0, \end{aligned}$$

as required. We also note that $\xi\{t\} = 0$ a.s. Q_t, by the corresponding property for the measures $Q_{s,t}$. Conversely, the latter property, along with (i), clearly determine Q_t whenever $p_t < \infty$.

If instead $a > 0$, then Lemma 6.4 (ii) shows that the measures in (iii) are versions of the Palm distributions Q_t. To prove (i) in this case, we see from Fubini's theorem and the regenerative property at τ_r and $t - s$ that, for any $r \in (0,t)$ and measurable $A, B \subset D$,

$$\begin{aligned} &p_t\, Q_t\Big\{X^{\tau_r} \in A,\ \theta_t X \in B\Big\} \\ &\quad = P\Big\{X^{\tau_r} \in A,\ t \in \Xi,\ \theta_t X \in B\Big\} \\ &\quad = E\Big(P\Big\{t - s \in \Xi,\ \theta_{t-s} X \in B\Big\}_{s=\tau_r};\, X^{\tau_r} \in A\Big) \\ &\quad = E\Big(P\{t - s \in \Xi\}_{s=\tau_r};\, X^{\tau_r} \in A\Big)\, P\{X \in B\} \\ &\quad = E\Big\{\,p(t - \tau_r);\, X^{\tau_r} \in A\Big\}\, Q_0(B), \end{aligned}$$

which extends to (i) by a monotone-class argument. The relation $\xi\{t\} = 0$ a.s. Q_t is obvious in this case, since $\xi \le a^{-1}\lambda$. Conversely, we see as before that the two conditions determine the Q_t.

(iv) When $\hat{\mu}_s \in L^1$ for all $s > 0$, fix any continuity point $t > 0$ of p, put $q_{s,t} = p_{s,t}/p_t$, and conclude from Lemmas 12.32 and 12.40 (iii) that, as $r \to t > 0$ for fixed $s > 0$,

$$\Big\| q_{s,t}\, Q_{s,t} - q_{s,r}\, Q_{s,r} \Big\| \le q_{s,t} \Big\| Q_{s,t} - Q_{s,r} \Big\| + |\, q_{s,t} - q_{s,r}| \to 0,$$

where the norm denotes total variation on $\mathcal{F}_{\tau_u} \vee \mathcal{G}_{\sigma_v}$, for arbitrary $(u,v) \ni t$. Here the left-hand side is bounded by $q_{s,t} + q_{s,r}$, which tends to $2q_{s,t}$ as $r \to t$, and satisfies

$$\int_0^\infty (q_{s,t} + q_{s,r})\, ds = 2 = \int_0^\infty 2\, q_{s,t}\, ds.$$

By extended dominated convergence (FMP 1.21), we obtain

$$\|\, Q_t - Q_r \| \le \int_0^\infty \Big\| q_{s,t}\, Q_{s,t} - q_{s,r}\, Q_{s,r} \Big\|\, ds \to 0,$$

which proves the asserted continuity in total variation. Similarly, when $f : D \to \mathsf{R}$ is bounded, measurable, and continuous in the uniform topology on D, the weak convergence part of Lemma 12.40 (iii) yields

$$|\, Q_t f - Q_r f| \le \int_0^\infty \Big|\, q_{s,t}\, Q_{s,t} f - q_{s,r}\, Q_{s,r} f \,\Big|\, ds \to 0,$$

which shows that $Q_r \xrightarrow{w} Q_t$ in D. Since also $\xi_t = s$ a.s. $Q_{s,t}$, by Lemma 12.40, formula (ii) shows that $Q_t \circ \xi_t^{-1}$ has density $q_{s,t}$. The last assertion in (iv) then follows from the fact that

$$\Big\| \, Q_t \circ \xi_t^{-1} - Q_r \circ \xi_r^{-1} \Big\| = \int_0^\infty \Big|\, q_{s,t} - q_{s,r} \Big|\, ds \to 0.$$

When $a > 0$, (iii) yields for any $s < t$

$$a\, \|\, Q_s - Q_t \| \le \Big|\, p_s^{-1} - p_t^{-1} \Big| + p_t^{-1} E \Big|\, 1_\Xi(s) - 1_\Xi(t) \,\Big|,$$

which tends to 0 as $s \to t$ or $t \to s$, by Theorem 12.33 (ii) and Corollary 12.34.

(v) When $\hat{\mu}_s \in L^1$ for all $s > 0$, the assertion is obvious from (ii) and Lemma 12.40 (iv), since the stated properties are preserved under mixing of distributions.

If instead $a > 0$, then for any $t > 0$, we may introduce the process $T_s^t = T(s\xi_t)$, $s \in [0,1]$, and note that $T_1^t = t$ a.s. Q_t, by the regenerative property at t. We need to show that, for any $t > 0$, the pair (ξ_t, T^t) is exchangeable under Q_t on the interval $[0,1]$. This holds for almost every $t > 0$, by Proposition 12.31. For a general $t > 0$, choose $t_n \downarrow t$ such that the measures Q_{t_n} have the desired property. Using the continuity of ξ_t and the right-continuity of T, we note that $\xi_{t_n} \to \xi_t$ and $T_s^{t_n} \to T_s^t$ for every $s \in [0,1]$.

Expressing the exchangeability of (ξ_t, T^t) in terms of increments, and using Lemma A2.7 below, we see that the exchangeability of ξ extends to Q_t. The conditional independence of the X-excursions, with distributions given by the kernel (ν_r), is an immediate consequence of (iii) and Lemma 6.7. □

For a first application, we prove a conditional property of suitably shifted distributions, needed for later purposes. To simplify the notation, we may take X to be the canonical process, defined on the relevant path space Ω, and choose $\mathcal{F}$ to be the right-continuous filtration on Ω induced by X.

Corollary 12.42 *(shift invariance) Let X be a regular, regenerative process, with local time ξ and associated Palm distributions Q_t. Then for any $t > 0$ with $p_t < \infty$,*

$$Q_t\Big\{\theta_{\tau_h}(X,\xi) \in \cdot \,\Big|\, \mathcal{F}_{\tau_h}\Big\} = Q_{t-\tau_h} \;\; a.s.\ Q_t, \quad h \in [0,t].$$

Proof: Using Theorem 12.41, the regenerative property at τ_h, and the disintegration theorem (FMP 6.4), we get for any times $h < v < t$, sets $A \in \mathcal{F}_{\tau_h}$, and measurable functions $f \geq 0$ on D

$$\begin{aligned} p_t\, Q_t\Big\{f(X^{\tau_h,\tau_v}); A\Big\} &= E\Big\{p(t-\tau_v)\, f(X^{\tau_h,\tau_v}); A\Big\} \\ &= E\Big(E\Big\{p(t-\tau_v)\, f(X^{\tau_h,\tau_v}) \,\Big|\, \mathcal{F}_{\tau_h}\Big\}\,; A\Big) \\ &= E\Big(E\Big\{p(t-u-\tau_{v-u})\, f(X^{\tau_{v-u}})\Big\}_{u=\tau_h}; A\Big) \\ &= E\Big(E\Big\{p(t-\tau_h)\, Q_{t-\tau_h} f(X^{\tau_{v-u}})\Big\}_{u=\tau_h}; A\Big), \end{aligned}$$

which extends to $v = t-$ by a monotone-class argument. Since also $Q_t(\,\theta_t X \in \cdot \,|\, \mathcal{F}_t) = Q_0$, we obtain

$$\begin{aligned} p_t\, Q_t\Big\{f(\theta_{\tau_h} X); A\Big\} &= E\Big\{p(t-\tau_h)\, Q_{t-\tau_h} f(X); A\Big\} \\ &= p_t\, Q_t\Big\{Q_{t-\tau_h} f(X); A\Big\}, \end{aligned}$$

which shows that

$$Q_t\Big\{f(\theta_{\tau_h} X) \,\Big|\, \mathcal{F}_{\tau_h}\Big\} = Q_{t-\tau_h} f(X) \;\text{ a.s. } Q_t.$$

The assertion now follows, since f was arbitrary. □

Next, we prove that Q_t is invariant under suitable time reversals on the interval $[0,t]$. Here we need some special notation, due to the subtlety of the subject. Form a process $\tilde{X}$ from X by reversing all excursions. Letting $[\sigma_n, \tau_n]$, $n \in \mathsf{N}$, be the excursion intervals of X, we define

$$\tilde{X}(t) = X(\sigma_n + \tau_n - t), \quad t \in [\sigma_n, \tau_n], \;\; n \in \mathsf{N},$$

and put $\tilde{X}_t = X_t$ otherwise. For any $t \geq 0$, write $\tilde{Q}_t$ for the distribution of the pair $(\tilde{X}, \xi)$, under the Palm distribution Q_t. On the interval $[0,t]$, we also introduce the reversals $R_t X$ and $R_t \xi$ of X and ξ, given by

$$(R_t X)_s = X_{t-s}, \qquad s \in [0,t],$$
$$(R_t \xi)B = \xi(t-B), \qquad B \in \mathcal{B}_{[0,t]},$$

where $t - B = \{t - s;\ s \in B\}$. In the regular case, the two kinds of reversal are equivalent under each Q_t:

Corollary 12.43 *(time reversal) Let X be a regular, regenerative process with Palm distributions Q_t. Then for any $t > 0$ with $p_t < \infty$,*

(i) $\tilde{Q}_t = Q_t \circ R_t^{-1}$ *on* $[0,t]$,

(ii) $\tilde{Q}_t\{X_{0+} = 0\} = 1$.

Proof: (i) Writing (ν_r) for the excursion kernel of X, we define $\tilde{\nu}_r = \nu_r \circ R_t^{-1}$ for each $r > 0$. By Theorem 12.41 (v) we see that, under Q_t, the excursions of $\tilde{X}$ are conditionally independent, given ξ, with distributions given by the kernel $(\tilde{\nu}_r)$. The same property clearly holds for the excursions of X itself, under the measure $\tilde{Q}_t$. It remains to note that

$$Q_t \circ (R_t\, \xi)^{-1} = Q_t \circ \xi^{-1} = \tilde{Q}_t \circ \xi^{-1} \ \text{ on } \ [0,t],$$

where the first equality holds by the exchangeability of ξ under Q_t, while the second one is part of the definition of $\tilde{Q}_t$.

(ii) For any $h > 0$, we have

$$\sup_{s \leq h} \rho(\tilde{X}_s, 0) \leq \sup_{s \leq \tau_h} \rho(\tilde{X}_s, 0) = \sup_{s \leq \tau_h} \rho(X_s, 0).$$

Now clearly $\tau_h \to 0$ a.s. Q_0 as $h \to 0$, which remains true under Q_t, by Theorem 12.41 (i). Since $X_{0+} = X_0 = 0$ a.s., we get $\rho(\tilde{X}_s, 0) \to 0$ a.s. Q_t as $s \to 0$, which means that $\tilde{X}_{0+} = 0$ a.s. Q_t. □

We proceed to some asymptotic continuity and independence properties of the Palm distributions Q_t of X on $[0,t]$. More specifically, we consider some asymptotic factorizations of the Palm measures Q_s on $[0,u] \cup [s-v, s]$, under each of the four conditions $u \to 0$, $v \to 0$, $u+v \to 0$, and $s-u-v \to \infty$. Here the required continuity of the densities p_t and associated Palm distributions Q_t is essential. Write $\|\cdot\|_{u,v}$ for the total variation of the restrictions to $[0,u] \times [0,v]$.

Theorem 12.44 *(Palm factorization) Let X be a regular, regenerative process with Palm distributions Q_t, and fix any continuity point $t > 0$ of p and an $r \in (0,t)$. Then as $s \to t$ and $h \to 0$,*

(i) $\left\| Q_s \circ (X, \tilde{X}_s)^{-1} - Q_0 \otimes \tilde{Q}_t \right\|_{h,r} \to 0,$

(ii) $\left\| Q_s \circ (X, \tilde{X}_s)^{-1} - Q_t \otimes \tilde{Q}_0 \right\|_{r,h} \to 0,$

(iii) $\left\| Q_s \circ (X, \tilde{X}_s)^{-1} - Q_0 \otimes \tilde{Q}_0 \right\|_{h,h} \to 0.$

If also $p_t \to (a + \nu l)^{-1} > 0$ *as* $t \to \infty$*, then as* $s - u - v \to \infty$*,*

(iv) $\left\| Q_s \circ (X, \tilde{X}_s)^{-1} - Q_0 \otimes \tilde{Q}_0 \right\|_{u,v} \to 0.$

When $a > 0$, Theorem 12.33 (ii) yields $p_t \to m^{-1}$ with $m = a + \nu l$. However, this convergence may fail when $\hat{\mu}_s^{-1} \in L^1$ for all $s > 0$. From (iv) we see in particular that, under the stated condition, $\|Q_t - Q_0\|_u$ as $t - u \to \infty$.

Proof: (i) By Corollary 12.42 and a monotone-class argument, we get for any measurable function $f \geq 0$ on D^2

$$p_t\, Q_t f\Big(X^{\tau_h}, \theta_{\tau_h} X\Big) = E\, p(t - \tau_h) \int f(X^{\tau_h}, y)\, Q_{t-\tau_h}(dy), \tag{18}$$

and so by Corollary 12.43 (i),

$$\begin{aligned} &p_s\, Q_s\Big\{ f(X^h, \tilde{X}_s^r);\ \tau_h \leq s - r \Big\} \\ &\qquad = E\Big\{ p(s - \tau_h) \int f(X^h, y^r)\, \tilde{Q}_{s-\tau_h}(dy);\ \tau_h \leq s - r \Big\}. \end{aligned}$$

By Fubini's theorem, we also have

$$(\, Q_0 \otimes \tilde{Q}_t) f = E \int f(X, y)\, \tilde{Q}_t(dy),$$

and so by combination,

$$\begin{aligned} &\left\| Q_s \circ (X, \tilde{X}_s)^{-1} - Q_0 \otimes \tilde{Q}_t \right\|_{h,r} \\ &\qquad \leq Q_s\{\tau_h > s - r\} + E\left| \frac{p(s - \tau_h)}{p(s)} - 1 \right| + E\left\| \tilde{Q}_{s-\tau_h} - \tilde{Q}_t \right\|_r. \end{aligned} \tag{19}$$

Now $\tau_h \to 0$ a.s. as $h \to 0$, and so $p(s - \tau_h) \to p_t$ and $1\{\tau_h > s - r\} \to 0$ a.s. Since $Ep(s - \tau_h) = p_s \to p_t$ by Lemma 12.38, we get by extended dominated convergence (FMP 1.21)

$$Q_s\{\tau_h > s - r\} = p_s^{-1} E\Big\{ p(s - \tau_h);\ \tau_h > s - r \Big\} \to 0.$$

The second term in (19) tends to 0 by FMP 4.12, since $p(s - \tau_h)/p_s \to 1$ a.s. and $Ep(s - \tau_h)/p_s \equiv 1$. Finally, the third term in (19) tends to 0 by dominated convergence, since the continuity in total variation of the distributions Q_s implies the corresponding property for the measures $\tilde{Q}_s$.

(ii) Equation (18) implies the reverse relation

$$p_t\, \tilde{Q}_t f\Big(X^{\tau_h}, \theta_{\tau_h} X\Big) = \tilde{Q}_0 p(t - \tau_h) \int f(X^{\tau_h}, y)\, \tilde{Q}_{t-\tau_h}(dy).$$

Proceeding as before, we may derive a reversal of (i), of the form

$$\left\| \tilde{Q}_s \circ (X, \tilde{X}_s)^{-1} - \tilde{Q}_0 \otimes Q_t \right\|_{h,r} \to 0.$$

which is equivalent to (ii), by Corollary 12.43.

(iii) Combine (i) and (ii).

(iv) Writing (19) with h, r replaced by u, v, and arguing as in (i), we may reduce our proof to the claims $P\{\tau_u > s - v\} \to 0$ and $\|Q_s - Q_0\|_u \to 0$. Here the former assertion follows, by Lemma 12.8, from the assumption $s - u - v \to \infty$. To prove the latter convergence, we see from Theorem 12.41 (i) that

$$Q_s\{X^u \in \cdot\} = p_s^{-1}\, E\Big\{p(s - \tau_u);\ X^u \in \cdot\Big\}, \quad 0 < u < s,$$

which shows that $\|Q_s - Q_0\|_u$ is bounded by the second term in (19). □

12.8 Local Hitting and Conditioning

A regenerative set Ξ is said to *hit* a fixed set B, if $\Xi \cap B \neq \emptyset$. Here we study the asymptotic hitting probabilities for small intervals I, along with the asymptotic distributions of X and ξ, given that such a hit occurs. More generally, we consider the corresponding probabilities and distributions, for the case where Ξ hits simultaneously some small neighborhoods $I_1, \dots, I_n$ around the given points $t_1, \dots, t_n$.

Our fundamental Theorems 12.45 and 12.49 show that, under suitable regularity conditions, the contributions to those sets are asymptotically stationary and independent. In other words, the local distribution is essentially the same as if X were replaced by n independent copies of the stationary version Y in Theorem 12.19. Furthermore, the conditional distributions outside $I_1, \dots, I_n$ are shown in Theorem 12.50 to agree asymptotically with the corresponding multi-variate Palm distributions. Note the close analogy with the general results for simple point processes in Section 6.6.

First we consider the asymptotic hitting probabilities. Truncating the interval lengths, if necessary, we may assume that the rate $m = a + \nu l$ is finite, so that a stationary version of X exists. For technical convenience, we may take Ω to be the path space of X, and write the stationary distribution in Theorem 12.2 as $\bar{\mathcal{L}}(X) = \bar{P}\{X \in \cdot\}$. The convergence $I_n \to t$ should be understood in the pointwise sense. We always assume ξ to be regular, in the sense of Theorem 12.33, to ensure the existence of a nice density $p > 0$ of $E\xi$. Write $\Delta t_j = t_j - t_{j-1}$ for $j = 1, \dots, n$, where $t_0 = 0$. In the asymptotic relations below, we assume the function $f \geq 0$ to be such that the associated integrals on both sides are finite. Here the excursion endpoints at s are often written as $\sigma_s^{\pm}$.

Theorem 12.45 *(local hitting factorizations) Let Ξ be a regular, regenerative set with local time ξ, and fix any $t_1 < \cdots < t_n$ such that p is continuous at each Δt_j. Then as $s_j \in I_j \to t_j$ for all j,*

(i) $$P\bigcap_{j\le n}\{\xi I_j > 0\} \sim m^n \prod_{j\le n} p(\Delta t_j)\, \bar{P}\{\xi I_j > 0\},$$

(ii) $$Ef(\sigma_{s_1}^{\pm}, \ldots, \sigma_{s_n}^{\pm}) \sim m^n \Big\{\prod\nolimits_i p(\Delta t_i)\Big\}\Big\{\bigotimes\nolimits_i \bar{\mathcal{L}}(\sigma_{s_i}^{\pm})\Big\} f,$$

uniformly over the class of measurable functions $f \ge 0$, supported by $I_1^2 \times \cdots \times I_{n-1}^2 \times I_n \times \mathsf{R}_+$.

The result follows immediately from the following estimates. For any intervals $I_1, \ldots, I_n$, we write $\Delta I_j = I_j - I_{j-1}$ for $j = 1, \ldots, n$, where $I_0 = \{0\}$. Put

$$p^+(B) = \sup_{s\in B} p(s), \qquad p^-(B) = \inf_{s\in B} p(s).$$

Lemma 12.46 *(hitting estimates) Consider a regular, regenerative set Ξ with local time ξ, some disjoint intervals $I_1, \ldots, I_n \subset \mathsf{R}_+$ and times $t_j \in I_j$, and a measurable function $f > 0$, supported by $I_1^2 \times \cdots \times I_{n-1}^2 \times I_n \times \mathsf{R}_+$. Then*

(i) $$m^n \prod_{j\le n} p^-(\Delta I_j) \le \frac{P\bigcap_j\{\xi I_j > 0\}}{\prod_j \bar{P}\{\xi I_j > 0\}} \le m^n \prod_{j\le n} p^+(\Delta I_j),$$

(ii) $$m^n \prod_{j\le n} p^-(\Delta I_j) \le \frac{Ef(\sigma_{t_1}^{\pm}, \ldots, \sigma_{t_n}^{\pm})}{\big\{\bigotimes_j \bar{\mathcal{L}}(\sigma_{t_j}^{\pm})\big\} f} \le m^n \prod_{j\le n} p^+(\Delta I_j).$$

Proof: Writing $\hat{\nu} = \nu \circ l^{-1}$, we see from Theorem 12.28 that

$$\begin{aligned} Ef(\sigma_t^{\pm}) &= a\, f(t,t)\, p(t) + \int f(s,u)\, p(s)\, \hat{\nu}(u-s, \infty)\, ds \\ &\le m\, p^+(I)\, \bar{E} f(\sigma_t^{\pm}), \end{aligned}$$

and similarly for the lower bound, proving (ii) for $n = 1$. Taking $f(s,t) = 1_{I_1}(s)$ yields (i) for $n = 1$.

Now suppose that (i) holds for up to $n-1$ intervals. Proceeding by induction, we turn to the case of n intervals $I_j = [u_j, v_j]$. Using the regenerative property at τ_{u_1}, the disintegration theorem, and the induction hypothesis, we get

$$\begin{aligned} P\bigcap_{j\le n}\{\xi I_j > 0\} &= \int_{I_1} P\{\tau_{u_1} \in ds\}\, P\bigcap_{j\ge 2}\Big\{\xi(I_j - s) > 0\Big\} \\ &\le P\{\xi I_1 > 0\}\, m^{n-1} \prod_{j\ge 2} p^+(\Delta I_j)\, \bar{P}\{\xi I_j > 0\} \\ &\le m^n \prod_{j\le n} p^+(\Delta I_j)\, \bar{P}\{\xi I_j > 0\}, \end{aligned}$$

as required. The proof of the lower bound is similar.

Next, suppose that (ii) holds for up to $n-1$ points t_k. In case of n points, we first consider tensor products $f = \bigotimes_k f_k$, such that each f_k is measurable and supported by I_k^2 for $k < n$, and by $I_n \times \mathsf{R}$ for $k = n$. Using the regenerative property at τ_{t_1}, the disintegration theorem, the induction hypothsis, and the stationarity under $\bar{P}$, we get

$$\begin{aligned} E \prod_{k\le n} f_k(\sigma_{t_k}^\pm) &= E f_1(\sigma_{t_1}^\pm)\, E\Big\{ \prod_{k\ge 2} f_k(r + \sigma_{t_k - r}^\pm) \Big\}_{r=\tau_{t_1}} \\ &\le m^{n-1} \prod_{k\ge 2} p^+(\Delta I_k)\, E f_1(\sigma_{t_1}^\pm) \prod_{k\ge 2} \bar{E}\Big\{ f_k(r + \sigma_{t_k-r}^\pm) \Big\}_{r=\tau_{t_1}} \\ &= m^{n-1} \prod_{k\ge 2} p^+(\Delta I_k)\, E f_1(\sigma_{t_1}^\pm) \prod_{k\ge 2} \bar{E} f_k(\sigma_{t_k}^\pm) \\ &\le m^{n} \prod_{k\le n} p^+(\Delta I_k) \prod_{k\le n} \bar{E} f_k(\sigma_{t_k}^\pm), \end{aligned}$$

and similarly for the lower bound. The relation extends, by additivity on both sides, to finite sums of tensor products, and then, by a monotone-class argument, to arbitrary functions $f \ge 0$ with the stated support. This completes the induction. □

Next, we compare the one- and two-sided hitting probabilities.

Lemma 12.47 *(hitting comparison) Let Ξ be a regular, regenerative set with local time ξ, and fix any $t_1 < \cdots < t_n$ such that p is continuous at each Δt_j. Then as $s_j \in I_j \to t_j$ for all j,*

(i) $E\Big(\max_i (\tau_{s_i} - \sigma_{s_i}) \wedge 1 \,\Big|\, \bigcap_i \{\xi I_i > 0\} \Big) \to 0,$

(ii) $P\Big(s_1, \ldots, s_n \in \Xi \,\Big|\, \bigcap_i \{\xi I_i > 0\} \Big) \to 0$ *when* $a > 0$.

Proof: (i) By Lemma 12.26 and Fubini's theorem, we have for any $h > 0$

$$\begin{aligned} m\, \bar{P}\Big\{ \xi[0,h] > 0 \Big\} &= a + \nu(l \wedge h) \\ &= a + \int_0^h \hat{\nu}(t,\infty)\, dt. \end{aligned}$$

Denoting the right-hand side by $m(h)$, we get for $h, k > 0$ by the same result

$$\begin{aligned} m\, \bar{P}\Big\{ \xi[-h,0] \wedge \xi[0,k] > 0 \Big\} &= a + \int_0^h \hat{\nu}(s, s+k)\, ds \\ &= m(h) + m(k) - m(h+k). \end{aligned}$$

Combining the two formulas, we get for any two adjacent intervals I and I' of lengths h and k

$$\bar{P}(\,\xi I' > 0 \,|\, \xi I > 0) = 1 - \frac{m(h+k) - m(k)}{m(h)},$$

which tends to 1 as $h \to 0$, since $m'(k) = \hat{\nu}(k,\infty) < \infty$, and by monotone convergence,

$$h^{-1}m(h) \to a \cdot \infty + \hat{\nu}(0,\infty) = \infty.$$

Now consider for every $j \le n$ some adjacent intervals I_j and I'_j with union U_j, and conclude from Lemma 12.46 (i)–(ii) that

$$P\Big(\bigcap_i \{\xi I'_i > 0\} \,\Big|\, \bigcap_i \{\xi I_i > 0\} \Big) \ge \prod_{j \le n} \frac{p^-(\Delta U_j)}{p^+(\Delta I_j)}\, \bar{P}(\,\xi I'_j > 0 \,|\, \xi I_j > 0),$$

which tends to 1 for all j, as $I_j \to t_j$ and then $I'_j \to t_j$. By monotonicity, the left-hand-side tends to 1, already as $I_j \to t_j$ for all j. Writing $I_j = [u_j, v_j]$, we get for any $\varepsilon > 0$

$$P\Big(\max_i (\tau_{v_i} - v_i) > \varepsilon \,\Big|\, \bigcap_i \{\xi I_i > 0\} \Big) \to 0,$$
$$P\Big(\max_i (u_i - \sigma_{u_i}) > \varepsilon \,\Big|\, \bigcap_i \{\xi I_i > 0\} \Big) \to 0,$$

and the assertion follows, since

$$\tau_{s_j} - \sigma_{s_j} \le (\tau_{v_j} - v_j) + (v_j - u_j) + (u_j - \sigma_{u_j}).$$

(ii) When $a > 0$, the regenerative property at $s_1, \ldots, s_n$ yields

$$P\{s_1, \ldots, s_n \in \Xi\} = a^n \prod_{j \le n} p(\Delta s_j) = \prod_{j \le n} p(\Delta s_j)\, m(0).$$

Combining with Lemma 12.46 (i), we obtain

$$P\Big(s_1, \ldots, s_n \in \Xi \,\Big|\, \bigcap_i \{\xi I_i > 0\} \Big) \ge \prod_{j \le n} \frac{p(\Delta s_j)}{p^+(\Delta I_j)}\, \frac{m(0)}{m(h_j)} \to 1.$$

□

The last two results may be combined into an approximation property for the conditional distributions of the excursion endpoints σ_s and τ_s, sometimes written as $\sigma_s^\pm$. As before, let $\gamma_s = \tau_s - \sigma_s$ denote the length of the excursion straddling s.

Corollary 12.48 *(excursion endpoints) Let Ξ be a regular, regenerative set with local time ξ, and fix any $t_1 < \cdots < t_n$ such that p is continuous at each Δt_j. Then as $s_j \in I_j \to t_j$ and $h_j \to 0$ for all j,*

(i) $\Big\| \mathcal{L}\Big(\sigma^\pm_{s_1}, \ldots, \sigma^\pm_{s_n} \,\Big|\, \bigcap_i \{\xi I_i > 0\} \Big) - \bigotimes_i \bar{\mathcal{L}}\Big(\sigma^\pm_{s_i} \,\Big|\, \xi I_i > 0 \Big) \Big\| \to 0,$

(ii) $E\Big(f(\sigma^\pm_{s_1}, \ldots, \sigma^\pm_{s_n}) \,\Big|\, \bigcap_i \{\gamma_{s_i} \le h_i\} \Big) \sim \Big(\bigotimes_i \bar{\mathcal{L}}\Big\{ \sigma^\pm_{s_i} \,\Big|\, \gamma_{s_i} \le h_i \Big\} \Big) f,$

uniformly for measurable functions $f \ge 0$ on R_+^{2n}.

Proof: (ii) Use Theorem 12.45.

(i) For any disjoint intervals $J_j \supset I_j$, we see from Lemma 12.46 (i)−(ii) that

$$\begin{aligned}
&\Big\| \mathcal{L}\Big(\sigma^\pm_{s_1}, \dots, \sigma^\pm_{s_n} \,\Big|\, \bigcap_i \{\xi I_i > 0\}\Big) - \bigotimes_i \bar{\mathcal{L}}\Big(\sigma^\pm_{s_i} \,\Big|\, \xi I_i > 0\Big)\Big\| \\
&\quad \le \prod_{i \le n} \frac{p^+(\Delta J_i)}{p^-(\Delta I_i)} - \prod_{i \le n} \frac{p^-(\Delta J_i)}{p^+(\Delta I_i)} \\
&\qquad + P\Big(\bigcup_i \{\sigma^\pm_{s_i} \notin J_i\} \,\Big|\, \bigcap_i \{\xi I_i > 0\}\Big) + \sum_{j \le n} \bar{P}\Big(\sigma^\pm_{s_j} \notin J_j \,\Big|\, \xi I_j > 0\Big).
\end{aligned}$$

By Lemma 12.47 and its proof, the last two terms tend to 0, as $I_j \to t_j$ for all j. Finally, let $J_j \to t_j$ for all j, and use the continuity of p at each Δt_j. □

Our next aim is to study the asymptotic behavior of the conditional distributions of X near the points $t_1, \dots, t_n$, given that Ξ hits some small neighborhoods $I_1, \dots, I_n$ of those points. The approximation holds in the sense of total variation, for the restrictions of X to the intervals $J_1, \dots, J_n$, here expressed in terms of the norm $\|\cdot\|_{J_1,\dots,J_n}$. To avoid repetitions, we give a combined statement for two kinds of hitting conditions H_j.

Theorem 12.49 *(local inner factorization) Let X be a regular, regenerative process with local time ξ, fix any $t_1 < \cdots < t_n$ such that p is continuous at each Δt_j, and put $H_j = \{\xi I_j > 0\}$ or $\{\gamma_{s_j} \le h_j\}$. Then as $s_j \in I_j \to t_j$, $J_j \to t_j$, and $h_j \to 0$ for all j, we have*

$$\Big\| \mathcal{L}\Big(X \,\Big|\, \bigcap_i H_i\Big) - \bigotimes_i \bar{\mathcal{L}}(\, X |\, H_i\,) \Big\|_{J_1,\dots,J_n} \to 0.$$

Proof: Let $s_j \in I_j$, and put $\beta_j = \sigma^-_{s_j} - \sigma^+_{s_{j-1}}$, where $s_0 = 0$. Assuming $J_j \subset [s_j - h, s_j + h]$ for all j, we get by Lemma 1.19 and Theorems 12.2 and 12.28

$$\begin{aligned}
&\Big\| \mathcal{L}\Big(X \,\Big|\, \bigcap_i H_i\Big) - \bigotimes_i \bar{\mathcal{L}}(\, X |\, H_i) \Big\|_{J_1,\dots,J_n} \\
&\quad \le \Big\| \mathcal{L}\Big(\sigma^\pm_1, \dots, \sigma^\pm_n \,\Big|\, \bigcap_i H_i\Big) - \bigotimes_i \bar{\mathcal{L}}\Big(\sigma^\pm_i \Big|\, H_i\Big) \Big\| \\
&\qquad + \sum_{k \le n} E\Big(\Big\| \Big\{Q_s \circ (X, \tilde{X}_s)^{-1}\Big\}_{s=\beta_k} - Q_0 \otimes \tilde{Q}_0 \Big\|_{h,h} \,\Big|\, \bigcap_i H_i\Big).
\end{aligned}$$

By Corollary 12.48 (i)−(ii), the first term on the right tends to 0, as $I_j \to t_j$ and $h_j \to 0$ for all j. To estimate the last sum, we see from Lemma 12.47 (i) that, as $I_j \to t_j$ or $h_j \to 0$, the β_j tend in probability to t_j, under the conditional distribution $P(\,\cdot \,|\, \bigcap_j H_j)$. By Theorem 12.44 (iii), we have in the same sense

$$\Big\| \Big\{Q_s \circ (X, \tilde{X}_s)^{-1}\Big\}_{s=\beta_j} - Q_0 \otimes \tilde{Q}_0 \Big\|_{h,h} \to 0, \quad j \le n,$$

as $I_j \to t_j$ or $h_j \to 0$. The desired convergence now follows by dominated convergence. □

Next, we show how the multi-variate Palm measures of X can be approximated by elementary conditional distributions. To ensure the desired weak convergence, we introduce the regularity condition

$$\liminf_{h\to 0} h^{-1}\Big(a + \nu\{\,l;\, l \le h\,\}\Big) > 0, \tag{20}$$

which may be compared with the conditions in Theorem 12.36.

Theorem 12.50 *(Palm approximation) Let X be a regular, regenerative process with local time ξ, fix any $t_1 < \cdots < t_n$ such that p is continuous at each Δt_j, and put $H_j = \{\xi I_j > 0\}$ or $\{\gamma_{s_j} \le h_j\}$. Then as $s_j \in I_j \to t_j$ and $h_j \to 0$ for all j,*

$$\mathcal{L}\Big(X, \xi \,\Big|\, \bigcap\nolimits_i H_i\Big) \to Q_{t_1,\dots,t_n},$$

in total variation, outside any neighborhood of $t_1, \dots, t_n$. When $H_j = \{\xi I_j > 0\}$, this holds even weakly for the uniform topology on D, and under (20) *it is also true when $H_j = \{\gamma_{s_j} \le h_j\}$.*

Our proof is based on two lemmas, beginning with the following continuity property under shifts. For $s < t$, let Q_s^t denote the distribution of $\theta_{-s}X$ under the measure Q_{t-s}, where $(\theta_{-s}X)_r = X_{r-s}$ for $r \in [s, t)$, and $(\theta_{-s}X)_r = 0$ otherwise.

Lemma 12.51 *(shifted Palm distributions) Let X be a regular, regenerative process with Palm distributions Q_t, and let the times $s < t$ be such that p is continuous at $t - s$. Then the kernel $Q_s^t = Q_{t-s} \circ \theta_{-s}^{-1}$ is jointly continuous in total variation, on every interval $[u, v] \subset (s, t)$, and also weakly for the uniform topology on D.*

Proof: Writing $\|\cdot\|_u^v$ for the total variation on $[u, v]$, we get for any $s' < u$ and $t' > v$

$$\begin{aligned}
\Big\|\, Q_{s'}^{t'} - Q_s^{t} \,\Big\|_u^v &\le \Big\|\, Q_{s'}^{t'} - Q_{s'}^{t} \Big\|_u^v + \Big\|\, Q_{s'}^{t} - Q_s^{t} \,\Big\|_u^v \\
&\le \|\, Q_{t'-s'} - Q_{t-s'} \|_{v-s'} + \|\, \tilde{Q}_{t-s'} - \tilde{Q}_{t-s} \|_{t-u} \\
&\le \|\, Q_{t'-s'} - Q_{t-s} \|_{v-s'} + \|\, Q_{t-s'} - Q_{t-s} \|_{v-s'} \\
&\qquad + \Big\|\, \tilde{Q}_{t-s'} - \tilde{Q}_{t-s} \Big\|_{t-u},
\end{aligned}$$

which tends to 0 as $s' \to s$ and $t' \to t$, by Theorem 12.44 (i)−(ii). To deduce the stated weak convergence, write $X_h^* = \sup_{s \le h} \rho(X_s, 0)$, and note that for any $h < r$,

$$\begin{aligned}
Q_r(X_h^* \wedge 1) &\le \|\, Q_r - Q_{t-s} \|_h + Q_{t-s}(X_h^* \wedge 1), \\
\tilde{Q}_r(X_h^* \wedge 1) &\le \|\, \tilde{Q}_r - \tilde{Q}_{t-s} \|_h + \tilde{Q}_{t-s}(X_h^* \wedge 1),
\end{aligned}$$

where the expressions on the right tend to 0 as $r \to t - s$ and $h \to 0$, by Theorem 12.44 (iii), and the facts that $X_{0+} = 0$ and $\tilde{X}_{0+} = 0$ a.s., by Lemma 12.43 (ii). The assertion now follows by Lemma 4.21. □

To estimate the excursions across $s_1, \dots, s_n$, we define the *height* of an excursion Y on the interval I as the random variable $\beta = \sup_{r \in I} \rho(Y_r, 0)$, where ρ is an arbitrary metrization of S.

Lemma 12.52 *(conditional length and height) Let X be a stationary, regenerative process with local time ξ, and write β and γ for the height and length of the excursion across* 0*. Then*

(i) $E(\beta \wedge 1 \,|\, \xi I > 0) \to 0$ *as* $I \downarrow \{0\}$,

(ii) *under* (20)*, we have* $E(\beta \wedge 1 \,|\, \gamma \le h) \to 0$ *as* $h \to 0$.

Proof: Fix any $r > 0$. For the non-delayed process X with associated inverse local time T, let κ_r be the number of excursions x up to time T_1 of height $x^* > r$, and note that $\kappa_r < \infty$ a.s., since X is rcll. The κ_r are Poisson by Theorem 12.15, and so

$$\nu\{x^* > r\} = E\kappa_r < \infty, \quad r > 0. \tag{21}$$

Then as $|I| = h \to 0$ for fixed $r > 0$, we get by Theorem 12.19

$$\begin{aligned} P(\beta > r \,|\, \xi I > 0) &= \frac{\nu(l \wedge h;\, x^* > r)}{a + \nu(l \wedge h)} \\ &\le \frac{h\,\nu\{x^* > r\}}{a + h\,\nu\{l \ge h\}} \to 0, \end{aligned}$$

since either $a > 0$ or $\nu(0, \infty) = \infty$. On the other hand, we see from (20), (21), and Theorem 12.19 that

$$\begin{aligned} P(\beta > r \,|\, \gamma \le h) &= \frac{\nu(l;\, x^* > r,\, l \le h)}{a + \nu(l;\, l \le h)} \\ &\le \frac{h\,\nu\{x^* > r,\, l \le h\}}{a + \nu(l;\, l \le h)} \\ &\lesssim \nu\{x^* > r,\, l \le h\} \to 0. \end{aligned}$$

Letting $A_h = \{\xi I > 0\}$ or $\{\gamma \le h\}$, we obtain $P(\beta > r \,|\, A_h) \to 0$, and so, by dominated convergence as $h \to 0$,

$$E(\beta \wedge 1 \,|\, A_h) = \int_0^1 P(\beta > r \,|\, A_h)\, dr \to 0.$$ □

Proof of Theorem 12.50: Let $H_j = \{\xi I_j > 0\}$ or $\{\gamma_{s_j} \le h_j\}$, and write $\sigma_{s_j} = \sigma_j$ and $\tau_{s_j} = \tau_j$, for simplicity. Fix any closed intervals

$$J_1 \subset [0, t_1),\ J_2 \subset (t_1, t_2),\ \dots,\ J_{n+1} \subset (t_n, \infty).$$

By Lemma 1.19 and Theorems 12.28 and 12.29, we have

$$\begin{aligned}\Big\| \mathcal{L}\Big(X, \xi \,\Big|\, \textstyle\bigcap_i H_i\Big) - Q_{t_1,\dots,t_n}\Big\|_{J_1,\dots,J_{n+1}} &\le \sum_{j\le n} E\Big(\Big\| Q^{\sigma_j}_{\tau_{j-1}} - Q^{t_j}_{t_{j-1}}\Big\|_{J_j};\, J_j \subset [\tau_{j-1},\sigma_j)\,\Big|\, \textstyle\bigcap_i H_i\Big)\\ &\quad + \sum_{j\le n} P\Big(J_j \not\subset [\tau_{j-1},\sigma_j)\,\Big|\, \textstyle\bigcap_i H_i\Big).\end{aligned}$$

By Corollary 12.39, Lemmas 12.47 and 12.51, and dominated convergence, the right-hand side tends to 0, as $s_j \in I_j \to t_j$ and $h_j \to 0$ for all j.

To prove the stated weak convergence, fix any $h > 0$, and assume that $J_j \subset [s_j - h,\, s_j + h]$ for all j. Then Theorem 12.28 yields

$$\begin{aligned}E\Big\{\big\|\rho(X,0)\wedge 1\big\|_{J_1,\dots,J_n}\Big|\textstyle\bigcap_i H_i\Big\} &\le \sum_{j\le n} E\Big\{(\,Q_{\beta_j} - \tilde{Q}_{\beta_j})(X^*_h \wedge 1);\, \max_i \gamma_i \le h\,\Big|\textstyle\bigcap_i H_i\Big\}\\ &\quad + \sum_{j\le n} E\Big\{\nu_{\gamma_j}(X^* \wedge 1)\,\Big|\textstyle\bigcap_i H_i\Big\} + P\Big\{\max_i \gamma_i > h\,\Big|\textstyle\bigcap_i H_i\Big\}.\end{aligned}$$

Using Theorem 12.49 and Lemmas 12.47 and 12.52, and arguing as in the proof of Lemma 12.51, we see that the right-hand side tends to 0, as $I_j \to t_j$ or $h_j \to 0$, and then $h \to 0$. Similarly, Theorem 12.29 yields for $J_j \subset [t_j - h,\, t_j + h]$

$$Q_{t_1,\dots,t_n}\Big\|\,\rho(X,0)\wedge 1\,\Big\|_{J_1,\dots,J_n} \le \sum_{j\le n}\Big(Q_{\Delta t_j} + \tilde{Q}_{\Delta t_j}\Big)\,(X^*_h \wedge 1),$$

which tends to 0 as $h \to 0$, by Corollary 12.43 (ii). Now use Lemma 4.21. □

Combining the last theorem with the Palm factorization in Theorem 12.29, we obtain the following asymptotic outer factorization, which may be compared with the inner factorization in Theorem 12.49.

Corollary 12.53 *(local outer factorization) Let X be a regular, regenerative process with local time ξ, fix any $t_1 < \cdots < t_n$ such that p is continuous at each Δt_j, and put $H_j = \{\xi I_j > 0\}$ or $\{\gamma_{s_j} \le h_j\}$. Then for any closed intervals $J_j \subset (t_{j-1}, t_j)$ with $j = 1,\dots,n+1$, we have as $s_j \in I_j \to t_j$ and $h_j \to 0$ for all j*

$$\Big\|\,\mathcal{L}\Big(X\,\Big|\,\textstyle\bigcap_i H_i\Big) - \textstyle\bigotimes_i \mathcal{L}(X|\,H_{i-1}\cap H_i\,)\,\Big\|_{J_1,\dots,J_{n+1}} \to 0.$$

Proof: By Lemma 1.19 and Theorem 12.29, the left-hand side is bounded by

$$\begin{aligned}&\Big\|\,\mathcal{L}\Big(X\,\Big|\,\textstyle\bigcap_i H_i\Big) - Q_{t_1,\dots,t_n}\Big\|_{J_1,\dots,J_{n+1}}\\ &\qquad + \sum_{j\le n}\Big\|\,Q_{t_{j-1},t_j} - \mathcal{L}(X|\,H_{j-1}\cap H_j\,)\,\Big\|_{J_j},\end{aligned}$$

which tends to 0 as $s_j \in I_j \to t_j$ and $h_j \to 0$, by Theorem 12.50. □

Chapter 13

Branching Systems and Super-processes

Branching systems of various kinds form another important application area of random measure theory. The literature on the subject is absolutely staggering in both volume and depth, which forces us to be very selective. Here we will focus on some aspects that are closely related to material in previous chapters, including the roles of clustering, Poisson and Cox processes, Palm distributions, and local approximations. In the case of super-processes, the basic existence and uniqueness theorems, as well as some key estimates, will be stated without proof.

Though super-processes are our main concern, we begin the chapter with a brief discussion of some discrete branching systems, mainly to motivate the subsequent results on genealogy and cluster structure. The simplest branching system is the classical *Bienaymé process*[1], which records only the total population size at integral times. Here each individual gives rise to a family of offspring in the next generation, independently and according to a fixed distribution. The process is said to be *critical* if the offspring distribution has mean 1.

As an obvious continuous-time counterpart, we may allow the life lengths of the individual particles to be i.i.d. random variables, independent of the branching. To ensure that the resulting process will be Markov, we need to choose the life-length distribution to be exponential. However, the precise nature of the branching mechanism makes little difference to the asymptotic behavior, as long as the first and second moments are finite.

The next step in complexity is to allow the particles to move around in a suitable state space, usually taken to be Euclidean. The most basic of such branching systems is arguably the *branching Brownian motion*, where the particles perform independent Brownian motions during their life spans. At the end of life, each particle either dies or splits into two new particles, with equal probability $\frac{1}{2}$. If we start with a Poisson process in R^d with bounded intensity μ, the entire branching system becomes infinitely divisible, with the individual clusters behaving like independent processes of the same kind starting from single individuals.

[1] traditionally known as a *Galton–Watson process*

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7_13

At this stage we may perform a scaling, on three different levels: increasing the initial intensity by a factor n, speeding up the branching by a factor n, and reducing the weight of each particle by a factor n^{-1}. The entire branching process, including the underlying historical structure, then converges in distribution toward a weakly continuous, measure-valued Markov process ξ starting at $\xi_0 = \mu$, called a *Dawson–Watanabe super-process*, or *DW-process* for short. We mostly consider DW-processes of dimension $d \geq 2$, where the random measures ξ_t with $t > 0$ are a.s. diffuse and singular with Hausdorff dimension 2. The limiting process inherits the fundamental branching property $P_{\mu_1+\mu_2} = P_{\mu_1} * P_{\mu_2}$, which allows us to extend the definition of ξ to a large class of unbounded initial measures μ.

In view of the singularity at fixed times, it is interesting to explore in further detail the local structure of the measures ξ_t. Here a major result is the *Lebesgue approximation*, established in Section 13.7, where we show how ξ_t can be approximated a.s. by suitable restrictions of Lebesgue measure λ^d. More precisely, letting Ξ_t^ε denote the ε-neighborhood of the support $\Xi_t = \operatorname{supp} \xi_t$, with associated Lebesgue restriction $\xi_t^\varepsilon = 1_{\Xi_t^\varepsilon} \lambda^d$, we show that $r_d(\varepsilon)\, \xi_t^\varepsilon \xrightarrow{v} \xi_t$ a.s. as $\varepsilon \to 0$, for some normalizing functions r_d, depending only on dimension d. The latter functions agree essentially with the normalizing factors, arising in the classical estimates of the hitting probabilities $P_\mu\{\xi_t B_x^\varepsilon > 0\}$, where B_x^ε denotes a ball around $x \in \mathsf{R}^d$ of radius $\varepsilon > 0$.

Another local property of interest concerns the conditional distribution of ξ_t in a small ball B_x^ε, given that the latter charges ξ_t. In Section 13.8, we prove the existence, for $d \geq 3$, of a space–time stationary version $\tilde{\xi}$ of the process, which asymptotically gives rise to the same conditional distribution within B_x^ε. Just as for the canonical cluster η, the approximating version $\tilde{\xi}$ exists only as a *pseudo-process*, which means that its "distribution" is unbounded but σ-finite. This causes no problem, since the mentioned conditioning reduces the distribution to a proper probability measure. Though no stationary version $\tilde{\xi}$ exists when $d = 2$, we may still prove a similar approximation in that case, involving instead a certain *stationary cluster* $\tilde{\eta}$, with related but weaker invariance properties.

Our discussion of super-processes concludes in Section 13.9 with some deep results about local hitting and conditioning, similar to those for simple point processes in Section 6.6, and for regenerative processes in Section 12.8. Here our main result is Theorem 13.56, which combines the diverse aspects of the theory into a single statement. The general context is that of multiple hits, where the random measure ξ_t charges every ball in a given collection $B_{x_1}^\varepsilon, \dots, B_{x_n}^\varepsilon$, and we are interested in the asymptotic distributions as $\varepsilon \to 0$.

Aspects covered by our master theorem include:

- asymptotic expressions for the multiple hitting probabilities, in terms of the associated higher order moment densities,
- asymptotic conditional independence between the restrictions to the n balls, as well as to compact subsets of the exterior region,

- approximation in total variation of the conditional distributions of ξ_t within the n balls, in terms of a similar conditioning for the stationary version $\tilde{\xi}$ or the stationary cluster $\tilde{\eta}$,
- approximation in total variation of the conditional distributions of ξ_t, within compact subsets of the exterior region, in terms of the associated multi-variate Palm distributions.

The mentioned results, regarding the local and conditional structure of a DW-process, rely on a wide range of auxiliary results and technical estimates, covered by earlier sections, some of which have considerable intrinsic interest. Foremost among those is the underlying genealogy and cluster structure, motivated in Section 13.1 by some elementary results for branching Brownian motion, and then extended to DW-processes in Section 13.2. Here the main property, constantly used throughout the chapter, is the fact that, for any fixed $t > 0$, the ancestors of ξ_t at an earlier time $s = t - h$ form a Cox process ζ_s on R^d directed by $h^{-1}\xi_s$. This allows us to represent ξ_t as a sum of many small and conditionally independent clusters of age h, originating at the points of ζ_s. As an additional feature, we prove the remarkable fact that, for fixed times $t > 0$, the ancestral processes ζ_s with $s \le t$ combine into a Brownian Yule process on the time scale $A_t = -\log(1 - s/t)$.

Higher order moment measures of a DW-process, studied in Section 13.3, play a crucial role throughout the remaining sections, for several reasons. Apart from their obvious use in providing technical estimates, they also appear explicitly in the asymptotic formulas for multiple hitting probabilities, and their continuity properties lead directly, via the duality theory of Section 6.7, to some crucial regularity properties of multi-variate Palm distributions, needed in Section 13.9.

Our discussion of general moment measures begins in Theorem 13.16 with a basic cluster decomposition, along with some forward and backward recursions. Those results lead in Theorem 13.18 to a remarkable and extremely useful representation, in terms of the endpoint distributions of certain *uniform Brownian trees*. The tree relationship is no coincidence, but reflects a basic tree structure of the higher order Campbell and Palm measures. Here our deep analysis, outlined in Theorem 13.21, is based on an extension of Le Gall's *Brownian snake*.

Basic roles are also played by the various hitting estimates, discussed in Section 13.6. As our starting point, we recall in Theorem 13.38 some classical estimates for hitting probabilities of the form $P_\mu\{\xi_t B_x^\varepsilon > 0\}$, quoted here without proof. They yield a variety of estimates for multiple hits, needed in subsequent sections. The asymptotic rates for $d \ge 3$ are the simple power functions ε^{d-2}, but for $d = 2$ the theory becomes much more subtle, as it involves an unspecified rate function $m(\varepsilon)$, whose boundedness and continuity properties play a crucial role in the sequel.

In the final Section 13.10, we return to the case of clustering in R^d, on the discrete time scale Z_+, where the offspring distribution is given by a non-arithmetic and invariant probability kernel $\nu\colon \mathsf{R}^d \to \mathcal{N}_d$, such that the total offspring size has mean 1. Assuming the initial point process ξ_0 to be stationary with a finite and positive intensity, and writing ξ_n for the generated process at time n, we show that either $\xi_n \xrightarrow{vd} 0$, or that $\xi_n \xrightarrow{vd} \tilde{\xi}$ along sub-sequences, for some non-trivial limiting processes $\tilde{\xi}$. The actual outcome depends only on ν, which in the latter case is said to be *stable*. In Theorem 13.64, we characterize the stability of ν in terms of certain univariate Palm trees[2], constructed recursively from the invariant Palm distribution of ν. This insight enables us to derive some more explicit stability criteria, in terms of the spatial motion and branching mechanism of the underlying process ξ.

13.1 Binary Splitting and Scaling Limits

The simplest branching system is the discrete-time *Bienaymé process* X_n, which records only the total population size at times $n \in \mathsf{Z}_+$. Here each individual begets a family of offspring in the next generation, whose size has a fixed distribution with generating function

$$f(s) = E_1 s^{X_1} = \sum_{k\geq 0} p_k s^k, \quad s \in [0,1],$$

independently for different individuals. This makes X into a discrete-time Markov chain in the state space Z_+, also satisfying the *branching property*

$$\mathcal{L}_{m+n}(X) = \mathcal{L}_m(X) * \mathcal{L}_n(X), \quad m,n \in \mathsf{Z}_+. \tag{1}$$

In fact, all branching systems in this chapter are essentially characterized by the Markov and branching properties (excluding the possibility of immigration).

The generating functions $f_n(s) = E_1 s^{X_n}$ of the n-step offspring distributions satisfy the recursive relation

$$f_{n+1}(s) = f \circ f_n(s), \quad s \in [0,1],\ n \in \mathsf{Z}_+,$$

where $f_0(s) \equiv s$. Assuming the offspring distribution to be non-degenerate, we note that 0 is the only absorbing state, and we say that *extinction* occurs if eventually $X_n = 0$. An elementary argument shows that the extinction probability $P_1\{X_n \to 0\}$ is the smallest root of the equation $f(s) = s$. In particular, extinction occurs a.s. when X is *critical*, in the sense that $E_1 X_1 = 1$.

The previous theory extends immediately to continuous time. To preserve the Markov property in that case, we need the life lengths to be exponentially distributed with a common rate $a > 0$ (FMP 12.16). At the end of life, each

[2]often referred to as *backward trees*

individual is replaced by a family of offspring, whose size is again governed by a fixed generating function f. Here the generating functions

$$F(s,t) = E_1 s^{X_t}, \quad s \in [0,1],\ t \geq 0,$$

are solutions to the *Kolmogorov backward equations* (FMP 12.22)

$$\frac{\partial}{\partial t} F(s,t) = u \circ F(s,t),$$

with initial condition $F(s,0) \equiv s$, where

$$u(s) = a\{f(s) - s\}, \quad s \in [0,1].$$

The solutions are unique, but can be determined explicitly only in very special cases.

The most basic branching systems of this kind are the birth & death processes with rates $n\lambda$ and $n\mu$, for some constants $\lambda, \mu \geq 0$ with $\lambda + \mu = a > 0$. Taking $\lambda = 1$ and $\mu = 0$ yields the *Yule process*, encountered already in Theorem 3.12 (iii). The process with $\lambda = \mu = 1$ is critical, and will be referred to as the *binary splitting* process. If we allow the transition rates λ and/or μ to depend on time, we obtain a branching process on the *time scale* $A_t = \int_0^t \lambda(s)\, ds$.

In addition to the branching, we may allow some spatial motion, in a suitable state space, usually taken to be R^d for some $d \in \mathsf{N}$. In discrete time, we then assume the individuals at points $s \in \mathsf{R}^d$ to give rise to independent families of offspring with distributions ν_s, for some typically invariant probability kernel ν from R^d to $\mathcal{N}_d$. The branching property still takes the form (1), but now for arbitrary $m, n \in \mathcal{N}_d$. The generated process ξ_n in $\mathcal{N}_d$ is said to be *critical*, if the corresponding property holds for the associated counting process $X_n = \|\xi_n\|$ with $X_0 < \infty$.

In continuous time, it may be more natural to let the individual motions of living particles form independent Markov processes with a common generator. To ensure space homogeneity, we need the latter to be Lévy processes (FMP 8.5), based on a common convolution semi-group. If the paths are further required to be continuous, we end up with a *branching Brownian motion*, arguably the most basic of all discrete branching systems, where the branching occurs according to a binary splitting process with rate 1, and the particles perform independent Brownian motions during their life spans.

In all the mentioned cases, the branching property allows us to regard the process ξ_t starting from a non-random measure $\xi_0 = \sum_k \delta_{s_k}$ as a sum of independent processes ξ_t^k with $\xi_0^k = \delta_{s_k}$, referred to below as the *cluster components* of ξ. If ξ_0 is instead a Poisson process, then ξ becomes infinitely divisible, and the associated clusters ξ_t^k agree with the components in the cluster decomposition[3] of Theorem 3.20.

[3]To be precise, this identification requires us to regard the entire process ξ as a point process on a suitable path space, the so-called *historical process*. In this chapter, we will use the same notation and move freely, without further comments, between the two descriptions.

For fixed $t > 0$, the origins of all clusters ξ^k with $\xi_t^k \neq 0$ form a point process ζ_0^t on R^d, consisting of all *ancestors* of ξ_t at time 0. This is clearly a p_t-thinning of ξ_0, where p_t denotes the survival probability of a single cluster. Likewise, by the Markov property, the ancestors of ξ_t at a time $s = t - h$ form a p_h-thinning of the process ξ_s. The family of ancestral processes ζ_s^t with $s \in [0, t]$ is itself a branching system, with the same spatial motion as the original process ξ, but with a different branching mechanism.

In the special case of a branching Brownian motion, the rates and asymptotic distributions can be calculated explicitly. Since the spatial motion is given by a set of independent Brownian motions, it is then enough to consider the underlying binary splitting process. The elementary results in this case motivate the much deeper properties of the DW-process in the next section.

Lemma 13.1 *(binary splitting) Let X be a critical binary splitting process with rate 1, and define $p_t = (1+t)^{-1}$. Then*

(i) $P_1\{X_t > 0\} = p_t, \quad t \geq 0,$

(ii) $P_1(X_t = n \,|\, X_t > 0) = p_t(1-p_t)^{n-1}, \quad t \geq 0,\ n \in \mathsf{N},$

(iii) *for fixed $t > 0$, the ancestors of X_t form a Yule process Z_s^t in $s \leq t$, on the time scale $A_s^t = -\log(1 - p_t s)$, $s \leq t$.*

Proof, (i)–(ii): The generating functions

$$F(s,t) = E_1 s^{X_t}, \quad s \in [0,1],\ t > 0,$$

satisfy the Kolmogorov backward equation

$$\frac{\partial}{\partial t} F(s,t) = \{1 - F(s,t)\}^2$$

with initial condition $F(s,0) = s$, which admits the unique solution

$$\begin{aligned} F(s,t) &= \frac{s + t - s\,t}{1 + t - s\,t} = q_t + \frac{p_t^2 s}{1 - q_t s} \\ &= q_t + \sum_{n \geq 1} p_t^2\, q_t^{n-1} s^n, \end{aligned}$$

where $q_t = 1 - p_t$. In particular, we get $P_1\{X_t = 0\} = F(0,t) = q_t$, which proves (i). To obtain (ii), we note that for any $n \in \mathsf{N}$,

$$\begin{aligned} P_1(X_t = n \,|\, X_t > 0) &= \frac{P_1\{X_t = n\}}{P_1\{X_t > 0\}} \\ &= \frac{p_t^2\, q_t^{n-1}}{p_t} = p_t\, q_t^{n-1}. \end{aligned}$$

(iii) Assuming $X_0 = 1$, and letting Z_s^t be the number of ancestors of X_t at time $s \le t$, we note that

$$\begin{aligned} P\{Z_s^t = 1\} &= EP(Z_s^t = 1 \,|\, X_s) \\ &= E\, X_s\, p_{t-s}\, q_{t-s}^{X_s - 1} \\ &= \sum_{n\ge 1} n\, p_s^2\, q_s^{n-1} p_{t-s}\, q_{t-s}^{n-1} \\ &= p_s^2\, p_{t-s} \sum_{n \ge 1} n (q_s\, q_{t-s})^{n-1} \\ &= \frac{p_s^2\, p_{t-s}}{(1 - q_s\, q_{t-s})^2} = \frac{1+t-s}{(1+t)^2}. \end{aligned}$$

Writing $\tau_1 = \inf\{s \le t;\ Z_s^t > 1\}$, we conclude that

$$\begin{aligned} P(\tau_1 > s \,|\, X_t > 0) &= P(Z_s^t = 1 \,|\, X_t > 0) \\ &= \frac{1+t-s}{1+t}, \end{aligned}$$

which shows that τ_1 has the defective probability density $(1+t)^{-1}$ on $[0,t]$. By Lemma 9.40, the associated compensator density becomes $(1+t-s)^{-1}$, which yields the time scale

$$\begin{aligned} A_s^t &= \int_0^s \frac{dr}{1+t-r} \\ &= -\log\left(1 - \frac{s}{1+t}\right) \\ &= -\log(1 - p_t s). \end{aligned}$$

The Markov and branching properties carry over to the process Z_s^t, by FMP 6.8. Using those properties, and proceeding recursively, we see that Z_s^t is a Yule process on $[0,t]$ with the given rate function A. □

For the general Bienaymé process, we have a similar asymptotic result:

Theorem 13.2 *(survival rate and distribution, Kolmogorov, Yaglom) Let X be a critical Bienaymé process with* $\mathrm{Var}_1(X_1) = 2c \in (0,\infty)$. *Then as $n \to \infty$,*

(i) $n\, P_1\{X_n > 0\} \to c^{-1}$,

(ii) $P_1\big(n^{-1} X_n > r \,\big|\, X_n > 0\big) \to e^{-r/c}, \quad r \ge 0.$

Our proof is based on a simple asymptotic property of the generating functions.

Lemma 13.3 *(generating functions) For X as above, define $f_n(s) = E_1 s^{X_n}$. Then as $n \to \infty$ we have, uniformly in $s \in [0,1)$,*

$$\frac{1}{n}\left\{ \frac{1}{1 - f_n(s)} - \frac{1}{1-s} \right\} \to c.$$

Proof (Spitzer): First we consider a discrete skeleton of the binary splitting process. Here Lemma 13.1 yields

$$P_1\{X_1 = k\} = (1-p)^2\, p^{k-1}, \quad k \in \mathsf{N},$$

for some constant $p \in (0,1)$. By an easy calculation,

$$\frac{1}{1-f_n(s)} - \frac{1}{1-s} = \frac{n\,p}{1-p} = n\,c,$$

which gives the asserted relation with equality. For a given $c > 0$, we need to choose

$$p = \frac{c}{1+c} \in (0,1).$$

In the general case, fix any $c_1 < c < c_2$, and introduce the associated processes $X^{(1)}$ and $X^{(2)}$, as above. By elementary estimates, there exist some constants $n_1, n_2 \in \mathsf{Z}$, such that the associated generating functions satisfy

$$f^{(1)}_{n+n_1} \le f_n \le f^{(2)}_{n+n_2}, \quad n > |n_1| \vee |n_2|.$$

Combining this with the identities for $f^{(1)}_n$ and $f^{(2)}_n$, we obtain

$$\begin{aligned} c_1\,\frac{(n+n_1)}{n} &\le \frac{1}{n}\left\{\frac{1}{1-f_n(s)} - \frac{1}{1-s}\right\} \\ &\le c_2\,\frac{(n+n_2)}{n}. \end{aligned}$$

Now let $n \to \infty$, and then $c_1 \uparrow c$ and $c_2 \downarrow c$. $\square$

Proof of Theorem 13.2: (i) Take $s = 0$ in Lemma 13.3, and note that $f_n(0) \to 1$.

(ii) Writing $s_n = e^{-r/n}$, and using the uniform convergence in Lemma 13.3, we get

$$\frac{1}{n\{1-f_n(s_n)\}} \approx \frac{1}{n(1-s_n)} + c \to r^{-1} + c,$$

where $a \approx b$ means $a - b \to 0$. Since also $n\{1-f_n(0)\} \to c^{-1}$, by part (i), we obtain

$$\begin{aligned} E_1\Big(e^{-rX_n/n}\,\Big|\,X_n > 0\Big) &= \frac{f_n(s_n) - f_n(0)}{1-f_n(0))} \\ &= 1 - \frac{n\{1-f_n(s_n)\}}{n\{1-f_n(0)\}} \\ &\to 1 - \frac{c}{r^{-1}+c} = \frac{1}{1+c\,r}, \end{aligned}$$

which agrees with the Laplace transform of the exponential distribution with mean c^{-1}. Now use the continuity theorem for Laplace transforms, in FMP 5.3. $\square$

To motivate the space-time scaling below, we consider the following classical result for the Bienaymé process:

Theorem 13.4 *(diffusion limit, Feller) Given a critical Bienaymé process X with* $\mathrm{Var}_1(X_1) = 2$, *and a constant $c \geq 0$, consider some versions $X^{(n)}$ with $X_0^{(n)} = \kappa_n$, where the variables κ_n are Poisson with means cn, and introduce the processes*

$$Y_t^{(n)} = n^{-1} X_{[nt]}^{(n)}, \quad t \geq 0.$$

Then $Y^{(n)} \xrightarrow{d} Y$ in $D(\mathsf{R}_+)$, where Y is a continuous martingale with $Y_0 = c$, satisfying the SDE

$$dY_t = Y_t^{1/2} dB_t, \quad t \geq 0.$$

This is a well-known example of an SDE satisfying the condition of pathwise uniqueness (FMP 23.3). The solution Y is known as a *squared Bessel process of order* 0.

Proof (outline): For each n, the process $Y^{(n)}$ is infinitely divisible with κ_n clusters, each of which originates from a single individual at time 0. By Theorem 13.2, the number of clusters surviving up to time $t > 0$ is again Poisson with mean $\sim c t^{-1}$, and each of them is asymptotically exponentially distributed with mean t. This implies $Y_t^{(n)} \xrightarrow{d} Y_t$, where Y_t is infinitely divisible on R_+, with constant component 0, and Lévy measure

$$\nu_t(dr) = c t^{-2} e^{-r/t} dr, \quad r, t > 0.$$

The Markov property of the processes $Y^{(n)}$ carries over to Y, which is also continuous, by the Kolmogorov–Chentsov criterion (FMP 3.23). A simple calculation yields $EY_t = c$ and $\mathrm{Var}(Y_t) = c t$, which suggests that Y is a diffusion process in R_+, with drift 0 and diffusion rate $\sigma_x^2 = x$ at state $x > 0$, so that Y satisfies the stated SDE. The asserted functional convergence $Y^{(n)} \xrightarrow{d} Y$ may now be inferred from FMP 19.28. □

Similar results hold for branching processes in continuous time, including the binary splitting process in Lemma 13.1. For branching systems with spatial motion, the asymptotic theory becomes more sophisticated. Here we consider only the scaling limits of a branching Brownian motion, which leads to the super-processes studied extensively in subsequent sections. Starting with the standard version of the process, considered earlier, we perform a scaling on three different levels:

- increasing the initial intensity by a factor n,
- increasing the branching rate by a factor n,
- reducing the particle weights by a factor n^{-1}.

The spatial motion is not affected by the scaling, and remains a standard Brownian motion for every n. We state without proof the asymptotic scaling theorem for branching Brownian motion. The result extends to large classes of branching systems in discrete or continuous time.

Theorem 13.5 *(scaling limit, Watanabe, Dawson) Fix any $\mu \in \hat{\mathcal{M}}_d$. For every $n \in \mathsf{N}$, consider a branching Brownian motion ξ^n in R^d, with branching rate $2n$ and particle weights n^{-1}, starting from a Poisson process ξ_0^n with intensity $n\mu$. Then $\xi^n \xrightarrow{d} \xi$ in the Skorohod space $D(\mathsf{R}_+, \hat{\mathcal{M}}_d)$, based on the weak topology in $\hat{\mathcal{M}}_d$, where ξ is a weakly continuous strong Markov process in $\hat{\mathcal{M}}_d$ with $\xi_0 = \mu$, satisfying the branching property $P_{\mu_1+\mu_2} = P_{\mu_1} * P_{\mu_2}$.*

The measure-valued diffusion process ξ above is known as a *Dawson–Watanabe super-process*, or simply a *DW-process*. Replacing R^d by an appropriate path space, we obtain a similar result for the associated *historical process*, which encodes the entire genealogy of the original particles. We shall mostly consider state spaces of dimension $d \geq 2$, where the random measures ξ_t with $t > 0$ are known to be a.s. diffuse and singular, with Hausdorff dimension 2. When the initial measure μ is bounded, the measures ξ_t with $t \geq h > 0$ are also known to have uniformly bounded supports.

We quote without proofs some basic characterizations of the DW-process, beginning with a powerful analytic result[4].

Lemma 13.6 *(log-Laplace and evolution equations, Dawson) Let ξ be a DW-process in R^d. Then for any $\mu \in \hat{\mathcal{M}}_d$ and $f \in \hat{C}_+(\mathsf{R}^d)$,*

$$E_\mu e^{-\xi_t f} = e^{-\mu v_t}, \quad t \geq 0,$$

where the function v on $\mathsf{R}_+ \times \mathsf{R}^d$ is the unique solution to the PDE

$$\dot{v} = \tfrac{1}{2}\,\Delta v - v^2, \qquad v_0 = f.$$

We also record some basic martingale properties, where ν_t denotes the symmetric normal distribution in R^d with variances $t > 0$.

Lemma 13.7 *(martingale properties) Let ξ be a DW-process in R^d, with $\xi_0 = \mu \in \hat{\mathcal{M}}_d$. Then*

(i) *for any $f \in \hat{C}_+^2$, the process*

$$M_t = \xi_t f - \mu f - \tfrac{1}{2} \int_0^t \xi_s(\Delta f)\, ds, \quad t \geq 0,$$

is a martingale on R_+ with quadratic variation

$$[M]_t = 2\int_0^t \xi_s f^2\, ds, \quad t \geq 0,$$

[4] Recall that $\mu f = \int f d\mu$. In the super-process literature, this is often written as $\langle \mu, f \rangle$.

(ii) *for fixed $t > 0$, and bounded, measurable functions $f \geq 0$, the process*

$$N_s = (\xi_s * \nu_{t-s})f - (\mu * \nu_t)f, \quad s \leq t,$$

is a martingale on $[0,t]$ with quadratic variation

$$[N]_s = 2\int_0^s \xi_u(\nu_{t-u} * f)^2\, du, \quad s \leq t.$$

In fact, ξ may be characterized as the unique solution to the *martingale problem* defined by condition (i). The statement shows that $\xi_t f$ solves the SDE

$$d\xi_t f = (\xi_t f^2)^{1/2}\, dB_t + \tfrac{1}{2}\, \xi_t(\Delta f)\, dt,$$

which extends the equation for the total mass $\|\xi_t\|$ in Theorem 13.4. Property (ii) can be derived from (i) by Itô's formula. Taking $t = u$ in (ii) yields $E_\mu \xi_t = \mu * \nu_t$, which agrees with the result for branching rate 0. Note, however, that the evolution in the latter case is purely deterministic and governed by the heat equation. In this sense, it is the branching rather than the spatial motion that gives rise to the randomness of the DW-process.

13.2 Finiteness, Extinction, and Genealogy

Here we consider some basic properties of the DW-process, including its cluster structure, along with criteria for local finiteness and extinction, and the continuity in total variation. We begin with some basic genealogy and cluster properties, used extensively in subsequent sections, often without further comments. Though the proofs are omitted, most results are suggested by the corresponding properties for the elementary splitting process in Lemma 13.1.

Theorem 13.8 *(genealogy and cluster structure) Let ξ be a DW-process in R^d. Then*

(i) *for non-random $\xi_0 = \mu$, the process ξ is infinitely divisible, and its cluster components are i.i.d. pseudo-processes, originating at single points,*

(ii) *for fixed $s < t$, the ancestors of ξ_t at time s form a Cox process ζ_s^t on R^d, directed by $h^{-1}\xi_s$ with $h = t - s$,*

(iii) *for fixed $s < t$, the random measure ξ_t is a sum of conditionally independent clusters η^i of age $t - s$, rooted at the points of ζ_s^t,*

(iv) *for fixed $t > 0$, the processes ζ_s^t with $s < t$ form a Brownian Yule process in R^d, on the time scale $A_s^t = -\log(1 - t^{-1}s)$,*

(v) *for fixed $t > 0$, we have $h^{-1}\zeta_{t-h}^t \xrightarrow{v} \xi_t$ a.s. as $h \to 0$.*

The infinite divisibility in (i) is clear from the basic branching property. It yields an associated cluster decomposition, similar to the one in Theorem 3.20. The individual clusters of the entire process ξ are *pseudo-processes*, in the sense that their "distributions" are unbounded on the appropriate path space. Since a cluster has a finite probability of surviving until time $t > 0$, it still makes sense to consider the conditional distribution of a cluster of age t, given that it survives until time t. The pseudo-clusters with different starting points are equally distributed apart from shifts, and we refer to the cluster $\tilde{\eta}$, originating at the origin and normalized by the condition $P\{\tilde{\eta}_1 \neq 0\} = 1$, as the *canonical cluster* of a DW-process.

Assertions (ii)$-$(iv) can be deduced from the results for binary splitting in Lemma 13.1, using the pathwise (historical) version of Theorem 13.5. In particular, Theorem 4.40 shows that the p_t-thinnings in Lemma 13.1 turn in the limit into appropriate Cox processes. We further note that the approximating Yule processes have time scales

$$\begin{aligned} A_n(s,t) &= -\log\left(1 - \frac{ns}{nt+1}\right) \\ &\to \log(1 - t^{-1}s), \quad s < t. \end{aligned}$$

In statement (v), the corresponding convergence in probability is elementary and follows from part (ii), by the law of large numbers and the continuity of ξ. The a.s. convergence, never needed below, may be more subtle.

We continue with an elementary comparison between the distributions of a DW-process ξ and its associated canonical cluster η.

Lemma 13.9 *(cluster comparison) Let the DW-process ξ in R^d with canonical cluster η be locally finite under P_μ, and fix any $B \in \mathcal{B}^d$. Then*

(i) $P_\mu\{\eta_t B > 0\} = -\log\big(1 - P_\mu\{\xi_t B > 0\}\big)$,

(ii) $P_\mu\{\xi_t B > 0\} = 1 - \exp\big(-P_\mu\{\eta_t B > 0\}\big)$,

(iii) $\big\|\mathcal{L}_\mu(\xi_t) - \mathcal{L}_\mu(\eta_t)\big\|_B \le \big(P_\mu\{\eta_t B > 0\}\big)^2$.

In particular, $P_\mu\{\xi_t B > 0\} \sim P_\mu\{\eta_t B > 0\}$, as either side tends to 0.

Proof, (i)$-$(ii): Under P_μ we have $\xi_t = \sum_i \eta_t^i$, where the η^i are conditionally independent clusters, rooted at the points of a Poisson process on R^d with intensity μ/t. For a cluster rooted at x, the hitting probability equals $b_x = tP_x\{\eta_t B > 0\}$, and so, by Theorem 3.2 (i), the number of clusters hitting B at time t is Poisson distributed with mean μb. Hence,

$$\begin{aligned} P_\mu\{\xi_t B = 0\} &= \exp(-\mu b) \\ &= \exp\big(-P_\mu\{\eta_t B > 0\}\big), \end{aligned}$$

proving (ii). Equations (i) and (ii) are clearly equivalent, and the last assertion is an immediate consequence.

(iii) Write the cluster representation as

$$1_B \xi_t = \int 1_B m \, \zeta_t(dm) = \int 1_B m \, \zeta_t^B(dm),$$

where ζ_t is a Poisson process on $\mathcal{M}_d$ with intensity $\mathcal{L}_\mu(\eta_t)$, and ζ_t^B denotes the restriction of ζ_t to the set of measures m with $mB > 0$. Using Lemma 3.10, we get for any measurable function $f \geq 0$ on $\mathcal{M}_d$ with $f(0) = 0$

$$\left| E_\mu f(1_B \xi_t) - E_\mu f(1_B \eta_t) \right| \leq \|f\| \left(P_\mu\{\eta_t B > 0\} \right)^2,$$

and the assertion follows since f was arbitrary. □

In standard constructions of the DW-process, the initial measure μ is usually taken to be bounded. Using the basic branching property, we can easily extend the definition to any σ-finite measure μ. Though the resulting process ξ is well-defined for any such μ, additional conditions are needed to ensure that ξ_t will be a.s. locally finite, for every $t > 0$.

Theorem 13.10 *(local finiteness) Let ξ be a DW-process in R^d with canonical cluster η, and fix a σ-finite measure μ. Then these four conditions are equivalent:*

(i) $\xi_t \in \mathcal{M}_d$ *a.s.* P_μ, $t > 0$,
(ii) $E_\mu \xi_t = E_\mu \eta_t = \mu * \nu_t \in \mathcal{M}_d$, $t > 0$,
(iii) $\mu p_t < \infty$, $t > 0$,
(iv) $\mu * p_t$ *is finite and continuous for all* $t > 0$.

The equivalence (i) ⇔ (ii) *holds even for fixed $t > 0$, and* (i)−(iv) *imply*

(v) $E_\mu \xi_t \theta_h$ *is h-continuous in total variation on B, for fixed $t > 0$ and bounded μ or B.*

Proof, (i) ⇔ (ii): The identities in (ii) follow from Theorems 13.16 (i) and 13.18 below, with $n = 1$. The implication (ii) ⇒ (i) is obvious. To prove the converse, we see from Theorem 13.16 (i)−(ii) below with $n = 2$ that, for any $\mu \in \mathcal{M}_d$, $t > 0$, and $B \in \hat{\mathcal{B}}^d$,

$$\begin{aligned} \mathrm{Var}_\mu(\xi_t B) &= (\mu * \nu_t^2) B^2 \\ &= 2 \int_0^t \left(\mu * \nu_s * \nu_{t-s}^{\otimes 2} \right) B^2 \, ds \\ &\leq 2 \int_0^t (\mu * \nu_s * \nu_{t-s}) B \, ds \\ &= 2t \, (\mu * \nu_t) B = 2t \, E_\mu \xi_t B, \end{aligned}$$

and so by FMP 4.1, for any $r \in (0, 1)$,

$$\begin{aligned} P_\mu\left\{ \xi_t B > r \, E_\mu \xi_t B \right\} &\geq (1 - r)^2 \, \frac{(E_\mu \xi_t B)^2}{E_\mu (\xi_t B)^2} \\ &\geq \frac{(1 - r)^2}{1 + 2t \, (E_\mu \xi_t B)^{-1}} \, . \end{aligned} \qquad (2)$$

If (ii) fails, we may choose a $B \in \hat{\mathcal{B}}^d$ with $E_\mu \xi_t B = \infty$, along with some measures $\mu_n \uparrow \mu$ in $\mathcal{M}_d$, such that $E_{\mu_n} \xi_t B \geq n$. Then by (2),

$$\begin{aligned} P_\mu\{\xi_t B > rn\} &\geq P_{\mu_n}\Big\{\xi_t B > r\, E_{\mu_n}\xi_t B\Big\} \\ &\geq \frac{(1-r)^2}{1+2t/n}. \end{aligned}$$

Letting $n \to \infty$ and then $r \to 0$, we obtain $\xi_t B = \infty$ a.s., which contradicts (i). This shows that (i) $\Leftrightarrow$ (ii), even for a fixed $t > 0$.

(ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv): Assume (iii). Fixing $x \in \mathsf{R}^d$, and choosing $r \geq t + 2|x|$, we see from Lemma A3.3 that $p_t(x-u) \lesssim p_r(u)$. Hence, $(\mu * p_t)(x) \lesssim \mu p_r < \infty$, which shows that the density $\mu * p_t$ of $E_\mu \xi_t$ is finite. We further note that

$$\Big|(\mu * p_t)(x+h) - (\mu * p_t)(x)\Big| \leq \int \Big|p_t(x+h-y) - p_t(x-y)\Big|\, \mu(dy),$$

where the integrand tends to 0 as $h \to 0$. Since also

$$\Big|p_t(x+h-y) - p_t(x-y)\Big| \lesssim p_{2t}(x-y), \quad |h| \leq t, \tag{3}$$

by Lemma A3.3, and $(\mu * p_{2t})(x) < \infty$ by (iii), the asserted continuity of $\mu * p_t$ follows by dominated convergence. This proves (iv), which in turn implies (ii) for every $t > 0$. Conversely, (ii) yields $(\mu * p_n)(x) < \infty$ for all $n \in \mathsf{N}$, and for $x \in \mathsf{R}^d$ a.e. λ^d. Fixing such an x, and using Lemma A3.3 as before, we obtain (iii).

(iii) $\Rightarrow$ (v): For any $h \in \mathsf{R}^d$ and $t > 0$, we have

$$\begin{aligned} \Big\|E_\mu \xi_t \theta_h - E_\mu \xi_t\Big\|_B &= \int_B \Big|(\mu * p_t)(x-h) - (\mu * p_t)(x)\Big|\, dx \\ &\leq \int \mu(dy) \int_B \Big|p_t(x-h-y) - p_t(x-y)\Big|\, dx, \end{aligned}$$

where the integrand tends to 0 as $h \to 0$. For $B = \mathsf{R}^d$ and bounded μ, we may use (3) again and note that $\int \mu(dy) \int p_{2t}(x-y)\, dx = \|\mu\| < \infty$, which justifies that we take limits under the integral sign. For μ as in (iii) and bounded B, choose $r > 0$ so large that $t + 2|x-h| \leq r$, for $x \in B$ and $|h| \leq 1$, and conclude from Lemma A3.3 that $p_t(x-h-y) \lesssim p_r(y)$ for such x and h. Since $\mu p_r < \infty$ by (iii), the dominated convergence is again justified, and (v) follows. $\square$

A DW-process ξ with bounded initial measure μ will always die out, and if μ has bounded support, then the range of ξ is bounded:

Theorem 13.11 *(global extinction and range) For a DW-process ξ in R^d, we have*

(i) $P_\mu\{\xi_t = 0\} = e^{-\|\mu\|/t}, \quad \mu \in \mathcal{M}_d,\ t > 0,$

(ii) $P_{a\delta_0}\Big\{\sup_{t>0} \xi_t(B_0^r)^c > 0\Big\} \lesssim ar^{-2}, \quad a, r > 0.$

Proof: (i) The ancestors of ξ_t form a Poisson process ζ_0 with intensity μ/t, and so the number of surviving clusters at time t is Poisson distributed with mean $\|\mu\|/t$.

(ii) Writing $\mu = a\delta_0$, and letting $t \le r^2/4$, we get by Theorem 13.38 (iv) below

$$\begin{aligned} P_\mu\Big\{\sup_{s>0} \xi_s(B_0^r)^c > 0\Big\} &\le P_\mu\Big\{\sup_{s>0} \xi_s(B_0^r)^c > 0,\ \xi_t = 0\Big\} + P_\mu\{\xi_t \ne 0\} \\ &\le P_\mu\Big\{\sup_{s\le t} \xi_s(B_0^r)^c > 0\Big\} + (1 - e^{-a/t}) \\ &\lesssim ar^d\, t^{-1-d/2}\, e^{-r^2/2t} + at^{-1}. \end{aligned}$$

The desired bound now follows as we take $t = r^2/4$. □

Though a DW-process with unbounded initial measure μ will never die out, extinction may still occur locally. To prepare for the precise result, we begin with a convergence criterion for fixed $t > 0$.

Lemma 13.12 *(null convergence) For any $\mu_1, \mu_2, \ldots \in \mathcal{M}_d$, these conditions are equivalent:*

(i) $\xi_t \xrightarrow{v} 0$ *in probability* $P_{\mu_n},\ t > 0,$

(ii) $E_{\mu_n}\xi_t \xrightarrow{v} 0,\ t > 0,$

(iii) $\operatorname{supp}\xi_t \xrightarrow{v} \emptyset$ *in probability* $P_{\mu_n},\ t > 0,$

(iv) $\mu_n p_t \to 0,\ t > 0.$

The equivalence (i) ⇔ (ii) *holds even for fixed* $t > 0$.

Proof: The implication (ii) ⇒ (i) is obvious. Conversely, if (ii) fails for some $t > 0$, we may choose some $B \in \hat{\mathcal{B}}^d$ and $\varepsilon > 0$ with $E_{\mu_n}\xi_t > \varepsilon$, and conclude from (2) that

$$\begin{aligned} P_{\mu_n}\{\xi_t B > r\varepsilon\} &\ge P_{\mu_n}\Big\{\xi_t B > r E_{\mu_n}\xi_t B\Big\} \\ &\ge \frac{(1-r)^2}{1 + 2t\,(E_{\mu_n}\xi_t B)^{-1}} \\ &\ge \frac{(1-r)^2}{1 + 2t/\varepsilon} > 0, \end{aligned}$$

contradicting (i). Next, Lemma A3.3 shows that (iv) is equivalent to $\mu_n * p_t \to 0$, uniformly on compacts, which is equivalent to (iii). Finally, (ii) ⇔ (iii) by Theorem 13.38 below. □

We may now give some precise criteria for local extinction. Note in particular that extinction in the vague sense is equivalent to a similar property for the supports. In both cases, extinction is understood in the sense of convergence in probability.

Theorem 13.13 *(local extinction) Let ξ be a locally finite DW-process in R^d with $d \geq 2$, and fix any sequences $\mu_n \in \mathcal{M}_d$ and $t_n \to \infty$. Then the following conditions are equivalent, where in each statement we consider convergence as $n \to \infty$, for any choice of sequence $t'_n \sim t_n$:*

(i) $\xi_{t'_n} \xrightarrow{v} 0$ *in* P_{μ_n},

(ii) $\operatorname{supp} \xi_{t'_n} \xrightarrow{v} \emptyset$ *in* P_{μ_n},

(iii) $\begin{cases} E_{\mu_n} \xi_{t'_n} \xrightarrow{v} 0, & d \geq 3, \\ (\log t_n)^{-1} E_{\mu_n} \xi_{t'_n} \xrightarrow{v} 0, & d = 2, \end{cases}$

(iv) $\begin{cases} \mu_n p_{t'_n} \to 0, & d \geq 3, \\ (\log t_n)^{-1} \mu_n p_{t'_n} \to 0, & d = 2. \end{cases}$

Proof: Here (ii) ⇔ (iii) holds by Theorem 13.38 below, whereas (iii) ⇔ (iv) by Lemma A3.3. The implication (ii) ⇒ (i) being obvious, it remains to prove that (i) ⇒ (ii). We may then introduce, on a common probability space, some independent processes $\xi^{(n)}$ with distributions P_{μ_n}. Assuming (i), we have $\xi^{(n)}_{t_n} \xrightarrow{v} 0$ a.s. along a sub-sequence, which remains true under conditioning on the sequence $\xi^{(n)}_{t_n - h}$, for fixed $h > 0$. Applying Lemma 13.12 with $t = h$, and proceeding to a further sub-sequence, we get under the same conditioning $\operatorname{supp} \xi^{(n)}_{t_n - rh} \xrightarrow{v} \emptyset$ a.s., for a suitable constant $r \in (0,1)$, which remains unconditionally true. By the sub-sequence criterion for convergence in probability, we obtain (ii) with $t'_n = t_n - rh$, which is clearly sufficient. □

We proceed to some basic scaling identities for a DW-process ξ and its canonical cluster η, involving a scaling in time, space, and mass. For special needs, we also consider a scaling relation for the *span* ρ of the canonical cluster η, defined by

$$\rho = \inf\Big\{ r > 0;\, \sup_{t>0} \eta_t (B_0^r)^c = 0 \Big\},$$

which is known to be a.s. finite. Here and below, we will use the scaling operators S_r on R^d, defined by $S_r x = rx$ for $x \in \mathsf{R}^d$ and $r > 0$, so that $(\mu S_r)B = \mu(S_r B)$ and $(\mu S_r) f = \mu(f \circ S_r^{-1})$. For any process X on R_+, we define the time-scaled process $\hat{X}_h$ on R_+ by $(\hat{X}_h)_t = X_{ht}$.

Theorem 13.14 *(scaling) Let ξ be a DW-process in R^d with canonical cluster η, and let ρ denote the span of η from 0. Then for any μ, $x \in \mathsf{R}^d$, and $r, c > 0$, we have*[5]

[5]Recall that $\mathcal{L}(\xi \,\|\, \eta)_x$ denotes the Palm distribution of ξ with respect to η at the point x. A possible subscript of $\mathcal{L}$ denotes the initial position or measure.

(i) $\mathcal{L}_{\mu S_r}(r^2\xi_1) = \mathcal{L}_{r^2\mu}(\xi_{r^2}S_r)$,

(ii) $\mathcal{L}_{\mu S_r}(r^2\eta) = r^2\mathcal{L}_\mu(\hat\eta_{r^2}S_r)$,

(iii) $\mathcal{L}_0(r^2\eta \,\|\, \eta_1)_x = \mathcal{L}_0\big(\hat\eta_{r^2}S_r \,\big\|\, \eta_{r^2}\big)_x$,

(iv) $P_0(\rho > c \,\|\, \eta_{r^2})_x \le P_0\big(r\rho + |x| > c \,\big\|\, \eta_1\big)_0$.

Proof: (i) If v solves the evolution equation for ξ, then so does the function

$$\tilde v(t,x) = r^2 v(r^2 t, rx), \quad t \in \mathsf{R}_+,\ x \in \mathsf{R}^d.$$

Writing

$$\tilde\xi_t = r^{-2}\,\xi_{r^2t}\,S_r, \quad \tilde\mu = r^{-2}\,\mu S_r, \quad \tilde f(x) = r^2\,f(rx),$$

we get

$$\begin{aligned} E_\mu e^{-\tilde\xi_t\tilde f} &= E_\mu e^{-\xi_{r^2t}f} \\ &= e^{-\mu v_{r^2t}} = e^{-\tilde\mu\tilde v_t} \\ &= E_{\tilde\mu} e^{-\xi_t\tilde f}, \end{aligned}$$

and so $\mathcal{L}_\mu(\tilde\xi) = \mathcal{L}_{\tilde\mu}(\xi)$, which is equivalent to (i).

(ii) Introduce the kernel $\nu_x = \mathcal{L}_x(\eta)$, $x \in \mathsf{R}^d$, and consider the cluster decomposition $\xi = \int m\,\zeta(dm)$, where ζ is a Poisson process on $\mathcal{M}_d$ with intensity $\mu\nu$, when $\xi_0 = \mu$. Here

$$r^{-2}\,\xi_{r^2t}S_r = \int \Big(r^{-2}\,m_{r^2t}S_r\Big)\,\zeta(dm), \quad r,t > 0.$$

Using (i) and the uniqueness of Lévy measure, we obtain

$$(r^{-2}\,\mu S_r)\nu = \mu\Big(\nu\{r^{-2}\,\hat m_{r^2}S_r \in \cdot\}\Big),$$

which is equivalent to

$$\begin{aligned} r^{-2}\mathcal{L}_{\mu S_r}(\eta) &= \mathcal{L}_{r^{-2}\mu S_r}(\eta) \\ &= \mathcal{L}_\mu\Big(r^{-2}\hat\eta_{r^2}S_r\Big). \end{aligned}$$

(iii) By Palm disintegration, we get from (ii)

$$\begin{aligned} \int E_0\eta_1(dx)\,E_0\Big\{f(x,r^2\eta)\,\Big\|\,\eta_1\Big\}_x &= E_0\int \eta_1(dx)\,f(x,r^2\eta) \\ &= r^2 E_0\int \eta_{r^2}(dx)\,f(x,\hat\eta_{r^2}S_r) \\ &= r^2\int E_0\eta_{r^2}(dx)\,E_0\Big\{f(x,\hat\eta_{r^2}S_r)\,\Big\|\,\eta_{r^2}\Big\}_x, \end{aligned}$$

and (iii) follows by the continuity of the Palm kernel.

(iv) By (iii), we have $\mathcal{L}_0(\rho \,\|\, \eta_{r^2})_x = \mathcal{L}_0(r\rho \,\|\, \eta_1)_x$, which implies (iv) for $x = 0$. For general x, it is enough to take $r = 1$. Then conclude from Corollary 13.22 below that, under $\mathcal{L}_0(\eta \,\|\, \eta_1)_x$, the cluster η is a countable sum of conditionally independent sub-clusters, rooted along the path of a Brownian bridge on $[0,1]$ from 0 to x. In view of Lemma 6.9, the evolution of the process after time 1 is then the same as under the original distribution $\mathcal{L}(\eta)$, and is therefore independent of x. We may now construct a cluster with distribution $\mathcal{L}_0(\eta \,\|\, \eta_1)_x$ from one with distribution $\mathcal{L}_0(\eta \,\|\, \eta_1)_0$, simply by shifting each sub-cluster born at time $s \le 1$ by an amount $(1-s)x$. Since all mass of η is then shifted by at most $|x|$, the span from 0 of the whole cluster is increased by at most $|x|$, and the assertion follows. □

By its very construction, a DW-process starting from a bounded measure μ is a.s. weakly continuous in the space of bounded measures on R^d. This property may be strengthened to local continuity in total variation, even for unbounded measures μ. The result will be used, in Section 13.5, to prove a similar continuity property for Palm distributions.

Lemma 13.15 *(strong time continuity) Consider a DW-process ξ in R^d, and some fixed $\mu \in \mathcal{M}_d$ and $B \in \mathcal{B}^d$, where either μ or B is bounded. Then $\mathcal{L}_\mu(1_B \xi_t)$ is continuous in total variation in $t > 0$.*

Proof: First let $\|\mu\| < \infty$. For any $t > 0$, the ancestors of ξ_t at time 0 form a Poisson process ζ_0 with intensity $t^{-1}\mu$. By Theorem 13.8, the ancestors splitting before time $s \in (0,t)$ form a Poisson process with intensity $st^{-2}\mu$, and so the probability of such a split equals $1 - e^{-st^{-2}\|\mu\|} \le st^{-2}\|\mu\|$. Hence, the process ζ_s of ancestors at time s agrees, up to a set of probability $st^{-2}\|\mu\|$, with a Poisson process with intensity $t^{-1}\mu * p_s$.

Replacing s and t by $s+h$ and $t+h$, where $|h| < s$, we see that the process ζ'_{s+h} of ancestors of ξ_{t+h} at time $s+h$ agrees, up to a probability $(s+h)(t+h)^{-2}\|\mu\|$, with a Poisson process with intensity $(t+h)^{-1}\mu * p_{s+h}$. Since ξ_t and ξ_{t+h} are both Cox cluster processes with the same cluster kernel, given by the normalized distribution of a $(t-s)$-cluster, the total variation distance between their distributions is bounded by the corresponding distance for the two ancestral processes. Since $\|\mathcal{L}(\eta_1) - \mathcal{L}(\eta_2)\| \le \|E\eta_1 - E\eta_2\|$ for any Poisson processes η_1 and η_2 on the same space, the total variation bound is of the order $\|\mu\|$ times

$$\frac{s}{t^2} + \frac{s+|h|}{(t+h)^2} + \left|\frac{1}{t} - \frac{1}{t+h}\right| + \frac{\|p_s - p_{s+h}\|_1}{t} \lesssim \frac{s+|h|}{t^2} + \frac{|h|}{st}.$$

Choosing $s = |h|^{1/2}$, we get convergence to 0 as $h \to 0$, uniformly for $t \in (0,\infty)$ in compacts, which proves the continuity in t.

Now let μ be arbitrary, and assume instead that B is bounded. Let μ_r and μ'_r denote the restrictions of μ to B_0^r and $(B_0^r)^c$. Then $P_{\mu'_r}\{\xi_t B > 0\} \lesssim (\mu'_r * \nu_t)B < \infty$, uniformly for $t \in (0,\infty)$ in compacts, by Theorem 13.42 below. As $r \to \infty$, we get $P_{\mu'_r}\{\xi_t B > 0\} \to 0$ by dominated convergence, in the same uniform sense. Finally, by the version for bounded μ, $\mathcal{L}_{\mu_r}(1_B \xi_t)$ is continuous in total variation, in $t > 0$ for fixed $r > 0$. □

13.3 Moment Measures and Palm Trees

The higher order moment measures of a DW-process in R^d play a crucial role, along with the associated densities, for the developments in subsequent sections. Here we establish a useful representation, in terms of certain uniform Brownian trees, reflecting a basic tree structure of the Campbell and Palm measures. We begin with a basic cluster decomposition, along with two recursive constructions and a Markov property, first stated in a concise and suggestive notation, and then spelled out explicitly in some detailed but less transparent formulas.

We use ν_t to denote the standard normal distribution in R^d with variances t, so that $\nu_t = p_t \cdot \lambda^d$ for the associated normal density p_t. Define $\nu_t^n = E_0 \eta_t^{\otimes n}$ and $\nu_t^J = E_0 \eta_t^{\otimes J}$ for the canonical cluster η. Let $\mathcal{P}_n$ denote the set of all partitions π of $\{1, \dots, n\}$ into subsets J, and similarly for $\mathcal{P}_J$. Given a finite set J, we write $\Sigma'_{I \subset J}$ for summation over all non-empty, proper subsets $I \subset J$. For any distinct elements $i, j \in J$, we form a new set $J_{ij} = J_{ji}$, by combining i and j into a single element $\{i, j\}$.

Theorem 13.16 *(decomposition and recursion) Let ξ be a DW-process in R^d with canonical cluster η, and write $\nu_t^J = E_0 \eta_t^{\otimes J}$. Then for any $t > 0$ and measure μ,*

(i) $$E_\mu \xi_t^{\otimes n} = \sum_{\pi \in \mathcal{P}_n} \bigotimes_{J \in \pi} (\mu * \nu_t^J), \quad n \in \mathsf{N},$$

(ii) $$\nu_t^J = \sum_{I \subset J}{}' \int_0^t \nu_s * \left(\nu_{t-s}^I \otimes \nu_{t-s}^{J \setminus I} \right) ds, \quad |J| \geq 2,$$

(iii) $$\nu_t^J = \sum_{i \neq j} \int_0^t \left(\nu_s^{J_{ij}} * \nu_{t-s}^{\otimes J} \right) ds, \quad |J| \geq 2,$$

(iv) $$\nu_{s+t}^n = \sum_{\pi \in \mathcal{P}_n} (\nu_s^\pi * \bigotimes_{J \in \pi} \nu_t^J), \quad s, t > 0, \ n \in \mathsf{N}.$$

Here (i) is the basic cluster decomposition, whereas (ii) and (iii) give forward and backward recursions for the individual clusters, and (iv) is a Markov property for the individual clusters. The symbol $*$ denotes convolution in the space variables, whereas temporal convolution is written out explicitly. To state the four relations more explicitly, let $f_1, \dots, f_n$ be any non-negative, measurable functions on R^d, and write $x_J = (x_j;\, j \in J) \in (\mathsf{R}^d)^J$. For $u_{J_{ij}} \in (\mathsf{R}^d)^{J_{ij}}$, take $u_k = u_{ij}$ when $k \in \{i, j\}$.

(i′) $$E_\mu \prod_{i \leq n} \xi_t f_i = \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} \int \mu(du) \int \nu_t^J(dx_J) \prod_{i \in J} f_i(u + x_i),$$

(ii′) $$\nu_t^J \bigotimes_{i \in J} f_i = \sum_{I \subset J}{}' \int_0^t ds \int \nu_s(du) \int \nu_{t-s}^I(dx_I) \int \nu_{t-s}^{J \setminus I}(dx_{J \setminus I}) \prod_{i \in J} f_i(u + x_i),$$

(iii′) $$\nu_t^J \bigotimes_{i \in J} f_i = \sum_{i \neq j} \int_0^t ds \int \nu_s^{J_{ij}}(du_{J_{ij}}) \prod_{k \in J} \int \nu_{t-s}(dx_k) f_k(u_k + x_k),$$

(iv′) $\nu_{s+t}^n \bigotimes_{i\le n} f_i = \sum_{\pi\in\mathcal{P}_n} \int \nu_s^\pi(du_\pi) \prod_{J\in\pi} \int \nu_t^J(dx_J) \prod_{i\in J} f_i(u_J + x_i).$

For the moment, we prove only properties (i) and (iii)−(iv), whereas the proof of (ii) is postponed until after Theorem 13.18 below.

Partial proof: (i) This is clear from Theorem 6.30 (i).

(iii) For any $t > 0$ and bounded, measurable functions $f_1, \dots, f_n \ge 0$ on R^d, we introduce the processes

$$M_s^i = (\xi_s * \nu_{t-s}) f_i, \quad s \le t,\ i \le n,$$

which are known from Lemma 13.7 (ii) to be martingales on $[0,t]$ with quadratic variation processes

$$[M^i]_s = 2\int_0^s \xi_u (\nu_{t-u} * f_i)^2 \, du, \quad s \le t,\ i \le n.$$

By polarization, we get for $i, j \le n$ the associated covariation processes

$$\begin{aligned} [M^i, M^j]_s &= 2\int_0^s \xi_u (\nu_{t-u} * f_i)(\nu_{t-u} * f_j) \, du \\ &= 2\int_0^s \left(\xi_u * \nu_{t-u}^{\otimes 2}\right)(f_i \otimes f_j) \, du. \end{aligned}$$

Applying Itô's formula to the product $X_s = M_s^1 \cdots M_s^n$, and noting that

$$X_0 = \prod_{k\le n} (\mu * \nu_t) f_k, \qquad X_t = \prod_{k\le n} \xi_t f_k,$$

we get

$$\begin{aligned} E_\mu \prod_{k\le n} \xi_t f_k &= \prod_{k\le n} (\mu * \nu_t) f_k \\ &\quad + \sum_{i\ne j} \int_0^t E_\mu \left(\xi_s * \nu_{t-s}^{\otimes 2}\right)(f_i \otimes f_j) \prod_{k\ne i,j} (\xi_s * \nu_{t-s}) f_k \, ds, \end{aligned}$$

where the first order terms vanish, by the boundedness of all moments. Extending from tensor products $f_1 \otimes \cdots \otimes f_n$ to general measurable functions $f \ge 0$ on $(\mathsf{R}^d)^n$, we may write the last relation, more concisely, as

$$E_\mu \xi_t^{\otimes n} = (\mu * \nu_t)^{\otimes n} + \sum_{i\ne j} \int_0^t \left(E_\mu \xi_s^{J_{ij}} * \nu_{t-s}^{\otimes n}\right) ds.$$

Using (i), we get

$$\sum_{\pi\in\mathcal{P}_n} \bigotimes_{J\in\pi} (\mu * \nu_t^J) = (\mu * \nu_t)^{\otimes n} + \sum_{i\ne j} \sum_{\pi\in\mathcal{P}_{ij}} \int_0^t ds \bigotimes_{J\in\pi} (\mu * \nu_s^J) * \nu_{t-s}^{\otimes n},$$

where $\mathcal{P}_{ij}$ denotes the class of partitions of J_{ij}. Taking $\mu = \varepsilon\,\delta_0$ with $\varepsilon > 0$, dividing by ε, and letting $\varepsilon \to 0$, we can eliminate all higher order terms, and (iii) follows.

(iv) Using (i) repeatedly, along with the Markov property of ξ at s, we obtain

$$\begin{aligned}\sum_{\pi\in\mathcal{P}_n}\bigotimes_{J\in\pi}(\mu*\nu_{s+t}^J) &= E_\mu\xi_{s+t}^{\otimes n} = E_\mu E_\mu\Big(\xi_{s+t}^{\otimes n}\,\Big|\,\xi_s\Big)\\ &= E_\mu E_{\xi_s}\xi_t^{\otimes n} = E_\mu\sum_{\pi\in\mathcal{P}_n}\bigotimes_{J\in\pi}(\xi_s*\nu_t^J)\\ &= \sum_{\pi\in\mathcal{P}_n}E_\mu\xi_s^{\otimes\pi}*\bigotimes_{J\in\pi}\nu_t^J\\ &= \sum_{\pi\in\mathcal{P}_n}\sum_{\kappa\in\mathcal{P}_\pi}\bigotimes_{I\in\kappa}(\mu*\nu_s^I)*\bigotimes_{J\in\pi}\nu_t^J.\end{aligned}$$

It remains to take $\mu=\varepsilon\,\delta_0$, as before, and equate the first order terms. □

To see how the previous moment formulas lead to interpretations in terms of Brownian trees, we first consider the underlying combinatorial tree structure. Say that a tree or branch is *defined on an interval* $[s,t]$, if it is rooted at time s, and all leaves extend to time t. It is said to be *simple* if it has only one leaf, and *binary* if exactly two branches are attached to every vertex. *Siblings* are defined as leaves attached to a common vertex. A tree is called *geometric*, if it can be displayed as a planar graph with no self-intersections. A set of trees is said to be *marked*, if distinct marks are attached to the leaves, and we say that the marks are *random*, if they are conditionally exchangeable, given the underlying tree structure.

Lemma 13.17 *(binary trees) There are $n!\,(n-1)!\,2^{1-n}$ marked, binary trees on $[0,n]$, with n leaves and splitting times $1,\dots,n-1$. The following constructions are equivalent, each giving the same probability to all such trees:*

(i) *Forward recursion: Starting with a simple tree on $[0,1]$, proceed recursively in $n-1$ steps, so that after $k-1$ steps we get a binary tree on $[0,k]$, with k leaves and splitting times $1,\dots,k-1$. Then split a randomly chosen leaf in two, and extend all leaves to time $k+1$. After the last step, assign random marks to all leaves.*

(ii) *Backward recursion: Starting with n simple, marked trees on $[n-1,n]$, proceed recursively in $n-1$ steps, so that after $k-1$ steps we get $n-k+1$ binary trees on $[n-k,n]$, with totally n leaves and splitting times $n-k+1,\dots,n-1$. Then join two randomly chosen roots, and extend all roots to time $n-k-1$. Continue until all trees are connected.*

(iii) *Sideways recursion: Let $\tau_1,\dots,\tau_{n-1}$ form an exchangeable permutation of $1,\dots,n-1$. Starting with a simple tree on $[0,n]$, proceed recursively in $n-1$ steps, so that after $k-1$ steps we get a binary, geometric tree on $[0,n]$, with k leaves and splitting times $\tau_1,\dots,\tau_{k-1}$. Then attach to the right a simple tree on $[\tau_k,n]$. After the last step, assign random marks to all leaves.*

Proof: (ii) By the obvious 1−1 correspondence between marked trees and selections of pairs, this construction gives the same probability to all possible trees. The total number of choices is clearly

$$\binom{n}{2}\binom{n-1}{2}\cdots\binom{2}{2} = \frac{n(n-1)}{2}\,\frac{(n-1)(n-2)}{2}\cdots\frac{2\cdot 1}{2} = \frac{n!\,(n-1)!}{2^{n-1}}.$$

(i) Before the final marking of leaves, the resulting tree can be realized as a geometric one, which yields a 1−1 correspondence between the $(n-1)!$ possible constructions, and the set of all geometric trees. Now any binary tree with n distinct splitting times, having m pairs of siblings, can be realized as a geometric tree in 2^{n-m-1} different ways. Furthermore, any geometric tree with n leaves and m pairs of siblings can be marked in $n!\,2^{-m}$ non-equivalent ways. Hence, every such tree has probability

$$\frac{2^{n-m-1}}{(n-1)!}\cdot\frac{2^m}{n!} = \frac{2^{n-1}}{n!\,(n-1)!},$$

which is independent of m, and hence is the same for all trees. (Note that this agrees with the probability in (ii).)

(iii) Before the final marking, this construction yields a binary, geometric tree with distinct splitting times $1,\dots,n$. Conversely, any such geometric tree can be realized in this way, for a suitable permutation $\tau_1,\dots,\tau_{n-1}$ of $1,\dots,n-1$. The correspondence is 1−1, since both sets of trees have the same cardinality $(n-1)!$. The proof can now be completed as in case of (i). □

We can now give the announced probabilistic description of moment measures, in terms of suitable *uniform Brownian trees*. To construct such a tree on $[0,t]$ of order n, we start from a uniform binary tree as in Lemma 13.17, then choose some random splitting times $\tau_1 < \cdots < \tau_{n-1}$, given by an independent, stationary binomial process on $[0,t]$, place the root at $0 \in \mathsf{R}^d$, and add a spatial motion, given by some independent Brownian motions in R^d along the branches.

Theorem 13.18 *(moment measures and Brownian trees) For any $t > 0$ and $n \in \mathsf{N}$, let ν_t^n be the n-th order moment measure of the canonical cluster η at time t, and let μ_t^n be the joint endpoint distribution of a uniform, n-th order Brownian tree in R^d, on the interval $[0,t]$. Then*

$$\nu_t^n = n!\,t^{n-1}\,\mu_t^n, \quad t > 0,\ n \in \mathsf{N}. \tag{4}$$

Proof: Defining $\tilde\nu_t^n$ as in (4) in terms of μ_t^n, we note that $\nu_t^1 = \tilde\nu_t^1$ by Lemma 13.7 (ii). It remains to show that the measures $\tilde\nu_t^n$ satisfy the same backward recursion as the ν_t^n in Theorem 13.16 (iii). Then let $\tau_1 < \cdots < \tau_{n-1}$ be the splitting times of the Brownian tree in Theorem 13.18. By Lemma 3.15, the time τ_{n-1} has density $(n-1)\,s^{n-2}\,t^{1-n}$ in s for fixed t, and conditionally on τ_{n-1}, the remaining times $\tau_1, \ldots, \tau_{n-2}$ form a stationary binomial process on $[0, \tau_{n-1}]$. By Lemma 13.17, the conditional structure up to time τ_{n-1} is then a uniform Brownian tree of order $n-1$, independent of the last branching and the motion up to time t. Defining μ_t^J as before, and conditioning on τ_{n-1}, we get

$$\mu_t^J = \binom{n}{2}^{-1} \sum_{\{i,j\} \subset J} \int_0^t (n-1) \frac{s^{n-2}}{t^{n-1}} \left(\mu_s^{J_{ij}} * \mu_{t-s}^{\otimes J} \right) ds.$$

To obtain the desired recursive relation, it remains to substitute

$$\mu_t^J = \frac{\tilde{\nu}_t^J}{n!\, t^{n-1}}, \qquad \mu_s^{J_{ij}} = \frac{\tilde{\nu}_s^{J_{ij}}}{(n-1)!\, s^{n-2}}, \qquad \mu_{t-s} = \tilde{\nu}_{t-s}.$$

Note that a factor 2 cancels out in the last computation, since every pair $\{i,j\} = \{j,i\}$ appears twice in the summation $\sum_{i,j \in J}$. □

Using the last result, we may now complete the proof of Theorem 13.16.

Proof of Theorem 13.16 (ii): The assertion is obvious for $n = 1$. Proceeding by induction, suppose that the statement holds for trees up to order $n-1$, and turn to trees of order n, marked by $J = \{1, \ldots, n\}$. For any $I \subset J$ with $|I| = k \in [1, n)$, Lemma 13.17 shows that the number of marked, discrete trees of order n, such that J first splits into I and $J \setminus I$, equals

$$\begin{aligned} k!\,(k-1)!\; 2^{1-k}\,(n-k)!\,(n-k-1)!\; 2^{1-n+k} \binom{n-2}{k-1} \\ = (n-2)!\; k!\,(n-k)!\; 2^{2-n}, \end{aligned}$$

where the last factor on the left arises from the choice of $k-1$ splitting times for the I-component, among the remaining $n-2$ splitting times for the original tree. Since the total number of trees is $n!\,(n-1)!\,2^{1-n}$, the probability that J first splits into I and $J \setminus I$ equals

$$\frac{(n-2)!\; k!\,(n-k)!\; 2^{2-n}}{n!\,(n-1)!\; 2^{1-n}} = \frac{2}{n-1} \binom{n}{k}^{-1}.$$

Since all genealogies are equally likely, the discrete subtrees marked by I and $J \setminus I$ are conditionally independent and uniformly distributed, and the remaining branching times $2, \ldots, n-1$ are uniformly divided between the two trees. Since the splitting times $\tau_1 < \cdots < \tau_{n-1}$ of the continuous tree form a stationary binomial process on $[0, t]$, Lemma 3.15 shows that $\mathcal{L}(\tau_1)$ has density $(n-1)\,(t-s)^{n-2}\,t^{1-n}$, whereas $\tau_2, \ldots, \tau_{n-1}$ form a stationary binomial process on $[\tau_1, t]$, conditionally on τ_1. Furthermore, Lemma 3.13 shows that the splitting times of the two subtrees form independent, stationary binomial processes on $[\tau_1, t]$, conditionally on τ_1 and the initial partition of J into I and $J \setminus I$. Combining these facts with the conditional independence of the spatial motion, we see that the continuous subtrees marked by I and $J \setminus I$

are conditionally independent, uniform Brownian trees on $[\tau_1, t]$, given the spatial motion up to time τ_1 and the subsequent splitting of the original index set into I and $J \setminus I$.

Conditioning as indicated, and using the induction hypothesis and Theorem 13.18, we get

$$\begin{aligned}
\nu_t^n &= n!\, t^{n-1} \mu_t^n \\
&= n! \sum_{I \subset J}{}' \binom{n}{k}^{-1} \int_0^t (t-s)^{n-2}\, \mu_s * \left(\mu_{t-s}^I \otimes \mu_{t-s}^{J \setminus I}\right) ds \\
&= n! \sum_{I \subset J}{}' \binom{n}{k}^{-1} \frac{1}{k!\,(n-k)!} \int_0^t \nu_s * \left(\nu_{t-s}^I \otimes \nu_{t-s}^{J \setminus I}\right) ds \\
&= \sum_{I \subset J}{}' \int_0^t \nu_s * \left(\nu_{t-s}^I \otimes \nu_{t-s}^{J \setminus I}\right) ds,
\end{aligned}$$

where $|I| = k$. Note that, once again, a factor 2 cancels out in the second step, since every partition $\{I, J \setminus I\} = \{J \setminus I, I\}$ is counted twice. This completes the induction. $\square$

The last theorem may be extended to a representation of the Campbell measures of a canonical DW-cluster, in terms of a Brownian tree, leading to an explicit representation of the associated Palm distributions. For this we need an extension of Le Gall's Brownian snake, along with an extension of the underlying Brownian excursion. We begin with some basic properties of the regular Brownian snake, which in turn rely on the path decomposition of a Brownian excursion, given in Lemma A2.2.

Lemma 13.19 *(Brownian snake) Given a Brownian excursion X, conditioned to reach height $t > 0$, let Y be a Brownian snake with contour process X. Let σ and τ be the first and last times that X visits t, and write ρ for the first time that X attains its minimum on $[\sigma, \tau]$. Then X_ρ is $U(0, t)$, and the processes Y_σ and Y_τ are Brownian motions on $[0, t]$, each independent of X_ρ and extending the process Y_ρ on $[0, X_\rho]$. Furthermore, σ, ρ, and τ are Markov times for Y.*

Proof: This may be proved by approximation with a similar discrete snake, based on a simple random walk. We omit the details. $\square$

Given a Brownian excursion X, conditioned to reach level $t > 0$, we may form an *extended Brownian excursion* X_n, by inserting n independent copies of the path between the first and last visits to t. Let $\tau_1 < \cdots < \tau_n$ be the times connecting those $n + 1$ paths. Proceeding as for the regular Brownian snake, we may form an *extended Brownian snake* Y_n, by taking X_n as our new contour process. To prepare for the main result, we consider first the values of Y_n at the random times $\tau_1, \ldots, \tau_n$, each of which is a Brownian motion on $[0, t]$.

Corollary 13.20 *(extended snake and uniform tree) Let Y_n be an extended Brownian snake at level $t > 0$, with connection times $\tau_1 < \cdots < \tau_n$. Then the processes $Y_n \circ \tau_k$, $k \le n$, form a discrete Brownian snake on $[0,t]$, which, when randomly ordered, traces out a uniform Brownian tree on $[0,t]$.*

Proof: The first assertion is clear from Lemma 13.19 and its proof. The second assertion then follows, by comparison with the recursive construction in Lemma 13.17 (iii). □

We may now give explicit formulas for the n-th order Campbell and Palm measures of η with respect to η_t, in terms of the generated measure-valued processes η^n. In view of the previous corollary, the Campbell formula extends our representation of moment measures in Theorem 13.18, which leads to an interesting probabilistic interpretation.

Theorem 13.21 *(Campbell and Palm measures) Let Y_n be an extended Brownian snake with connection times $\tau_1, \ldots, \tau_n$, generating a measure-valued process η^n in R^d. Choose a random permutation $\pi_1, \ldots, \pi_n$ of $1, \ldots, n$, and define $\chi_k = Y_n(\tau_{\pi_k}, t)$, $k \le n$. Then*

(i) $\quad E_\mu \displaystyle\int f(x,\eta)\, \eta_t^{\otimes n}(dx) = n!\, t^{n-1} E_\mu f(\chi, \eta^n),$

(ii) $\quad \mathcal{L}_\mu\Big(\eta \,\Big\|\, \eta_t^{\otimes n}\Big)_\chi = \mathcal{L}_\mu(\eta^n | \, \chi)$ *a.s.*

Proof: (i) Let τ_0 and τ_1 be the first and last times that X visits t, and let $\tau_0, \ldots, \tau_{n+1}$ be the endpoints of the corresponding $n+1$ paths for the extended process X_n. Introduce the associated local times $\sigma_0, \ldots, \sigma_{n+1}$ of X_n at t, so that $\sigma_0 = 0$ and the remaining times form the beginning of a Poisson process with rate $c = t^{-1}$. Writing $(\sigma_{\pi_1}, \ldots, \sigma_{\pi_n}) = \sigma \circ \pi$, we get by Lemma 3.16

$$n!\, t^n E f\Big(\sigma \circ \pi,\, \sigma_{n+1}\Big) = E \int_{[0,\sigma_1]^n} f(s, \sigma_1)\, ds. \tag{5}$$

By Theorem 12.15, the shifted path $\theta_{\tau_0} X$ on $[0, \tau_1 - \tau_0]$ is generated by a Poisson process $\zeta \perp\!\!\!\perp \sigma_1$ of excursions from t, restricted to the set of paths not reaching level 0. By the strong Markov property, we can use the same process ζ to encode the excursions of X_n on the extended interval $[\tau_0, \tau_{n+1}]$, provided we choose $\zeta \perp\!\!\!\perp (\sigma_1, \ldots, \sigma_{n+1})$. By Lemma A2.2, the restrictions of X to the intervals $[0, \tau_0]$ and $[\tau_1, \infty)$ are independent of the intermediate path, and so, by construction, the corresponding property holds for the restrictions of X_n to $[0, \tau_0]$ and $[\tau_{n+1}, \infty)$. For convenience, we may extend ζ to a point process ζ' on $[0, \infty]$, by adding suitable points at 0 and ∞, encoding the initial and terminal paths of X and X_n. Then by independence, we get from (5)

$$n!\, t^n E f\Big(\sigma \circ \pi,\, \sigma_{n+1}, \zeta'\Big) = E \int_{[0,\sigma_1]^n} f(s, \sigma_1, \zeta')\, ds. \tag{6}$$

The inverse local time processes T and T_n of X and X_n can be obtained from the pairs (σ_1, ζ') and (σ_{n+1}, ζ'), respectively, by a common measurable construction, and we note that $T \circ \sigma_k = \tau_k$ for $k = 0, 1$, and $T_n \circ \sigma_k = \tau_k$ for $0 \le k \le n+1$. Furthermore, Theorem 12.18 shows that $\xi = \lambda \circ T^{-1}$ and $\xi_n = \lambda \circ T_n^{-1}$ are local time random measures of X and X_n at height t. Since the entire excursions X and X_n can be recovered from the same pairs by a common measurable map, we get from (6)

$$\begin{aligned} n!\, t^n E f\Big(\tau \circ \pi, X_n\Big) &= E \int_{[0,\sigma_1]^n} f(T \circ s, X)\, ds \\ &= E \int f(r, X)\, \xi^{\otimes n}(dr), \end{aligned} \tag{7}$$

where $T \circ s = (T_{s_1}, \ldots, T_{s_n})$, and the second equality holds by the elementary substitution rule in FMP 1.22.

Now introduce some Brownian snakes Y and Y_n, with contour processes X and X_n, with initial distribution μ, and with Brownian spatial motion in R^d. By Le Gall's path-wise construction, or by a suitable discrete approximation, the conditional distributions $\mathcal{L}_\mu(Y \mid X)$ and $\mathcal{L}_\mu(Y_n \mid X_n)$ are given by a common probability kernel. The same constructions justify the conditional independence $Y_n \perp\!\!\!\perp_{X_n} (\tau \circ \pi)$, and so by (7)

$$n!\, t^n E_\mu f\Big(\tau \circ \pi, Y_n\Big) = E_\mu \int f(r, Y)\, \xi^{\otimes n}(dr).$$

Since $\chi_k = Y_n(\tau \circ \pi_k, t)$ for all k, and η_t is the image of ξ under the map $Y(\cdot, t)$, the substitution rule yields

$$\begin{aligned} n!\, t^n E_\mu f(\chi, Y_n) &= E_\mu \int f\Big\{Y(r,t), Y\Big\}\, \xi^{\otimes n}(dr) \\ &= E_\mu \int f(x, Y)\, \eta_t^{\otimes n}(dx). \end{aligned}$$

Finally, the entire clusters η and η^n are generated by Y and Y_n, respectively, through a common measurable mapping, and so

$$n!\, t^n E_\mu f(\chi, \eta^n) = E_\mu \int f(x, \eta)\, \eta_t^{\otimes n}(dx),$$

which extends by monotone convergence to any initial measure μ. This agrees with the asserted formula, apart from an extra factor t, arising from the fact that we have been working with the conditional distribution, given that the cluster η reaches level t. The discrepancy disappears for the canonical cluster η, where the conditioning event has pseudo-probability t^{-1}.

(ii) This follows from (i), by disintegration of both sides. □

Using properties of the extended Brownian snake Y_n, along with the underlying contour process X_n, we obtain an explicit description of the multivariate Palm trees, in terms of conditional Brownian trees, adorned with some independent clusters sprouting from the branches. We omit the standard but subtle details of the proof.

Corollary 13.22 *(Palm trees) Given a canonical cluster η of a DW-process in R^d, we may form its Palm distribution with respect to $\eta_t^{\otimes n}$ at the points $x_1, \ldots, x_n \in \mathsf{R}^d$, by starting from a uniform, n-th order Brownian tree on $[0,t]$, conditioned to reach $x_1, \ldots, x_n$, and attach some independent clusters, rooted along the branches at a constant rate 2.*

In particular, the first order Brownian tree on $[0,t]$ is simply a standard Brownian motion, which makes the corresponding historical Palm tree at $x \in \mathsf{R}^d$ a Brownian bridge from 0 to x on the time scale $[0,t]$. This can also be seen more directly. A related discrete-time Palm tree is used in Theorem 13.64 below to characterize stability. The present multi-variate result seems to be much deeper.

13.4 Moment Densities

Using our representation of the moment measures of a DW-process, given in Theorem 13.18, in terms of some uniform Brownian trees, we can now derive a variety of estimates of the densities, required in subsequent sections. We begin with some estimates of the second order moment densities, covering our needs in Sections 13.7 and 13.8. For a DW-process ξ in R^d, we define the associated *covariance measure* on R^{2d} by

$$\mathrm{Cov}_\mu \xi_t = E_\mu \xi_t^{\otimes 2} - (E_\mu \xi_t)^{\otimes 2}.$$

Recalling our notation p_t for the symmetric density of the normal distribution in R^d with variances $t > 0$, we introduce the functions

$$q_t = 2 \int_0^t \left(p_s * p_{t-s}^{\otimes 2} \right) ds, \quad t > 0,$$

where the spatial convolutions $*$ are defined, for $x, y \in \mathsf{R}^d$, by

$$\begin{aligned}(\mu * q_t)(x,y) &= \int \mu(du)\, q_t(x-u, y-u),\\ \left(p_s * p_{t-s}^{\otimes 2}\right)(x,y) &= \int p_s(u)\, p_{t-s}(x-u)\, p_{t-s}(y-u)\, du.\end{aligned}$$

We shall further write $\simeq$ for exact or asymptotic equality up to a constant factor.

Lemma 13.23 *(covariance measure) Let ξ be a DW-process in R^d with canonical cluster η. Then for any $t > 0$, and for $x = (x_1, x_2)$ in R^{2d} with $x_i = \bar{x} \pm r$, we have*

(i) $\quad \mathrm{Cov}_\mu \xi_t = E_\mu \eta_t^2 = (\mu * q_t) \cdot \lambda^{2d}$,

(ii) $$(\mu * q_t)(x) \lesssim (\mu * p_t)(\bar{x}) \begin{cases} p_t(r)\, |r|^{2-d}\, t^{d/2}, & d \geq 3, \\ \left\{ t p_{t/2}(r) + \log_+(t|r|^{-2}) \right\}, & d = 2, \end{cases}$$

(iii) $\quad (\mu * q_t)(x) \simeq (\mu * p_t)(a) \Big| \log |r| \Big| \;\; \textit{as } x_1, x_2 \to a, \quad d = 2.$

Proof: (i) This is clear from Theorem 13.16 (i)−(ii) with $n = 2$.

(ii) By the definition of p_t and the parallelogram identity,

$$\begin{aligned} p_t(x_1)\, p_t(x_2) &\simeq t^{-d} \exp\Big\{ -(\, |x_1|^2 + |x_2|^2)/2t \Big\} \\ &= t^{-d} \exp\Big\{ -(\, |\bar{x}|^2 + |r|^2)/t \Big\} \\ &\simeq p_{t/2}(\bar{x})\, p_{t/2}(r), \end{aligned}$$

and so, by the semigroup property of the p_t,

$$\begin{aligned} q_t(x_1, x_2) &= 2 \int_0^t ds \int p_s(u)\, p_{t-s}(x_1 - u)\, p_{t-s}(x_2 - u)\, du \\ &\simeq \int_0^t ds \int p_s(u)\, p_{(t-s)/2}(\bar{x} - u)\, p_{(t-s)/2}(r)\, du \\ &= \int_0^t p_{(t-s)/2}(\bar{x})\, p_{(t-s)/2}(r)\, ds \\ &\lesssim p_t(\bar{x}) \int_0^t p_{s/2}(r)\, ds \\ &\lesssim p_t(\bar{x}) \int_0^t s^{-d/2}\, e^{-|r|^2/s}\, ds \\ &= p_t(\bar{x})\, |r|^{2-d} \int_{|r|^2/t}^\infty v^{d/2-2}\, e^{-v}\, dv. \end{aligned}$$

For $d \geq 3$, we obtain

$$\begin{aligned} q_t(x_1, x_2) &\lesssim p_t(\bar{x})\, |r|^{2-d}\, e^{-|r|^2/2t} \\ &\simeq p_t(\bar{x})\, p_t(r)\, |r|^{2-d}\, t^{d/2}, \end{aligned}$$

and the asserted estimate follows by convolution with μ. If instead $d = 2$, we get

$$q_t(x_1, x_2) \lesssim p_t(\bar{x}) \int_{|r|^2/t}^\infty v^{-1}\, e^{-v}\, dv.$$

When $|r|^2 \leq t/2$, we have

$$\begin{aligned} \int_{|r|^2/t}^\infty v^{-1}\, e^{-v}\, dv &\lesssim \int_{|r|^2/t}^1 v^{-1}\, dv \\ &= \log(t/|r|^2), \end{aligned}$$

whereas for $|r|^2 \geq t/2$, we get

$$\begin{aligned} \int_{|r|^2/t}^\infty v^{-1}\, e^{-v}\, dv &\leq \int_{|r|^2/t}^\infty e^{-v}\, dv \\ &= \exp\Big(-|r|^2/t \Big) \lesssim t\, p_{t/2}(r). \end{aligned}$$

Again the required estimate follows by convolution with μ.

(iii) For fixed $\varepsilon > 0$, we have

$$\begin{aligned} q_t(x_1, x_2) &\simeq \int_0^t p_{(t-s)/2}(r)\, p_{(t+s)/2}(\bar{x})\, ds \\ &\sim \int_{t-\varepsilon}^t p_{(t-s)/2}(r)\, p_{(t+s)/2}(\bar{x})\, ds, \end{aligned}$$

since

$$\begin{aligned} \int_0^{t-\varepsilon} p_{(t-s)/2}(r)\, p_{(t+s)/2}(\bar{x})\, ds \;&\lesssim\; p_t(\bar{x}) \int_{|r|^2/t}^{|r|^2/\varepsilon} v^{-1}\, dv \\ &\to\; p_t(x)\, \log(t/\varepsilon) < \infty. \end{aligned}$$

Noting that $p_{(t+s)/2}(\bar{x}) \to p_t(\bar{x}) \to p_t(x)$, as $s \to t$ and then $x_1, x_2 \to x$, we get for fixed $b > 0$

$$\begin{aligned} q_t(x_1, x_2) &\simeq p_t(x) \int_0^t p_{s/2}(r)\, ds \\ &= p_t(x) \int_{|r|^2/t}^{\infty} v^{-1} e^{-v} dv \\ &\sim p_t(x) \int_{|r|^2/t}^{b} v^{-1} e^{-v} dv, \end{aligned}$$

where the last relation holds, since $\int_b^\infty v^{-1} e^{-v} dv < \infty$. Since $e^{-v} \to 1$ as $v \to 0$, we obtain

$$\begin{aligned} q_t(x_1, x_2) &\simeq p_t(x) \int_{|r|^2/t}^{1} v^{-1}\, dv \\ &= p_t(x)\, \log\left(t/|r|^2\right) \\ &\simeq p_t(x) \Big| \log |r| \Big|. \end{aligned}$$

This proves the assertion for $\mu = \delta_0$. For general μ, let $c > 0$ be such that $q_t(x_1, x_2) \sim c p_t(x) |\log |r||$. We need to show that

$$\Big| \log |r| \Big|^{-1} \int \mu(du)\, q_t(x_1 - u, x_2 - u) \to c \int \mu(du)\, p_t(x - u),$$

as $x_1, x_2 \to x$ for fixed μ and t. Then note that, by (ii) and Lemma A3.3,

$$\begin{aligned} \Big| \log |r| \Big|^{-1} q_t(x_1 - u, x_2 - u) \;&\lesssim\; p_t(\bar{x} - u) \\ &\lesssim\; p_{t+h}(x - u), \end{aligned}$$

as long as $|\bar{x} - x| \leq h$. Since $(\mu * p_{t+h})(x) < \infty$, the desired relation follows by dominated convergence. □

Part (iii) of the last lemma yields a useful asymptotic scaling property, needed in the proof of Theorem 13.53. Recall the definition of the scaling operators S_r, prior to the statement of Theorem 13.14.

Corollary 13.24 *(diagonal rate) Let ξ be a DW-process in R^2 with canonical cluster η. Consider a measurable function $f \geq 0$ on R^4, such that $f(x,y)\log(|x-y|^{-1} \vee e)$ is integrable, and suppose that either μ or* supp f *is bounded. Then for fixed $t > 0$ and μ, we have as $\varepsilon \to 0$*

$$E_\mu(\xi_t S_\varepsilon)^2 f \simeq \varepsilon^4 |\log \varepsilon|\, \lambda^4 f\, \mu p_t.$$

This holds in particular when f is bounded with bounded support, and it remains true for the clusters η_t.

Proof: When $x_1 \approx x_2$ in R^2, Lemma 13.23 (iii) shows that the density g of $E_\mu \xi_t^2$ satisfies

$$g(x_1, x_2) \sim c\,\Big|\log|x_1 - x_2|\Big|\,(\mu * p_t)\Big\{\tfrac{1}{2}(x_1 + x_2)\Big\},$$

for some constant $c > 0$, and is otherwise bounded for bounded μ. Furthermore,

$$\begin{aligned} E_\mu \xi_t^2 f &= \int f(u/\varepsilon)\, g(u)\, du \\ &= \varepsilon^4\, |\log \varepsilon| \int f(x)\, \frac{g(\varepsilon x)}{|\log \varepsilon|}\, dx. \end{aligned}$$

Here the ratio in the last integrand tends to $c\,\mu p_t$ as $\varepsilon \to 0$, and so for bounded μ or supp f, the integral tends to $c\,\mu p_t\, \lambda^4 f$, by dominated convergence. To verify the stated integrability when f is bounded, we may change (x_1, x_2) into new coordinates $x_1 \pm x_2$, express $x_1 - x_2$ in terms of polar coordinates (r, θ), and note that $\int_0^1 r|\log r| dr < \infty$. $\square$

Turning to the higher order moment densities, required for the multivariate approximations in Section 13.9, we begin with a density estimate for a general Brownian tree, defined as a random tree in R^d, with spatial motion given by independent Brownian motions along the branches. For any $x \in (\mathsf{R}^d)^n$, let r_x denote the orthogonal distance from x to the diagonal set $D_n = \{(\mathsf{R}^d)^{(n)}\}^c$.

Lemma 13.25 *(Brownian tree density) For any marked Brownian tree on $[0, s]$, with n leaves and paths in R^d, the joint distribution at time s has a continuous density q on $(\mathsf{R}^d)^{(n)}$, satisfying*

$$q(x) \leq \Big(1 \vee tdn^2 r_x^{-2}\Big)^{nd/2} p_{nt}^{\otimes nd}(x), \quad x \in (\mathsf{R}^d)^{(n)},\ s \leq t.$$

Proof: Conditionally on tree structure and splitting times, the joint distribution is a convolution of centered Gaussian distributions $\mu_1, \ldots, \mu_n$, supported by some linear subspaces $S_1 \subset \cdots \subset S_n = \mathsf{R}^{nd}$ of dimensions $d, 2d, \ldots, nd$. The tree structure is determined by a nested sequence of partitions $\pi_1, \ldots, \pi_n$ of the index set $\{1, \ldots, n\}$. Writing $h_1, \ldots, h_n$ for the times between splits, we see from Lemma A4.2 that μ_k has principal variances

$h_k|J|$, $J \in \pi_k$, each with multiplicity d. Writing $\nu_t = p_t \cdot \lambda^{\otimes d}$, and noting that $|J| \le n-k+1$ for $J \in \pi_k$, we get by Lemma A3.1

$$\mu_k \le (n-k+1)^{(k-1)d/2}\, \nu^{kd}_{(n-k+1)h_k} \otimes \delta_0^{\otimes (n-k)d}, \quad k \le n.$$

Putting

$$c = \prod_{k \le n} (n-k+1)^{(k-1)d/2} \le n^{n^2 d/2},$$
$$s_k = (n-k+1)h_k,$$
$$t_k = s_k + \cdots + s_n, \qquad k \le n,$$

we get

$$c^{-1} \mathop{(*)}_{k \le n} \mu_k \le \mathop{(*)}_{k \le n} \Big(\nu^{kd}_{s_k} \otimes \delta_0^{\otimes (n-k)d} \Big)$$
$$= \bigotimes_{k \le n} \mathop{(*)}_{j \ge k} \nu^d_{s_j} = \bigotimes_{k \le n} \nu^d_{t_k}.$$

Now consider the orthogonal decomposition $x = x_1 + \cdots + x_n$ in R^{nd} with $x_k \in S_k \ominus S_{k-1}$, and write $x' = x - x_n$. Since $|x_n|$ equals the orthogonal distance of x to the subspace $S_{n-1} \subset D_n$, we get $|x_n| \ge r_x$. Using Lemma A6.2, and noting that $h_n = t_n \le t_k \le nt$, we see that the continuous density of $c^{-1}(\mu_1 * \cdots * \mu_n)$ at x is bounded by

$$\begin{aligned}
\prod_{k \le n} p^{\otimes d}_{t_k}(x_k) &= \prod_{k \le n} (2\pi t_k)^{-d/2}\, e^{-|x_k|^2/2t_k} \\
&\le (2\pi h_n)^{-nd/2}\, e^{-|x_n|^2/2h_n} \prod_{k<n} e^{-|x_k|^2/2nt} \\
&= p^{\otimes nd}_{h_n}(|x_n|)\, e^{-|x'|^2/2nt} \\
&\le \Big(1 \vee td\, n^2 |x_n|^{-2}\Big)^{nd/2} p^{\otimes nd}_{nt}(|x_n|)\, e^{-|x'|^2/2nt} \\
&\le \Big(1 \vee td\, n^2 r_x^{-2}\Big)^{nd/2} p^{\otimes nd}_{nt}(x),
\end{aligned}$$

where $|x_n|$ also denotes the vector $(|x_n|, 0, \dots, 0)$. Since the right-hand side is independent of branching structure and splitting times, the unconditional density $q(x)$ has the same bound, and the desired estimate follows. The stated continuity follows by dominated convergence from the continuity of the normal density. □

The last result leads to a useful estimate, for the moment densities of a single cluster of a DW-process, needed throughout the remainder of this section.

Lemma 13.26 *(cluster moment densities) For a DW-process in R^d, the cluster moment measures $\nu^n_t = E_0 \eta_t^{\otimes n}$ have densities $q^n_t(x)$, which are jointly continuous in $(x,t) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$, and satisfy the uniform bounds*

$$\sup_{s \le t} q^n_s(x) \lesssim \Big(1 \vee r_x^{-2}\, t\Big)^{nd/2} p^{\otimes n}_{nt}(x), \quad x \in (\mathsf{R}^d)^{(n)},\ t > 0.$$

Proof: By Theorem 13.18, it is equivalent to consider the joint endpoint distribution of a uniform, n-th order Brownian tree in R^d, on the interval $[0,t]$, where the stated estimate holds by Lemma 13.25. To prove the asserted continuity, we may condition on the tree structure and splitting times, to get a non-singular Gaussian distribution, for which the assertion is obvious. The unconditional statement then follows by dominated convergence, based on the uniform bound in Lemma 13.25. □

Using the last estimate, we can easily derive some density versions of the basic moment identities of Theorem 13.16, needed throughout the remaining sections.

Corollary 13.27 *(identities for moment densities) For a DW-process in* R^d *with initial measure* $\mu \neq 0$, *the moment measures* $E_\mu \xi_t^{\otimes n}$ *and* $\nu_t^n = E_0 \eta_t^{\otimes n}$ *have positive, jointly continuous densities on* $(\mathsf{R}^d)^{(n)} \times (0,\infty)$, *satisfying density versions of the identities in Theorem 13.16* (i)$-$(iv).

Proof: Let q_t^n denote the jointly continuous densities of ν_t^n, considered in Lemma 13.26. As mixtures of normal densities, they are again strictly positive. Inserting the versions q_t^n into the convolutions of Theorem 13.16 (i), we get some strictly positive densities of the measures $E_\mu \xi_t^{\otimes n}$, and their joint continuity follows, by extended dominated convergence (FMP 1.21), from the estimates in Lemma 13.26, and the joint continuity in Lemma A3.4.

Inserting the continuous densities q_t^n into the expressions on the right of Theorem 13.16 (ii)$-$(iv), we obtain densities of the measures on the left. If the latter functions can be shown to be continuous on $(\mathsf{R}^d)^{(J)}$ or $(\mathsf{R}^d)^{(n)}$, respectively, they must agree with the continuous densities q_t^J or q_{s+t}^n, and the desired identities follow. By Lemma 13.26 and extended dominated convergence, it is enough to prove the required continuity, with q_s and q_s^n replaced by the normal densities p_t and $p_{nt}^{\otimes n}$, respectively. Hence, we need to show that the convolutions

$$p_t * p_{nt}^{\otimes n}, \qquad p_{(n-1)t}^{\otimes(n-1)} * p_{nt}^{\otimes n}, \qquad p_{nt}^{\otimes \pi} * p_{nt}^{\otimes n}$$

are continuous, which is clear, since they are all non-singular Gaussian. □

We also consider the conditional moment densities of a DW-process, for the purpose of deriving related regularity properties of the multi-variate Palm distributions, via the duality theory in Section 6.7.

Theorem 13.28 *(conditional moment densities) Let* ξ *be a DW-process in* R^d *with* $\xi_0 = \mu$. *Then for every* n, *there exist some processes* M_s^t *on* R^{nd}, $0 \le s < t$, *such that*

(i) $E_\mu\big(\xi_t^{\otimes n}\big|\,\xi_s\big) = M_s^t \cdot \lambda^{\otimes nd}$ *a.s.*, $0 \le s < t$,

(ii) M_s^t *is a martingale in* $s \in [0,t)$, *for fixed* $x \in (\mathsf{R}^d)^{(n)}$ *and* $t > 0$,

(iii) *$M_s^t(x)$ is continuous, a.s. and in L^1, in $(x,t) \in (\mathsf{R}^d)^{(n)} \times (s,\infty)$ for fixed $s \geq 0$.*

Proof: Write $S_n = (\mathsf{R}^d)^{(n)}$, let q_t^n denote the continuous densities in Lemma 13.26, and let x_J be the projection of $x \in \mathsf{R}^{nd}$ onto $(\mathsf{R}^d)^J$. By the Markov property of ξ and Theorem 13.16 (i), the random measures $E_\mu(\xi_t^{\otimes n} \,|\, \xi_s)$ have a.s. densities

$$M_s^t(x) = \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} \Big(\xi_s * q_{t-s}^J\Big)(x_J), \quad x \in S_n, \tag{8}$$

which are a.s. continuous in $(x,t) \in S_n \times (s,\infty)$ for fixed $s \geq 0$, by Corollary 13.27. Indeed, the previous theory applies with μ replaced by ξ_s, since $E_\mu \xi_s p_t = \mu p_{s+t} < \infty$, and hence $\xi_s p_t < \infty$ for every $t > 0$ a.s.

To prove the L^1-continuity in (iii), it suffices, by FMP 1.32, to show that $E_\mu M_s^t(x)$ is continuous in $(x,t) \in S_n \times (s,\infty)$. By Lemma 13.26 and extended dominated convergence, it is then enough to prove the a.s. and L^1 continuity in $x \in S_n$ alone, for the processes in (8), with q_{t-s}^J replaced by $p_t^{\otimes J}$. Here the a.s. convergence holds by Lemma A3.4, and so, by Theorem 13.16 (i), it remains to show that $\mu * q_s^n * p_t^{\otimes n}$ is continuous on S_n, for fixed s, t, μ, and n. Since $q_s^n * p_t^{\otimes n} = \nu_s^n * p_t^{\otimes n}$ is continuous on S_n, by Theorem 13.18 and Lemma 13.25, it suffices, by Lemma 13.26 and extended dominated convergence, to show that $\mu * p_t^{\otimes n}$ is continuous on R^{nd} for fixed t, μ, and n, which holds by Lemma A3.4.

To prove (ii), let $B \subset \mathsf{R}^{nd}$ be measurable, and note that

$$\begin{aligned} \lambda^{\otimes nd}(M_0^t; B) &= E_\mu \xi_t^{\otimes n} B = E_\mu E_\mu\Big(\xi_t^{\otimes n} B \Big|\, \xi_s\Big) \\ &= E_\mu \lambda^{\otimes nd}(M_s^t;\, B) \\ &= \lambda^{\otimes nd}\Big(E_\mu M_s^t;\, B\Big), \end{aligned}$$

which implies $M_0^t = E_\mu M_s^t$ a.e. Since both sides are continuous on S_n, they agree identically on the same set, and so by (8),

$$E_\mu \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} \Big(\xi_s * q_{t-s}^J\Big)(x_J) = \sum_{\pi \in \mathcal{P}_n} \prod_{J \in \pi} (\mu * q_t^J)(x_J), \quad s < t.$$

Replacing μ by ξ_r for arbitrary $r > 0$, and using the Markov property at r, we obtain

$$E_\mu\Big(M_{t+s}^{r+t}(x)\Big|\, \xi_r\Big) = M_r^{r+t}(x) \ \text{ a.s.}, \quad x \in S_n,\ r > 0,\ 0 \leq s < t,$$

which yields the martingale property in (ii). □

For some purposes of subsequent sections, we often need some more careful estimates of the moment densities near the diagonals. Here we write $q_{\mu,t}^n$ for the continuous density of $E_\mu \xi_t^{\otimes n}$, considered in Corollary 13.27.

Lemma 13.29 *(densities near diagonals) Let ξ be a DW-process in R^d. Then*

(i) $E_\mu \xi_s^{\otimes n} * p_h^{\otimes n}$ *is continuous on* R^{nd},

(ii) *for fixed $t > 0$, we have, uniformly on R^{nd} in $s \le t$, $h > 0$, and μ,*

$$E_\mu \xi_s^{\otimes n} * p_h^{\otimes n} \lesssim \left(1 \vee h^{-1} t\right)^{nd/2} \sum_{\pi \in \mathcal{P}_n} \bigotimes_{J \in \pi} \left(\mu * p_{nt+h}^{\otimes J}\right) < \infty,$$

(iii) $E_\mu \xi_s^{\otimes n} * p_h^{\otimes n} \to q_{\mu,t}^n$ *on* $(\mathsf{R}^d)^{(n)}$, *as $s \to t$ and $h \to 0$.*

Proof, (i)$-$(ii): By Theorem 13.16 (i), it suffices to show that

$$\nu_s^n * p_h^{\otimes n} \lesssim (1 \vee h^{-1} t)\, p_{nt+h}^{\otimes n},$$

uniformly in $s < t$. By Theorem 13.18, we may then replace ν_s^n by the distribution of the endpoint vector γ_s^n of a uniform Brownian tree. Conditioning on tree structure and splitting times, we see from Lemma A4.2 that γ_s^n becomes centered Gaussian with principal variances bounded by nt. Convolving with $p_h^{\otimes n}$ gives a centered Gaussian density with principal variances in $[h, nt + h]$, and Lemma A3.1 yields the required bound for the latter density, in terms of the rotationally symmetric version $p_{nt+h}^{\otimes n}$. Taking expected values yields the corresponding unconditional bound. The asserted continuity may now be proved as in case of Lemma A3.4.

(iii) First we consider the corresponding statement for a single cluster. Here both sides are mixtures of similar normal densities, obtained by conditioning on splitting times and branching structure in the equivalent Brownian trees of Theorem 13.18, and the statement results from an elementary approximation of a stationary binomial process on $[0, t]$ by a similar process on $[0, s]$. The general result now follows by dominated convergence, from the density version of Theorem 13.16 (i), established in Corollary 13.27. □

Yet another moment estimate will be needed in the next section. For any $x = (x_1, \ldots, x_n) \in (\mathsf{R}^d)^n$ and $k \le n$, we write $x^k = (x_1, \ldots, x_k)$.

Lemma 13.30 *(rates of moment densities) For any μ and $1 \le k \le n$ we have, uniformly for $0 < h \le r \le (t \wedge \frac{1}{2})$ and $(x, t) \in (\mathsf{R}^d)^{(n)} \times (0, \infty)$ in compacts,*

$$\left\{E_\mu \xi_t^{\otimes (n+k)} * (p_h^{\otimes n} \otimes p_r^{\otimes k})\right\}(x, x^k) \lesssim \begin{cases} r^{k(1-d/2)}, & d \ge 3, \\ |\log r|^k, & d = 2. \end{cases}$$

Proof: First we prove a similar estimate for the moment measures ν_t^{n+k} of a single cluster. By Theorem 13.18, it is equivalent to consider the distribution μ_t^{n+k} of the endpoint vector $(\gamma_1, \ldots, \gamma_{n+k})$, for a uniform Brownian tree on $[0, t]$. Then let τ_i and α_i be the time and place of attachment of leaf number $n + i$, and put $\tau = (\tau_i)$ and $\alpha = (\alpha_i)$. Let $\mu_{t|\tau}^n$ and $\mu_{t|\tau,\alpha}^n$ denote the conditional distributions of $(\gamma_1, \dots, \gamma_n)$, given τ or (τ, α), respectively, and put $u = t + \tau$. Then we have, uniformly for h and r as above and $x \in (\mathsf{R}^d)^{(n)}$,

$$\begin{aligned}
&\Big\{ \mu_t^{n+k} * (p_h^{\otimes n} \otimes p_r^{\otimes k}) \Big\}(x, x^k) \\
&\quad = E\Big(\mu^n_{t|\tau,\alpha} * p_h^{\otimes n}\Big)(x) \prod_{i \le k} p_{u-\tau_i}(x_i - \alpha_i) \\
&\quad \le E\Big(\mu^n_{t|\tau,\alpha} * p_h^{\otimes n}\Big)(x) \prod_{i \le k} (u - \tau_i)^{-d/2} \\
&\quad \lesssim q_u^n(x) \Big\{ \int_r^u s^{-d/2}\, ds \Big\}^k \\
&\quad \lesssim q_u^n(x) \begin{cases} r^{k(1-d/2)}, \\ |\log r|^k, \end{cases}
\end{aligned}$$

when $d \ge 3$ or $d = 2$, respectively. Here the first equality holds, since $(\gamma_1, \dots, \gamma_n)$ and $\gamma_{n+1}, \dots, \gamma_{n+k}$ are conditionally independent given τ and α. The second relation holds since $\|p_r\| \le r^{-d/2}$. We may now use the chain rule for conditional expectations to replace $\mu^n_{t|\tau,\alpha}$ by $\mu^n_{t|\tau}$. Next, we apply Lemma 3.13 twice, first to replace $\mu^n_{t|\tau}$ by μ^n_t, and then to replace $\tau_{n+1}, \dots, \tau_{n+k}$ by a stationary binomial process on $[0, t]$. We also note that $\mu_t^n * \nu_h^{\otimes n} \lesssim \mu_u^n$, by the corresponding property of the binomial process. This implies $\mu_t^n * p_h^{\otimes n} \lesssim q_u^n$, since both sides have continuous densities outside the diagonals, by Lemma 13.26, justifying the third step. The last step holds by elementary calculus.

Since the previous estimate is uniform in $x \in (\mathsf{R}^d)^{(n)}$, it remains valid for the measures $\mu * \nu_t^{n+k}$, with q_u^n replaced by the convolution $\mu * q_u^n$, which is bounded on compacts in $(\mathsf{R}^d)^{(n)} \times (0, \infty)$ by Lemma 13.26. Finally, Theorem 13.16 (i) shows as before that

$$\mu * \nu_t^n * p_h^{\otimes n}(x) \lesssim \mu * q_u^n(x), \quad h \le t,$$

which is again locally bounded. □

Two more technical results will be needed, for the proof of our crucial Lemma 13.33 below.

Lemma 13.31 *(uniform density estimate) For any initial measure $\mu \in \mathcal{M}_d$ and $\pi \in \mathcal{P}_n$ with $|\pi| < n$,*

$$\mu * \nu_t^\pi * \bigotimes_{J \in \pi} \nu_h^J = q_{t,h} \cdot \lambda^{\otimes nd}, \quad t, h > 0,$$

where $q_{t,h}(x) \to 0$ as $h \to 0$, uniformly for $(x, t) \in (\mathsf{R}^d)^{(n)} \times (0, \infty)$ in compacts.

Proof: For $x = (x_1, \dots, x_n) \in (\mathsf{R}^d)^{(n)}$, write $\Delta = \min_{i \ne j} |x_i - x_j|$, and note that

$$\inf_{u \in (\mathsf{R}^d)^\pi} \sum_{i \in J \in \pi} |x_i - u_J|^2 \ge (\Delta/2)^2 \sum_{J \in \pi} (|J| - 1) \ge \Delta^2/4.$$

Letting q_h denote the continuous density of $\bigotimes_{J\in\pi}\nu_h^J$ on $(\mathsf{R}^d)^{(n)}$, and using Lemma 13.26, we get as $h \to 0$

$$\begin{aligned}\sup_{u\in(\mathsf{R}^d)^\pi} q_h(x-u) &\lesssim \sup_{u\in(\mathsf{R}^d)^\pi} \prod_{i\in J\in\pi} p_{nh}(x_i - u_J)\\ &\lesssim h^{-nd/2}\, e^{-\Delta^2/8nhd} \to 0,\end{aligned}$$

uniformly for $x \in (\mathsf{R}^d)^{(n)}$ in compacts. Since $\|\nu_t^\pi\| = |\pi|!\, t^{|\pi|-1}$ by Theorem 13.18, we conclude that

$$\sup_{u\in(\mathsf{R}^d)^\pi} (\nu_t^\pi * q_h)(x-u) \to 0, \quad h \to 0,$$

uniformly for $(x,t) \in (\mathsf{R}^d)^{(n)} \times \mathsf{R}_+$ in compacts.

Since the densities q_t^n of ν_t^n satisfy $\nu_t^\pi * \bigotimes_{J\in\pi} q_h^J \le q_{t+h}^n$, by Theorem 13.16 (iv) and Corollary 13.27, we have $\nu_t^\pi * q_h \le q_{t+h}^n$. Using Lemmas A3.3 and 13.26, and writing $\bar u = (u,\dots,u)$, we get

$$(\nu_t^\pi * q_h)(x-u) \lesssim p_b^{\otimes n}(-\bar u) = p_b^{\otimes n}(\bar u), \quad u \in \mathsf{R}^d,$$

for some constant $b > 0$, uniformly for $(x,t) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$ in compacts. Here $\int p_b^{\otimes n}(\bar u)\,\mu(du) < \infty$, since $\mu p_t < \infty$ for all $t > 0$. Letting $h = h_n \to 0$, and restricting $(x,t) = (x_n,t_n)$ to a compact subset of $(\mathsf{R}^d)^{(n)} \times (0,\infty)$, we get by dominated convergence

$$\begin{aligned} q_{t,h}(x) &= (\mu * \nu_t^\pi * q_h)(x)\\ &= \int \mu(du)\,(\nu_t^\pi * q_h)(x-u) \to 0,\end{aligned}$$

which yields the required uniform convergence. □

We finally estimate the contribution from remote clusters, as needed for the proof of Lemma 13.33. Writing η for the canonical cluster of a DW-process in R^d, we define

$$\nu_{h,\varepsilon}(dx) = E_0\Big\{\eta_h(dx);\, \sup_t \eta_t(B_x^\varepsilon)^c > 0\Big\}, \quad h,\varepsilon > 0,\ x \in \mathsf{R}^d.$$

Lemma 13.32 *(remote cluster densities) For any initial measure $\mu \in \mathcal{M}_d$,*

$$\mu * \nu_t^n * \Big(\nu_{h,\varepsilon} \otimes \nu_h^{\otimes(n-1)}\Big) = q_{t,h}^\varepsilon \cdot \lambda^{\otimes nd}, \quad t,h,\varepsilon > 0,$$

where $q_{t,h}^\varepsilon(x) \to 0$, as $\varepsilon^{2+r} \ge h \to 0$ for some $r > 0$, uniformly for $(x,t) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$ in compacts.

Proof: Let ρ be the span of η from 0, and put $T(r) = P_0(\rho > r \,\|\, \eta_1)_0$ and $h' = h^{1/2}$. By Palm disintegration and Theorem 13.14 (iv), the density of $\nu_{h,\varepsilon}$ is bounded by

$$p_h(x)\,T\Big(\frac{\varepsilon-|x|}{h'}\Big) \le \begin{cases} p_h(x)\,T(\varepsilon/2h'), & |x|\le \varepsilon/2,\\ p_h(x), & |x|>\varepsilon/2.\end{cases}$$

Writing $\nu'_{h,\varepsilon}$ and $\nu''_{h,\varepsilon}$ for the restrictions of $\nu_{h,\varepsilon}$ to $B_0^{\varepsilon/2}$ and $(B_0^{\varepsilon/2})^c$, respectively, we conclude that $\nu'_{h,\varepsilon}\otimes\nu_h^{\otimes(n-1)}$ has a density bounded by $T(\varepsilon/2h')\times p_h^{\otimes n}(x)$. Hence, Theorem 13.18 shows that, for $0<h\le t$, even $\nu_t*(\nu'_{h,\varepsilon}\otimes\nu_h^{\otimes(n-1)})$ has a density of the order $T(\varepsilon/2h')\,q^n_{t+h}(x)$. Here $T(\varepsilon/2h')\to 0$ as $h/\varepsilon^2\to 0$, since $\rho<\infty$ a.s., and Lemma 13.26 gives $\sup_u q^n_{t+h}(x-u)\lesssim 1$, uniformly for $(x,t)\in(\mathsf{R}^d)^{(n)}\times(0,\infty)$ in compacts.

Next, we note that the density of $\nu''_{h,\varepsilon}\otimes\nu_h^{\otimes(n-1)}$ is bounded by

$$p_h^{\otimes n}(x)\,1\Big\{|x_1|>\varepsilon/2\Big\}\lesssim h^{-nd/2}\,e^{\varepsilon^2/8h}\to 0.$$

Since $\|\nu_t^n\|\lesssim 1$ for bounded $t>0$, by Theorem 13.18, even $\nu_t^n*(\nu''_{h,\varepsilon}\otimes\nu_h^{\otimes(n-1)})$ has a density that tends to 0, uniformly for $x\in\mathsf{R}^{nd}$ and bounded $t>0$. Combining the results for $\nu'_{h,\varepsilon}$ and $\nu''_{h,\varepsilon}$, we conclude that $\nu_t^n*(\nu_{h,\varepsilon}\otimes\nu_h^{\otimes(n-1)})$ has a density $q^\varepsilon_{t,h}$ satisfying $\sup_u q^\varepsilon_{t,h}(x-u)\to 0$, uniformly for $(x,t)\in(\mathsf{R}^d)^{(n)}\times(0,\infty)$ in compacts. To deduce the stated result for general μ, we may argue as in the previous proof, using dominated convergence based on the relations $\nu_{h,\varepsilon}\le\nu_h$ and $\nu_t^n*\nu_h^{\otimes n}\le\nu^n_{t+h}$, with associated density versions, valid by Theorem 13.16 (iv) and Corollary 13.27. □

13.5 Regular Palm Distributions

Here we use the density estimates of the previous section, along with the duality theory in Section 6.7, to construct versions of the multi-variate Palm distributions of a DW-process with strong regularity properties. Such regular versions will in turn be needed for the local and conditional approximations of a DW-process in Section 13.9.

The following uniform bound for the shifted Palm distributions will play a crucial role in the sequel.

Lemma 13.33 *(strongly continuous Palm versions) For fixed μ, $t>0$, and open $G\subset\mathsf{R}^d$, there exist some functions p_h on $G^{(n)}$ with $p_h\to 0$ as $h\to 0$, uniformly for $(x,t)\in G^{(n)}\times(0,\infty)$ in compacts, such that a.e. $E_\mu\xi_s^{\otimes n}$ on $G^{(n)}$, and for $r<s\le t$ with $2s>t+r$,*

$$\Big\|\,\mathcal{L}_\mu\Big(1_{G^c}\xi_t\,\Big\|\,\xi_s^{\otimes n}\Big)-E_\mu\Big\{\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r})\,\Big\|\,\xi_s^{\otimes n}\Big\}\Big\|\le p_{t-r}\,.$$

Proof: The random measures ξ_s and ξ_t may be regarded as Cox cluster processes, generated by the random measure $h^{-1}\xi_r$ and the probability kernel $h\mathcal{L}_u(\eta_h)$ from R^d to $\mathcal{M}_d$, where $h=s-r$ or $t-r$, respectively. To keep track of the cluster structure, we introduce some marked versions $\tilde\xi_s$ and $\tilde\xi_t$ on $\mathsf{R}^d\times[0,1]$, where each cluster ξ_i is replaced by $\tilde\xi_i=\xi_i\otimes\delta_{\sigma_i}$, for some i.i.d.

$U(0,1)$ random variables σ_i, independent of the ξ_i. Note that $\tilde{\xi}_s$ and $\tilde{\xi}_t$ are again cluster processes, with generating kernels $\tilde{\nu}_h = \mathcal{L}(\xi_h \otimes \delta_\sigma)$ from R^d to $\mathcal{M}(\mathsf{R}^d \times [0,1])$, where $\mathcal{L}(\xi_h, \sigma) = \nu_h \otimes \lambda$. By the transfer theorem (FMP 6.10), we may assume that $\tilde{\xi}_t(\cdot \times [0,1]) = \xi_t$ a.s.

For any $v = (v_1, \dots, v_n) \in [0,1]^n$, we may write $\tilde{\xi}_t = \tilde{\xi}_{t,v} + \tilde{\xi}'_{t,v}$, where $\tilde{\xi}_{t,v}$ denotes the restriction of $\tilde{\xi}_t$ to $\mathsf{R}^d \times \{v_1, \dots, v_n\}^c$, which is product-measurable in (ω, v) by Lemma 1.11. Writing $D = ([0,1]^{(n)})^c$, we get for any $B \in \hat{\mathcal{B}}^{nd}$, and for measurable functions $f\colon (\mathsf{R}^d \times [0,1])^n \times \mathcal{M}(\mathsf{R}^d \times [0,1]) \to [0,1]$, with $f_{x,v} = 0$ for $x \in B^c$,

$$\begin{aligned}
&\left| \iint E_\mu \tilde{\xi}_s^{\otimes n}(dx\,dv) \left(E_\mu\Big\{ f_{x,v}(1_{G^c}\tilde{\xi}_t) \Big\| \, \tilde{\xi}_s^{\otimes n} \Big\}_{x,v} - E_\mu\Big\{ f_{x,v}(1_{G^c}\tilde{\xi}_{t,v}) \Big\| \, \tilde{\xi}_s^{\otimes n} \Big\}_{x,v} \right) \right| \\
&\quad \le E_\mu \iint 1_B(x)\, \tilde{\xi}_s^{\otimes n}(dx\,dv) \left| f_{x,v}(1_{G^c}\tilde{\xi}_t) - f_{x,v}(1_{G^c}\tilde{\xi}_{t,v}) \right| \\
&\quad \le E_\mu \iint 1_B(x)\, \tilde{\xi}_s^{\otimes n}(dx\,dv)\, 1\Big\{ \tilde{\xi}'_{t,v} G^c > 0 \Big\} \\
&\quad \le E_\mu \tilde{\xi}_s^{\otimes n}(B \times D) + E_\mu \iint_{B \times D^c} \tilde{\xi}_s^{\otimes n}(dx\,dv) \sum_{i \le n} 1\Big\{ \tilde{\xi}'_{t,v_i} G^c > 0 \Big\}.
\end{aligned}$$

To estimate the first term on the right, define $\mathcal{P}'_n = \{\pi \in \mathcal{P}_n;\, |\pi| < n\}$. For any $\kappa, \pi \in \mathcal{P}_n$, we mean by $\kappa \prec \pi$ that every set in κ is a union of sets in π, and put $\pi I = \{J \in \pi;\, J \subset I\}$. Let ζ_r be the Cox process of ancestors to ξ_s at time $r = s - h$. Using the definition of $\tilde{\xi}_s$, the conditional independence of clusters η_u, and Theorem 13.16 (i), we get

$$\begin{aligned}
E_\mu \tilde{\xi}_s^{\otimes n}(\cdot \times D) &= \sum_{\pi \in \mathcal{P}'_n} E_\mu \int \zeta_r^{(\pi)}(du) \bigotimes_{J \in \pi} \eta_{h,u_J}^{\otimes J} \\
&= \sum_{\pi \in \mathcal{P}'_n} E_\mu \xi_s^{\otimes \pi} * \bigotimes_{J \in \pi} \nu_h^J \\
&= \sum_{\pi \in \mathcal{P}'_n} \sum_{\kappa \prec \pi} \bigotimes_{I \in \kappa} (\mu * \nu_s^{\pi I}) * \bigotimes_{J \in \pi I} \nu_h^J\,,
\end{aligned}$$

and similarly for the associated densities, where $\zeta_s^{(n)}$ denotes the factorial measure of ζ_s on $(\mathsf{R}^d)^{(n)}$. For each term on the right, we have $|\pi I| < |I|$ for at least one $I \in \kappa$, and then Lemma 13.31 yields a corresponding density, which tends to 0 as $h \to 0$, uniformly for $(x, r) \in (\mathsf{R}^d)^{(I)} \times (0, \infty)$ in compacts. The remaining factors have locally bounded densities on $(\mathsf{R}^d)^{(I)} \times (0,\infty)$, e.g. by Lemma 13.29. Hence, by combination, $E_\mu \tilde{\xi}_s^{\otimes n}(\cdot \times D)$ has a density that tends to 0 as $h \to 0$, uniformly for $(x, s) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$ in compacts.

Turning to the second term on the right, let $B = B_1 \times \cdots \times B_n \subset G^{(n)}$ be compact, and write $B_J = \prod_{i \in J} B_i$ for $J \subset \{1, \dots, n\}$. Using the previous notation, and defining $\nu_{h,\varepsilon}$ as in Lemma 13.32, we get for small enough $\varepsilon > 0$

$$\begin{aligned}
&E_\mu \iint_{B \times D^c} \tilde{\xi}_s^{\otimes n}(dx\,dv)\, 1\Big\{ \tilde{\xi}'_{t,v_1} G^c > 0 \Big\} \\
&\quad = E_\mu \int \zeta_r^{(n)}(du) \int_{B_1} \eta_h^{u_1}(dx_1)\, 1\{\eta_h^{u_1} G^c > 0\} \prod_{i > 1} \eta_h^{u_i} B_i
\end{aligned}$$

$$\begin{aligned}&\le E_\mu \xi_r^{\otimes n} * \Big(\nu_{h,\varepsilon} \otimes \nu_h^{\otimes(n-1)}\Big)B\\ &= \sum_{\pi\in\mathcal{P}_n} \Big\{\mu * \nu_r^{J_1} * \Big(\nu_{h,\varepsilon}\otimes \nu_h^{\otimes J_1'}\Big)B_{J_1}\Big\} \prod_{J\in\pi'} \Big(\mu * \nu_t^J * \nu_h^{\otimes J}\Big)B_J\,,\end{aligned}$$

where $1 \in J_1 \in \pi$, $J_1' = J_1\setminus\{1\}$, and $\pi' = \pi\setminus\{J_1\}$. For each term on the right, Lemma 13.32 yields a density of the first factor, which tends to 0 as $h \to 0$ for fixed $\varepsilon > 0$, uniformly for $(x_{J_1}, t) \in (\mathsf{R}^d)^{J_1} \times (0,\infty)$ in compacts. Since the remaining factors have locally bounded densities on the sets $(\mathsf{R}^d)^J \times (0,\infty)$, by Lemma 13.29, the entire sum has a density that tends to 0 as $h \to 0$ for fixed $\varepsilon > 0$, uniformly for $(x,t) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$ in compacts. Combining the previous estimates, and using Lemma 1.20, we obtain

$$\Big\|\,\mathcal{L}_\mu\Big(1_{G^c}\tilde\xi_t\,\Big\|\,\tilde\xi_s^{\otimes n}\Big)_{x,v} - \mathcal{L}_\mu\Big(1_{G^c}\tilde\xi_{t,v}\,\Big\|\,\tilde\xi_s^{\otimes n}\Big)_{x,v}\Big\| \le p_h \ \text{ a.e. } E_\mu\tilde\xi_s^{\otimes n}, \qquad (9)$$

for some measurable functions p_h on $(\mathsf{R}^d)^{(n)}$ with $p_h \to 0$ as $h \to 0$, uniformly on compacts.

We now apply Theorem 6.30 to the pair $(\tilde\xi_s, \tilde\xi_t)$, regarded as a Cox cluster process generated by ξ_r. Under $\mathcal{L}(\tilde\xi_t \| \,\tilde\xi_s^{\otimes n})_{x,v}$, the leading term agrees with the non-diagonal component $\tilde\xi_{t,v}$. To see this, we first condition on ξ_r, so that $\tilde\xi_t$ becomes a Poisson cluster process, generated by a non-random measure at time r. By Fubini's theorem, the leading term is a.e. restricted to $\mathsf{R}^d \times \{v_1,\dots,v_n\}^c$, whereas by Lemma 6.13, the remaining terms are a.e. restricted to $\mathsf{R}^d \times \{v_1,\dots,v_n\}$. Hence, Theorem 6.30 yields a.e.

$$\mathcal{L}_\mu\Big(\tilde\xi_{t,v}\,\Big\|\,\tilde\xi_s^{\otimes n}\Big)_{x,v} = E_\mu\Big\{\mathcal{L}_{\xi_r}(\tilde\xi_{t-r})\,\Big\|\,\tilde\xi_s^{\otimes n}\Big\}_{x,v}, \quad x \in (\mathsf{R}^d)^{(n)},\ v \in [0,1]^n,$$

and so by (9),

$$\Big\|\,\mathcal{L}_\mu\Big(1_{G^c}\tilde\xi_t\,\Big\|\,\tilde\xi_s^{\otimes n}\Big)_{x,v} - E_\mu\Big\{\mathcal{L}_{\xi_r}(1_{G^c}\tilde\xi_{t-r})\,\Big\|\,\tilde\xi_s^{\otimes n}\Big\}_{x,v}\Big\| \le p_h \ \text{ a.e. } E_\mu\tilde\xi_s^{\otimes n}.$$

The assertion now follows by Lemma 1.22. □

We may now establish the basic regularity theorem for the multi-variate Palm distributions of a DW-process. The resulting regular versions will be needed for the local approximations in Section 13.9.

Theorem 13.34 *(regular Palm kernels) The kernels $\mathcal{L}_\mu(\xi_t \| \,\xi_t^{\otimes n})$ have versions satisfying*

(i) *$\mathcal{L}_\mu(\xi_t \| \,\xi_t^{\otimes n})_x$ is tight and weakly continuous in $(x,t) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$,*

(ii) *for any $t > 0$ and open $G \subset \mathsf{R}^d$, the kernel $\mathcal{L}_\mu(1_{G^c}\xi_t \| \,\xi_t^{\otimes n})_x$ is continuous is total variation in $x \in G^{(n)}$,*

(iii) *for open $G \subset \mathsf{R}^d$ and bounded μ or G^c, the kernel $\mathcal{L}_\mu(1_{G^c}\xi_t \| \,\xi_t^{\otimes n})_x$ is continuous in total variation in $(x,t) \in G^{(n)} \times (0,\infty)$.*

Proof: (ii) For the conditional moment measure $E_\mu(\xi_t^{\otimes n}|\,\xi_r)$, with $r \geq 0$ fixed, Theorem 13.28 yields a Lebesgue density that is L^1-continuous in $(x,t) \in (\mathsf{R}^d)^{(n)} \times (r,\infty)$. Since the continuous density of $E_\mu \xi_t^{\otimes n}$ is even strictly positive, by Corollary 13.27, the L^1-continuity extends to the density of $E_\mu(\xi_t^{\otimes n}|\,\xi_r)$ with respect to $E_\mu \xi_t^{\otimes n}$. Hence, by Theorem 6.42, the Palm kernels $\mathcal{L}_\mu(\xi_r \| \,\xi_t^{\otimes n})_x$ have versions that are continuous in total variation in $(x,t) \in (\mathsf{R}^d)^{(n)} \times (r,\infty)$, for fixed $r \geq 0$. Fixing any $t > r \geq 0$ and $G \subset \mathsf{R}^d$, we see in particular that the kernel $E_\mu\{\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r}) \,\|\, \xi_t^{\otimes n}\}_x$ has a version that is continuous in total variation in $x \in (\mathsf{R}^d)^{(n)}$. Choosing any $r_1, r_2, \ldots \in (0,t)$ with $r_k \to t$, using Lemma 13.33 with $r = r_k$ and $s = t$, and invoking Lemma 1.21, we obtain a similar continuity property for the kernel $\mathcal{L}_\mu(1_{G^c}\xi_t \| \,\xi_t^{\otimes n})_x$.

(iii) Let μ and G^c be bounded, and fix any $r \geq 0$. As before, we may choose the kernels $\mathcal{L}_\mu(\xi_r \| \,\xi_t^{\otimes n})_x$ to be continuous in total variation in $(x,t) \in (\mathsf{R}^d)^{(n)} \times (r,\infty)$. For any $x, x' \in (\mathsf{R}^d)^{(n)}$ and $t, t' > r$, we may write

$$\begin{aligned}
&\Big\| E_\mu\Big\{\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r}) \,\Big\|\, \xi_t^{\otimes n}\Big\}_x - E_\mu\Big\{\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t'-r}) \,\Big\|\, \xi_{t'}^{\otimes n}\Big\}_{x'}\Big\| \\
&\quad \leq E_\mu\Big\{ \|\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r}) - \mathcal{L}_{\xi_r}(1_{G^c}\xi_{t'-r})\| \,\Big\|\, \xi_t^{\otimes n}\Big\}_x \\
&\qquad + \Big\| \mathcal{L}_\mu(\xi_r \| \,\xi_t^{\otimes n})_x - \mathcal{L}_\mu(\xi_r \| \,\xi_{t'}^{\otimes n})_{x'}\Big\|.
\end{aligned}$$

As $x' \to x$ and $t' \to t$ for fixed r, the first term on the right tends to 0 in total variation, by Lemma 13.15 and dominated convergence, whereas the second term tends to 0 in the same sense, by the continuous choice of kernels. This shows that $E_\mu[\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r}) \,\|\, \xi_t^{\otimes n}]_x$ is continuous in total variation in $(x,t) \in (\mathsf{R}^d)^{(n)} \times (r,\infty)$, for fixed r, μ, and G.

Now let $h_1, h_2, \ldots > 0$ be rationally independent with $h_n \to 0$, and define

$$r_k(t) = h_k[h_k^{-1}t-], \quad t > 0,\ k \in \mathsf{N}.$$

Then Lemma 13.33 applies, with $r = r_k(t)$ and $s = t$, for some functions p_k with $p_k \to 0$, uniformly for $(x,t) \in G^{(n)} \times (0,\infty)$ in compacts. Since the sets $U_k = h_k \bigcup_j (j-1, j)$ satisfy $\limsup_k U_k = (0,\infty)$, Lemma 1.21 yields a version of the kernel $\mathcal{L}_\mu(1_{G^c}\xi_t \| \,\xi_t^{\otimes n})_x$, which is continuous in total variation in $(x,t) \in G^{(n)} \times (0,\infty)$.

(i) Writing $U_x^r = \bigcup_i B_{x_i}^r$, and using Corollary 13.27 and Lemma 13.30, we get

$$\begin{aligned}
\frac{E_\mu\, \xi_t^{\otimes n} B_x^\varepsilon\,(\xi_t U_x^r \wedge 1)}{E_\mu \xi_t^{\otimes n} B_x^\varepsilon} &\lesssim r^d \sum_{i \leq n} \Big\{E_\mu \xi_t^{\otimes(n+1)} * \Big(p_{\varepsilon^2}^{\otimes n} \otimes p_{r^2}\Big)\Big\}(x, x_i) \\
&\lesssim \begin{cases} r^2, & d \geq 3, \\ r^2|\log r|, & d = 2, \end{cases}
\end{aligned}$$

uniformly for $(x,t) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$ in compacts. Now use part (iii), along with a uniform version of Lemma 6.45, for random measures ξ_t. $\square$

The spatial continuity of the Palm distributions, established in the preceding theorem, is easily extended to the forward version $\mathcal{L}_\mu(\xi_t \| \xi_s^{\otimes n})_x$ of the Palm kernel, for arbitrary $s < t$.

Corollary 13.35 *(forward Palm distributions) For fixed $t > s > 0$, the kernel $\mathcal{L}_\mu(\xi_t \| \xi_s^{\otimes n})_x$ has a version that is continuous in total variation in $x \in (\mathsf{R}^d)^{(n)}$.*

Proof: Let ζ_s denote the ancestral process of ξ_t at time $s = t - h$. Since $\xi_t \perp\!\!\!\perp_{\zeta_s} \xi_s^{\otimes n}$, it suffices by Lemma 6.9 to prove the continuity in total variation of $\mathcal{L}_\mu(\zeta_s \| \xi_s^{\otimes n})_x$. Since ζ_s is a Cox process directed by $h^{-1}\xi_s$, we see from Theorem 6.27 that $\mathcal{L}_\mu(\zeta_s \| \xi_s^{\otimes n})_x$ is a.e. the distribution of a Cox process directed by $\mathcal{L}_\mu(h^{-1}\xi_s \| \xi_s^{\otimes n})_x$. Hence, for any $G \subset \mathsf{R}^d$,

$$\begin{aligned}
&\Big\| \mathcal{L}_\mu(\zeta_s \| \xi_s^{\otimes n})_x - \mathcal{L}_\mu(1_{G^c}\zeta_s \| \xi_s^{\otimes n})_x \Big\| \\
&\quad \le P_\mu\Big(\zeta_s G > 0 \,\Big\|\, \xi_s^{\otimes n}\Big)_x \\
&\quad = E_\mu\Big(1 - e^{-h^{-1}\xi_s G} \,\Big\|\, \xi_s^{\otimes n}\Big)_x \\
&\quad \le E_\mu\Big(h^{-1}\xi_s G \wedge 1 \Big\| \xi_s^{\otimes n}\Big)_x.
\end{aligned}$$

By Theorem 13.34 (i), we may choose versions of the Palm distributions of ζ_s, such that $\mathcal{L}_\mu(\zeta_s \| \xi_s^{\otimes n})_x$ can be approximated in total variation by some kernels $\mathcal{L}_\mu(1_{G^c}\zeta_s \| \xi_s^{\otimes n})_x$ with open $G \subset \mathsf{R}^d$, uniformly for $x \in G^{(n)}$ in compacts. It is then enough to choose the latter kernels to be continuous in total variation on $G^{(n)}$. Since $1_{G^c}\zeta_s \perp\!\!\!\perp_{1_{G^c}\xi_s} 1_G\xi_s$, such versions exist by Lemma 6.9, given the corresponding property of $\mathcal{L}_\mu(1_{G^c}\xi_s \| \xi_s^{\otimes n})_x$ in Theorem 13.34 (ii). □

We proceed to establish a joint continuity property of the Palm distributions, under shifts in both space and time. It is also needed for the developments in Section 13.9.

Lemma 13.36 *(space-time continuity of Palm kernel) Fix any μ, $t > 0$, and open $G \subset \mathsf{R}^d$. Then as $s \uparrow t$ and $u \to x \in G^{(n)}$, we have*

$$\Big\| \mathcal{L}_\mu(\xi_t \,\|\, \xi_s^{\otimes n})_u - \mathcal{L}_\mu(\xi_t \,\|\, \xi_s^{\otimes n})_x \Big\| \to 0.$$

Proof: Letting $r < s \le t$, we may write

$$\begin{aligned}
&\Big\| \mathcal{L}_\mu(1_{G^c}\xi_t \,\|\, \xi_s^{\otimes n})_x - \mathcal{L}_\mu(1_{G^c}\xi_t \,\|\, \xi_t^{\otimes n})_x \Big\| \\
&\qquad \le \Big\| \mathcal{L}_\mu(1_{G^c}\xi_t \,\|\, \xi_s^{\otimes n})_x - E_\mu\Big\{\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r}) \,\Big\|\, \xi_s^{\otimes n}\Big\}_x \Big\| \\
&\qquad\quad + \Big\| \mathcal{L}_\mu(1_{G^c}\xi_t \,\|\, \xi_t^{\otimes n})_x - E_\mu\Big\{\mathcal{L}_{\xi_r}(1_{G^c}\xi_{t-r}) \,\Big\|\, \xi_t^{\otimes n}\Big\}_x \Big\| \\
&\qquad\quad + \Big\| \mathcal{L}_\mu(\xi_r \,\|\, \xi_s^{\otimes n})_x - \mathcal{L}_\mu(\xi_r \,\|\, \xi_t^{\otimes n})_x \Big\|. \qquad (10)
\end{aligned}$$

By Corollary 13.27 and Theorems 6.42 and 13.28, the kernels $\mathcal{L}_\mu(\xi_r \| \xi_s^{\otimes n})_x$ and $\mathcal{L}_\mu(\xi_r \| \xi_t^{\otimes n})_x$ have versions that are continuous in total variation in

$x \in (\mathsf{R}^d)^{(n)}$. With such versions, and for $2s > t + r$, Lemma 13.33 shows that the first two terms on the right of (10) are bounded by some functions p_{t-r}, where $p_h \downarrow 0$ as $h \to 0$, uniformly for $(x,t) \in G^{(n)} \times (0,\infty)$ in compacts. Next, we see from Theorems 6.42 and 13.28 that the last term in (10) tends to 0, as $s \to t$ for fixed r and t, uniformly for $(x,t) \in (\mathsf{R}^d)^{(n)} \times (0,\infty)$ in compacts. Letting $s \to t$ and then $r \to t$, we conclude that the left-hand side of (10) tends to 0 as $s \uparrow t$, uniformly for $x \in G^{(n)}$ in compacts. Since $\mathcal{L}_\mu(1_{G^c}\xi_t \| \xi_t^{\otimes n})_x$ is continuous in total variation in $x \in G^{(n)}$, by Theorem 13.34 (ii), the required joint convergence follows, as $s \uparrow t$ and $u \to x$. $\square$

We finally prove a continuity property of univariate Palm distributions, under joint spatial shifts. This result will be needed in Section 13.8.

Lemma 13.37 *(joint shift continuity) Let ξ be a DW-process in R^d with canonical cluster η. Then for fixed $t > 0$ and bounded μ or $B \in \mathcal{B}^d$, the shifted Palm distributions $\mathcal{L}_\mu(\,\theta_{-x}\xi_t \,\|\, \xi_t)_x$ and $\mathcal{L}_\mu(\,\theta_{-x}\eta_t \,\|\, \eta_t)_x$ are x-continuous in total variation on B.*

Proof: First let μ be bounded, write ζ_s for the ancestral process at time $s \le t$, and put $\tau = \inf\{s > 0;\, \|\zeta_s\| > \|\zeta_0\|\}$. Then ζ_0 is Poisson with intensity μ/t, and Theorem 13.8 shows that each point in ζ_0 splits before time $h \in (0,t)$ with probability h/t. Hence, the number of such branching particles is Poisson with mean $\|\mu\|h/t^2$, and so $P\{\tau > h\} = e^{-\|\mu\|h/t^2}$. Conditionally on $\tau > h$, the process ζ_h is again Poisson with intensity $t^{-1}(\mu * p_h)\cdot\lambda^d = E_\mu\xi_h/t$, and moreover $\xi_t \perp\!\!\!\perp_{\zeta_h} \{\tau > h\}$. Hence,

$$\begin{aligned}
&\Big\|\mathcal{L}_\mu(\xi_t\theta_r) - \mathcal{L}_\mu(\xi_t)\Big\| \\
&\qquad \le P_\mu\{\tau \le h\} + \Big\|\mathcal{L}_\mu(\zeta_h\theta_r \,|\, \tau > h) - \mathcal{L}_\mu(\zeta_h \,|\, \tau > h)\Big\| \\
&\qquad \le \Big(1 - e^{-\|\mu\|h/t^2}\Big) + t^{-1}\Big\|E_\mu\xi_h\theta_r - E_\mu\xi_h\Big\|,
\end{aligned}$$

which tends to 0, as $r \to 0$ and then $h \to 0$, by Theorem 13.10 (v).

For general μ, we may choose some bounded measures $\mu_n \uparrow \mu$, so that $\mu_n' = \mu - \mu_n \downarrow 0$. Fixing any $B \in \hat{\mathcal{B}}^d$, we have

$$\begin{aligned}
&\Big\|\mathcal{L}_\mu(\xi_t\theta_r) - \mathcal{L}_\mu(\xi_t)\Big\|_B \\
&\qquad \le \Big\|\mathcal{L}_{\mu_n}(\xi_t\theta_r) - \mathcal{L}_{\mu_n}(\xi_t)\Big\| + 2P_{\mu_n'}\Big\{\xi_t(B \cup \theta_r B) > 0\Big\},
\end{aligned}$$

which tends to 0, as $r \to 0$ and then $n \to \infty$, by the previous case and Lemma 13.12. This yields the continuity of $\mathcal{L}_\mu(\xi_t\theta_r)$.

By Lemma 6.16, the Palm distribution $\mathcal{L}^0_\mu(\xi_t)$ is the convolution of $\mathcal{L}_\mu(\xi_t)$ with the Palm distribution at 0, for the Lévy measure $\mathcal{L}_\mu(\eta_t) = \int \mu(dx)\,\mathcal{L}_x(\eta_t)$. By the previous result and Fubini's theorem, it remains to show that the latter factor is continuous in total variation under shifts in μ. By Corollary 13.22, the corresponding historical path is a Brownian bridge X on $[0,t]$, from α to 0, where $\mathcal{L}(\alpha) = (p_t \cdot \mu)/\mu p_t$. The measure η_t is a sum of independent clusters, rooted along X, whose birth times form an independent Poisson process ζ on $(0,t)$ with rate $g(s) = 2(t-s)^{-1}$.

Let τ be the first point of ζ. Since $P\{\tau \le h\} \to 0$ as $h \to 0$, and since the event $\tau > h$ is independent of the restriction of ζ to the interval $[h,t]$, it suffices to prove that, for fixed $h > 0$, the sum of clusters born after time h is continuous in total variation. Since X remains a Brownian bridge on $[h,t]$, conditionally on α and X_h, the mentioned sum is conditionally independent of α, given X_h, and it suffices to prove that $\mathcal{L}_\mu(X_h)$ is continuous in total variation under shifts in μ.

Then put $s = t - h$, and note that X_h is conditionally $N(s\alpha, sh)$, given $\alpha = X_0$. Thus, the conditional density of X_h equals $p_{sh}(x - s\alpha)$. By the distribution of α, the unconditional density of X_h equals

$$f_\mu(x) = (\mu p_t)^{-1} \int p_{sh}(x - su)\, p_t(u)\, \mu(du), \quad x \in \mathsf{R}^d.$$

Replacing μ by the shifted measure $\mu\theta_r$ leads to the density

$$f_{\mu\theta_r}(x) = \big\{(\mu * p_t)(r)\big\}^{-1} \int p_{sh}(x - su + sr)\, p_t(u - r)\, \mu(du),$$

and we need to show that $f_{\mu\theta_r} \to f_\mu$ in L^1, as $r \to 0$. Since $(\mu * p_t)(r) \to \mu p_t$, by Theorem 13.10 (iv), it suffices to prove convergence of the μ-integrals. Here the L^1-distance is bounded by

$$\int dx \int \mu(du)\, \Big| p_{sh}(x - su + sr)\, p_t(u - r) - p_{sh}(x - su)\, p_t(u) \Big|,$$

which tends to 0 as $r \to 0$, by FMP 1.32, since the integrand tends to 0 by continuity, and

$$\begin{aligned} \int dx \int \mu(du)\, p_{sh}(x - su + sr)\, p_t(u - r) & \\ &= (\mu * p_t)(r) \to \mu p_t \\ &= \int dx \int \mu(du)\, p_{sh}(x - su)\, p_t(u), \end{aligned}$$

by Fubini's theorem and Theorem 13.10 (iv). □

13.6 Hitting Rates and Approximation

Hitting estimates of various kinds, for a DW-process ξ_t and its associated clusters η_t, will play a basic role for our developments in Sections 13.7–9. All subsequent results in this section depend ultimately on some deep estimates of the hitting probabilities of a single ball B_x^ε, with center $x \in \mathsf{R}^d$ and radius $\varepsilon > 0$, stated here without proofs. The present versions of parts (ii)−(iii), quoted from K(08), are easy consequences of the original results. Statement (iv), regarding the range of a single cluster, is only needed occasionally. It implies in particular the remarkable fact that the range of η is a.s. finite.

Theorem 13.38 *(hitting balls, Dawson/Iscoe/Perkins, Le Gall) Let ξ be a DW-process in R^d with canonical cluster η. Then*

(i) *for $d \geq 3$, there exists a constant $c_d > 0$, such that as $\varepsilon \to 0$, uniformly in x and bounded μ and t^{-1},*

$$\varepsilon^{2-d} P_\mu\{\xi_t B_x^\varepsilon > 0\} \to c_d(\mu * p_t)(x),$$

(ii) *for $d \geq 3$, we have with $t(\varepsilon) = t + \varepsilon^2$, uniformly in μ, t, and $\varepsilon > 0$ with $\varepsilon^2 \leq t$,*

$$\mu p_t \lesssim \varepsilon^{2-d} P_\mu\{\eta_t B_0^\varepsilon > 0\} \lesssim \mu p_{t(\varepsilon)},$$

(iii) *for $d = 2$, there exist some functions $t(\varepsilon)$ and $l(\varepsilon)$ with*

$$t(\varepsilon) = t\, l(\varepsilon t^{-1/2}), \qquad 0 \leq l(\varepsilon) - 1 \lesssim |\log \varepsilon|^{-1/2},$$

such that, uniformly in μ and $t, \varepsilon > 0$ with $\varepsilon^2 < t/4$,

$$\mu p_t \lesssim \log(t\varepsilon^{-2})\, P_\mu\{\eta_t B_0^\varepsilon > 0\} \lesssim \mu p_{t(\varepsilon)},$$

(iv) $$P_{a\delta_0}\Big\{\sup_{s \leq t} \xi_s (B_0^r)^c > 0\Big\} \lesssim a r^d\, t^{-1-d/2}\, e^{-r^2/2t}, \quad t < r^2/4.$$

We proceed to establish some useful estimates for multiple hits, needed frequently in subsequent proofs. A multi-variate extension will be given in Lemma 13.44 below.

Lemma 13.39 *(multiple hits) For a DW-process ξ in R^d, let κ_h^ε be the number of h-clusters of ξ_t hitting B_0^ε at time t. Then*

(i) *for $d \geq 3$, we have with $t(\varepsilon) = t + \varepsilon^2$, as $\varepsilon^2 \ll h \leq t$,*

$$E_\mu \kappa_h^\varepsilon(\kappa_h^\varepsilon - 1) \lesssim \varepsilon^{2(d-2)} \Big\{h^{1-d/2}\, \mu p_t + (\mu p_{t(\varepsilon)})^2\Big\},$$

(ii) *for $d = 2$, there exist some functions $t_{h,\varepsilon} > 0$ with $t_{h,\varepsilon} \lesssim h|\log \varepsilon|^{-1/2}$, such that as $\varepsilon \ll h \leq t$,*

$$E_\mu \kappa_h^\varepsilon(\kappa_h^\varepsilon - 1) \lesssim |\log \varepsilon|^{-2}\Big\{\log(t/h)\, \mu p_t + (\mu p_{t(h,\varepsilon)})^2\Big\}.$$

Proof: (i) Let ζ_s be the Cox process of ancestors to ξ_t at time $s = t - h$, and write η_h^i for the associated h-clusters. Using Theorem 13.38 (ii), the conditional independence of the clusters, and the fact that $E_\mu \zeta_s^2 = h^{-2} E_\mu \xi_s^2$ outside the diagonal, we get with $p_h^\varepsilon(x) = P_x\{\eta_h B_0^\varepsilon > 0\}$

$$\begin{aligned} E_\mu \kappa_h^\varepsilon(\kappa_h^\varepsilon - 1) &= E_\mu \sum_{i \neq j} 1\Big\{\eta_h^i B_0^\varepsilon \wedge \eta_h^j B_0^\varepsilon > 0\Big\} \\ &= \iint_{x \neq y} p_h^\varepsilon(x)\, p_h^\varepsilon(y)\, E_\mu \zeta_s^2(dx\, dy) \\ &\lesssim \varepsilon^{2(d-2)} \iint p_{h(\varepsilon)}(x)\, p_{h(\varepsilon)}(y)\, E_\mu \zeta_s^2(dx\, dy). \end{aligned}$$

By Theorem 13.10, Fubini's theorem, and the semigroup property of the densities p_t, we get

$$\begin{aligned}\int p_{h(\varepsilon)}(x)\, E_\mu \xi_s(dx) &= \int p_{h(\varepsilon)}(x)\,(\mu * p_s)(x)\, dx\\ &= \int \mu(du)\,(p_{h(\varepsilon)} * p_s)(u) = \mu p_{t(\varepsilon)}.\end{aligned}$$

Next, we get by Theorem 13.16 (i), Fubini's theorem, the form of the densities p_t, and the relations $t \le t_\varepsilon \le 2t - s$

$$\begin{aligned}&\iint p_{h(\varepsilon)}(x)\, p_{h(\varepsilon)}(y)\, \mathrm{Cov}_\mu \xi_s(dx\, dy)\\ &\quad = 2\iint p_{h(\varepsilon)}(x)\, p_{h(\varepsilon)}(y)\, dx\, dy \int \mu(du) \int_0^s dr\\ &\qquad\qquad \times \int p_r(v-u)\, p_{s-r}(x-v)\, p_{s-r}(y-v)\, dv\\ &\quad = 2\int \mu(du) \int_0^s dr \int p_r(u-v)\, \big\{p_{t(\varepsilon)-r}(v)\big\}^2\, dv\\ &\quad \lesssim \int \mu(du) \int_0^s (t-r)^{-d/2} \big(p_r * p_{(t(\varepsilon)-r)/2}\big)(u)\, dr\\ &\quad = \int \mu(du) \int_0^s (t-r)^{-d/2}\, p_{(t(\varepsilon)+r)/2}(u)\, dr\\ &\quad \lesssim \int p_t(u)\, \mu(du) \int_h^t r^{-d/2}\, dr \lesssim \mu p_t\, h^{1-d/2}.\end{aligned}$$

The assertion follows by combination of those estimates.

(ii) Here we may proceed as before, with the following changes: Using Theorem 13.38 (iii), we see that the factor $\varepsilon^{2(d-2)}$ should be replaced by $|\log(\varepsilon/\sqrt{h})|^{-2} \lesssim |\log \varepsilon|^{-2}$. In the last computation, we now get $\int_h^t r^{-1} dr = \log(t/h)$. Since $h_\varepsilon = h\, l_{\varepsilon/\sqrt{h}}$ with $0 \le l_\varepsilon - 1 \lesssim |\log \varepsilon|^{-1/2}$, we may choose $t_{h,\varepsilon} = t + (h_\varepsilon - h)$ in the second term on the right. As for the estimates leading up to the first term, we note that the bound $t_{h,\varepsilon} + s \le 2t$ remains valid, for sufficiently small ε/h. □

Returning to the single hitting probabilities of a DW-cluster η in R^d, we define

$$p_h^\varepsilon(x) = P_x\{\eta_h B_0^\varepsilon > 0\}, \quad h, \varepsilon > 0,\ x \in \mathsf{R}^d, \tag{11}$$

so that for $d \ge 3$, we have by Theorem 13.38 (i)

$$\varepsilon^{2-d} p_t^\varepsilon(x) \to c_d\, p_t(x), \quad \varepsilon \to 0,$$

uniformly in $x \in \mathsf{R}^d$ and bounded t^{-1}. Part (iii) of the same result suggests that, when $d = 2$, we introduce the normalizing function

$$m(\varepsilon) = |\log \varepsilon|\, P_{\lambda^2}\{\eta_1 B_0^\varepsilon > 0\} = |\log \varepsilon|\, \lambda^2 p_1^\varepsilon, \quad \varepsilon > 0.$$

Though the possible convergence of $m(\varepsilon)$ seems elusive, we may establish some strong boundedness and continuity properties, sufficient for our needs:

Lemma 13.40 *(rate function) For a DW-process in* R^2*, the function* $t \mapsto \log m\{\exp(-e^t)\}$ *is bounded and uniformly continuous on* $[1, \infty)$*.*

Proof: The boundedness of $\log m$ is clear from Theorem 13.38 (ii). For any $h \in (0, 1]$, let ζ_s be the process of ancestors to ξ_1 at time $s = 1 - h$, and denote the generated h-clusters by η_h^i. Then for $0 < r \ll 1$ and $0 < \varepsilon \ll h$, we get the following chain of relations, subsequently explained and justified:

$$\begin{aligned} m(\varepsilon)\,|\log\varepsilon|^{-1} &\approx r^{-1} P_{r\lambda^2}\Big\{\xi_1 B_0^\varepsilon > 0\Big\} \\ &\approx r^{-1} E_{r\lambda^2} \textstyle\sum_i 1\Big\{\eta_h^i B_0^\varepsilon > 0\Big\} \\ &= r^{-1} E_{r\lambda^2} \zeta_s p_h^\varepsilon \\ &= h^{-1} P_{\lambda^2}\Big\{\eta_h B_0^\varepsilon > 0\Big\} \\ &= P_{\lambda^2}\Big\{\eta_1 B_0^{\varepsilon/\sqrt{h}} > 0\Big\} \\ &= m(\varepsilon/\sqrt{h})\,\Big|\log(\varepsilon/\sqrt{h})\Big|^{-1} \\ &\approx m(\varepsilon/\sqrt{h})\,|\log\varepsilon|^{-1}. \end{aligned}$$

Here the first two steps are suggested by Lemmas 13.9 and 13.39 (ii), respectively, the third step holds by the conditional independence of the clusters, the fourth step holds by the Cox property of ζ_s, the fifth step holds by Theorem 13.14 (ii), the sixth step holds by the definition of m, and the last step is suggested by the relation $\varepsilon \ll h$.

To estimate the approximation errors, we see from Lemma 13.9 and Theorem 13.38 (ii) that

$$\begin{aligned} &\Big| m(\varepsilon) - r^{-1}|\log\varepsilon|\, P_{r\lambda^2}\{\xi_1 B_0^\varepsilon > 0\}\Big| \\ &\quad = \; r^{-1}|\log\varepsilon|\, \Big| P_{r\lambda^2}\{\eta_1 B_0^\varepsilon > 0\} - P_{r\lambda^2}\{\xi_1 B_0^\varepsilon > 0\}\Big| \\ &\quad \lesssim \; r^{-1}|\log\varepsilon|\, (P_{r\lambda^2}\{\eta_1 B_0^\varepsilon > 0\})^2 \\ &\quad \lesssim \; r^{-1}|\log\varepsilon|^{-1}(r\lambda^2 p_{l(\varepsilon)})^2 = r|\log\varepsilon|^{-1}. \end{aligned}$$

Next, Lemma 13.39 (ii) yields

$$\begin{aligned} r^{-1}|\log\varepsilon|\,&\Big| E_{r\lambda^2} \textstyle\sum_i 1\{\eta_h^i B_0^\varepsilon > 0\} - P_{r\lambda^2}\{\xi_1 B_0^\varepsilon > 0\}\Big| \\ &= \; r^{-1}|\log\varepsilon|\, E_{r\lambda^2}(\kappa_h^\varepsilon - 1)_+ \\ &\lesssim \; \frac{|\log h|\, r\lambda^2 p_1 + (r\lambda^2 p_{t(h,\varepsilon)})^2}{r|\log\varepsilon|} \\ &= \; \frac{|\log h| + r}{|\log\varepsilon|}. \end{aligned}$$

Finally, we note that

$$m(\varepsilon/\sqrt{h})\left|\frac{|\log\varepsilon|}{|\log(\varepsilon/\sqrt{h})|} - 1\right| \lesssim \frac{|\log h|}{|\log\varepsilon|},$$

by the boundedness of m. Combining these estimates, and letting $r \to 0$, we obtain

$$\left| m(\varepsilon) - m(\varepsilon/\sqrt{h}) \right| \lesssim \frac{|\log h|}{|\log \varepsilon|}.$$

Taking $\varepsilon = e^{-u}$ and $\varepsilon/\sqrt{h} = e^{-v}$, with $|u - v| \ll u$, gives

$$\begin{aligned} \left| \log \frac{m(e^{-u})}{m(e^{-v})} \right| &\lesssim \left| \frac{m(e^{-u})}{m(e^{-v})} - 1 \right| \\ &\lesssim |m(e^{-u}) - m(e^{-v})| \\ &\lesssim \frac{|u-v|}{u} \lesssim |\log(u/v)|, \end{aligned}$$

which extends immediately to arbitrary $u, v \geq 1$. Substituting $u = e^s$ and $v = e^t$ yields

$$\left| \log m\left\{\exp(-e^t)\right\} - \log m\left\{\exp(-e^s)\right\} \right| \lesssim |t - s|,$$

which implies the asserted uniform continuity. □

We proceed to approximate the hitting probabilities p_t^ε in (11) by suitably normalized Dirac functions. The result will play a crucial role below and in the next section.

Lemma 13.41 *(Dirac approximation) For a DW-cluster in* R^d*, we have*

(i) *when* $d \geq 3$*, and as* $\varepsilon^2 \ll h \to 0$*,*

$$\left\| \varepsilon^{2-d} (p_h^\varepsilon * f) - c_d f \right\| \to 0,$$

(ii) *when* $d = 2$*, and as* $\varepsilon \leq h \to 0$ *with* $|\log h| \ll |\log \varepsilon|$*,*

$$\left\| |\log \varepsilon| \, (p_h^\varepsilon * f) - m(\varepsilon) f \right\| \to 0,$$

uniformly over any class of uniformly bounded and equi-continuous functions $f \geq 0$ *on* R^d*.*

Proof: (i) Using Lemma 13.9 and Theorems 13.14 (ii) and 13.38 (i)−(ii), we get by dominated convergence

$$\begin{aligned} \lambda^d p_h^\varepsilon &= h^{d/2} \, \lambda^d p_1^{\varepsilon/\sqrt{h}} \\ &\sim c_d \, h^{d/2} (\varepsilon/\sqrt{h})^{d-2} \, \lambda^d p_1 \\ &= c_d \, \varepsilon^{d-2} \, h. \end{aligned} \tag{12}$$

Similarly, Theorem 13.38 (iii) yields for fixed $r > 0$, in terms of a standard normal random vector γ in R^d,

$$\begin{aligned} \varepsilon^{2-d} \, h^{-1} \int_{|x|>r} p_h^\varepsilon(x) \, dx &\lesssim \int_{|u|>r/\sqrt{h}} p_{l(\varepsilon)}(u) \, du \\ &= P\left\{ |\gamma| \, l_\varepsilon^{1/2} > r/\sqrt{h} \right\} \to 0. \end{aligned} \tag{13}$$

By (12), it is enough to show that $\|\hat{p}_h^\varepsilon * f - f\| \to 0$ as $h, \varepsilon^2/h \to 0$, where $\hat{p}_h^\varepsilon = p_h^\varepsilon/\lambda^d p_h^\varepsilon$. Writing w_f for the modulus of continuity of f, we get

$$\begin{aligned}\left\| \hat{p}_h^\varepsilon * f - f \right\| &= \sup_x \left| \int \hat{p}_h^\varepsilon(u) \left\{ f(x-u) - f(x) \right\} du \right| \\ &= \int \hat{p}_h^\varepsilon(u)\, w_f(|u|)\, du \\ &= w_f(r) + 2\|f\| \int_{|u|>r} \hat{p}_h^\varepsilon(u)\, du,\end{aligned}$$

which tends to 0, as $h, \varepsilon^2/h \to 0$ and then $r \to 0$, by (13) and the uniform continuity of f.

(ii) By Theorem 13.14 (ii) and Lemma 13.40, we have

$$\begin{aligned}\lambda^2 p_h^\varepsilon &= h\,\lambda^2 p_1^{\varepsilon/\sqrt{h}} \\ &= h\, m(\varepsilon/\sqrt{h}) \left| \log(\varepsilon/\sqrt{h}) \right|^{-1} \\ &\sim h\, m(\varepsilon)\, |\log \varepsilon|^{-1}.\end{aligned}$$

We also see that, with t_ε as in Theorem 13.38 (iii),

$$h^{-1}\, |\log \varepsilon| \int_{|x|>r} p_h^\varepsilon(x)\, dx \lesssim \int_{|u|>r/\sqrt{h}} p_{t(\varepsilon)}(u)\, du \to 0.$$

The proof may now be completed as in case of (i). The last assertion is clear from the estimates in the preceding proofs. □

We may now establish the basic approximation property, for suitably normalized hitting probabilities. The result will play a fundamental role in subsequent sections.

Theorem 13.42 *(uniform approximation) Let ξ be a DW-process in R^d. Then for fixed $t > 0$ and bounded μ or B, we have as $\varepsilon \to 0$*

(i) $\displaystyle\sup_{x\in B} \left| \varepsilon^{2-d} P_\mu\{\, \xi_t B_x^\varepsilon > 0\} - c_d\, (\mu * p_t)(x) \right| \to 0, \qquad d \ge 3,$

(ii) $\displaystyle\sup_{x\in B} \Big| |\log \varepsilon|\, P_\mu\{\, \xi_t B_x^\varepsilon > 0\} - m_\varepsilon (\mu * p_t)(x) \Big| \to 0, \quad d = 2,$

and similarly for the clusters η_t.

Proof: (i) For bounded μ, this is just the uniform version of Theorem 13.38 (i). In general, we may write $\mu = \mu' + \mu''$ for bounded μ', and let $\xi = \xi' + \xi''$ be the corresponding decomposition of ξ. Then

$$\begin{aligned}P_\mu\{\xi_t B_x^\varepsilon > 0\} &\le P_\mu\{\xi_t' B_x^\varepsilon > 0\} + P_\mu\{\xi_t'' B_x^\varepsilon > 0\} \\ &= P_{\mu'}\{\xi_t B_x^\varepsilon > 0\} + P_{\mu''}\{\xi_t B_x^\varepsilon > 0\},\end{aligned}$$

and so by Lemma 13.9 and Theorem 13.38 (ii),

$$\begin{aligned}\Big|P_\mu\{\xi_t B_x^\varepsilon>0\}-P_{\mu'}\{\xi_t B_x^\varepsilon>0\}\Big| &\le P_{\mu''}\{\xi_t B_x^\varepsilon>0\}\\ &\lesssim t\,\varepsilon^{d-2}(\mu''*p_{t(\varepsilon)})(x).\end{aligned}$$

For any $r>0$ and small enough $\varepsilon_0>0$, Lemma A3.3 yields a $t'>0$, such that

$$p_{t(\varepsilon)}(u-x)\lesssim p_{t'}(u),\quad |x|\le r,\ \varepsilon<\varepsilon_0,\ u\in\mathsf{R}^d,$$

which implies $(\mu''*p_{t(\varepsilon)})(x)\le\mu''p_{t'}$, for the same x and ε. Hence,

$$\begin{aligned}&\Big\|\,\varepsilon^{2-d}P_\mu\Big\{\xi_t B^\varepsilon_{(\cdot)}>0\Big\}-c_d(\mu*p_t)\,\Big\|_{B_0^r}\\ &\quad\lesssim\Big\|\,\varepsilon^{2-d}P_{\mu'}\Big\{\xi_t B^\varepsilon_{(\cdot)}>0\Big\}-c_d(\mu'*p_t)\,\Big\|+\mu''p_{t'}\,,\end{aligned}$$

which tends to 0 as $\varepsilon\to0$ and then $\mu'\uparrow\mu$, by the result for bounded μ, and dominated convergence.

(ii) First let μ be bounded. Let $\varepsilon,h\to0$ with $|\log h|\ll|\log\varepsilon|$, and write ζ_s for the ancestral process at time $s=t-h$. Then we have, uniformly on R^2,

$$\begin{aligned}P_\mu\Big\{\xi_t B^\varepsilon_{(\cdot)}>0\Big\}&\approx E_\mu(\zeta_s*p_h^\varepsilon)\\ &=h^{-1}E_\mu(\xi_s*p_h^\varepsilon)\\ &=h^{-1}(\mu*p_s*p_h^\varepsilon)\\ &\approx m(\varepsilon)\,|\log\varepsilon|^{-1}\,(\mu*p_s)\\ &\approx m(\varepsilon)\,|\log\varepsilon|^{-1}\,(\mu*p_t).\end{aligned}$$

To justify the first approximation, we conclude from Lemma 13.39 (ii) that

$$\begin{aligned}&|\log\varepsilon|\,\Big\|P_\mu\{\xi_t B_x^\varepsilon>0\}-E_\mu(\zeta_s*p_h^\varepsilon)\Big\|\\ &\quad\lesssim\frac{|\log h|\,\|\mu*p_t\|+\|\mu*p_{t(h,\varepsilon)}\|^2}{|\log\varepsilon|}\lesssim\frac{|\log h|}{|\log\varepsilon|}\to0.\end{aligned}$$

For the second approximation, Lemma 13.41 (ii) yields

$$\begin{aligned}&\Big\|\,h^{-1}\,|\log\varepsilon|\,(\mu*p_s*p_h^\varepsilon)-m(\varepsilon)\,(\mu*p_s)\Big\|\\ &\quad\le\|\mu\|\,\Big\|\,h^{-1}\,|\log\varepsilon|\,(p_s*\tilde p_h^\varepsilon)-m(\varepsilon)\,p_s\Big\|\to0,\end{aligned}$$

since the functions $p_s=p_{t-h}$ are uniformly bounded and equi-continuous, for small enough $h>0$. The third approximation holds, since m is bounded, and

$$\Big\|\mu*p_s-\mu*p_t\Big\|\le\|\mu\|\,\|p_s-p_t\|\to0.$$

This completes the proof for bounded μ. The result extends to the general case, by the same argument as for (i).

To prove the indicated version of (i) for the clusters η_t, we see from Lemma 13.9 and Theorem 13.38 (ii) that, for the canonical cluster η,

$$\begin{aligned}\varepsilon^{2-d}\Big|P_\mu\{\eta_t B_x^\varepsilon>0\}-P_\mu\{\xi_t B_x^\varepsilon>0\}\Big| &\lesssim \varepsilon^{2-d}\Big(P_\mu\{\eta_t B_x^\varepsilon>0\}\Big)^2\\ &\lesssim \varepsilon^{d-2}\Big\{(\mu*p_{t(\varepsilon)})(x)\Big\}^2.\end{aligned}$$

For bounded μ, this clearly tends to 0 as $\varepsilon\to 0$, uniformly in x. In general, we see from Theorem 13.10 (iv) and Lemma A3.3 that the right-hand side tends to 0, uniformly for bounded x, which proves the cluster version of (i). The proof in case of (ii) is similar. □

The hitting estimates and approximations, derived so far, are sufficient for our needs in Sections 13.7–8. We turn to some multi-variate estimates, which will only be required in Section 13.9. For fixed $t>h>0$, we write η_h^i for the clusters in ξ_t of age h, and let ζ_s denote the ancestral process of ξ_t at time $s=t-h$. When $\zeta_s=\sum_i\delta_{u_i}$ is simple, we may also write $\eta_h^{u_i}$ for the h-cluster rooted at u_i. Let $(\mathsf{N}^n)'$ denote the set $\mathsf{N}^n\setminus\mathsf{N}^{(n)}$ of sequences $k_1,\dots,k_n\in\mathsf{N}$ with at least one repetition. Define $h(\varepsilon)$ as in Theorem 13.38, though with t replaced by h.

First, we estimate the probability that some h-cluster in ξ_t will hit several balls $B_{x_j}^\varepsilon$, $j\le n$. We display the required conditions, needed repeatedly below:

$$\begin{cases}\varepsilon^2\ll h\le\varepsilon, & d\ge 3,\\ h\le|\log\varepsilon|^{-1}\ll|\log h|^{-1}, & d=2.\end{cases}\tag{14}$$

Lemma 13.43 *(clusters hitting several balls) Let ξ be a DW-process in R^d, and fix any μ, $t>0$, and $x\in(\mathsf{R}^d)^{(n)}$. Then as $\varepsilon,h\to 0$ subject to (14),*

$$P_\mu\bigcup_{k\in(\mathsf{N}^n)'}\bigcap_{j\le n}\Big\{\eta_h^{k_j}B_{x_j}^\varepsilon>0\Big\}\ll\begin{cases}\varepsilon^{n(d-2)}, & d\ge 3,\\ |\log\varepsilon|^{-n}, & d=2.\end{cases}$$

Proof: We need to show that, for any $i\ne j$ in $\{1,\dots,n\}$,

$$P_\mu\bigcup_{k\in\mathsf{N}}\Big\{\eta_h^kB_{x_i}^\varepsilon\wedge\eta_h^kB_{x_j}^\varepsilon>0\Big\}\ll\begin{cases}\varepsilon^{n(d-2)}, & d\ge 3,\\ |\log\varepsilon|^{-n}, & d=2.\end{cases}$$

Writing $\bar x=(x_i+x_j)/2$ and $\Delta x=|x_i-x_j|$, and using Cauchy's inequality, Theorems 13.10 and 13.38 (ii), and the parallelogram identity, we get for $d\ge 3$

$$\begin{aligned}&P_\mu\bigcup_{k\in\mathsf{N}}\Big\{\eta_h^kB_{x_i}^\varepsilon\wedge\eta_h^kB_{x_j}^\varepsilon>0\Big\}\\ &\quad\le E_\mu\sum_{k\in\mathsf{N}}1\Big\{\eta_h^kB_{x_i}^\varepsilon\wedge\eta_h^kB_{x_j}^\varepsilon>0\Big\}\\ &\quad= E_\mu\int\zeta_s(du)\,1\Big\{\eta_h^uB_{x_i}^\varepsilon\wedge\eta_h^uB_{x_j}^\varepsilon>0\Big\}\\ &\quad=\int E_\mu\zeta_s(du)\,P_u\Big\{\eta_hB_{x_i}^\varepsilon\wedge\eta_hB_{x_j}^\varepsilon>0\Big\}\end{aligned}$$

$$\begin{aligned}
&\leq \int (\mu * p_s)_u \, du \left(P_u\{\eta_h B^\varepsilon_{x_i} > 0\} \, P_u\{\eta_h B^\varepsilon_{x_j} > 0\} \right)^{1/2} \\
&\lesssim \varepsilon^{d-2} \int (\mu * p_s)_u \, du \left\{ p_{h_\varepsilon}(x_i - u) \, p_{h_\varepsilon}(x_j - u) \right\}^{1/2} \\
&= \varepsilon^{d-2} e^{-|\Delta x|^2/8h'} \int (\mu * p_s)_u \, p_{h_\varepsilon}(\bar{x} - u) \, du \\
&\lesssim \varepsilon^{d-2} e^{-|\Delta x|^2/8h'} (\mu * p_{2t})(\bar{x}),
\end{aligned}$$

which tends to 0, faster than any power of ε. If instead $d = 2$, we get from Theorem 13.38 (iii) the bound

$$\begin{aligned}
&\left\{ \log(h/\varepsilon^2) \right\}^{-1} e^{-|\Delta x|^2/8h'} (\mu * p_{2t})(\bar{x}) \\
&\qquad \lesssim |\log \varepsilon|^{-1} e^{-|\Delta x|^2/8h'} (\mu * p_{2t})(\bar{x}),
\end{aligned}$$

which tends to 0, faster than any power of $|\log \varepsilon|^{-1}$. □

We turn to the case where a single ball $B^\varepsilon_{x_j}$ is hit by several h-clusters in ξ_t. The result extends the univariate estimates in Lemma 13.39.

Lemma 13.44 *(balls hit by several clusters) Let ξ be a DW-process in R^d, and fix any μ, $t > 0$, and $x \in (\mathsf{R}^d)^{(n)}$. Then as $\varepsilon, h \to 0$ subject to* (14),

$$E_\mu \Big(\sum\nolimits_{k \in \mathsf{N}^{(n)}} \prod\nolimits_{i \leq n} 1\Big\{ \eta_h^{k_i} B^\varepsilon_{x_i} > 0 \Big\} - 1 \Big)_+ \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

To compare with Lemma 13.39, we note that, for $n = 1$, the stated estimates reduce to

$$E_\mu(\kappa_h^\varepsilon - 1)_+ \ll \begin{cases} \varepsilon^{d-2}, & d \geq 3, \\ |\log \varepsilon|^{-1}, & d = 2. \end{cases}$$

Proof: On the sets

$$A_{h,\varepsilon} = \bigcap_{j \leq n} \Big(\sum\nolimits_{k \in \mathsf{N}} 1\Big\{ \eta_h^k B^\varepsilon_{x_j} > 0 \Big\} \leq n \Big), \quad h, \varepsilon > 0,$$

we have

$$\begin{aligned}
&\Big(\sum\nolimits_{k \in \mathsf{N}^{(n)}} \prod\nolimits_{i \leq n} 1\Big\{ \eta_h^{k_i} B^\varepsilon_{x_i} > 0 \Big\} - 1 \Big)_+ \\
&\qquad \leq \sum_{k \in \mathsf{N}^n} \prod_{j \leq n} 1\Big\{ \eta_h^{k_j} B^\varepsilon_{x_j} > 0 \Big\} \\
&\qquad = \prod_{j \leq n} \sum_{k \in \mathsf{N}} 1\Big\{ \eta_h^k B^\varepsilon_{x_j} > 0 \Big\} \leq n^n.
\end{aligned}$$

On $A^c_{h,\varepsilon}$, we note that $\prod_{j \leq n} \eta_h^{k_j} B^\varepsilon_{x_j} > 0$ implies $\eta_h^l B^\varepsilon_{x_i} > 0$, for some $i \leq n$ and $l \neq k_1, \ldots, k_n$, and so

$$\Big(\sum\nolimits_{k \in \mathsf{N}^{(n)}} \prod\nolimits_{i \leq n} 1\Big\{ \eta_h^{k_i} B^\varepsilon_{x_i} > 0 \Big\} - 1 \Big)_+$$

$$\begin{aligned}
&\le \sum_{k\in\mathsf{N}^{(n)}} \prod_{j\le n} 1\Big\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\Big\}\\
&\le \sum_{i\le n} \sum_{(k,l)\in\mathsf{N}^{(n+1)}} 1\Big\{\eta_h^{l} B_{x_i}^\varepsilon > 0\Big\} \prod_{j\le n} 1\Big\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\Big\}\\
&= \sum_{i\le n} \iint \zeta_s^{(n+1)}(du\,dv)\, 1\Big\{\eta_h^{v} B_{x_i}^\varepsilon > 0\Big\} \prod_{j\le n} 1\Big\{\eta_h^{u_j} B_{x_j}^\varepsilon > 0\Big\},
\end{aligned}$$

where $u = (u_1, \ldots, u_n) \in \mathsf{R}^{nd}$ and $v \in \mathsf{R}^d$. Finally, writing $U_{h,\varepsilon}$ for the union in Lemma 13.43, and using Lemma A4.1, we get on $U_{h,\varepsilon}^c$

$$\begin{aligned}
&\Big(\textstyle\sum_{k\in\mathsf{N}^{(n)}} \prod_{i\le n} 1\Big\{\eta_h^{k_i} B_{x_i}^\varepsilon > 0\Big\} - 1\Big)_+\\
&\quad = \Big(\textstyle\prod_{i\le n} \sum_{k\in\mathsf{N}} 1\Big\{\eta_h^{k} B_{x_i}^\varepsilon > 0\Big\} - 1\Big)_+\\
&\quad \le \sum_{i\le n} \Big(\textstyle\sum_{l\in\mathsf{N}} 1\Big\{\eta_h^{l} B_{x_i}^\varepsilon > 0\Big\} - 1\Big) \prod_{j\le n} \sum_{k\in\mathsf{N}} 1\Big\{\eta_h^{k} B_{x_j}^\varepsilon > 0\Big\}\\
&\quad = \sum_{i\le n} \sum_{(k,l)\in\mathsf{N}^{(n+1)}} 1\Big\{\eta_h^{l} B_{x_i}^\varepsilon > 0\Big\} \prod_{j\le n} 1\Big\{\eta_h^{k_j} B_{x_j}^\varepsilon > 0\Big\},
\end{aligned}$$

which agrees with the preceding bound.

Now let $q_{\mu,s}^m$ denote the continuous density of $E_\mu \xi_s^{\otimes m}$ in Corollary 13.27. Since

$$\Omega = (U_{h,\varepsilon} \cap A_{h,\varepsilon}) \cup U_{h,\varepsilon}^c \cup A_{h,\varepsilon}^c,$$

we may combine the previous estimates, and use Lemmas 13.29 and 13.38 (ii), to get for $d \ge 3$

$$\begin{aligned}
&E_\mu\Big(\textstyle\sum_{k\in\mathsf{N}^{(n)}} \prod_{i\le n} 1\Big\{\eta_h^{k_i} B_{x_i}^\varepsilon > 0\Big\} - 1\Big)_+ - n^n P_\mu U_{h,\varepsilon}\\
&\quad \le \sum_{i\le n} E_\mu \iint \zeta_s^{(n+1)}(du\,dv)\, 1\Big\{\eta_h^{v} B_{x_i}^\varepsilon > 0\Big\} \prod_{j\le n} 1\Big\{\eta_h^{u_j} B_{x_j}^\varepsilon > 0\Big\}\\
&\quad = \sum_{i\le n} \iint E_\mu \xi_s^{\otimes(n+1)}(du\,dv)\, P_v\Big\{\eta_h B_{x_i}^\varepsilon > 0\Big\} \prod_{j\le n} P_{u_j}\Big\{\eta_h B_{x_j}^\varepsilon > 0\Big\}\\
&\quad \lesssim \varepsilon^{(n+1)(d-2)} \sum_{i\le n} \iint q_{\mu,s}^{n+1}(u,v)\, p_{h_\varepsilon}(x_i - v)\, du\, dv \prod_{j\le n} p_{h_\varepsilon}(x_j - u_j)\\
&\quad = \varepsilon^{(n+1)(d-2)} \sum_{i\le n} \Big(q_{\mu,s}^{n+1} * p_{h_\varepsilon}^{\otimes(n+1)}\Big)(x, x_i)\\
&\quad \lesssim \varepsilon^{(n+1)(d-2)}\, h^{1-d/2} \ll \varepsilon^{n(d-2)}.
\end{aligned}$$

The term $n^n P_\mu U_{h,\varepsilon}$ on the left is of the required order, by Lemma 13.43. For $d = 2$, a similar argument based on Theorem 13.38 (iii) yields the bound

$$|\log \varepsilon|^{-n-1} |\log h| \ll |\log \varepsilon|^{-n}.$$

□

Next, we estimate the contribution to the balls $B_{x_j}^\varepsilon$ from distantly rooted h-clusters of ξ_t.

Lemma 13.45 *(remote clusters 1) Fix any $t, r > 0$, $x \in (\mathsf{R}^d)^{(n)}$, and μ, and put $B_x = \prod_j B_{x_j}^r$. Then as $\varepsilon, h \to 0$, with $\varepsilon^2 \leq h$ for $d \geq 3$ and $\varepsilon \leq h$ for $d = 2$, we have*

$$E_\mu \int_{B_x^c} \zeta_s^{(n)}(du) \prod_{j \leq n} 1\Big\{\eta_h^{u_j} B_{x_j}^\varepsilon > 0\Big\} \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

Proof: Let $q_{\mu,s}^n$ denote the jointly continuous density of $E_\mu \xi_s^{\otimes n}$ in Corollary 13.27. For $d \geq 3$, we may use the conditional independence and Theorem 13.38 (ii), to write the ratio between the two sides as

$$\begin{aligned} \varepsilon^{n(2-d)} \int_{B_x^c} E_\mu \xi_s^{\otimes n}(du) \prod_{j \leq n} P_{u_j}\Big\{\eta_h B_{x_j}^\varepsilon > 0\Big\} \\ \lesssim \int_{B_x^c} q_{\mu,s}^n(u)\, p_{h_\varepsilon}^{\otimes n}(x-u)\, du \\ = \Big(q_{\mu,s}^n * p_{h_\varepsilon}^{\otimes n}\Big)(x) - \int_{B_0} q_{\mu,s}^n(x-u)\, p_{h_\varepsilon}^{\otimes n}(u)\, du, \end{aligned}$$

where $h_\varepsilon = h + \varepsilon^2$. Here the first term on the right tends to $q_{\mu,t}^n(x)$ as $h_\varepsilon \to 0$, by Lemma 13.29. The same limit is obtained for the second term, by the joint continuity in Corollary 13.27 and elementary estimates, and so the difference tends to 0. For $d = 2$, the same argument yields a similar bound, though with $\varepsilon^{n(d-2)}$ replaced by $|\log(\varepsilon^2/h)|^{-n}$. Since $0 < \varepsilon \leq h \to 0$, we have $|\log(\varepsilon^2/h)| \geq |\log \varepsilon|$, and the assertion follows. □

We finally estimate the probability that, for any open set $G \subset \mathsf{R}^d$ with a compact subset B, the complement G^c is hit by some h-cluster of ξ_t rooted in B. For our needs in Section 13.9, it suffices to give a bound that tends to 0 as $h \to 0$, faster than any power of h.

Lemma 13.46 *(remote clusters 2) Fix any $t = s + h > 0$ and $B \subset G \subset \mathsf{R}^d$, with G open and B compact. Let r denote the minimum distance between B and G^c, and write ζ_s for the ancestral process of ξ_t at time s. Then*

$$P_\mu\Big\{\int_B \zeta_s(du)\, \eta_h^u G^c > 0\Big\} \lesssim r^d\, h^{-1-d/2}\, e^{-r^2/2h}\, (\mu * \nu_t) B.$$

Proof: Letting $h < r^2/4$, and using Theorem 13.38 (iv), we get

$$\begin{aligned} P_\mu\Big\{\int_B \zeta_s(du)\, \eta_h^u G^c > 0\Big\} &\leq P_\mu\Big\{\int_B \zeta_s(du)\, \eta_h^u (B_u^r)^c > 0\Big\} \\ &= E_\mu P_{(\xi_s B)\delta_0}\Big\{\xi_h (B_0^r)^c > 0\Big\} \\ &\lesssim E_\mu \xi_s B\, r^{-2} (r^2/h)^{1+d/2}\, e^{-r^2/2h} \\ &\lesssim (\mu * \nu_t) B\, r^d h^{-1-d/2}\, e^{-r^2/2h}. \end{aligned}$$

Here the first relation holds by the definition of r, the second one holds by conditional independence and shift invariance, and the last one holds since $E_\mu \xi_s = \mu * \nu_s$ and $p_s \lesssim p_t$ when $s \in [t/2, t]$. □

13.7 Lebesgue Approximation

Here we establish the basic *Lebesgue approximation* of a DW-process ξ in R^d with $d \geq 2$. The idea is to approximate ξ_t, for a fixed $t > 0$, by a suitably normalized restriction of d-dimensional Lebesgue measure λ^d. More precisely, given any measure μ on R^d and constant $\varepsilon > 0$, we define the *neighborhood measure* μ^ε as the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp}\mu$, so that μ^ε has λ^d-density $1\{\mu B_x^\varepsilon > 0\}$. Then ξ_t can be approximated, a.s. and in L^1, by a suitable normalization of ξ_t^ε, for any initial measure μ and time $t > 0$.

Recall the definitions of the constants c_d and function $m(\varepsilon)$, as in Theorem 13.42. For convenience, we will often write $m_\varepsilon^{-1} = 1/m_\varepsilon$. Also recall from (11) the definition of $p_h^\varepsilon(x)$.

Theorem 13.47 *(Lebesgue approximation) Let ξ be a DW-process in R^d. Then under P_μ for fixed $t > 0$ we have, a.s. and in L^1 as $\varepsilon \to 0$,*

(i) $c_d^{-1}\,\varepsilon^{2-d}\,\xi_t^\varepsilon \xrightarrow{v} \xi_t, \qquad d \geq 3,$

(ii) $m_\varepsilon^{-1}|\log\varepsilon|\,\xi_t^\varepsilon \xrightarrow{v} \xi_t, \quad d = 2.$

This remains true in the weak sense when μ is bounded, and the weak versions hold even for the clusters η_t when $\|\mu\| = 1$.

For the intensity measures, we have even convergence in total variation.

Corollary 13.48 *(mean convergence) Let ξ be a DW-process in R^d. Then for fixed $t > 0$ and bounded μ,*

(i) $\left\| \varepsilon^{2-d}E_\mu\xi_t^\varepsilon - c_d\,E_\mu\xi_t \right\| \to 0, \qquad d \geq 3,$

(ii) $\left\| |\log\varepsilon|\,E_\mu\xi_t^\varepsilon - m(\varepsilon)\,E_\mu\xi_t \right\| \to 0, \quad d = 2.$

This remains true for the clusters η_t, and for ξ_t it holds even locally, when ξ is locally finite under P_μ.

Several lemmas will be needed for the proofs. We begin with some approximations of the first and second moments of a single cluster.

Lemma 13.49 *(mean approximation) Let η_1 be the unit cluster of a DW-process in R^d. Then as $\varepsilon \to 0$,*

(i) $\left\| \varepsilon^{2-d}E\eta_1^\varepsilon - c_d\,p_1 \cdot \lambda^d \right\| \to 0, \qquad d \geq 3,$

(ii) $\left\| |\log\varepsilon|\,E\eta_1^\varepsilon - m(\varepsilon)\,p_1 \cdot \lambda^2 \right\| \to 0, \qquad d = 2,$

(iii) $$E\|\eta_1^\varepsilon\|^2 \asymp (E\|\eta_1^\varepsilon\|)^2 \asymp \begin{cases} \varepsilon^{2(d-2)}, & d \geq 3, \\ |\log\varepsilon|^{-2}, & d = 2. \end{cases}$$

Proof: (i) Fubini's theorem yields $E_0\eta_1^\varepsilon = p_1^\varepsilon \cdot \lambda^d$, and so for $d \geq 3$,

$$\left\| \varepsilon^{2-d} E_0 \eta_1^\varepsilon - c_d(p_1 \cdot \lambda^d) \right\| = \lambda^d \left| \varepsilon^{2-d} p_1^\varepsilon - c_d\, p_1 \right|. \tag{15}$$

Here the integrand on the right tends to 0 as $\varepsilon \to 0$, by Theorem 13.42 (i), and by Theorem 13.38 (ii) it is bounded by $C_d p_{1'} + c_d p_1 \to (C_d + c_d) p_1$, for some constant $C_d > 0$, where $1' = 1 + \varepsilon^2$. Since both sides have the same integral $C_d + c_d$, the integral in (15) tends to 0, by FMP 1.21.

(ii) Here we may use a similar argument based on Theorems 13.38 (iii) and 13.42 (ii).

(iii) For a DW-process ξ, let ζ_s be the process of ancestors of ξ_1 at time $s = 1 - h$, where $\varepsilon^2 \leq h \leq 1$, and denote the generated h-clusters by η_h^i. For $x_1, x_2 \in \mathsf{R}^d$, we write $x_i = \bar{x} \pm r$. Using Lemmas 13.23 (i)−(ii) and Theorem 13.38 (ii), the conditional independence of sub-clusters, the Cox property of ζ_s, and the semi-group property of p_t, we obtain with $h' = h + \varepsilon^2$ and $1' = 1 + \varepsilon^2$

$$\begin{aligned}
E_{\delta_0} \sum_{i \neq j} 1\Big\{\eta_h^i B_{x_1}^\varepsilon \wedge \eta_h^j B_{x_2}^\varepsilon > 0\Big\}
&= \iint_{u_1 \neq u_2} p_h^\varepsilon(x_1 - u_1)\, p_h^\varepsilon(x_2 - u_2)\, E_{\delta_0} \zeta_s^2(du_1\, du_2) \\
&\lesssim \varepsilon^{2(d-2)} \iint p_{h'}(x_1 - u_1)\, p_{h'}(x_2 - u_2)\, E_{\delta_0} \xi_s^2(du_1\, du_2) \\
&= \varepsilon^{2(d-2)} \Big\{ (p_s^{\otimes 2} + q_s) * p_{h'}^{\otimes 2} \Big\}(x_1, x_2) \\
&\leq \varepsilon^{2(d-2)} \Big(p_{1'}^{\otimes 2} + q_{1'} \Big)(x_1, x_2) \\
&\lesssim \varepsilon^{2(d-2)}\, p_{1'}(\bar{x})\, p_{1'}(r)\, |r|^{2-d}.
\end{aligned}$$

Combining the previously mentioned properties with Theorems 13.10 (iv) and 13.38 (ii), Cauchy's inequality, the parallelogram identity, and the special form of the densities p_t, we obtain

$$\begin{aligned}
E_{\delta_0} \textstyle\sum_i 1\Big\{\eta_h^i B_{x_1}^\varepsilon \wedge \eta_h^i B_{x_2}^\varepsilon > 0\Big\}
&= \int P_u\Big\{\eta_h B_{x_1}^\varepsilon \wedge \eta_h B_{x_2}^\varepsilon > 0\Big\} E_{\delta_0} \zeta_s(du) \\
&\leq h^{-1} \int \Big\{ p_h^\varepsilon(x_1 - u)\, p_h^\varepsilon(x_2 - u) \Big\}^{1/2} E_{\delta_0} \xi_s(du) \\
&\lesssim \varepsilon^{d-2} \int \Big\{ p_{h'}(x_1 - u)\, p_{h'}(x_2 - u) \Big\}^{1/2} p_s(u)\, du \\
&\lesssim \varepsilon^{d-2} \int \Big\{ p_{h'/2}(\bar{x} - u)\, p_{h'/2}(r) \Big\}^{1/2} p_s(u)\, du \\
&\lesssim \varepsilon^{d-2}\, h^{d/2} (p_{h'} * p_s)(\bar{x})\, p_{h'}(r) \\
&= \varepsilon^{d-2}\, h^{d/2}\, p_{1'}(\bar{x})\, p_{h'}(r).
\end{aligned}$$

Since ξ_1 is the sum of κ independent unit clusters, where κ is Poisson under P_{δ_0} with mean 1, the previous estimates remain valid for the sub-clusters of η of age h. Since η_1^ε has Lebesgue density $1\{\eta_1 B_x^\varepsilon > 0\}$, Fubini's theorem yields

$$\begin{aligned}
E_0\|\eta_1^\varepsilon\|^2 &= \iint P_0\Big\{\eta_1 B^\varepsilon_{x_1}\wedge \eta_1 B^\varepsilon_{x_2}>0\Big\}\,dx_1\,dx_2\\
&\lesssim \iint dx_1\,dx_2\,E_{\delta_0}\sum_{i,j}1\Big\{\eta_h^i B^\varepsilon_{x_1}\wedge \eta_h^j B^\varepsilon_{x_2}>0\Big\}\\
&\lesssim \iint \Big\{\varepsilon^{2(d-2)}\,p_{1'}(r)\,|r|^{2-d}+\varepsilon^{d-2}\,h^{d/2}\,p_{h'}(r)\Big\}\,p_{1'}(\bar x)\,d\bar x\,dr\\
&\lesssim \varepsilon^{2(d-2)}+\varepsilon^{d-2}h^{d/2},
\end{aligned}$$

where, in the last step, we used the fact that

$$\int p_1(r)\,|r|^{2-d}dr \lesssim \int_0^\infty v\,e^{-v^2/2}dv<\infty.$$

Taking $h=\varepsilon^2$, we get by (i) and Jensen's inequality

$$\begin{aligned}
\varepsilon^{2(d-2)} &\asymp \|E_0\eta_1^\varepsilon\|^2\le E_0\|\eta_1^\varepsilon\|^2\\
&\lesssim \varepsilon^{2(d-2)}+\varepsilon^{2d-2}\asymp \varepsilon^{2(d-2)}.
\end{aligned}$$

When $d=2$, let $\varepsilon^2\ll h\to 0$. Using Lemma 13.23 (ii) and Theorem 13.38 (iii), we get as before

$$\begin{aligned}
E_{\delta_0}\sum_{i\neq j}1\Big\{\eta_h^i B^\varepsilon_{x_1}\wedge \eta_h^j B^\varepsilon_{x_2}>0\Big\} &\lesssim \Big\{\log(h/\varepsilon^2)\Big\}^{-2}\Big(p_{1'}^{\otimes 2}+q_{1'}\Big)(x)\\
&\lesssim |\log\varepsilon|^{-2}\,p_{1'}(\bar x)\,p_1(r)\log\Big(|r|^{-1}\vee e\Big),\\
1_{\delta_0}\textstyle\sum_i 1\Big\{\eta_h^i B^\varepsilon_{x_1}\wedge \eta_h^i B^\varepsilon_{x_2}>0\Big\} &\lesssim h|\log\varepsilon|^{-1}p_{1'}(\bar x)\,p_{h'}(r),
\end{aligned}$$

where $1'-1=h'-h\lesssim h|\log\varepsilon|^{-1/2}$. Noting that

$$\int p_1(r)\log\Big(|r|^{-1}\vee e\Big)\,dr\lesssim \int_{|r|e<1}\Big|\log|r|\Big|\,dr+\int p_1(r)\,dr<\infty,$$

we get by combination

$$\begin{aligned}
E_0\|\eta_1^\varepsilon\|^2 &\lesssim \iint\Big\{|\log\varepsilon|^{-2}p_1(r)\log\Big(|r|^{-1}\vee e\Big)+h\,|\log\varepsilon|^{-1}p_{h'}(r)\Big\}p_{1'}(\bar x)\,d\bar x\,dr\\
&\lesssim |\log\varepsilon|^{-2}+h\,|\log\varepsilon|^{-1}.
\end{aligned}$$

Choosing $h=|\log\varepsilon|^{-1}\gg\varepsilon^2$, and combining with (ii), we get

$$|\log\varepsilon|^{-2}\asymp\|E_0\eta_1^\varepsilon\|^2\le E_0\|\eta_1^\varepsilon\|^2\lesssim|\log\varepsilon|^{-2}. \qquad \square$$

Our next step is to derive similar moment estimates for a Poisson "forest" of clusters.

Lemma 13.50 *(cluster sums) Consider some conditionally independent h-clusters η_{ih} in R^d, rooted at the points of a Poisson process ξ with $E\xi=\mu$. Fix any measurable function $f\ge 0$ on R^d, and let $h\ge\varepsilon\to 0$. Then*

(i) $E_\mu \sum_i \eta^\varepsilon_{ih} = (\mu * p^\varepsilon_h) \cdot \lambda^d$,

(ii) $\mathrm{Var}_\mu \sum_i \eta^\varepsilon_{ih} f \lesssim h^2 \, \|f\|^2 \, \|\mu\| \begin{cases} \varepsilon^{2(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-2}, & d = 2. \end{cases}$

Proof: (i) By Fubini's theorem and the definitions of η^ε_h and p^ε_h, we have

$$\begin{aligned} E_x \eta^\varepsilon_h f &= E_x \int 1\Big\{\eta_h B^\varepsilon_u > 0\Big\} f(u)\, du \\ &= (p^\varepsilon_h * f)(x), \end{aligned}$$

and so by independence,

$$E\Big(\sum\nolimits_i \eta^{i\varepsilon}_h f \,\Big|\, \xi\Big) = \int \xi(dx)\, E_x \eta^\varepsilon_h f = \xi(p^\varepsilon_h * f). \tag{16}$$

Hence, Fubini's theorem yields

$$\begin{aligned} E_\mu \sum\nolimits_i \eta^{i\varepsilon}_h f &= E_\mu \xi(p^\varepsilon_h * f) = \mu(p^\varepsilon_h * f) \\ &= \Big((\mu * p^\varepsilon_h) \cdot \lambda^d\Big) f. \end{aligned}$$

(ii) By Theorem 13.14 (ii), we have

$$\begin{aligned} \|\eta^\varepsilon_h\| &= \int 1\Big\{\eta_h B^\varepsilon_x > 0\Big\}\, dx \\ &\overset{d}{=} \int 1\Big\{\eta_1 B^{\varepsilon/\sqrt{h}}_{x/\sqrt{h}} > 0\Big\}\, dx \\ &= h^{d/2} \int 1\Big\{\eta_1 B^{\varepsilon/\sqrt{h}}_x > 0\Big\}\, dx \\ &= h^{d/2}\, \|\eta^{\varepsilon/\sqrt{h}}_1\|, \end{aligned}$$

and so by Lemma 13.49 (iii),

$$\begin{aligned} \mathrm{Var}_x(\eta^\varepsilon_h f) &\leq E_x(\eta^\varepsilon_h f)^2 \leq E\|\eta^\varepsilon_h\|^2 \, \|f\|^2 \\ &= h^d\, E \Big\| \, \eta^{\varepsilon/\sqrt{h}}_1 \Big\|^2 \, \|f\|^2 \\ &\lesssim h^d\, (\varepsilon/\sqrt{h})^{2(d-2)}\, \|f\|^2 \\ &= \varepsilon^{2(d-2)}\, h^2\, \|f\|^2. \end{aligned}$$

Hence, by independence,

$$\begin{aligned} E_\mu \mathrm{Var}\Big(\sum\nolimits_i \eta^{i\varepsilon}_h f \,\Big|\, \xi\Big) &= E_\mu \int \xi(dx)\, \mathrm{Var}_x(\eta^\varepsilon_h f) \\ &\lesssim \varepsilon^{2(d-2)}\, h^2\, \|f\|^2\, \|\mu\|. \end{aligned}$$

Since $\lambda^d p^\varepsilon_h \lesssim \varepsilon^{d-2} h$ by Theorem 13.38 (ii), and $\mathrm{Var}_\mu(\xi f) = \mu f^2$, we get from (16)

$$\begin{aligned} \mathrm{Var}_\mu\, E\Big(\sum\nolimits_i \eta^{i\varepsilon}_h f \,\Big|\, \xi\Big) &= \mathrm{Var}_\mu \xi(p^\varepsilon_h * f) \\ &= \mu(p^\varepsilon_h * f)^2 \\ &\leq \|f\|^2\, \|\mu\|\, (\lambda^d p^\varepsilon_h)^2 \\ &\lesssim \varepsilon^{2(d-2)}\, h^2\, \|f\|^2\, \|\mu\|. \end{aligned}$$

Combining these estimates gives

$$\begin{aligned}\mathrm{Var}_\mu \sum\nolimits_i \eta_h^{i\varepsilon} f &= E_\mu \mathrm{Var}\Big(\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi \Big) + \mathrm{Var}_\mu E\Big(\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi \Big) \\ &\lesssim \varepsilon^{2(d-2)}\, h^2\, \|f\|^2\, \|\mu\|.\end{aligned}$$

For $d = 2$, let $h \geq \varepsilon$, and conclude from Lemma 13.49 (iii) that

$$\begin{aligned}\mathrm{Var}_x(\eta_h^\varepsilon f) &\lesssim h^2 \Big| \log(\varepsilon/\sqrt{h}) \Big|^{-2} \|f\|^2 \\ &\lesssim h^2\, |\log \varepsilon|^{-2}\, \|f\|^2,\end{aligned}$$

which implies

$$E_\mu \mathrm{Var}\Big(\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi \Big) \lesssim h^2\, |\log \varepsilon|^{-2}\, \|f\|^2\, \|\mu\|.$$

Next, Theorem 13.38 (iii) yields $\lambda^2 p_h^\varepsilon \lesssim h |\log \varepsilon|^{-1}$, and so as before

$$\mathrm{Var}_\mu E\Big(\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi \Big) \lesssim h^2\, |\log \varepsilon|^{-2}\, \|f\|^2\, \|\mu\|.$$

The stated estimate now follows by combination. □

For fixed $t > 0$, we may also use Lemma 13.39 to estimate the overlap between the sub-clusters in ξ_t of age h.

Lemma 13.51 *(cluster overlap) Let ξ be a DW-process in R^d, fix any $\mu \in \hat{\mathcal{M}}_d$ and $t > 0$, and let η_{ih} denote the sub-clusters in ξ_t of age $h > 0$. Assuming $\varepsilon^2 \leq h \to 0$ when $d \geq 3$, and $\varepsilon \leq h \to 0$ when $d = 2$, we have*

$$E_\mu \Big\| \sum\nolimits_i \eta_{ih}^\varepsilon - \xi_t^\varepsilon \Big\| \lesssim \begin{cases} \big(\varepsilon^2/\, h^{1/2}\big)^{d-2}, & d \geq 3, \\ |\log h| \,/\, |\log \varepsilon|^2, & d = 2. \end{cases}$$

Proof: Let $\kappa_h^\varepsilon(x)$ be the number of sub-clusters of age h, hitting B_x^ε at time t. For $d \geq 3$, Lemma 13.39 (i) yields with $t' = t + \varepsilon^2$

$$\begin{aligned}E_\mu \Big\| \sum\nolimits_i \eta_h^{i\varepsilon} - \xi_t^\varepsilon \Big\| &= E_\mu \int \Big| \sum\nolimits_i 1\{\eta_h^i B_x^\varepsilon > 0\} - 1\{\xi_t B_x^\varepsilon > 0\} \Big|\, dx \\ &= \int E_\mu \Big\{ \kappa_h^\varepsilon(x) - 1 \Big\}_+ dx \\ &\lesssim \varepsilon^{2(d-2)}\, \lambda^d \Big\{ h^{1-d/2}\, (\mu * p_t) + (\mu * p_{t'})^2 \Big\} \\ &\lesssim \varepsilon^{2(d-2)} \Big(h^{1-d/2}\, \|\mu\| + t^{-d/2}\, \|\mu\|^2 \Big).\end{aligned}$$

If $d = 2$, we may use Lemma 13.39 (ii) instead to get

$$\begin{aligned}E_\mu \Big\| \sum\nolimits_i \eta_h^{i\varepsilon} - \xi_t^\varepsilon \Big\| &\lesssim |\log \varepsilon|^{-2}\, \lambda^2 \Big\{ \log(t/h)(\mu * p_t) + (\mu * p_{t'})^2 \Big\} \\ &\lesssim |\log \varepsilon|^{-2} \Big(|\log h|\, \|\mu\| + t^{-1}\, \|\mu\|^2 \Big),\end{aligned}$$

for suitable $t' \geq t$. □

We are now ready to prove the main results of the section.

Proof of Theorem 13.47: (i) Fix any $t > 0$, $\mu \in \hat{\mathcal{M}}_d$, and $f \in C_K^d$, and write η_h^i for the sub-clusters in ξ_t of age h. Since the ancestors of ξ_t at time $s = t - h$ form a Cox process directed by $h^{-1}\xi_s$, Lemma 13.50 (i) yields

$$E_\mu\Big(\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi_s \Big) = h^{-1}\, \xi_s(p_h^\varepsilon * f),$$

and so by Lemma 13.50 (ii),

$$\begin{aligned} E_\mu\Big\{ \sum\nolimits_i \eta_h^{i\varepsilon} f - h^{-1}\, \xi_s(p_h^\varepsilon * f) \Big\}^2 &= E_\mu \mathrm{Var}\Big(\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi_s \Big) \\ &\lesssim \varepsilon^{2(d-2)}\, h^2\, \|f\|^2\, E_\mu \|\xi_s/h\| \\ &\lesssim \varepsilon^{2(d-2)}\, h\, \|f\|^2\, \|\mu\|. \end{aligned}$$

Combining with Lemma 13.51 gives

$$\begin{aligned} &E_\mu\Big| \xi_t^\varepsilon f - h^{-1}\xi_s(p_h^\varepsilon * f) \Big| \\ &\quad \le E_\mu\Big| \xi_t^\varepsilon f - \sum\nolimits_i \eta_h^{i\varepsilon} f \Big| + E_\mu\Big| \sum\nolimits_i \eta_h^{i\varepsilon} f - h^{-1}\xi_s(p_h^\varepsilon * f) \Big| \\ &\quad \lesssim \varepsilon^{2(d-2)}\, h^{1-d/2}\, \|f\| + \varepsilon^{d-2}\, h^{1/2}\, \|f\| \\ &\quad = \varepsilon^{d-2}\Big\{ \sqrt{h} + (\varepsilon/\sqrt{h})^{d-2} \Big\}\, \|f\|. \end{aligned}$$

Taking $h = \varepsilon = r^n$ for a fixed $r \in (0,1)$, and writing $s_n = t - r^n$, we obtain

$$\begin{aligned} &E_\mu \sum\nolimits_n r^{n(2-d)} \Big| \xi_t^{r^n} f - r^{-n}\, \xi_{s_n}(p_{r^n}^{r^n} * f) \Big| \\ &\qquad \lesssim \sum\nolimits_n \Big(r^{n/2} + r^{n(d-2)/2} \Big)\, \|f\| < \infty, \end{aligned}$$

which implies

$$r^{n(2-d)} \Big| \xi_t^{r^n} f - r^{-n}\, \xi_{s_n}(p_{r^n}^{r^n} * f) \Big| \to 0 \ \text{ a.s. } P_\mu. \tag{17}$$

Now write

$$\begin{aligned} \Big| \varepsilon^{2-d}\, \xi_t^\varepsilon f - c_d\, \xi_t f \Big| \le\ & \varepsilon^{2-d} \Big| \xi_t^\varepsilon f - h^{-1}\xi_s(p_h^\varepsilon * f) \Big| \\ & + c_d\, |\, \xi_s f - \xi_t f\,| \\ & + \|\xi_s\|\, \Big\| \varepsilon^{2-d}\, h^{-1}(p_h^\varepsilon * f) - c_d\, f \Big\|. \end{aligned}$$

Using (17), Lemma 13.41 (i), and the a.s. weak continuity of ξ, we see that the right-hand side tends a.s. to 0 as $n \to \infty$, which implies $\varepsilon^{2-d}\xi_t^\varepsilon f \to c_d \xi_t f$ a.s., as $\varepsilon \to 0$ along (r^n), for every fixed $r \in (0,1)$. Since this holds simultaneously for all rational $r \in (0,1)$, outside a fixed P-null set, the a.s. convergence extends by Lemma A6.2 to the entire interval $(0,1)$.

Now let $\mu \in \mathcal{M}_d$ be arbitrary with $\mu p_t < \infty$ for all $t > 0$. Write $\mu = \mu' + \mu''$ for bounded μ', and let $\xi = \xi' + \xi''$ be the corresponding decomposition of ξ, into independent components with initial measures μ' and μ''. Fixing $r > 1$ with $\operatorname{supp} f \subset B_0^{r-1}$, and using the result for bounded μ, we get, a.s. on $\{\xi'' B_0^r = 0\}$,

$$\varepsilon^{2-d}\,\xi_t^\varepsilon f = \varepsilon^{2-d}\,\xi_t'^\varepsilon f \to c_d\,\xi_t' f = c_d\,\xi_t f.$$

As $\mu' \uparrow \mu$, we get by Lemma 13.12

$$P_\mu\{\xi_t'' B_0^r = 0\} = P_{\mu''}\{\xi_t B_0^r = 0\} \to 1,$$

and the a.s. convergence extends to μ. Applying this to a countable, convergence-determining class of functions f, we obtain the required a.s. vague convergence. If μ is bounded, then ξ_t has a.s. bounded support, by Theorem 13.38 (iv), and the a.s. convergence remains valid in the weak sense.

To prove the convergence in L^1, we note that for any $f \in C_K^d$

$$\begin{aligned} \varepsilon^{2-d}\,E_\mu \xi_t^\varepsilon f \;&=\; \varepsilon^{2-d}\int P_\mu\{\xi_t B_x^\varepsilon > 0\}\, f(x)\, dx \\ &\to \int c_d\,(\mu * p_t)(x)\, f(x)\, dx = c_d\, E_\mu \xi_t f, \end{aligned} \tag{18}$$

by Theorem 13.42 (i). Combining this with the a.s. convergence under P_μ, and using FMP 4.12, we obtain $E_\mu|\varepsilon^{2-d}\xi_t^\varepsilon f - c_d \xi_t f| \to 0$. For bounded μ, (18) extends to any $f \in C_b^d$, by dominated convergence based on Lemma 13.9 and Theorem 13.38 (ii), along with the fact that $\lambda^d(\mu * p_t) = \|\mu\| < \infty$, by Fubini's theorem.

(ii) Fix any t, μ, and f, as before. Using Lemma 13.50 (iii), we see as in case (i) that

$$E_\mu \Big| \textstyle\sum_i \eta_h^{i\varepsilon} f - h^{-1}\xi_s(p_h^\varepsilon * f)\Big|^2 \lesssim h\,|\log\varepsilon|^{-2}\,\|f\|^2\,\|\mu\|.$$

Combining with Lemma 13.51, we get for fixed μ and f

$$E_\mu\Big|\,\xi_t^\varepsilon f - h^{-1}\xi_s(p_h^\varepsilon * f)\Big| \lesssim h^{1/2}\,|\log\varepsilon|^{-1} + |\log h|\,|\log\varepsilon|^{-2}.$$

Choosing $\sqrt{h} = |\log\varepsilon|^{-1} = r^n$ for fixed $r \in (0,1)$, we get

$$\begin{aligned} |\log\varepsilon|\,E_\mu\Big|\,\xi_t^\varepsilon f - h^{-1}\xi_s(p_h^\varepsilon * f)\Big| \;&\lesssim\; h^{1/2} + |\log h|\,|\log\varepsilon|^{-1} \\ &=\; r^n + 2n\,|\log r|\,r^n \lesssim r^{n/2}. \end{aligned} \tag{19}$$

Now we write

$$\begin{aligned} \Big| m_\varepsilon^{-1}|\log\varepsilon|\,\xi_t^\varepsilon f - \xi_t f\,\Big| \;&\lesssim\; |\log\varepsilon|\,\Big|\,\xi_t^\varepsilon f - h^{-1}\xi_s(p_h^\varepsilon * f)\Big| \\ &\quad + |\,\xi_s f - \xi_t f| \\ &\quad + \|\xi_s\|\,\Big\|\,h^{-1}m_\varepsilon^{-1}\,|\log\varepsilon|\,(p_h^\varepsilon * f) - f\Big\|. \end{aligned}$$

As $n \to \infty$, we see from (19), Lemma 13.41 (ii), and the weak continuity of ξ that the right-hand side tends to 0 a.s. Writing $\varepsilon = e^{-1/s}$, and putting $\tilde\xi_t^s = \xi_t^\varepsilon$, we conclude that

$$\Big(m \circ e^{-1/s}\Big)^{-1} s^{-1}\,\tilde\xi_t^s f \to \xi_t f \;\text{ a.s. } P_\mu, \tag{20}$$

as $s \to 0$ along (r^n), for any $r \in (0,1)$. Since the function $\log m\{\exp(-e^t)\}$ is bounded and uniformly continuous on R_+, by Lemma 13.40, (20) remains true along $(0,1)$, by Lemma A6.2. Hence,

$$m_\varepsilon^{-1}|\log\varepsilon|\,\xi_t^\varepsilon f \to \xi_t f \ \text{ a.s. } P_\mu,$$

for fixed f and bounded μ, which extends as before to $m_\varepsilon^{-1}|\log\varepsilon|\,\xi_t^\varepsilon \xrightarrow{v} \xi_t$ a.s., even for unbounded μ.

To prove the corresponding L^1-convergence, let $f \in C_K^d$, and conclude from Theorem 13.42 (ii) that

$$\begin{aligned} m_\varepsilon^{-1}|\log\varepsilon|\,E_\mu\xi_t^\varepsilon f &= m_\varepsilon^{-1}|\log\varepsilon| \int P_\mu\{\xi_t B_x^\varepsilon > 0\}\, f(x)\,dx \\ &\to \int (\mu * p_t)(x)\, f(x)\,dx = E_\mu\xi_t f. \end{aligned}$$

For bounded μ, this extends by dominated convergence to any $f \in C_b^d$. The assertion now follows as before, by combination with the corresponding a.s. convergence.

To extend (i) and (ii) to the individual clusters η_t, let ζ_0 be the process of ancestors of ξ_t at time 0, and note that

$$\begin{aligned} \mathcal{L}_0(\eta_t) &= \mathcal{L}_{\delta_0}\Big\{\xi_t \,\Big|\, \|\zeta_0\| = 1\Big\}, \\ P_{\delta_0}\{\|\zeta_0\| = 1\} &= t^{-1}e^{-1/t} > 0. \end{aligned}$$

The a.s. convergence then follows from the corresponding statement for ξ_t. To obtain the weak L^1-convergence in this case, we note that, for $f \in C_b^d$ and $d \geq 3$ or $d = 2$, respectively,

$$\begin{aligned} \varepsilon^{2-d}\, E_0\eta_t^\varepsilon f &= \varepsilon^{2-d}\,\lambda^d(p_t^\varepsilon f) \\ &\to c_d\, t\,\lambda^d(p_t f) = c_d\, E_0\eta_t f, \\ m_\varepsilon^{-1}|\log\varepsilon|\, E_0\eta_t^\varepsilon f &= m_\varepsilon^{-1}|\log\varepsilon|\,\lambda^d(p_t^\varepsilon f) \\ &\to t\,\lambda^d(p_t f) = E_0\eta_t f, \end{aligned}$$

by dominated convergence based on Theorems 13.38 and 13.42. $\square$

Proof of Corollary 13.48: The assertions are equivalent to the statements

$$\begin{aligned} &\int \Big| \varepsilon^{2-d}\, P_\mu\{\xi_t B_x^\varepsilon > 0\} - c_d\,(\mu * p_t)(x) \Big|\, dx \to 0, \\ &\int \Big| |\log\varepsilon|\, P_\mu\{\xi_t B_x^\varepsilon > 0\} - m(\varepsilon)\,(\mu * p_t)(x) \Big|\, dx \to 0, \end{aligned}$$

which are L^1-versions of Theorem 13.42, and follow as before by dominated convergence. $\square$

13.8 Local Stationarity and Invariant Cluster

Here we establish the local approximation of a DW-process by the space-time stationary version $\tilde{\xi}$ in Theorem 13.52, along with a similar approximation by the stationary cluster $\tilde{\eta}$ in Theorem 13.53. Since $\tilde{\xi}$ exists only for $d \geq 3$, the latter approximation is more general. However, the strong invariance properties of $\tilde{\xi}$ hold only approximately for $\tilde{\eta}$, and the difference between the two cases is highlighted by the last two theorems of the section.

For any $B \in \hat{\mathcal{B}}^d$, we write $\|\cdot\|_B$ for total variation on the set $H_B = \{\mu;\, \mu B > 0\}$, equipped with the σ-field $\mathcal{H}_B$ generated by the restriction map $\mu \mapsto 1_B\mu$. Recall that the prefix "pseudo" refers to the fact that $\tilde{\xi}$ and $\tilde{\eta}$ can only be defined on a "probability" space with unbounded but σ-finite measure $\tilde{P}$. This causes no problem below, since $\tilde{P}\{\tilde{\xi}B > 0\} < \infty$, and similarly for $\tilde{\eta}$. For ξ_t and η_t, we introduce the Palm distributions

$$\mathcal{L}^0_\mu(\xi_t) = \mathcal{L}_\mu(\xi_t \| \, \xi_t)_0, \qquad \mathcal{L}^0_\mu(\eta_t) = \mathcal{L}_\mu(\eta_t \| \, \eta_t)_0,$$

where $\mathcal{L}_\mu(\xi_t \| \, \xi_t)_x$ and $\mathcal{L}_\mu(\eta_t \| \, \eta_t)_x$ are defined by continuity, as before. Since $\tilde{\xi}$ and $\tilde{\eta}$ are stationary, by definition, their Palm distributions $\mathcal{L}^0(\tilde{\xi})$ and $\mathcal{L}^0(\tilde{\eta})$ may be defined as in Section 5.1.

Theorem 13.52 *(local stationarity) For a DW-process ξ in R^d with $d \geq 3$, there exists a pseudo-random measure $\tilde{\xi}$ on R^d such that*

(i) *as $\varepsilon \to 0$ for fixed μ and $t > 0$,*

$$\Big\| \, \mathcal{L}_\mu(\xi_t) - \mu p_t \, \mathcal{L}(\tilde{\xi}) \Big\|_{B_0^\varepsilon} \ll \varepsilon^{d-2},$$
$$\Big\| \, \mathcal{L}^0_\mu(\xi_t) - \mathcal{L}^0(\tilde{\xi}) \Big\|_{B_0^\varepsilon} \to 0,$$

and similarly with ξ_t replaced by η_t,

(ii) *for any $r > 0$,*

$$\mathcal{L}(\tilde{\xi} S_r) = r^{d-2} \mathcal{L}(r^2 \tilde{\xi}),$$
$$\mathcal{L}^0(\tilde{\xi} S_r) = \mathcal{L}^0(r^2 \tilde{\xi}),$$

(iii) *$\tilde{\xi}$ is stationary with $E\tilde{\xi} = \lambda^d$,*

(iv) *$\mathcal{L}(\tilde{\xi})$ is an invariant measure for ξ,*

(v) *as $r \to \infty$ for bounded B,*

$$\Big\| \, r^{d-2} \mathcal{L}_{r^{2-d}\lambda^d}(\xi_{r^2}) - \mathcal{L}(\tilde{\xi}) \Big\|_{H_B} \to 0.$$

Proof of (i)–(ii) *for $\mathcal{L}_\mu(\tilde{\xi})$:* Fix any $t > h > 0$ and $B \in \hat{\mathcal{B}}^d$, and consider any $\mathcal{H}_B$-measurable function $f \geq 0$ on $\mathcal{M}_d$ with $f \leq 1_{H_B}$. Consider the process ζ_s of ancestors of ξ_t at time $s = t - h$, and let η^i_h denote the associated

h-clusters. As $h \to 0$ and $r = \varepsilon/\sqrt{h} \to 0$, we have the following chain of relations, explained in further detail below:

$$\begin{aligned}
E_\mu f(\varepsilon^{-2}\xi_t S_\varepsilon) &= E_\mu\Big(\varepsilon^{-2}\textstyle\sum_i \eta_h^i S_\varepsilon\Big) \\
&\approx E_\mu \textstyle\sum_i f\Big(\varepsilon^{-2}\eta_h^i S_\varepsilon\Big) \\
&= \int E_x f\Big(\varepsilon^{-2}\eta_h S_\varepsilon\Big)\, E_\mu \zeta_s(dx) \\
&= h^{-1}\int \mu(dy)\int p_s(x-y)\, E_x f\Big(\varepsilon^{-2}\eta_h S_\varepsilon\Big)\, dx \\
&\approx h^{-1}\mu p_t \int E_x f\Big(\varepsilon^{-2}\eta_h S_\varepsilon\Big)\, dx \\
&= h^{-1}\,\mu p_t \int E_{x/\sqrt{h}}\, f\Big(h\varepsilon^{-2}\eta_1 S_{\varepsilon/\sqrt{h}}\Big)\, dx \\
&= (\varepsilon/r)^{d-2}\,\mu p_t \int E_x f\Big(r^{-2}\eta_1 S_r\Big)\, dx. \qquad (21)
\end{aligned}$$

Here the third relation holds by the conditional independence of clusters, the fourth one holds since

$$E_\mu \zeta_s = h^{-1} E_\mu \xi_s = h^{-1}(\mu * p_s)\cdot \lambda^d,$$

and the sixth one holds by Theorem 13.14.

To justify the first approximation in (21), define κ_h^ε as in Lemma 13.39, and fix a $c > 0$ with $B \subset B_0^c$. Then the mentioned lemma yields

$$\begin{aligned}
&\varepsilon^{2-d}\, E_\mu \Big| f\Big(\varepsilon^{-2}\textstyle\sum_i \eta_h^i S_\varepsilon\Big) - \textstyle\sum_i f\Big(\varepsilon^{-2}\eta_h^i S_\varepsilon\Big)\Big| \\
&\qquad \le\ \varepsilon^{2-d}\, E_\mu\Big(\kappa_h^{c\varepsilon};\ \kappa_h^{c\varepsilon} > 1\Big) \\
&\qquad \lesssim\ \varepsilon^{2-d}\,(c\varepsilon)^{2(d-2)}\Big\{h^{1-d/2}\,\mu p_t + (\mu p_{t(c\varepsilon)})^2\Big\} \\
&\qquad \lesssim\ r^{d-2} \to 0. \qquad (22)
\end{aligned}$$

The second approximation in (21) amounts to replacing $p_s(x-y)$ by $p_t(y)$ in the inner integral. To estimate the resulting error, we note that, by Theorem 13.38 (ii),

$$\begin{aligned}
&\varepsilon^{2-d}\, h^{-1}\Big|\int \mu(dy)\int \Big\{p_s(y-x) - p_t(y)\Big\}\, E_x f\Big(\varepsilon^{-2}\eta_h S_\varepsilon\Big)\, dx\Big| \\
&\qquad \lesssim\ \int \mu(dy)\int \Big|p_s(y-x) - p_t(y)\Big|\, p_{h(c\varepsilon)}(x)\, dx \\
&\qquad =\ \int \mu(dy)\, E\Big|p_s\Big(y - \gamma\, h_{c\varepsilon}^{1/2}\Big) - p_t(y)\Big|, \qquad (23)
\end{aligned}$$

where $h_\varepsilon = h + \varepsilon^2$, and γ denotes a standard normal random vector in R^d. As $\varepsilon^2 \le h \to 0$, we get $p_s(y - \gamma h_\varepsilon^{1/2}) \to p_t(y)$ a.s., by the joint continuity of $p_t(x)$. Since also

$$\begin{aligned}
E p_s\Big(y - \gamma\, h_{c\varepsilon}^{1/2}\Big) &= \Big(p_s * p_{\,h(c\varepsilon)}\Big)(y) \\
&= p_{\,t(c\varepsilon)}(y) \to p_t(y),
\end{aligned}$$

the last expectation in (23) tends to 0, by FMP 1.32. Finally, since

$$E\Big|\, p_s\Big(y-\gamma\, h_{c\varepsilon}^{1/2}\Big) - p_t(y)\Big| \;\le\; p_{t(c\varepsilon)}(y) + p_t(y) \\ \lesssim\; p_{2t}(y),$$

where $\mu p_{2t} < \infty$, the right-hand side of (23) tends to 0, by dominated convergence.

This proves that, as $\varepsilon \ll r \to 0$ for fixed $\mu \in \mathcal{M}_d$, $B \in \hat{\mathcal{B}}^d$, and $t > 0$, we have

$$\Big\|\, \varepsilon^{2-d}\,\mathcal{L}_\mu\Big(\varepsilon^{-2}\,\xi_t S_\varepsilon\Big) - r^{2-d}\,\mu p_t \int \mathcal{L}_x\Big(r^{-2}\,\eta_1 S_r\Big)\,dx\Big\|_B \to 0. \tag{24}$$

In particular, the first term on the left is uniformly Cauchy convergent on H_B, as $\varepsilon \to 0$. Hence, both terms converge, as $\varepsilon \to 0$ and $r \to 0$, respectively, to a common limit of the form $\mu p_t\,\varphi_B$, where the set function φ_B on H_B is independent of μ and t. Thus,

$$\Big\|\, \varepsilon^{2-d}\,\mathcal{L}_\mu\Big(\varepsilon^{-2}\,\xi_t S_\varepsilon\Big) - \mu p_t\,\varphi_B\Big\|_B \to 0, \tag{25}$$

where the uniformity of the convergence ensures that φ_B is a bounded measure on $(H_B, \mathcal{H}_B)$.

Comparing the statements in (25) for different sets B, we see that the φ_B are restrictions of a common set function φ on $\bigcup_B \mathcal{H}_B$. We need to prove that φ can be extended to a measure $\hat\varphi$ on

$$H = \bigcup\nolimits_B H_B = \mathcal{M}_d \setminus \{0\} = \mathcal{M}_d',$$

endowed with the σ-field $\mathcal{H} = \bigvee_B \mathcal{H}_B$, generated by all projection maps $\mu \mapsto \mu B$. Choosing $\tilde P = \hat\varphi$, and letting $\tilde\xi$ denote the identity map on $\mathcal{M}_d'$, we may then write (25) in the form (i).

To construct $\hat\varphi$, it suffices to form, for every fixed $B \in \hat{\mathcal{B}}^d$, the restriction $\hat\varphi_B$ of $\hat\varphi$ to H_B, with trace σ-field $H_B \cap \mathcal{H}$, since the measure $\hat\varphi = \sup_B \hat\varphi_B$ has then the desired properties. Writing $S = \mathcal{M}_d$ and $S_n = \mathcal{M}_{B_0^n}$, for all n with $B_0^n \supset B$, we introduce the restriction maps $\pi_n\colon S \to S_n$ and $\pi_{n,k}\colon S_n \to S_k$, $n \ge k$. Put $\varphi_n' = \varphi_{B_0^n}(H_B \cap \cdot)$, and form the bounded measures $\psi_n = \varphi_n' \circ \pi_n^{-1}$ on S_n. Since $\psi_n \circ \pi_{n,k}^{-1} = \psi_k$ for all $n \ge k$, and since the measures in $\mathcal{M}_d$ are measurably determined by their restrictions to the balls B_0^n, there exists, by FMP 6.15, a measure ψ on S with $\psi_n = \psi \circ \pi_n^{-1}$ for all n. Since the ψ_n are restricted to H_B, so is ψ, and therefore $\hat\varphi_B = \psi$ has the desired properties.

To see that (i) remains true for the clusters η_t, with p_t replaced by tp_t, we may apply the first four relations in (21)—as justified by (22)—with $h = t$ and $s = 0$, to get, as $\varepsilon \to 0$ for fixed $B \in \hat{\mathcal{B}}^d$,

$$\Big\|\, t\,\varepsilon^{2-d}\,\mathcal{L}_\mu\Big(\varepsilon^{-2}\,\xi_1 S_\varepsilon\Big) - \varepsilon^{2-d}\,\mathcal{L}_\mu\Big(\varepsilon^{-2}\,\eta_t S_\varepsilon\Big)\Big\|_B \to 0.$$

The required convergence now follows from (i).

Using the shift and semi-group properties of the operators S_r, and the shift invariance of λ^d, we get for any $r, \varepsilon > 0$ and $a \in \mathsf{R}^d$

$$\varepsilon^{2-d} \int \mathcal{L}_x\Big(\varepsilon^{-2}\eta_1 S_\varepsilon S_r \theta_a\Big)\, dx = r^{d-2}\,(r\varepsilon)^{2-d} \int \mathcal{L}_x\Big\{r^2 (r\varepsilon)^{-2}\eta_1 S_{r\varepsilon}\Big\}\, dx.$$

Letting $\varepsilon \to 0$ for fixed r, and applying the cluster version of (i) to each side, we obtain (ii) on $(H_B, \mathcal{H}_B)$, for every $B \in \hat{\mathcal{B}}^d$, and the general result follows by a monotone-class argument.

Proof of (i)$-$(ii) *for* $\mathcal{L}^0_\mu(\tilde{\xi})$*:* Noting that

$$E_\mu \xi_t = t^{-1} E_\mu \eta_t = (\mu * p_t) \cdot \lambda^d,$$

and using the continuity in Theorem 13.10 (iv), we get as $\varepsilon \to 0$ for fixed $B \in \hat{\mathcal{B}}^d$

$$\varepsilon^{-d} E_\mu \xi_t(\varepsilon B) = t^{-1} \varepsilon^{-d} E_\mu \eta_t(\varepsilon B) \to \mu p_t \lambda^d B. \tag{26}$$

Using Theorem 13.16 (i)$-$(ii) and Lemma 13.23 (i), we easily obtain

$$\begin{aligned} \mathrm{Var}_\mu\, \xi_t B_0^\varepsilon \;&\lesssim\; E_\mu\, \xi_t B_0^\varepsilon \int_0^t \Big(\varepsilon^d s^{-d/2} \wedge 1\Big)\, ds \\ &\lesssim\; \varepsilon^d\, \mu p_t\, \lambda^d B_0^1 \left\{ \int_0^{\varepsilon^2} ds + \varepsilon^d \int_{\varepsilon^2}^t s^{-d/2}\, ds \right\} \\ &\lesssim\; \varepsilon^{d+2}\, \mu p_t. \end{aligned}$$

Combining this with (26) and Theorem 13.42 (i), we get

$$\begin{aligned} E_\mu\Big\{\Big(\varepsilon^{-2}\, \xi_t B_0^\varepsilon\Big)^2 \,\Big|\, \xi_t B_0^\varepsilon > 0\Big\} \;&=\; \frac{\Big(\varepsilon^{-2}\, E_\mu \xi_t B_0^\varepsilon\Big)^2 + \mathrm{Var}_\mu\Big(\varepsilon^{-2}\xi_t B_0^\varepsilon\Big)}{P_\mu\{\xi_t B_0^\varepsilon > 0\}} \\ &\lesssim\; \frac{(\varepsilon^{d-2}\, \mu p_t)^2 + \varepsilon^{d-2}\, \mu p_t}{\varepsilon^{d-2}\, \mu p_t} \lesssim 1. \end{aligned} \tag{27}$$

Next, the result for $\mathcal{L}_\mu(\tilde{\xi})$ yields, for $B_0^1 \subset B \in \hat{\mathcal{B}}^d$,

$$\Big\| \mathcal{L}_\mu\Big(\varepsilon^{-2}\, \xi_t S_\varepsilon \,\Big|\, \xi_t B_0^\varepsilon > 0\Big) - \tilde{\mathcal{L}}\Big(\tilde{\xi} \,\Big|\, \tilde{\xi} B_0^1 > 0\Big)\Big\|_B \to 0. \tag{28}$$

By (27), the random variables $\varepsilon^{-2}\xi_t B_0^\varepsilon$ are uniformly integrable, conditionally on $\xi_t B_0^\varepsilon > 0$. By a uniform version of FMP 4.11, we may then extend (28) to

$$\Big\|E_\mu\Big(\varepsilon^{-2}\, \xi_t B_0^\varepsilon;\, \varepsilon^{-2}\, \xi_t S_\varepsilon \in \cdot \,\Big|\, \xi_t B_0^\varepsilon > 0\Big) - \tilde{E}\Big(\tilde{\xi} B_0^1;\, \tilde{\xi} \in \cdot \,\Big|\, \tilde{\xi} B_0^1 > 0\Big)\Big\|_B \to 0.$$

Combining this with the version for $\mathcal{L}_\mu(\tilde{\xi})$ yields

$$\Big\|\, \varepsilon^{2-d}\, E_\mu\Big(\varepsilon^{-2}\, \xi_t B_0^\varepsilon;\, \varepsilon^{-2}\, \xi_t S_\varepsilon \in \cdot\Big) - \mu p_t\, \tilde{E}\Big(\tilde{\xi} B_0^1;\, \tilde{\xi} \in \cdot\Big)\Big\|_B \to 0. \tag{29}$$

Since $B_0^1 \subset B$, we see from (26) and (29) that

$$\begin{aligned} t^{-1}\varepsilon^{-d} E_\mu \eta_t B_0^\varepsilon \;&=\; \varepsilon^{-d} E_\mu \xi_t B_0^\varepsilon \\ &\to \mu p_t\, \tilde{E}\tilde{\xi} B_0^\varepsilon = \mu p_t\, \lambda^d B_0^\varepsilon. \end{aligned} \tag{30}$$

By stationarity, we obtain $\tilde{E}\tilde{\xi} = \lambda^d$, which justifies the definition of $\tilde{P}^0$. Combining (30) with the the version for $\mathcal{L}_\mu(\xi_t)$, we obtain

$$E_\mu\Big(\varepsilon^{-2}\eta_t B_0^\varepsilon \,\Big|\, \eta_t B_0^\varepsilon > 0\Big) \to \tilde{E}\Big(\tilde{\xi} B_0^1 \,\Big|\, \tilde{\xi} B_0^1 > 0\Big),$$
$$\Big\| \mathcal{L}_\mu\Big(\varepsilon^{-2}\, \eta_t S_\varepsilon \Big|\, \eta_t B_0^\varepsilon > 0\Big) - \tilde{\mathcal{L}}\Big(\tilde{\xi} \,\Big|\, \tilde{\xi} B_0^1 > 0\Big)\Big\|_B \to 0.$$

Invoking FMP 4.11 in the opposite direction, we conclude that the variables $\varepsilon^{-2}\eta_t B_0^\varepsilon$ are uniformly integrable, conditionally on $\eta_t B_0^\varepsilon > 0$. Hence, by the uniform version of the same lemma,

$$\Big\| \varepsilon^{2-d}\, E_\mu\Big(\varepsilon^{-2}\, \eta_t B_0^\varepsilon;\, \varepsilon^{-2}\, \eta_t S_\varepsilon \in \cdot\,\Big) - t\, \mu p_t\, \tilde{E}\Big(\tilde{\xi} B_0^1;\, \tilde{\xi} \in \cdot\,\Big)\Big\|_B \to 0. \tag{31}$$

The asserted convergence now follows by Lemma 6.8, adapted to the case of a limiting pseudo-random measure $\tilde{\xi}$ with $\tilde{P}\{\tilde{\xi}B > 0\} < \infty$. Here (i) and (ii) hold by (29) and (31), and Lemma 13.37 yields (iii), for the shifted Palm distributions of ξ_t and η_t, based on an arbitrary initial measure μ.

(iii) Taking $B = B_0^1$ and $\mu = \lambda^{\otimes d}$ in (29), we get in particular

$$\varepsilon^{-d} E_{\lambda^{\otimes d}}\, \xi_t S_\varepsilon B_0^1 \to E\tilde{\xi} B_0^1,$$

which extends by stationarity to arbitrary B. Hence,

$$\lambda^{\otimes d} = \varepsilon^{-d}\, \lambda^{\otimes d} S_\varepsilon = \varepsilon^{-d}\, E_{\lambda^{\otimes d}}(\xi_t S_\varepsilon) \to E\tilde{\xi},$$

and so $E\tilde{\xi} = \lambda^{\otimes d}$.

(iv) Let $(\tilde{\xi}_t)$ denote the DW-process ξ with initial measure $\tilde{\xi}$. Using (ii) and Theorem 13.14 (i), we get for any $r > 0$

$$\begin{aligned}
\mathcal{L}(\tilde{\xi}_{r^2} S_r) &= E\mathcal{L}_{\tilde{\xi}}(\xi_{r^2} S_r) \\
&= E\mathcal{L}_{r^{-2}\tilde{\xi}S_r}(r^2 \xi_1) \\
&= r^{d-2}\, E\mathcal{L}_{\tilde{\xi}}(r^2 \xi_1) \\
&= r^{d-2}\, \mathcal{L}(r^2 \tilde{\xi}_1) = \mathcal{L}(\tilde{\xi} S_r),
\end{aligned}$$

which implies $\tilde{\xi}_{r^2} S_r \overset{d}{=} \tilde{\xi} S_r$. Hence, $\tilde{\xi}_t \overset{d}{=} \tilde{\xi}$ for all $t \geq 0$.

(v) Using (i) and Theorem 13.14 (i), and noting that $\lambda^{\otimes d} S_r = r^d \lambda^{\otimes d}$, we get as $r \to \infty$

$$r^{d-2} \mathcal{L}_{r^{2-d}\lambda^{\otimes d}}(\xi_{r^2}) = r^{d-2} \mathcal{L}_{\lambda^{\otimes d}}\Big(r^2\, \xi_1 S_{1/r}\Big) \to \mathcal{L}(\tilde{\xi}). \qquad \square$$

The last theorem fails for $d = 2$, already because no stationary version $\tilde{\xi}$ of the DW-process exists in that case. However, we have the following more general result, where $\tilde{\xi}$ is replaced by the *stationary cluster* $\tilde{\eta} = \tilde{\eta}_1$, defined by

$$\mathcal{L}(\tilde{\eta}_t) = \mathcal{L}_{\lambda^d}(\eta_t) = \int \mathcal{L}_x(\eta_t)\, dx, \quad t > 0.$$

Theorem 13.53 *(cluster approximation) Let ξ be a DW-process in R^d with stationary cluster $\tilde{\eta}$, where $d \geq 2$. Then as $\varepsilon \to 0$ for fixed μ and $t > 0$, we have*

$$\text{(i)} \qquad \left\| \mathcal{L}_\mu(\xi_t) - \mu p_t \, \mathcal{L}(\tilde{\eta}) \right\|_{B_0^\varepsilon} \ll \begin{cases} \varepsilon^{d-2}, & d \geq 3, \\ |\log \varepsilon|^{-1}, & d = 2, \end{cases}$$

$$\text{(ii)} \qquad \left\| \mathcal{L}^0_\mu(\xi_t) - \mathcal{L}^0(\tilde{\eta}) \right\|_{B_0^\varepsilon} \to 0,$$

and similarly with ξ_t replaced by η_t. Furthermore,

(iii) *for any $r > 0$,*

$$\mathcal{L}(\tilde{\eta}_{r^2} S_r) = r^{d-2} \mathcal{L}(r^2 \tilde{\eta}),$$
$$\mathcal{L}^0(\tilde{\eta}_{r^2} S_r) = \mathcal{L}^0(r^2 \tilde{\eta}),$$

(iv) *for $d \geq 3$, and as $t \to \infty$ for bounded B,*

$$\left\| \mathcal{L}(\tilde{\eta}_t) - \mathcal{L}(\tilde{\xi}) \right\|_{H_B} \to 0.$$

Proof: (i) The assertion for $d \geq 3$ follows from Theorem 13.52 (i), applied to ξ_t under P_μ and to η_1 under $P_{\lambda^{\otimes d}}$. Now let $d = 2$. Fixing $t > 0$, $\mu \in \mathcal{M}_2$, and $B \in \hat{\mathcal{B}}^2$, writing η_h^i for the h-clusters of ξ_t, with associated point processes ζ_s of ancestors at times $s = t - h$, and letting $\varepsilon, h \to 0$ with $|\log h| \ll |\log \varepsilon|$, we get as in (21), for any $\mathcal{H}_B$-measurable function f with $0 \leq f \leq 1_{H_B}$,

$$\begin{aligned} E_\mu f(\xi_t S_\varepsilon) &= E_\mu f\Big(\sum\nolimits_i \eta_h^i S_\varepsilon\Big) \approx E_\mu \sum\nolimits_i f(\eta_h^i S_\varepsilon) \\ &= E_\mu \int \zeta_s(dx) \, f(\eta_h^x S_\varepsilon) \\ &= \int E_\mu \xi_s(dx) \, E_x f(\eta_h S_\varepsilon) \\ &= \int \mu(dy) \int p_s(x-y) \, E_x f(\eta_h S_\varepsilon) \, dx \\ &\approx \mu p_t \int E_x f(\eta_h S_\varepsilon) \, dx \\ &= \mu p_t E f\Big(h \, \tilde{\eta} S_{\varepsilon/\sqrt{h}}\Big), \end{aligned}$$

where the third equality holds, by the conditional independence of the clusters and the Cox nature of ζ_s, and the last equality holds by Theorem 13.14 (ii).

As for the first approximation, we get by Lemma 13.39

$$\begin{aligned} E_\mu \Big| f\Big(\sum\nolimits_i \eta_h^i S_\varepsilon\Big) - \sum\nolimits_i f(\eta_h^i S_\varepsilon) \Big| &\leq E_\mu \Big\{ \kappa_h^{c\varepsilon}; \, \kappa_h^{c\varepsilon} > 1 \Big\} \\ &\lesssim \frac{\log(t/h) \, \mu p_t + \Big(\mu p_{t(h,\varepsilon)}\Big)^2}{|\log \varepsilon|^2} \\ &\lesssim \frac{|\log h| + 1}{|\log \varepsilon|^2} \ll |\log \varepsilon|^{-1}, \end{aligned}$$

where κ_h^ε denotes the number of clusters η_h^i hitting B_0^ε. For the second approximation, we get by Theorem 13.38 (ii), as $\varepsilon \le h \to 0$,

$$\begin{aligned}
&\left| \int \mu(dy) \int \Big\{ p_s(y-x) - p_t(y) \Big\} \, E_x f(\eta_h S_\varepsilon) \, dx \right| \\
&\qquad \lesssim |\log(\varepsilon^2/h)|^{-1} \int \mu(dy) \int \Big| p_s(y-x) - p_t(y) \Big| \, p_{h_{c\varepsilon}}(x) \, dx \\
&\qquad \lesssim |\log \varepsilon|^{-1} \int \mu(dy) \, E \Big| \, p_s(y - \gamma \, h_{c\varepsilon}^{1/2}) - p_t(y) \, \Big| \, ,
\end{aligned}$$

where γ denotes a standard normal random vector in R^d. Since $p_s(y - \gamma h_{c\varepsilon}^{1/2}) \to p_t(y)$, by the joint continuity of $p_t(x)$, and

$$\begin{aligned}
E p_s\Big(y - \gamma \, h_{c\varepsilon}^{1/2}\Big) &= \Big(p_s * p_{h_{c\varepsilon}}\Big)(y) \\
&= p_{s+h_{c\varepsilon}}(y) \to p_t(y),
\end{aligned}$$

the last expectation tends to 0, by FMP 1.32, and so the integral on the right tends to 0, by dominated convergence.

In summary, noting that both approximations are uniform in f, we get as $\varepsilon, h \to 0$, with $|\log h| \ll |\log \varepsilon|$,

$$\Big\| \mathcal{L}_\mu(\xi_t S_\varepsilon) - \mu p_t \, \mathcal{L}\Big(h \, \tilde\eta S_{\varepsilon/\sqrt{h}}\Big) \Big\|_B \ll |\log \varepsilon|^{-1}, \tag{32}$$

which extends to unbounded μ, by an easy truncation argument. Furthermore, Lemma 13.9 and Theorem 13.42 (ii) yield, as $\varepsilon \to 0$,

$$\begin{aligned}
\Big| E_\mu f(\xi_t S_\varepsilon) - E_\mu f(\eta_t S_\varepsilon) \Big| \;&\lesssim\; \Big(P_\mu\{\eta_t B_0^{c\varepsilon} > 0\} \Big)^2 \\
&\lesssim\; |\log \varepsilon|^{-2}.
\end{aligned} \tag{33}$$

Hence, (32) remains valid with ξ_t replaced by η_t. Now (i) follows, as we take $t = 1$ and $\mu = \lambda^{\otimes 2}$, and combine with (32). The corresponding result for η_t follows by means of (33).

(ii) Once again, the statement for $d \ge 3$ follows from Theorem 13.52 (i). For $d = 2$, we see from Corollary 13.24 and Theorem 13.42 (ii) that, as $\varepsilon \to 0$ for fixed $t > 0$,

$$\begin{aligned}
P_\mu\{\xi_t B_0^\varepsilon > 0\} \;&\asymp\; |\log \varepsilon|^{-1} \mu p_t, \\
E_\mu(\xi_t B_0^\varepsilon)^2 \;&\asymp\; \varepsilon^4 \, |\log \varepsilon| \, \Big(\lambda^{\otimes 2} B_0^1\Big)^2 \mu p_t \\
&\asymp\; \varepsilon^4 \, |\log \varepsilon| \, \mu p_t,
\end{aligned}$$

and similarly for η_t. Hence,

$$E_\mu\Big\{ (\xi_t B_0^\varepsilon)^2 \, \Big| \, \xi_t B_0^\varepsilon > 0 \Big\} = \frac{E_\mu(\xi_t B_0^\varepsilon)^2}{P_\mu\{\xi_t B_0^\varepsilon > 0\}} \asymp \varepsilon^4 \, |\log \varepsilon|^2,$$

and similarly for η_t. Thus, $\xi_t B_0^\varepsilon / \varepsilon^2 |\log \varepsilon|$ is uniformly integrable, conditionally on $\xi_t B_0^\varepsilon > 0$, and correspondingly for η_t under both P_μ and $P_{\lambda^{\otimes 2}}$. Noting that by (i),

$$\Big\| \, \mathcal{L}_\mu\Big(\xi_t S_\varepsilon \Big| \, \xi_t B_0^\varepsilon > 0 \Big) - \mathcal{L}\Big(\tilde\eta S_\varepsilon \Big| \, \tilde\eta B_0^\varepsilon > 0 \Big) \Big\|_{B_0^1} \to 0,$$

we obtain

$$\Big\| E_\mu\Big(\xi_t B_0^\varepsilon;\, \xi_t S_\varepsilon \in \cdot \,\Big|\, \xi_t B_0^\varepsilon > 0\Big) - E\Big(\tilde\eta B_0^\varepsilon;\, \tilde\eta S_\varepsilon \in \cdot \,\Big|\, \tilde\eta B_0^\varepsilon > 0\Big) \Big\|_{B_0^1} \ll \varepsilon^2\, |\log \varepsilon|,$$

and so by (i)

$$\Big\| E_\mu\Big(\xi_t B_0^\varepsilon;\, \xi_t S_\varepsilon \in \cdot\Big) - \mu p_t\, E\Big(\tilde\eta B_0^\varepsilon;\, \tilde\eta S_\varepsilon \in \cdot\Big) \Big\|_{B_0^1} \ll \varepsilon^2,$$

and similarly for η_t. Next, Theorem 13.10 yields

$$E_\mu \xi_t B_0^\varepsilon = \lambda^{\otimes 2}(\mu * p_t) 1_{B_0^\varepsilon} \asymp \varepsilon^2 \mu p_t.$$

Combining the last two estimates with Lemma 13.37, and using a version of Lemma 6.9 for pseudo-random measures, we obtain the desired convergence.

(iii) Use Theorem 13.14 (ii)−(iii).

(iv) Using (iii) and Theorem 13.52 (i)−(ii), we get as $r \to \infty$

$$\mathcal{L}(\tilde\eta_{r^2}) = r^{d-2} \mathcal{L}\Big(r^2\, \tilde\eta S_{1/r}\Big) \to \mathcal{L}(\tilde\xi). \qquad \Box$$

Though for $d \ge 3$ the scaling properties of $\tilde\eta$ are weaker than those of $\tilde\xi$, the former measure does satisfy a strong continuity property under scaling, even for $d = 2$. The result is reminiscent of Lemma 13.40.

Theorem 13.54 *(scaling continuity) Let $\tilde\eta$ be the stationary cluster of a DW-process in R^2, and define a kernel $\nu : (0,\infty) \to \mathcal{M}_2$ by*

$$\nu(r) = |\log r|\, \mathcal{L}\Big(r^{-2}\, \tilde\eta S_r\Big), \quad r > 0.$$

Then the kernel $t \mapsto \nu\{\exp(-e^t)\}$ is continuous in total variation on H_B for every bounded B, uniformly on $[1,\infty)$.

Proof: For any $\varepsilon, r, h \in (0,1)$, let ζ_s denote the ancestral process of ξ_1 at times $s = 1 - h$, and let η_h^u be the h-clusters rooted at the associated atoms at u. Then

$$\begin{aligned} |\log \varepsilon|^{-1} \nu(\varepsilon) &\approx r^{-1} \mathcal{L}_{r\lambda^{\otimes 2}}\Big(\varepsilon^{-2}\, \xi_1 S_\varepsilon\Big) \\ &\approx r^{-1} E_{r\lambda^{\otimes 2}} \int \zeta_s(du)\, 1\Big\{\varepsilon^{-2}\, \eta_h^u S_\varepsilon \in \cdot\Big\} \\ &= \int \mathcal{L}_u\Big(\varepsilon^{-2}\, \eta_h S_\varepsilon\Big)\, du \\ &= \mathcal{L}\Big(\varepsilon^{-2} h\, \tilde\eta S_{\varepsilon/\sqrt{h}}\Big) \\ &\approx |\log \varepsilon|^{-1}\, \nu(\varepsilon/\sqrt{h}), \end{aligned}$$

where all relations are explained and justified below. The first equality holds, by the conditional independence of the clusters, along with the fact that

$E_{r\lambda^{\otimes 2}}\zeta_s = (r/h)\lambda^{\otimes 2}$. The second equality follows from Theorem 13.14 (i), by an elementary substitution.

To estimate the error in the first approximation, we see from Lemma 13.9 and Theorem 13.38 (ii) that, for $\varepsilon < 1/2$,

$$\begin{aligned} \Big\| r\,|\log\varepsilon|^{-1}\,\nu(\varepsilon) - \mathcal{L}_{r\lambda^{\otimes 2}}\big(\varepsilon^{-2}\,\xi_1 S_\varepsilon\big)\Big\|_B \\ &= \Big\|\mathcal{L}_{r\lambda^{\otimes 2}}(\eta_1 S_\varepsilon) - \mathcal{L}_{r\lambda^{\otimes 2}}(\xi_1 S_\varepsilon)\Big\|_B \\ &\lesssim \big(r\,P\{\tilde\eta(\varepsilon B) > 0\}\big)^2 \\ &\lesssim r^2\,|\log\varepsilon|^{-2}, \end{aligned}$$

where the η_h^i are h-clusters of ξ at time t. As for the second approximation, we get by Lemma 13.39 (ii), for small enough ε/h,

$$\begin{aligned} \Big\| E_{r\lambda^{\otimes 2}}\textstyle\sum_k 1\{\eta_h^k S_\varepsilon \in \cdot\} - \mathcal{L}_{r\lambda^{\otimes 2}}(\xi_1 S_\varepsilon)\Big\|_B \\ &\lesssim E_{r\lambda^{\otimes 2}}\Big(\textstyle\sum_k 1_+\big\{\eta_h^k(\varepsilon B)\big\} - 1\Big)_+ \\ &\lesssim \frac{|\log h|\,r\lambda^{\otimes 2}p_1 + \big(r\lambda^{\otimes 2}p_{t(h,\varepsilon)}\big)^2}{|\log\varepsilon|^2} \\ &= r\,\frac{|\log h| + r}{|\log\varepsilon|^2}. \end{aligned}$$

The third approximation relies on the estimate

$$\Big\|\nu(\varepsilon/\sqrt{h})\Big\|_B\left|\frac{|\log\varepsilon|}{|\log(\varepsilon/\sqrt{h})|} - 1\right| \lesssim \frac{|\log h|}{|\log\varepsilon|},$$

which holds for $\varepsilon \le h$, by the boundedness of ν. Combining these estimates and letting $r \to 0$, we obtain

$$\Big\|\nu(\varepsilon) - \nu(\varepsilon/\sqrt{h})\Big\|_B \lesssim \frac{|\log h|}{|\log\varepsilon|}, \quad \varepsilon \ll h < 1. \tag{34}$$

Putting $\varepsilon = e^{-u}$ and $\varepsilon/\sqrt{h} = e^{-v}$, and writing $\nu_A(x) = \nu(x, A)$ for measurable $A \subset H_B$, we get for $u - v \ll u$ (with $0/0 = 1$)

$$\begin{aligned} \left|\log\frac{\nu_A(e^{-u})}{\nu_A(e^{-v})}\right| &\lesssim \left|\frac{\nu_A(e^{-u})}{\nu_A(e^{-v})} - 1\right| \\ &\lesssim \big|\nu_A(e^{-u}) - \nu_A(e^{-v})\big| \\ &\lesssim \frac{u-v}{u} \lesssim \left|\log\frac{u}{v}\right|, \end{aligned}$$

and so for $u = e^s$ and $v = e^t$ (with $\infty - \infty = 0$),

$$\Big|\log\nu_A\big\{\exp(-e^t)\big\} - \log\nu_A\big\{\exp(-e^s)\big\}\Big| \lesssim |t - s|,$$

which extends immediately to arbitrary $s,t \ge 1$. Since ν is bounded on H_B, by Lemma 13.42, the function $\nu_A \circ \exp(-e^t)$ is again uniformly continuous on $[1,\infty)$, and the assertion follows, since all estimates are uniform in A. $\Box$

We finally prove a property of asymptotic age invariance for $\tilde\eta$, which should be compared with the exact invariance for $\tilde\xi$, in Theorem 13.52 (iv).

Corollary 13.55 *(asymptotic age invariance) Let $\varepsilon^2 \ll h \ll \varepsilon^{-2}$ when $d \ge 3$, and $|\log\varepsilon| \gg |\log h|$ when $d=2$. Then as $\varepsilon \to 0$,*

$$\left\|\mathcal{L}_\mu(\tilde\eta_h) - \mathcal{L}(\tilde\eta_1)\right\|_{B_0^\varepsilon} \ll \begin{cases} \varepsilon^{d-2}, & d \ge 3, \\ |\log\varepsilon|^{-1}, & d=2. \end{cases}$$

Proof: Letting $d \ge 3$, and fixing any $B \in \hat{\mathcal{B}}^d$, we get by Theorems 13.52 and 13.53

$$\begin{aligned} &\left\|\mathcal{L}(\tilde\eta S_\varepsilon) - r^{d-2}\mathcal{L}\Big(r^2\,\tilde\eta S_{\varepsilon/r}\Big)\right\|_B \\ &\quad \le \left\|\mathcal{L}(\tilde\eta S_\varepsilon) - \mathcal{L}(\tilde\xi S_\varepsilon)\right\|_B + r^{d-2}\left\|\mathcal{L}(\tilde\eta S_{\varepsilon/r}) - \mathcal{L}(\tilde\xi S_{\varepsilon/r})\right\|_B \\ &\quad \ll \varepsilon^{d-2} + r^{d-2}(\varepsilon/r)^{d-2} \lesssim \varepsilon^{d-2}. \end{aligned}$$

When $d=2$, we may use (34) instead, to get

$$\begin{aligned} &|\log\varepsilon|\left\|\mathcal{L}(\tilde\eta S_\varepsilon) - \mathcal{L}(r^2\,\tilde\eta S_{\varepsilon/r})\right\|_B \\ &\quad \lesssim \|\nu(\varepsilon) - \nu(\varepsilon/r)\|_B + \left|\frac{|\log\varepsilon|}{|\log(\varepsilon/r)|} - 1\right| \|\nu(\varepsilon/r)\|_B \\ &\quad \lesssim \frac{|\log r|}{|\log\varepsilon|} + \frac{|\log r|}{|\log\varepsilon| - |\log r|} \to 0. \end{aligned}$$

It remains to note that

$$r^{d-2}\mathcal{L}\Big(r^2\tilde\eta S_{1/r}\Big) = \mathcal{L}(\tilde\eta_{r^2}), \quad r>0,$$

by Theorem 13.53 (iii). $\Box$

13.9 Local Hitting and Conditioning

We may now establish some multi-variate local approximation and conditioning properties of the DW-process, similar to those for simple point processes in Section 6.6, and for regenerative sets and processes in Section 12.8. Our main result gives a multi-variate extension of the local approximations in Section 13.8, and provides at the same time an approximation of multi-variate Palm distributions by elementary conditional distributions.

We continue to use the notation, introduced at various points in previous sections. Thus, $\tilde\xi$ and $\tilde\eta$ denote the pseudo-random measures from Theorems

13.52 and 13.53, $q_{\mu,t}$ denotes the continuous density of $E_\mu \xi_t^{\otimes n}$, from Corollary 13.27, and $\mathcal{L}_\mu(\xi_t \,\|\, \xi_t^{\otimes n})_x$ denotes the regular, multi-variate Palm distributions, introduced in Theorem 13.34. The constants c_d and function $m(\varepsilon)$ are defined as in Theorem 13.42.

We write $f \sim g$ for $f/g \to 1$, $f \approx g$ for $f - g \to 0$, and $f \ll g$ for $f/g \to 0$. The notation $\|\cdot\|_B$, with associated terminology, was explained in Section 13.8. To achieve the desirable combination of conciseness and transparency, we will also use the scaling and shifting operators S_x^r, defined by $S_x^r B = rB + x$, so that $S_r = S_0^r$. Thus, μS_x^r is the measure obtained from μ by a magnification around x by a factor r^{-1}.

We may now state our main approximation theorem, which extends the univariate results in Section 13.8.

Theorem 13.56 *(multiple hitting and conditioning) Consider a DW-process ξ in R^d with $d \ge 2$, and let $\varepsilon \to 0$ for fixed $\mu, t > 0$ and open $G \subset \mathsf{R}^d$. Then*

(i) *for any $x \in (\mathsf{R}^d)^{(n)}$,*

$$P_\mu\{\xi_t^n B_x^\varepsilon > 0\} \sim q_{\mu,t}(x) \begin{cases} c_d^n \, \varepsilon^{n(d-2)}, & d \ge 3, \\ m_\varepsilon^n |\log \varepsilon|^{-n}, & d = 2, \end{cases}$$

(ii) *for any $x \in G^{(n)}$, and in total variation on $(B_0^1)^n \times G^c$,*

$$\mathcal{L}_\mu\Big(\{\xi_t S_{x_j}^\varepsilon\}_{j\le n}, \xi_t \,\Big|\, \xi_t^{\otimes n} B_x^\varepsilon > 0\Big) \approx \mathcal{L}^{\otimes n}\Big(\tilde\eta S_\varepsilon \Big| \tilde\eta B_0^\varepsilon > 0\Big) \otimes \mathcal{L}_\mu(\,\xi_t \| \,\xi_t^{\otimes n})_x,$$

(iii) *when $d \ge 3$, we have in the same sense*

$$\mathcal{L}_\mu\Big(\{\varepsilon^{-2}\xi_t S_{x_j}^\varepsilon\}_{j\le n}, \xi_t \,\Big|\, \xi_t^{\otimes n} B_x^\varepsilon > 0\Big) \approx \mathcal{L}^{\otimes n}\Big(\tilde\xi \,\Big|\, \tilde\xi B_0^1 > 0\Big) \otimes \mathcal{L}_\mu(\,\xi_t \| \,\xi_t^{\otimes n})_x.$$

Here (i) extends the asymptotic results for $n = 1$, in Theorems 13.52 and 13.53. Parts (ii)−(iii) show that the contributions of ξ_t to the sets $B_{x_1}^\varepsilon, \dots, B_{x_1}^\varepsilon$ and G^c are conditionally independent, asymptotically as $\varepsilon \to 0$. They further contain the multi-variate Palm approximation

$$\mathcal{L}_\mu\Big(1_{G^c}\xi_t \,\Big|\, \xi_t^{\otimes n} B_x^\varepsilon > 0\Big) \to \mathcal{L}_\mu\Big(1_{G^c}\xi_t \,\Big\|\, \xi_t^{\otimes n}\Big)_x, \quad x \in G^{(n)},$$

as well as the asymptotic equivalence or convergence on B_0^1,

$$\mathcal{L}_\mu\Big(\xi_t S_{x_j}^\varepsilon \,\Big|\, \xi_t^{\otimes n} B_x^\varepsilon > 0\Big) \begin{cases} \approx \mathcal{L}^{\otimes n}(\,\tilde\eta S_\varepsilon \,|\, \tilde\eta B_0^\varepsilon > 0), & d \ge 2, \\ \to \mathcal{L}(\,\tilde\xi \,|\, \tilde\xi B_0^1 > 0), & d \ge 3, \end{cases}$$

for any $x \in (\mathsf{R}^d)^{(n)}$, extending the versions for $n = 1$ in Section 13.8.

Several lemmas are needed for the proof. The following cluster approximation follows easily from the results in Section 13.6. As before, we write ζ_s for the ancestral process of ξ_t at time $s = t - h$, and let $\eta_h^{u_j}$ denote the h-cluster in ξ_t, rooted at the point u_j of ζ_s. To avoid repetitions, we restate from Section 13.6 the conditions

$$\begin{cases} \varepsilon^2 \ll h \le \varepsilon, & d \ge 3, \\ h \le |\log \varepsilon|^{-1} \ll |\log h|^{-1}, & d = 2. \end{cases} \tag{35}$$

Lemma 13.57 *(cluster approximation) For fixed t, x, and μ, consider a DW-process ξ in R^d, with h-clusters η_h^k at time t, along with a random element γ in T, and put $B = (B_0^1)^n$. Letting $\varepsilon, h \to 0$ subject to (35), we have for $d \geq 3$ and $d = 2$, respectively,*

$$\Big\| \mathcal{L}_\mu\Big(\{\xi_t S_{x_j}^\varepsilon\}_1^n, \gamma\Big) - E_\mu \int \zeta_s^{(n)}(du)\, 1\Big\{ \Big(\{\eta_h^{u_j} S_{x_j}^\varepsilon\}_1^n, \gamma\Big) \in \cdot \Big\} \Big\|_B \ll \begin{cases} \varepsilon^{n(d-2)}, \\ |\log \varepsilon|^{-n}. \end{cases}$$

Proof: It is enough to establish the corresponding bounds for

$$E_\mu \Big| f\Big(\{\xi_t S_{x_j}^\varepsilon\}_{j\le n}, \gamma\Big) - \int \zeta_s^{(n)}(du)\, f\Big(\{\eta_h^{u_j} S_{x_j}^\varepsilon\}_{j\le n}, \gamma\Big) \Big|,$$

uniformly for H_B-measurable functions f on $\mathcal{M}_d^n \times T$, with $0 \le f \le 1_{H_B}$. Writing Δ_h^ε for the stated absolute value, $U_{h,\varepsilon}$ for the union in Lemma 13.43, and κ_h^ε for the sum in Lemma 13.44, we note that

$$\Delta_h^\varepsilon \le \kappa_h^\varepsilon\, 1\{\kappa_h^\varepsilon > 1\} + 1\Big\{ \kappa_h^\varepsilon \le 1;\, U_{h,\varepsilon} \Big\}.$$

Since $k1\{k > 1\} \le 2(k-1)_+$ for any $k \in \mathsf{Z}_+$, we get

$$E_\mu \Delta_h^\varepsilon \le 2E_\mu(\kappa_h^\varepsilon - 1)_+ + P_\mu U_{h,\varepsilon},$$

which is of the required order, by Lemmas 13.43 and 13.44. □

We proceed with a property of asymptotic conditional independence.

Lemma 13.58 *(asymptotic independence) Given a DW-process ξ in R^d and some times t and $s = t - h$, choose $\tilde\xi_t \perp\!\!\!\perp_{\xi_s} \xi_t$ with $(\xi_s, \xi_t) \overset{d}{=} (\xi_s, \tilde\xi_t)$. For any open set $G \subset \mathsf{R}^d$, put $B = (B_0^1)^n \times G^c$, and fix any $x \in G^{(n)}$. Letting $\varepsilon, h \to 0$ subject to (35), we have*

$$\Big\| \mathcal{L}_\mu\Big(\{\xi_t S_{x_j}^\varepsilon\}_1^n, \xi_t\Big) - \mathcal{L}_\mu\Big(\{\xi_t S_{x_j}^\varepsilon\}_1^n, \tilde\xi_t\Big) \Big\|_B \ll \begin{cases} \varepsilon^{n(d-2)}, & d \ge 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

Proof: Putting $U_x^r = \bigcup_j B_{x_j}^r$, fix any $r > 0$ with $U_x^{2r} \subset G$, and write $\xi_t = \xi_t' + \xi_t''$ and $\tilde\xi_t = \tilde\xi_t' + \tilde\xi_t''$, where ξ_t' is the sum of all clusters in ξ_t rooted in U_x^r, and similarly for $\tilde\xi_t'$. Putting $D_n = \mathsf{R}^{nd} \setminus (\mathsf{R}^d)^{(n)}$, and letting ζ_s be the ancestral process of ξ_t at time s, we get

$$\begin{aligned}
&\Big\| \mathcal{L}_\mu\Big(\{\xi_t S_{x_j}^\varepsilon\}_1^n, \xi_t\Big) - \mathcal{L}_\mu\Big(\{\xi_t' S_{x_j}^\varepsilon\}_1^n, \xi_t''\Big) \Big\|_B \\
&\quad \le P_\mu\Big\{ \textstyle\prod_{i\le n} \xi_t B_{x_i}^\varepsilon > \prod_{i\le n} \xi_t' B_{x_i}^\varepsilon \Big\} + P_\mu\{\xi_t' G^c > 0\} \\
&\quad \le P_\mu\Big\{ \displaystyle\int_{D_n} \zeta_s^{\otimes n}(du) \textstyle\prod_{i\le n} \eta_h^{u_i} B_{x_i}^\varepsilon > 0 \Big\} \\
&\qquad + E_\mu \displaystyle\int_{(B_x^r)^c} \zeta_s^{(n)}(du) \prod_{j\le n} 1\{\eta_h^{u_j} B_{x_j}^\varepsilon > 0\} \\
&\qquad + P_\mu\Big\{ \displaystyle\int_{U_x^r} \zeta_s(du)\, \eta_h^u G^c > 0 \Big\} \ll \begin{cases} \varepsilon^{n(d-2)}, \\ |\log \varepsilon|^{-n}, \end{cases}
\end{aligned}$$

by Lemmas 13.43, 13.45, and 13.46. Since the last estimate depends only on the marginal distributions of the pairs $(\{\xi_t S^\varepsilon_{x_j}\}_{j\le n}, \xi_t)$ and $(\{\xi'_t S^\varepsilon_{x_j}\}_{j\le n}, \xi''_t)$, we get the same bound for

$$\Big\|\mathcal{L}_\mu\Big(\{\xi_t S^\varepsilon_{x_j}\}_{j\le n}, \tilde{\xi}_t\Big) - \mathcal{L}_\mu\Big(\{\xi'_t S^\varepsilon_{x_j}\}_{j\le n}, \tilde{\xi}''_t\Big)\Big\|_B.$$

It remains to note that $(\xi'_t, \xi''_t) \stackrel{d}{=} (\xi'_t, \tilde{\xi}''_t)$. □

We may now combine the various preliminary results, here and in previous sections, to establish our basic approximation lemma, which will lead rather easily to the main theorem. For brevity, we write $q_{\mu,t}(x) = q^x_{\mu,t}$.

Lemma 13.59 *(approximation) Consider a DW-process ξ in R^d, fix any μ, $t > 0$, and open $G \subset \mathsf{R}^d$, and put $B = (B^1_0)^n \times G^c$. Then as $\varepsilon \to 0$ for fixed $x \in G^{(n)}$, we have*

$$\Big\|\mathcal{L}_\mu\Big\{(\xi_t S^\varepsilon_{x_j})^n_1, \xi_t\Big\} - q^x_{\mu,t}\, \mathcal{L}^{\otimes n}(\tilde{\eta} S_\varepsilon) \otimes \mathcal{L}_\mu\Big(\xi_t \Big\| \, \xi_t^{\otimes n}\Big)_x\Big\|_B \ll \begin{cases} \varepsilon^{n(d-2)}, & d \ge 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

Proof: We may regard ξ_t as a sum of conditionally independent clusters η^u_h of age $h \in (0,t)$, rooted at the points u, of the ancestral process ζ_s at time $s = t - h$. Choose the random measure ξ'_t to satisfy

$$\xi'_t \perp\!\!\!\perp_{\xi_s} \Big\{\xi_t, \zeta_s, (\eta^u_h)\Big\}, \qquad (\xi_s, \xi_t) \stackrel{d}{=} (\xi_s, \xi'_t).$$

Our argument may be summarized as follows:

$$\begin{aligned} \mathcal{L}_\mu\Big\{(\xi_t S^\varepsilon_{x_j})_{j\le n}, \xi_t\Big\} &\approx E_\mu \int \zeta^{(n)}_s(du)\, 1_{(\cdot)}\Big\{(\eta^{u_j}_h S^\varepsilon_{x_j})_{j\le n}, \xi'_t\Big\} \\ &= \int E_\mu \xi^{\otimes n}_s(du) \bigotimes_{j\le n} \mathcal{L}_{u_j}(\,\eta_h S^\varepsilon_{x_j}) \otimes \mathcal{L}_\mu\Big(\xi_t \Big\| \, \xi^{\otimes n}_s\Big)_u \\ &\approx \Big(E_\mu \xi^{\otimes n}_s * p^{\otimes n}_h\Big)_x \mathcal{L}^{\otimes n}(\,\tilde{\eta}_h S_\varepsilon) \otimes \mathcal{L}_\mu\Big(\xi_t \Big\| \, \xi^{\otimes n}_t\Big)_x \\ &\approx q^x_{\mu,t}\, \mathcal{L}^{\otimes n}(\tilde{\eta} S_\varepsilon) \otimes \mathcal{L}_\mu\Big(\xi_t \Big\| \, \xi^{\otimes n}_t\Big)_x, \end{aligned} \tag{36}$$

where h and ε are related as in (35), and the approximations hold in the sense of total variation on H_B, of the order $\varepsilon^{n(d-2)}$ or $|\log \varepsilon|^{-n}$, respectively. We proceed with some detailed justifications:

The first relation in (36) is immediate from Lemmas 13.57 and 13.58. To justify the second relation, we may insert some intermediate steps:

$$\begin{aligned} &E_\mu \int \zeta^{(n)}_s(du)\, 1_{(\cdot)}\Big\{(\eta^{u_j}_h S^\varepsilon_{x_j})_{j\le n}, \xi'_t\Big\} \\ &\quad = E_\mu \int \zeta^{(n)}_s(du)\, \mathcal{L}_\mu\Big\{(\eta^{u_j}_h S^\varepsilon_{x_j})_{j\le n}, \xi'_t \,\Big|\, \xi_s, \zeta_s\Big\} \\ &\quad = h^n E_\mu \int \zeta^{(n)}_s(du) \bigotimes_{j\le n} \mathcal{L}_{u_j}(\eta_h S^\varepsilon_{x_j}) \otimes \mathcal{L}_\mu(\,\xi'_t |\, \xi_s) \end{aligned}$$

$$= E_\mu \int \xi_s^{(n)}(du) \bigotimes_{j\le n} \mathcal{L}_{u_j}(\eta_h S^\varepsilon_{x_j}) \otimes \mathcal{L}_\mu(\,\xi_t|\,\xi_s)$$
$$= E_\mu \int \xi_s^{(n)}(du) \bigotimes_{j\le n} \mathcal{L}_{u_j}(\eta_h S^\varepsilon_{x_j}) \otimes 1_{(\cdot)}(\xi_t)$$
$$= \int E_\mu \xi_s^{(n)}(du) \bigotimes_{j\le n} \mathcal{L}_{u_j}(\eta_h S^\varepsilon_{x_j}) \otimes \mathcal{L}_\mu\Big(\,\xi_t \,\Big\|\, \xi_s^{\otimes n}\Big)_u.$$

Here the first and fourths equalities hold by disintegration and Fubini's theorem. The second relation follows from the conditional independence for the h-clusters and the process ξ'_t, along with the normalization of $\mathcal{L}(\eta)$. The third relation holds by the choice of ξ'_t, and the moment relation $E_\mu(\zeta_s^{(n)}|\,\xi_s) = h^{-n}\xi_s^{\otimes n}$, from Theorem 6.27. The fifth relation holds by Palm disintegration.

To justify the third relation in (36), we may first examine the effect of changing the last factor. By Lemmas 1.20 and 13.38,

$$\Big\|\int E_\mu \xi_s^{\otimes n}(du) \bigotimes\nolimits_{j\le n} \mathcal{L}_{u_j}(\eta_h S^\varepsilon_{x_j}) \otimes \Big\{\mathcal{L}_\mu(\xi_t \|\, \xi_s^{\otimes n})_u - \mathcal{L}_\mu(\xi_t \|\, \xi_t^{\otimes n})_x\Big\}\Big\|_B$$
$$\lesssim \int E_\mu \xi_s^{\otimes n}(du)\, p_{h_\varepsilon}^{\otimes n}(x-u)\, \Big\|\mathcal{L}_\mu(\xi_t \|\, \xi_s^{\otimes n})_u - \mathcal{L}_\mu(\xi_t \|\, \xi_t^{\otimes n})_x\Big\|_{G_c} \begin{cases} \varepsilon^{n(d-2)}, \\ |\log\varepsilon|^{-n}, \end{cases}$$

where h_ε is defined as in Lemma 13.38, with t replaced by h. Choosing $r > 0$ with $B_x^{2r} \subset G^{(n)}$, we may estimate the integral on the right by

$$\Big(E_\mu \xi_s^{\otimes n} * p_{h_\varepsilon}^{\otimes n}(x)\Big) \sup_{u\in B_x^r} \Big\|\mathcal{L}_\mu(\,\xi_t \|\, \xi_s^{\otimes n})_u - \mathcal{L}_\mu(\,\xi_t \|\, \xi_t^{\otimes n})_x\Big\|_{G_c}$$
$$+ \int_{(B_x^r)^c} E_\mu \xi_s^{\otimes n}(du)\, p_{h_\varepsilon}^{\otimes n}(x-u).$$

Here the first term tends to 0, by Lemmas 13.29 and 13.36, whereas the second term tends to 0, as in the proof of Lemma 13.45. Hence, in the second line of (36), we may replace $\mathcal{L}_\mu(\xi_t \,\|\, \xi_s^{\otimes n})_u$ by $\mathcal{L}_\mu(\xi_t \,\|\, \xi_t^{\otimes n})_x$.

A similar argument, based on Lemma 1.20 and Corollary 13.55, allows us to replace $\mathcal{L}(\tilde\eta S_\varepsilon)$ in the last line by $\mathcal{L}(\tilde\eta_h S_\varepsilon)$. It then remains to prove that

$$\int E_\mu \xi_s^{\otimes n}(du) \bigotimes_{j\le n} \mathcal{L}_{u_j}(\eta_h S^\varepsilon_{x_j}) \approx q_t(x)\, \mathcal{L}^{\otimes n}(\tilde\eta_h S_\varepsilon),$$

where q_t denotes the continuous density of $E_\mu \xi_t^{\otimes n}$, from Corollary 13.27. Here the total variation distance may be expressed in terms of densities, as

$$\Big\|\int \Big\{q_s(x-u) - q_t(x)\Big\} \bigotimes\nolimits_{i\le n} \mathcal{L}_{u_i}(\eta_h S_\varepsilon)\, du\Big\|_B$$
$$\lesssim \int \Big|q_s(x-u) - q_t(x)\Big|\, p_{h_\varepsilon}^{\otimes n}(u)\, du \begin{cases} \varepsilon^{n(d-2)}, \\ |\log\varepsilon|^{-n}. \end{cases}$$

Letting γ be a standard normal random vector in R^{nd}, we may write the integral on the right as $E|q_s(x-\gamma h_\varepsilon^{1/2}) - q_t(x)|$. Here $q_s(x-\gamma h_\varepsilon^{1/2}) \to q_t(x)$ a.s., by the joint continuity of $q_s(u)$, and Lemma 13.29 yields

$$E q_s\Big(x - \gamma h_\varepsilon^{1/2}\Big) = \Big(q_s * p_{h_\varepsilon}^{\otimes n}\Big)(x) \to q_t(x).$$

The former convergence then extends to L^1, by FMP 1.32, and the required approximation follows. □

Proof of Theorem 13.56: (i) For any $x \in (\mathsf{R}^d)^{(n)}$, Lemma 13.59 yields

$$\left| P_\mu\left\{\xi_t^{\otimes n} B_x^\varepsilon > 0\right\} - q_{\mu,t}^x \left(P\{\tilde\eta B_0^\varepsilon > 0\}\right)^n \right| \ll \begin{cases} \varepsilon^{n(d-2)}, & d \geq 3, \\ |\log \varepsilon|^{-n}, & d = 2. \end{cases}$$

It remains to note that, by Theorem 13.42,

$$P\{\tilde\eta B_0^\varepsilon > 0\} \sim \begin{cases} c_d\, \varepsilon^{d-2}, & d \geq 3 \\ m(\varepsilon)\, |\log \varepsilon|^{-1}, & d = 2. \end{cases}$$

(ii) Assuming $x \in G^{(n)}$, and using (i) and Lemma 13.59, we get in total variation on $(B_0^1)^n \times G^c$

$$\begin{aligned} \mathcal{L}_\mu\left\{ (\, \xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t \,\middle|\, \xi_t^{\otimes n} B_x^\varepsilon > 0 \right\} &= \frac{\mathcal{L}_\mu\left\{ (\xi_t S_{x_j}^\varepsilon)_{j \leq n}, \xi_t \right\}}{P_\mu\{\xi_t^{\otimes n} B_x^\varepsilon > 0} \\ &\approx \frac{q_{\mu,t}^x\, \mathcal{L}^{\otimes n}(\tilde\eta S_\varepsilon) \otimes \mathcal{L}_\mu\left(\xi_t \parallel \xi_t^{\otimes n}\right)_x}{q_{\mu,t}^x \left(P\{\tilde\eta B_0^\varepsilon > 0\}\right)^n} \\ &= \mathcal{L}^{\otimes n}\left(\tilde\eta S_\varepsilon \,\middle|\, \tilde\eta B_0^\varepsilon > 0\right) \otimes \mathcal{L}_\mu\left(\xi_t \,\middle\|\, \xi_t^{\otimes n}\right)_x . \end{aligned}$$

(iii) When $d \geq 3$, Theorem 13.52 (i) yields

$$\varepsilon^{2-d} \mathcal{L}\left(\varepsilon^{-2} \tilde\eta S_\varepsilon\right) \to \mathcal{L}(\tilde\xi),$$

in total variation on B_0^1. Hence,

$$\mathcal{L}\left(\varepsilon^{-2}\, \tilde\eta S_\varepsilon \,\middle|\, \tilde\eta B_0^\varepsilon > 0\right) \to \mathcal{L}\left(\tilde\xi \,\middle|\, \tilde\xi B_0^1 > 0\right),$$

and the assertion follows by means of (ii). □

13.10 Stability of Discrete Clustering

Returning to the discrete-time case, we will now study the repeated clustering of a stationary point process on R^d, generated by a probability kernel $\nu\colon \mathsf{R}^d \to \mathcal{N}_d$. Recall that ν is said to be *shift-invariant*, if

$$\nu_r = \nu_0 \circ \theta_r^{-1}, \quad r \in \mathsf{R}^d.$$

For convenience, we may then introduce a point process χ on R^d with distribution ν_0, and say that ν is *critical*, if $E\|\chi\| = 1$. The simplest choice is to take χ to be a mixed binomial process, so that the spatial motion becomes independent of family size, though most of our results cover even the general case.

Let ν^n denote the n-th iterate of ν, and choose a point process χ_n with distribution ν^n. Given a point process ξ on R^d, let $\xi_0, \xi_1, \ldots$ be the generated branching process. Let κ_n^r denote the number of clusters in ξ_n, contributing to the ball B_0^r. If ξ is stationary with intensity c, then clearly

$$E\kappa_n^r = c\int P\{\chi_n B_x^r > 0\}\, dx < \infty, \quad n \in \mathsf{N},\ r > 0. \tag{37}$$

Lemma 13.60 *(monotonicity) Let ν be critical, and fix any $r > 0$. Then $E\kappa_n^r$ is non-increasing in n, and the limit is positive, simultaneously for all r.*

Proof: We may form χ_{n+1} by iterated clustering in n steps, starting from χ. Assuming χ to be simple, we may write χ_n^u for the associated clusters rooted at u. Using the conditional independence, Fubini's theorem, the invariance of λ^d, and the criticality of ν, we obtain

$$\begin{aligned}
\int P\Big\{\chi_{n+1} B_x^r > 0\Big\}\, dx &= \int P\Big\{\int \chi(du)\, \chi_n^u B_x^r > 0\Big\}\, dx \\
&\le \int dx\, E\int \chi(du)\, 1\{\chi_n^u B_x^r > 0\} \\
&= \int dx \int E\chi(du)\, P\{\chi_n B_{x-u}^r > 0\} \\
&= \int E\chi(du) \int P\{\chi_n B_{x-u}^r > 0\}\, dx \\
&= \int E\chi(du) \int P\{\chi_n B_x^r > 0\}\, dx \\
&= \int P\{\chi_n B_x^r > 0\}\, dx.
\end{aligned}$$

For general χ, the same argument applies to any uniform randomizations of the processes χ_n. The last assertion is clear, since $E\chi_n^r$ is non-decreasing in r, and B_0^r is covered by finitely many balls B_x^1. □

Motivated by the last lemma, we say that ν is *stable*, if

$$\lim_{n\to\infty} \int P\{\chi_n B_x^r > 0\}\, dx > 0, \quad r > 0, \tag{38}$$

and *unstable* otherwise. The importance of stability is due in part to the following facts. Here we say that a point process ξ on R^d is ν-*invariant* (in distribution), if the generated cluster process ξ_n is stationary in n.

Theorem 13.61 *(persistence/extinction dichotomy) Fix a critical and invariant cluster kernel ν on R^d. Then each of these conditions is equivalent to the stability of ν:*

(i) *For every $c > 0$, there exists a stationary and ν-invariant point process ξ on R^d with intensity c.*

(ii) *For any stationary point process ξ on R^d with finite intensity, the cluster process (ξ_n) generated by ξ and ν is locally uniformly integrable.*

If the stability fails, then no point process ξ as in (i) *exists, and the processes ξ_n in* (ii) *satisfy $\xi_n \xrightarrow{vd} 0$.*

Property (ii) means that the random variables $\xi_n B$ are uniformly integrable, for every bounded set $B \in \mathcal{B}^d$. To appreciate this condition, we note that if $\xi_n \xrightarrow{vd} \xi_\infty$ along a sub-sequence, then $E\xi_\infty = E\xi$ when ν is stable, whereas otherwise $\xi_\infty = 0$ a.s. Under suitable regularity conditions, there is even convergence along the entire sequence. For the moment, we can only give a partial proof, to be completed after the next lemma.

Partial proof: Let ξ be a stationary point process on R^d with intensity $c \in (0, \infty)$, generating a cluster process (ξ_n). If ν is unstable, then (37) yields $\kappa_n^r \xrightarrow{P} 0$, and so $\xi_n \xrightarrow{vd} 0$. In particular, no ν-invariant point process exists with intensity in $(0, \infty)$. Furthermore, ξ_n is not uniformly integrable, since that would imply $E\xi = E\xi_n \xrightarrow{v} 0$, and therefore $\xi = 0$ a.s.

Now let ν be stable. Letting (ξ_n) be generated by a stationary Poisson process ξ on R^d with intensity 1, we introduce some processes η_n with

$$\mathcal{L}(\eta_n) = n^{-1} \sum_{k=1}^{n} \mathcal{L}(\xi_k), \quad n \in \mathsf{N}.$$

Then the η_n are again stationary with intensity 1, and so by Theorem 4.10 we have convergence $\eta_n \xrightarrow{vd} \tilde{\eta}$, along a sub-sequence $N' \subset \mathsf{N}$. The limit $\tilde{\eta}$ is again stationary, and $E\tilde{\eta} \le \lambda^d$, by FMP 4.11. Since

$$\left\| \mathcal{L}(\eta_n) - \mathcal{L}(\eta_{n-1}) \right\| \le (n-1)\left(\frac{1}{n-1} - \frac{1}{n} \right) + \frac{1}{n} = \frac{2}{n} \to 0,$$

we have even $\eta_{n-1} \xrightarrow{vd} \tilde{\eta}$, and so $\tilde{\eta}$ is ν-invariant.

To see that $E\tilde{\eta} \neq 0$, fix any $r > 0$ with $\tilde{\eta}\, \partial B_0^r = 0$ a.s., and let $p > 0$ be the corresponding limit in (38). Then Theorem 3.2 (i) yields

$$P\{\xi_n B_0^r = 0\} = \exp\Big(-\int P\{\chi_n B_x^r > 0\}\, dx \Big) \to e^{-p},$$

and so

$$P\{\eta_n B_0^r = 0\} = n^{-1} \sum_{k=1}^{n} P\{\xi_k B_0^r = 0\} \to e^{-p}.$$

Letting $n \to \infty$ along N', we obtain $P\{\tilde{\eta} B_0^r = 0\} = e^{-p} < 1$, which shows that $E\tilde{\eta} \neq 0$. Since the intensities of ξ and $\tilde{\eta}$ are proportional, the assertion in (i) follows. $\Box$

To complete the proof of (ii), we need the following lemma, which will also play a crucial role in the sequel. For any measure $\mu \in \mathcal{M}_d$ and constants $r, k > 0$, we introduce the truncated measure

$$\mu^{r,k}(dx) = 1\{\mu B_x^r \le k\}\, \mu(dx), \quad r, k > 0.$$

Lemma 13.62 *(truncation criteria, Debes et al.) Let ν be a critical and invariant cluster kernel on R^d, and fix any $r > 0$. Then ν is stable iff*

$$\lim_{k\to\infty} \inf_{n\ge 1} E\|\chi_n^{r,k}\| = 1,$$

and otherwise

$$\lim_{n\to\infty} E\|\chi_n^{r,k}\| = 0, \quad k > 0.$$

Proof: If ν is stable, then Theorem 13.61 (i) yields a stationary and ν-invariant point process ξ on R^d with intensity in $(0,\infty)$. By stationarity and truncation of ξ_n or χ_n, we have for any $r, k, n > 0$

$$E\xi^{r,k} = E\xi_n^{r,k} \le E\|\chi_n^{r,k}\|\, E\xi.$$

As $k \to \infty$, we get $E\xi^{r,k} \xrightarrow{v} E\xi$, by monotone convergence, and so

$$E\xi \le \lim_{k\to\infty} \inf_{n\ge 1} E\|\chi_n^{r,k}\|\, E\xi.$$

The first condition now follows, as we cancel the measure $E\xi$ on both sides.

Using Fubini's theorem, the invariance of λ^d, and the fact that $u \in B_x^r$ iff $x \in B_u^r$, we get for any $r, k, n > 0$

$$\begin{aligned}
\int P\{\chi_n B_x^r > 0\}\, dx &\ge \int P\Big\{\chi_n^{r,k} B_x^r > 0\Big\}\, dx \\
&\ge k^{-1}\int E\chi_n^{r,k} B_x^r\, dx \\
&= k^{-1} E\int \chi_n^{r,k}(du) \int 1\{x \in B_u^r\}\, dx \\
&= k^{-1} E\|\chi_n^{r,k}\|\, \lambda^d B_0^r.
\end{aligned}$$

If ν is unstable, then the left-hand side tends to 0 as $n \to \infty$, and the second condition follows. □

End of proof of Theorem 13.61: It remains to prove (ii) when ν is stable. By monotone convergence, we may replace ξ by $\xi^{r,k}$, for some fixed $r, k > 0$, and by Lemma 13.62, we may also replace χ_n by $\chi_n^{r,k}$. Letting κ_n be the number of truncated clusters in ξ_n hitting B_0^r, we have $\xi_n B_0^r \le k\,\kappa_n$, and so it suffices to prove the uniform integrability of (κ_n). By conditional independence, κ_n is the total mass of a p_n-thinning of the measure $\xi'(dx) = \xi(-dx)$, where

$$p_n^{r,k}(x) = P\Big\{\chi_n^{r,k} B_x^r > 0\Big\}, \quad x \in \mathsf{R}^d,$$

and so

$$E(\kappa_n)^2 \le E(\xi' p_n^{r,k})^2 + E\xi' p_n^{r,k}.$$

Writing g_r for the uniform probability density on B_0^r, we note that $p_n^{r,k} \le p_n^{2r,k} * g_r$, and so

$$\begin{aligned}
\xi' p_n^{r,k} &\le \xi'(p_n^{2r,k} * g_r) \\
&= (\xi' * g_r) p_n^{2r,k} \lesssim \lambda^d p_n^{2r,k} \\
&\lesssim E\xi_n B_0^{2r} \lesssim \lambda^d B_0^{2r}.
\end{aligned}$$

Thus, $E(\kappa_n)^2$ is bounded, which implies the required uniform integrability. □

Letting $\nu_n = \mathcal{L}(\chi_n)$, we may write the associated cluster iteration as

$$\nu_{m+n} = \nu_m \circ \nu_n, \quad m, n \in \mathsf{N}.$$

For every $n \in \mathsf{N}$, we introduce a point process η_n, with the centered Palm distribution of ν_n, given by

$$Ef(\eta_n) = E\int f(\theta_{-x}\chi_n)\,\chi_n(dx), \quad n \in \mathsf{N}.$$

The corresponding reduced Palm distributions $\nu_n^0 = \mathcal{L}(\eta_n - \delta_0)$ satisfy a similar recursive property:

Lemma 13.63 *(Palm recursion) For every $n \in \mathsf{N}$, let $\nu_n = \mathcal{L}(\chi_n)$, with associated reduced, centered Palm distribution ν_n^0. Then*

$$\nu_{m+n}^0 = (\nu_m^0 \circ \nu_n) * \nu_n^0, \quad m, n \in \mathsf{N}.$$

Proof: For any cluster kernels ν_1 and ν_2 on R^d, we need to show that

$$(\nu_1 \circ \nu_2)^0 = (\nu_1^0 \circ \nu_2) * \nu_2^0. \tag{39}$$

When $d = 0$, the measures ν_i are determined by the generating functions $f_i(s) = Es^{\chi_i}$ on $[0, 1]$, and the kernel composition $\nu = \nu_1 \circ \nu_2$ becomes equivalent to $f = f_1 \circ f_2$. The reduced Palm measure ν_i^0 has generating function $E\eta_i s^{\eta_i - 1} = f'(s)$, and so, in this case, (39) is equivalent to the chain rule for differentiation, in the form $(f_1 \circ f_2)' = (f_1' \circ f_2) f_2'$. The general case follows by a suitable conditioning. □

By the last result, we may construct recursively an increasing sequence of Palm trees η_n, along with a random walk $\zeta = (\zeta_n)$ in R^d, based on the distribution $\rho = E\chi$, such that each η_n is rooted at $-\zeta_n$, and $\Delta\eta_n \perp\!\!\!\perp_{\zeta_n} \eta_{n-1}$ for every n. The resulting tree $\eta = (\eta_n)$ is essentially a discrete-time counterpart of the univariate Palm tree from Corollary 13.22. We proceed to express the stability criterion of Lemma 13.62 in terms of the infinite Palm tree η.

Theorem 13.64 *(Palm tree criteria) Let ν be a critical and invariant cluster kernel on R^d, with associated Palm processes η_n, and fix any $r > 0$. Then ν is stable iff*

$$\sup_n \eta_n B_0^r < \infty \quad a.s.,$$

and otherwise

$$\eta_n B_0^r \to \infty \quad a.s.$$

Proof: From the definitions of $\chi_n^{r,k}$ and η_n, we see that

$$\begin{aligned} E\|\chi_n^{r,k}\| &= E\int 1\{\chi_n B_x^r \le k\}\,\chi_n(dx) \\ &= P\{\eta_n B_0^r \le k\}, \end{aligned}$$

and so, by the monotonicity of η_n, we get as $k \to \infty$

$$\begin{aligned} \inf_n E\|\chi_n^{r,k}\| &= \inf_n P\{\eta_n B_0^r \le k\} \\ &= P\bigcap\nolimits_n \{\eta_n B_0^r \le k\} \\ &= P\{\eta_\infty B_0^r \le k\} \\ &\to P\{\eta_\infty B_0^r < \infty\}. \end{aligned}$$

The stated criteria now follow by Lemma 13.62. □

Using the Palm tree structure in Lemma 13.63, we can derive some more explicit criteria for stability. Here we introduce the intensity measure $\rho = E\chi$, along with the generating function $f(s) = Es^\kappa$ of the offspring distribution, where $\kappa = \|\chi\|$, and define

$$g(s) = 1 - f'(1-s), \quad s \in [0,1]. \tag{40}$$

For every n, we may write

$$\Delta\eta_n = \sum_{j \le \kappa_n} \theta_{-\zeta_{n-1}+\gamma_{nj}}\, \chi_{n-1}^j, \qquad \beta_n = \sum_{j \le \kappa_n} \delta_{\gamma_{nj}},$$

where $\mathcal{L}(\beta_n) = \nu_1^0$, and the χ_{n-1}^j are independent with distribution ν_{n-1}. In the mixed binomial case, even the displacements γ_{nj} are clearly independent with distribution ρ.

Theorem 13.65 *(conditioning criteria) Let ν be a critical and invariant cluster kernel on R^d, and fix any $r > 0$. Then*

(i) *ν is stable iff*

$$\sum_n E\Big(1 - \prod\nolimits_{j \le \kappa_n} P\Big\{\chi_{n-1} B_{(\cdot)}^r = 0\Big\}_{\zeta_{n-1}-\gamma_{nj}} \,\Big|\, \zeta_{n-1}\Big) < \infty \;\; a.s.,$$

(ii) *when ν is mixed binomial, it is equivalent that*

$$\sum_n g \circ P\Big\{\theta_\gamma \chi_{n-1}(B_0^r + \zeta_n) > 0 \,\Big|\, \zeta_n\Big\} < \infty \;\; a.s.$$

If the stability fails, then the series in (i)−(ii) *are a.s. divergent.*

Proof: (i) Theorem 13.64 shows that ν is stable iff

$$\sum\nolimits_n \Delta\eta_n B_0^r < \infty \;\text{ a.s.}$$

Since the $\Delta\eta_n$ are integer valued, this is equivalent to

$$P\Big\{\Delta\eta_n B_0^r > 0 \;\text{ i.o.}\Big\} = 0. \tag{41}$$

Noting that the sequence $(\Delta\eta_n)$ is adapted to the filtration $(\mathcal{F}_n)$, induced by the sequence of pairs (η_n, ζ_n), we may use Lévy's extended version of the

Borel–Cantelli lemma (FMP 7.20), along with the conditional independence $\Delta\eta_n \perp\!\!\!\perp_{\zeta_{n-1}} \mathcal{F}_{n-1}$, to obtain the equivalent condition

$$\sum_n P\Big(\Delta\eta_n B_0^r > 0 \,\Big|\, \zeta_{n-1}\Big) < \infty \text{ a.s.}$$

By the chain rule for conditioning, this is in turn equivalent to

$$\sum_n E\Big(1 - P\Big\{\Delta\eta_n B_0^r = 0 \,\Big|\, \zeta_{n-1}, \beta_n\Big\} \,\Big|\, \zeta_{n-1}\Big) < \infty \text{ a.s.}$$

Using the conditional independence and Fubini's theorem, we get

$$\begin{aligned} P\Big(\Delta\eta_n B_0^r = 0 \,\Big|\, \zeta_{n-1}, \beta_n\Big) &= \prod_{j\le\kappa} P\Big\{\chi_{n-1}(B_0 + \zeta_{n-1} - \gamma_{nj}) = 0 \,\Big|\, \zeta_{n-1}, \gamma_{nj}\Big\} \\ &= \prod_{j\le\kappa} P\Big\{\chi_{n-1} B_{(\cdot)}^r = 0\Big\}_{\zeta_{n-1}-\gamma_{nj}}. \end{aligned}$$

(ii) When ν is mixed binomial, the increments $\Delta\eta_n$ are conditionally independent, given $\zeta = (\zeta_n)$. Applying the elementary Borel–Cantelli lemma to the conditional distributions, given ζ, and using the conditional independence $\Delta\eta_n \perp\!\!\!\perp_{\zeta_n} \zeta$, we see that (41) is now equivalent to

$$\sum_n P\Big(\Delta\eta_n B_0^r > 0 \,\Big|\, \zeta_n\Big) < \infty \text{ a.s.,}$$

which can be written as

$$\sum_n E\Big(1 - P\Big\{\Delta\eta_n B_0^r = 0 \,\Big|\, \zeta_n, \kappa_n\Big\} \,\Big|\, \zeta_n\Big) < \infty \text{ a.s.}$$

Since in this case

$$P\Big(\Delta\eta_n B_0^r = 0 \,\Big|\, \zeta_{n-1}, \kappa_n\Big) = \Big(P\Big\{\theta_\gamma \chi_{n-1}(B_0^r + \zeta_n) = 0 \,\Big|\, \zeta_n\Big\}\Big)^{\kappa_n},$$

the desired criterion follows.

If ν is unstable, Theorem 13.64 shows that the probability in (41) equals 1 instead of 0. Using the converse of either version of the Borel–Cantelli lemma, and proceeding as above, we see that the stated series diverges a.s. Alternatively, we may express the processes η and ζ, in the obvious way, in terms of some i.i.d. random elements, and then apply the Hewitt–Savage $0-1$ law (FMP 3.15). $\square$

When the spatial motion is independent of family size, the last result yields a simple sufficient condition for stability:

Corollary 13.66 *(random walk criterion) Let ν be a critical and invariant cluster kernel on R^d, such that ν_0 is mixed binomial. Consider a random walk (ζ_n) based on ρ, and fix any $r > 0$. Then ν is stable, whenever*

$$\sum_n g \circ \rho_n(B_0^r + \zeta_n) < \infty \text{ a.s.}$$

Proof: Since $\gamma \perp\!\!\!\perp \chi_{n-1}$ with $\mathcal{L}(\gamma) = \rho$ and $E\chi_{n-1} = \rho_{n-1}$, Fubini's theorem yields

$$\begin{aligned} P\Big\{\theta_\gamma \chi_{n-1} B_x^r > 0\Big\} &\leq E\theta_\gamma \chi_{n-1} B_x^r \\ &= (\rho * \rho_{n-1}) B_x^r = \rho_n B_x^r. \end{aligned}$$

Since also $\zeta_n \perp\!\!\!\perp (\gamma, \chi_{n-1})$, we may use Fubini's theorem again to get a.s.

$$P\Big\{\theta_\gamma \chi_{n-1}(B_0^r + \zeta_n) > 0 \,\Big|\, \zeta_n\Big\} \leq \rho_n(B^r + \zeta_n). \qquad \Box$$

Though in general, the stability of ν can not be characterized in terms of ρ and f only, we may still give some sufficient conditions, involving only ρ and g. Let $\tilde{\rho}$ denote the symmetrization of $\rho = E\chi$, and define $\rho_n = \rho^{*n}$ and $\tilde{\rho}_n = \tilde{\rho}^{*n}$. Let $(\tilde{\zeta}_n)$ be a symmetrized random walk, based on $\tilde{\rho}$. Writing φ for the characteristic function of ρ, we define

$$I_n(\varepsilon) = \int_{B_0^\varepsilon} |\varphi(t)|^n \, dt, \quad \varepsilon > 0, \; n \in \mathsf{Z}_+.$$

Corollary 13.67 *(series criterion) Let ν be a critical and invariant cluster kernel on R^d, and fix any $r > 0$. Then ν is stable, whenever*

(i) $\sum_n g(\sup_x \rho_n B_x^r) < \infty$.

If even $E\kappa^2 < \infty$, then $g(s) \asymp s$, and (i) *becomes equivalent to each of the conditions*

(ii) $\sum_n \tilde{\rho}_n B_0^r < \infty$,

(iii) *the random walk $(\tilde{\zeta}_n)$ is transient,*

(iv) $\displaystyle\int_{B_0^\varepsilon} \frac{dt}{1 - |\varphi(t)|} < \infty$.

Proof: Noting that, for any $x \in \mathsf{R}^d$,

$$\begin{aligned} P\{\chi_{n-1} B_x^r > 0\} &\leq E\chi_{n-1} B_x^r \\ &\leq \sup_x \rho_{n-1} B_x^r, \end{aligned}$$

and further that $\kappa \perp\!\!\!\perp \zeta_{n-1}$, we see that the n-th term in Lemma 13.65 (i) is bounded by

$$1 - E\Big(1 - \sup_x \rho_{n-1} B_x^r\Big)^\kappa = g\Big(\sup_x \rho_{n-1} B_x^r\Big).$$

Thus, (i) implies the stability of ν.

If $E\kappa^2 < \infty$, then

$$g'(0) = f''(1) = E\kappa(\kappa - 1) < \infty.$$

Since g is also concave with $g(0) = 0$, we have $g(s) \asymp s$ in this case, and so (i) is equivalent to the corresponding condition without g. Even (ii) is then equivalent, since by Lemma A1.5,

$$I_{2n}(\varepsilon) \lesssim \tilde{\rho}_n B_0^r \leq \sup_x \rho_n B_x^r \lesssim I_n(\varepsilon),$$

uniformly for fixed $r, \varepsilon > 0$ and $d \in \mathsf{N}$. The same relations show that even (iv) is equivalent, since

$$\begin{aligned}\sum_{n \geq 0} I_n(\varepsilon) &= \int_{B_0^\varepsilon} dt \sum_{n \geq 0} |\varphi(t)|^n \\ &= \int_{B_0^\varepsilon} \frac{dt}{1 - |\varphi(t)|},\end{aligned}$$

by Fubini's theorem. Finally, Theorem 12.3 yields (ii) $\Leftrightarrow$ (iii). $\square$

Appendices

To avoid some technical distractions in the main text, we list here some auxiliary results needed from different areas, ranging from the elementary to the profound. For the latter, we are of course omitting the proofs.

A1. Measure and Probability

We begin with some deep results from general measure theory. Recall that a set or function is said to be *universally measurable* (or *u-measurable* for short), if it is measurable with respect to every completion of the underlying σ-field.

Theorem A1.1 *(projection and section) For any Borel spaces* $(S, \mathcal{S})$ *and* $(T, \mathcal{T})$*, let* $A \in \mathcal{S} \otimes \mathcal{T}$ *be arbitrary with projection* πA *onto* S*. Then*

(i) πA *is u-measurable,*

(ii) *there exists a u-measurable function* $f\colon S \to T$*, such that* $\{s, f(s)\} \in A$ *for all* $s \in \pi A$*.*

Proof: See Dudley (1989), Theorems 13.2.1, 13.2.6, and 13.2.7. □

Corollary A1.2 *(graph, range, and inverse) Let* S *and* T *be Borel. Then for any measurable function* $f\colon S \to T$*, we have*

(i) *the graph* $A = \{(s, f(s));\ s \in S\} \subset S \times T$ *is product-measurable,*

(ii) *the range* $f(S) = \{f(s);\ s \in S\} \subset T$ *is u-measurable,*

(iii) *there exists a u-measurable function* $g\colon T \to S$ *with* $f \circ g \circ f = f$*.*

Proof: (i) The diagonal $D = \{(t, t);\ t \in T\} \subset T^2$ is product-measurable since T is Borel, and $A = \tilde{f}^{-1} D$, where $\tilde{f}(s, t) = \{f(s), t\}$.

(ii) Since A has T-projection $f(S)$, this holds by (i) and Theorem A1.1 (i).

(iii) By (i) and Theorem A1.1 (ii), there exists a u-measurable function $g\colon T \to S$, such that $\{g(t), t\} \in A$ for all $t \in f(S)$. Then for any $s \in S$, we have $\{g \circ f(s), f(s)\} \in A$, which means that $f \circ g \circ f(s) = f(s)$. □

For the next result, recall that the random variables ξ_n are said to be *uniformly integrable*, if

$$\lim_{r \to \infty} \sup_{n \geq 1} E(|\xi_n|;\ |\xi_n| > r) = 0.$$

They are further said to *converge weakly in* L^1 toward a limit ξ, if $E\xi_n \alpha \to E\xi\alpha$ for every bounded random variable α.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7

Lemma A1.3 *(weak compactness, Dunford) Any uniformly integrable sequence of random variables has a sub-sequence that converges weakly in L^1.*

Proof: Let (ξ_n) be uniformly integrable. For any fixed $k \in \mathsf{N}$, the sequence $\xi_n^k = \xi_n 1\{|\xi_n| \leq k\}$ is bounded, and hence weakly compact in L^2. By a diagonal argument, we may choose a sub-sequence $N' \subset \mathsf{N}$, such that ξ_n^k converges weakly in L^2, in $n \in N'$ for fixed k, toward a random variable η_k. In particular,

$$\begin{aligned}
\|\eta_k - \eta_l\|_1 &= E(\eta_k - \eta_l)_+ + E(\eta_l - \eta_k)_+ \\
&\leftarrow E\Big(\xi_n^k - \xi_n^l;\, \eta_k > \eta_l\Big) + E\Big(\xi_n^l - \xi_n^k;\, \eta_l > \eta_k\Big) \\
&\leq \|\xi_n^k - \xi_n^l\|_1 \\
&\leq E\Big(|\xi_n|;\, |\xi_n| > k \wedge l\Big),
\end{aligned}$$

which tends to 0 as $k, l \to \infty$, uniformly in n. Hence, the sequence (η_k) is Cauchy in L^1, and so it converges in L^1 toward some random variable ξ. For any random variable α with $|\alpha| \leq 1$, we get

$$\Big|E\xi_n\alpha - E\xi\alpha\Big| \leq E|\xi_n - \xi_n^k| + \Big|E\xi_n^k\alpha - E\eta_k\alpha\Big| + E|\eta_k - \xi|,$$

which tends to 0, as $n \to \infty$ along N', and then $k \to \infty$. □

We proceed with a general statement about conditional independence[1], needed in Section 12.4.

Lemma A1.4 *(conditional independence) Consider some σ-fields $\mathcal{F}_i, \mathcal{G}_i, \mathcal{H}_i$ and sets $H_i \in \mathcal{H}_i$, $i = 1, 2$, satisfying*

$$\begin{aligned}
&(\mathcal{F}_1, \mathcal{G}_1, \mathcal{H}_1) \perp\!\!\!\perp (\mathcal{F}_2, \mathcal{G}_2, \mathcal{H}_2), \\
&\mathcal{F}_i \underset{\mathcal{H}_i}{\perp\!\!\!\perp} \mathcal{G}_i \ \text{ on } H_i, \quad i = 1, 2.
\end{aligned}$$

Then $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}_1, \mathcal{G}_2$ are conditionally independent on $H_1 \cap H_2$, given $\mathcal{H}_1 \vee \mathcal{H}_2$.

Proof: By a suitable randomization on the complements H_i^c, we may reduce to the case where the relations $\mathcal{F}_i \perp\!\!\!\perp_{\mathcal{H}_i} \mathcal{G}_i$ hold everywhere. Proceeding recursively, as in FMP 3.8, it suffices to show that

$$\mathcal{F}_1 \underset{\mathcal{H}_1, \mathcal{H}_2}{\perp\!\!\!\perp} \mathcal{G}_1, \qquad (\mathcal{F}_1, \mathcal{G}_1) \underset{\mathcal{H}_1, \mathcal{H}_2}{\perp\!\!\!\perp} \mathcal{F}_2, \qquad (\mathcal{F}_1, \mathcal{G}_1, \mathcal{F}_2) \underset{\mathcal{H}_1, \mathcal{H}_2}{\perp\!\!\!\perp} \mathcal{G}_2.$$

Here it is enough to prove the last relation, since the first two are equivalent to special cases. We may then start from the formulas

$$\mathcal{F}_2 \underset{\mathcal{H}_2}{\perp\!\!\!\perp} \mathcal{G}_2, \qquad (\mathcal{F}_1, \mathcal{G}_1, \mathcal{H}_1) \underset{\mathcal{F}_2, \mathcal{H}_2}{\perp\!\!\!\perp} \mathcal{G}_2,$$

[1] Recall that $\mathcal{F} \perp\!\!\!\perp_{\mathcal{H}} \mathcal{G}$ denotes conditional independence of $\mathcal{F}$ and $\mathcal{G}$, given $\mathcal{H}$. By $\mathcal{F} \vee \mathcal{G}$ we mean the σ-field generated by $\mathcal{F}$ and $\mathcal{G}$.

where the first relation holds by hypothesis, and the second one follows, by the claim (i) ⇒ (ii) in FMP 6.8, from the independence of the two triples. Using the reverse implication (ii) ⇒ (i) in FMP 6.8, we conclude that

$$(\mathcal{F}_1, \mathcal{G}_1, \mathcal{H}_1, \mathcal{F}_2) \underset{\mathcal{H}_2}{\perp\!\!\!\perp} \mathcal{G}_2,$$

and the desired relation follows, by yet another application of FMP 6.8. □

The *concentration function* of a probability measure μ on R^d is given by $f(r) = \sup_x \mu B_x^r$, $r > 0$, where B_x^r denotes the r-ball centered at x. For any distribution μ on R^d with symmetrization $\tilde{\mu}$, we introduce the characteristic function $\hat{\mu}$, and define

$$I_n(\varepsilon) = \int_{t<\varepsilon} |\hat{\mu}(t)|^n \, dt, \quad \varepsilon > 0, \ n \in \mathsf{N}.$$

Lemma A1.5 *(concentration functions, Esseen) For fixed $r, \varepsilon > 0$ and $d \in \mathsf{N}$, we have, uniformly in μ and n,*

$$I_{2n}(\varepsilon) \lesssim \tilde{\mu}^{*n} B_0^r \le \sup_x \mu^{*n} B_x^r \lesssim I_n(\varepsilon).$$

Proof: See Petrov (1995), pp. 25–26. □

We conclude with a classical estimate of concentration functions, for sums of independent random variables.

Theorem A1.6 *(concentration of sums, Kolmogorov, Rogozin) Let $\xi_1, \ldots, \xi_n$ be independent random variables with symmetrizations $\tilde{\xi}_1, \ldots, \tilde{\xi}_n$. Then for any constants $t, t_1, \ldots, t_n > 0$, we have the uniform bound*

$$P\Big\{ \Big| \sum\nolimits_k \xi_k \Big| \le t \Big\} \lesssim t \Big(\sum\nolimits_k t_k^2 P\big\{ |\tilde{\xi}_k| \ge t_k \big\} \Big)^{-1/2}.$$

Proof: See Petrov (1995), pp. 63–68. □

A2. Stochastic Processes

The general *Doléans exponential* plays a key role in stochastic calculus, for semi-martingales with jump discontinuities. Here we need only the elementary special case for non-decreasing processes.

Theorem A2.1 *(Doléans exponential) For any locally finite measure μ on $(0, \infty)$ with $\Delta\mu \le 1$, the equation*[2] *$X = 1 - X_- \cdot \mu$ has the unique solution*

$$X_t = \exp(-\mu_t^c) \prod_{s \le t} (1 - \mu\{s\}), \quad t \ge 0. \tag{1}$$

[2]Recall that $f \cdot \mu$ denotes the function $\int_0^{t+} f \, d\mu$, and that $f_-(s) = f(s-)$ whenever the left-hand limits exist.

Proof: For X as in (1), an easy computation gives $\Delta X_t = -X_{t-}\mu\{t\}$. When μ is diffuse, (1) simplifies to $X_t = e^{-\mu_t}$, which yields $dX_t = -X_{t-}d\mu_t$. Hence, by combination, $X = 1 - X_- \cdot \mu$ when μ has finitely many atoms, and the general result follows by dominated convergence. To prove the uniqueness, let X^1 and X^2 be any two solutions, and note that $X = X^1 - X^2$ has finite variation, and satisfies the homogeneous equation $X = -X_- \cdot \mu$ with $X_0 = 0$. If $\tau = \inf\{t \geq 0;\, X_t \neq 0\}$ is finite, we may choose $h > 0$ so small that $p_h = \mu(\tau, \tau + h] < 1$, and put $M_h = \sup_{s \leq \tau + h} |X_s|$. Then $M_h \leq p_h M_h$, and so by iteration $M_h = 0$. This contradicts the definition of τ, and $X \equiv 0$ follows. □

Given a process X in a space S, and some random times $\sigma \leq \tau$, we define the *restriction* Y of X to $[\sigma, \tau]$ as the process $Y_s = X_{\sigma+s}$ for $s \leq \tau - \sigma$, and $Y_s = \Delta$ for $s > \tau - \sigma$, where $\Delta \notin S$. By a *Markov time* for X, we mean a random time $\sigma \geq 0$, such that the restrictions of X to $[0, \sigma]$ and $[\sigma, \infty)$ are conditionally independent, given X_σ.

Theorem A2.2 *(Brownian excursion, Williams) Let X be a Brownian excursion, conditioned to reach height $t > 0$, let σ and τ be the first and last times that X visits t, and write ρ for the first time that X attains its minimum on $[\sigma, \tau]$. Then X_ρ is $U(0, t)$, and σ, ρ, and τ are Markov times for X.*

Proof: See Williams (1974) or Le Gall (1986). □

By a *local sub-martingale* we mean a process X, which reduces to a uniformly integrable sub-martingale through a suitable localization. It is said to be of *class* (D), if the family $\{X_\tau\}$ is uniformly integrable, where τ ranges over the class of all optional times $\tau < \infty$.

Lemma A2.3 *(uniform integrability) Any local sub-martingale X with $X_0 = 0$ is locally of class (D).*

Proof: By localization, we may assume that X is a true sub-martingale. For any $n \in \mathsf{N}$, consider the optional time $\tau = n \wedge \inf\{t > 0;\, |X_\tau| > n\}$. Then $|X^\tau| \leq n \vee |X_\tau|$, where X_τ is integrable. Hence, X^τ is of class (D). □

We proceed with a version of the multi-variate ergodic theorem. For any convex set $B \subset \mathsf{R}^d$, we define the *inner radius* $r(B)$ as the maximum radius of all open balls contained in B.

Theorem A2.4 *(multi-variate ergodic theorem, Wiener) Let R^d act measurably on a space S, let $\mathcal{I}$ denote the shift-invariant σ-field in S, and let ξ be a stationary random element in S. Consider an increasing sequence of bounded, open, convex sets $B_n \subset \mathsf{R}^d$ with $r(B_n) \to \infty$. Then for any measurable function $f \geq 0$ on S,*

$$(\lambda^d B_n)^{-1} \int_{B_n} f(\theta_r \xi)\, dr \to E\{f(\xi)\,|\,\mathcal{I}_\xi\} \;\; a.s.$$

This holds even in L^p with $p \geq 1$, whenever $f(\xi) \in L^p$.

Proof: See FMP 10.14. □

The stated L^p-version holds under much weaker conditions. Say that the probability measures $\mu_1, \mu_2, \ldots$ on R^d are *asymptotically invariant*, if $\|\mu_n - \theta_r \mu_n\| \to 0$ for every $r \in \mathsf{R}^d$, where the norm $\|\cdot\|$ denotes total variation.

Theorem A2.5 *(mean ergodic theorem) Let $\mu_1, \mu_2, \ldots$ be asymptotically invariant probability measures on R^d. For ξ and f as before, fix a $p \geq 1$ with $f(\xi) \in L^p$. Then*

$$\int f(\theta_r \xi)\, \mu_n(dr) \to E\{f(\xi)\,|\,\mathcal{I}_\xi\} \ \text{ in } L^p.$$

Proof: See FMP 10.20. □

We proceed with the powerful group coupling theorem.

Theorem A2.6 *(shift coupling, Thorisson) Let G be an lcscH group, acting measurably on a space S, and consider some random elements ξ and η in S, such that $\xi \stackrel{d}{=} \eta$ on the G-invariant σ-field $\mathcal{I}$ in S. Then $\gamma\xi \stackrel{d}{=} \eta$, for some random element γ in G.*

Proof: See FMP 10.28. □

We conclude with an elementary extension property for exchangeable sequences, needed in Section 12.7.

Lemma A2.7 *(exchangeable sequences) Let P and $P_1, P_2, \ldots$ be probability measures on Ω with $\|P_n - P\| \to 0$, and consider some finite or infinite sequences $\xi^n = (\xi_j^n)$ of random variables, each exchangeable under P_n, such that $\xi_j^n \to \xi_j$, as $n \to \infty$ for fixed j. Then $\xi = (\xi_j)$ is exchangeable under P.*

Proof: It is enough to consider sequences of fixed length $m < \infty$. Then the P-exchangeability of ξ is equivalent to the relations

$$E \exp \sum_{k \leq m} i t_k \xi_k = E \exp \sum_{k \leq m} i t_{p_k} \xi_k,$$

for arbitrary $t_1, \ldots, t_m \in \mathsf{R}$ and permutations $p_1, \ldots, p_m$ of $1, \ldots, m$. Every such equation can be written as $Ef(\xi) = 0$, for some bounded, continuous function f on R^m. Writing E_n for expectation under P_n, we get by dominated convergence

$$\begin{aligned} |Ef(\xi)| &= \Big|Ef(\xi) - E_n f(\xi_n)\Big| \\ &\leq \|f\|\, \|P - P_n\| + E\Big|f(\xi) - f(\xi_n)\Big| \to 0, \end{aligned}$$

which implies $Ef(\xi) = 0$. □

A3. Normal Distributions

Here we list some elementary estimates for normal densities, needed in Chapter 13. Let p_t denote the symmetric normal density in R^d with variances $t > 0$.

Lemma A3.1 *(normal comparison) Let ν_Λ denote the centered normal distribution on R^n with covariance matrix Λ. Then*

$$\nu_\Lambda \le \left(\frac{\|\Lambda\|^n}{\det\Lambda}\right)^{1/2} \nu_{\|\Lambda\|}^{\otimes n}.$$

Proof: Let $\lambda_1 \le \cdots \le \lambda_n$ be the eigenvalues of Λ, and write $(x_1, \ldots, x_n)$ for the associated coordinate representation of $x \in \mathsf{R}^n$. Then ν_Λ has density

$$\begin{aligned}
\prod_{k\le n} p_{\lambda_k}(x_k) &= \prod_{k\le n} (2\pi\lambda_k)^{-1/2}\, e^{-x_k^2/2\lambda_k} \\
&\le \prod_{k\le n} (\lambda_n/\lambda_k)^{1/2}\, p_{\lambda_n}(x_k) \\
&= \left(\frac{\lambda_n^n}{\lambda_1\cdots\lambda_n}\right)^{1/2} p_{\lambda_n}^{\otimes n}(x),
\end{aligned}$$

and the assertion follows since $\|\Lambda\| = \lambda_n$ and $\det(\Lambda) = \lambda_1 \cdots \lambda_n$. □

Lemma A3.2 *(maximum inequality) The normal densities p_t on R^d satisfy*

$$p_s(x) \le \left(1 \vee td\,|x|^{-2}\right)^{d/2} p_t(x), \quad 0 < s \le t, \ \ x \in \mathsf{R}^d \setminus \{0\}.$$

Proof: For fixed $t > 0$ and $x \ne 0$, the maximum of $p_s(x)$ for $s \in (0, t]$ occurs when $s = (|x|^2/d) \wedge t$. This gives $p_s(x) \le p_t(x)$ for $|x|^2 \ge td$, whereas for $|x|^2 \le td$, we have

$$\begin{aligned}
p_s(x) &\le \left(2\pi|x|^2/d\right)^{-d/2} e^{-d/2} \\
&\le \left(2\pi|x|^2/d\right)^{-d/2} e^{-|x|^2/2t} \\
&= \left(td\,|x|^{-2}\right)^{d/2} p_t(x).
\end{aligned}$$ □

Lemma A3.3 *(density comparison) Let the p_t be standard normal densities on R^d. Then for fixed $d \in \mathsf{N}$ and $T > 0$,*

$$p_t(x + y) \lesssim p_{t+h}(x), \quad x \in \mathsf{R}^d, \ |y| \le h \le t \le T.$$

Proof: When $|x| \ge 4t$ and $|y| \le h$, we have $|y|/|x| \le h/4t$, and so for $r = h/t \le 1$,

$$\begin{aligned}
\frac{|x+y|^2}{t} \cdot \frac{(t+h)}{|x|^2} &\ge \left(1 - \frac{|y|}{|x|}\right)^2 \left(1 + \frac{h}{t}\right) \\
&\ge \left(1 - \frac{r}{2}\right)(1 + r) \ge 1,
\end{aligned}$$

which implies $p_t(x+y) \lesssim p_{t+h}(x)$ when $h \le t$. The same relation holds trivially for $|x| \le 4t$ and $|y| \le h \le t$. □

For any measure $\mu \in \mathcal{M}_d$ and measurable functions $f_1, \dots, f_n \ge 0$ on R^d, we consider the convolution

$$\Big\{\mu * \bigotimes_{k\le n} f_k\Big\}(x) = \int \mu(du) \prod_{k\le n} f_k(x_k - u), \quad x \in (\mathsf{R}^d)^n.$$

Lemma A3.4 *(convolution density) Let $\mu \in \mathcal{M}_d$ with $\mu p_t < \infty$ for all $t > 0$. Then for every $n \in \mathsf{N}$, the function $(\mu * p_t^{\otimes n})(x)$ is finite and jointly continuous in $x \in (\mathsf{R}^d)^n$ and $t > 0$.*

Proof: Fixing any $c > 1$, and letting $t \in [c^{-1}, c]$ and $x \in (\mathsf{R}^d)^n$ with $|x| < c$, we get by Lemma A3.3

$$\begin{aligned}\prod_{k\le n} p_t(x_k - u) &\lesssim \prod_{k\le n} p_c(x_k - u)\\ &\lesssim p_{2c}^n(u) \lesssim p_{2c/n}(u),\end{aligned}$$

uniformly in $u \in \mathsf{R}^d$. Since $\mu p_{2c/n} < \infty$, we obtain $(\mu * p_t^{\otimes n})(x) < \infty$ for any $t > 0$, and the asserted continuity follows, by dominated convergence, from the fact that $p_t(x)$ is jointly continuous in $x \in \mathsf{R}^d$ and $t > 0$. □

A4. Algebra and Geometry

The following elementary estimate is needed in Chapter 13.

Lemma A4.1 *(product estimate) For any $n \in \mathsf{N}$ and $k_1, \dots, k_n \in \mathsf{Z}_+$, we have*

$$2\Big\{\prod_{i\le n} k_i - 1\Big\}_+ \le \sum_{i\le n} (k_i - 1) \prod_{j\le n} k_j.$$

Proof: Clearly,

$$(hk-1)_+ \le h(k-1)_+ + k(h-1)_+, \quad h, k \in \mathsf{Z}_+.$$

Proceeding by induction on n, we obtain

$$\Big\{\prod_{i\le n} k_i - 1\Big\}_+ \le \sum_{i\le n} (k_i - 1)_+ \prod_{j\ne i} k_j.$$

It remains to note that

$$(k-1)_+ \le \frac{k(k-1)}{2}, \quad k \in \mathsf{Z}_+.$$ □

By the *principal variances* of a random vector, we mean the positive eigenvalues of the associated covariance matrix.

Lemma A4.2 *(principal variances) Given a partition $\pi \in \mathcal{P}_n$, consider some uncorrelated random vectors ξ_J, $J \in \pi$, in R^d, with uncorrelated entries of variance σ^2, and put $\xi_j = \xi_J$ for $j \in J \in \pi$. Then the array $(\xi_1, \dots, \xi_n)$ has principal variances $\sigma^2|J|$, $J \in \pi$, each with multiplicity d.*

Proof: By scaling, we may take $\sigma^2 = 1$, and since the ξ_J have uncorrelated components, we may further take $d = 1$. Writing "$\sim$" for the equivalence relation induced by π, we get $\mathrm{Cov}(\xi_i, \xi_j) = 1\{i \sim j\}$. It remains to note that the $m \times m$ matrix with entries $a_{ij} = 1$ has eigenvalues $m, 0, \dots, 0$. □

For any convex set $B \subset \mathsf{R}^d$, let $r(B)$ be the maximum radius of all open balls contained in B. Writing $\partial_\varepsilon B$ for the ε-neighborhood of the boundary ∂B of B, we put $\partial_\varepsilon^+ B = \partial_\varepsilon B \setminus B$ and $\partial_\varepsilon^- B = \partial_\varepsilon B \cap B$.

Lemma A4.3 *(convex sets) For any bounded, convex set $B \subset \mathsf{R}^d$ with $r(B) > 0$, we have*

$$\lambda^d(\partial_\varepsilon B) \le 2\left(\left\{1 + \frac{\varepsilon}{r(B)}\right\}^d - 1\right)\lambda^d B.$$

Proof: By scaling and centering, we may assume that $r(B) = 1$ and $B_0^1 \subset B$. Then by convexity,

$$\begin{aligned} B^\varepsilon &= B + \varepsilon B_0^1 \\ &\subset B + \varepsilon B = (1+\varepsilon)B, \end{aligned}$$

and so $\lambda^d B^\varepsilon \le (1+\varepsilon)^d \lambda^d B$. The relation $\lambda^d \partial_\varepsilon^- B \le \lambda^d \partial_\varepsilon^+ B$ is obvious for convex polyhedra, and it follows in general by an outer approximation. This yields a similar estimate for the inner neighborhood, and the assertion follows. □

We proceed with a classical covering theorem in R^d, later to be extended to Riemannian manifolds.

Theorem A4.4 *(covering, Besicovitch) For any $d \in \mathsf{N}$, there exists an $m_d \in \mathsf{N}$, such that for any class $\mathcal{C}$ of non-degenerate closed balls in R^d with uniformly bounded radii, there exist some countable classes $\mathcal{C}_1, \dots, \mathcal{C}_{m_d} \subset \mathcal{C}$ of disjoint balls, such that the centers of all balls in $\mathcal{C}$ lie in $\bigcup_k \mathcal{C}_k$.*

Proof: See Bogachev (2007), pp. 361–366. □

This implies a basic differentiation theorem, which will also be extended in the next section.

Corollary A4.5 *(differentiation basis) The centered balls in R^d form a differentiation basis, for any locally finite measure on R^d.*

Proof: See Bogachev (2007), pp. 367–369. □

A5. Differential Geometry

A *Riemannian manifold* is defined as a differential manifold M, equipped with a bilinear form, given in local coordinates $x = (x^1, \ldots, x^d)$ by a smooth family of symmetric, positive definite $d \times d$ matrices $g_{ij}(x)$. The length of a smooth curve in M is obtained by integration of the *length element ds* along the curve, where ds^2 is given in local coordinates by

$$ds^2(x) = \sum\nolimits_{i,j} g_{ij}(x)\, dx^i dx^j. \tag{2}$$

The distance between two points $x, y \in M$ is defined as the minimum length of all smooth curves connecting x and y. This yields a metric ρ on M, which in turn determines the open balls B_x^r in M of radius r, centered at x. When M is orientable, we may also form a *volume element dv* in M, given in local coordinates by

$$dv(x) = \Big\{ \det g_{ij}(x) \Big\}^{1/2} (dx^1 \cdots dx^d). \tag{3}$$

Integrating dv over Borel sets yields a measure λ on M with Lebesgue density $(\det g_{ij})^{1/2}$ on local charts. To summarize:

Theorem A5.1 *(Riemannian manifold) Let M be a d-dimensional Riemannian manifold with Riemannian metric g, given in local coordinates by the matrices $g_{ij}(x)$. Then the length and volume elements in* (2) *and* (3) *determine a metric ρ on M with associated open balls B_x^r, and a volume measure λ on M with Lebesgue density* $(\det g_{ij})^{1/2}$*. Both ρ and λ are independent of the choice of local coordinates.*

Proof: See Boothby (1986), pp. 186, 188–89, 219. □

The covering and differentiation properties in Theorem A4.4 and its corollary extend to any Riemannian manifold. This fact is needed in Section 7.5.

Theorem A5.2 *(covering and differentiation) Besicovitch's covering Theorem A4.4 and its corollary extend to any compact subset B of a Riemannian manifold M. Furthermore, the centered balls form a differentiation basis, for any locally finite measure on B.*

Proof: To extend the covering theorem to M, fix any $x_0 \in M$, and consider any relatively compact neighborhood $G \subset M$ of x_0, covered by a single chart. Choosing local coordinates in G with $g_{ij}(x_0) = \delta_{ij}$, we may approximate the Riemannian balls in G by Euclidean balls of the same radius. The proof in Bogachev (2007) remains valid with obvious changes, for covers by Riemannian balls. See also Heinonen (2001), Thm. 1.14 and Ex. 1.15c, for the restrictions to any compact subset $B \subset M$. The differentiation property now follows, as in the Euclidean case. □

A *Lie group* is a topological group G, endowed with a differentiable manifold structure, such that the group operations are smooth functions on G.

Theorem A5.3 *(Riemannian metric on a Lie group) Every Lie group G is an orientable manifold. Any inner product on the basic tangent space generates a left-invariant Riemannian metric g on G. The induced volume measure λ is a left Haar measure on G.*

Proof: See Boothby (1986), p. 247. □

Lie groups, of fundamental importance in their own right, are also significant as the basic building blocks of general locally compact groups. This is crucial for our developments in Sections 7.5–6. A topological group G is said to be a *projective limit* of Lie groups, if every neighborhood of the identity element ι contains a compact, invariant subgroup H, such that the quotient group G/H is isomorphic to a Lie group.

Theorem A5.4 *(projective limits) Let the topological group G be a projective limit of Lie groups. Then G contains some compact, invariant subgroups $H_n \downarrow \{\iota\}$, such that the quotient groups G/H_n are isomorphic to Lie groups. The projection maps $\pi_n : r \mapsto rH_n$ are then continuous.*

Proof: See Montgomery & Zippin (1955), pp. 27, 177. □

We conclude with the powerful representation of locally compact groups in terms of Lie groups.

Theorem A5.5 *(representation of lcscH groups) Every lcscH group G contains an open subgroup H, which is a projective limit of Lie groups. The cosets of H are again open, the projection map $\pi : r \mapsto rH$ is both continuous and open, and the coset space G/H is discrete and countable.*

Proof: See Montgomery & Zippin (1955), pp. 28, 54, 175. □

A6. Analysis

A function f on R_+ is said to be *completely monotone* if

$$(-1)^n \Delta_h^n f(t) \geq 0, \quad t, h, n \geq 0,$$

where $\Delta_h f(t) = f(t+h) - f(t)$, and then recursively

$$\Delta_h^0 f(t) = f(t), \quad \Delta_h^{n+1} f(t) = \Delta_h \Big\{ \Delta_h^n f(t) \Big\}, \quad t, h, n \geq 0,$$

all differences being with respect to t. The definition is the same for functions on $[0, 1]$, apart from the obvious restrictions on t, h, and n. The following characterizations are classical:

Theorem A6.1 *(completely monotone functions, Hausdorff, Bernstein) Let f be a function on $I = \mathsf{R}_+$ or $[0,1]$ with $f(0) = f(0+) = 1$. Then f is completely monotone iff*

(i) *when $I = \mathsf{R}_+$, $f(t) \equiv Ee^{-\rho t}$ for some random variable $\rho \ge 0$,*

(ii) *when $I = [0,1]$, $f(t) \equiv E(1-t)^\kappa$ for some random variable κ in Z_+.*

Proof: See Feller (1971), pp. 223–224, 439–440. □

The following simple approximation is needed in Chapter 13.

Lemma A6.2 *(interpolation) Fix any functions $f, g > 0$ on $(0,1]$ and constants $p, c > 0$, such that f is non-decreasing, $\log g(e^{-t})$ is bounded and uniformly continuous on R_+, and $t^{-p} f(t)\, g(t) \to c$, as $t \to 0$ along any sequence (r^n), with r restricted to some dense set $D \subset (0,1)$. Then the convergence remains valid along $(0,1)$.*

Proof: Writing w for the modulus of continuity of $\log g(e^{-t})$, we get

$$\begin{aligned} e^{-w(h)} g(e^{-t}) &\le g(e^{-t-h}) \\ &\le e^{w(h)} g(e^{-t}), \quad t, h \ge 0. \end{aligned}$$

Putting $b_r = \exp w(-\log r)$, we obtain

$$b_r^{-1} g(t) \le g(rt) \le b_r\, g(t), \quad t, r \in (0,1).$$

For any $r, t \in (0,1)$, we define $n = n(r,t)$ by $r^{n+1} < t \le r^n$. Then the monotonicity of f yields

$$\begin{aligned} r^p (r^{n+1})^{-p} f(r^{n+1})\, b_r^{-1} g(r^{n+1}) &\le t^{-p} f(t)\, g(t) \\ &\le r^{-p} (r^n)^{-p} f(r^n)\, b_r\, g(r^n). \end{aligned}$$

Letting $t \to 0$ for fixed $r \in D$, we get by the hypothesis

$$\begin{aligned} r^p\, b_r^{-1} c &\le \liminf_{t\to 0} t^{-p} f(t)\, g(t) \\ &\le \limsup_{t\to 0} t^{-p} f(t)\, g(t) \\ &\le r^{-p}\, b_r\, c. \end{aligned}$$

It remains to note that $r^{-p}\, b_r \to 1$, as $r \to 1$ along D. □

Historical and Bibliographical Notes

Here I am only commenting on work directly related to material in the main text. No completeness is claimed, and I apologize for any unintentional errors or omissions. The following abbreviations are often used for convenience:

$\mathrm{MKM}_{1,2,3}$	Matthes et al. (1974/78/82)
$\mathrm{DVJ}_{1,2}$	Daley & Vere-Jones (2003/08)
K(- -)	Kallenberg (19- -) or (20- -)
$\mathrm{RM}_{1,2}$	K(75/76) and K(83/86)
$\mathrm{FMP}_{1,2}$	K(97/02)

1. Spaces, Kernels, and Disintegration

1.1. Mazurkiewicz (1916) proved that a topological space is Polish iff it is homeomorphic to a G_δ (a countable intersection of open sets) in $[0,1]^\infty$. The fact that that every uncountable Borel set in a Polish space is Borel isomorphic to 2^∞ was proved independently by Alexandrov (1916) and Hausdorff (1916). Careful proofs of both results appear in Dudley (1989), along with further remarks and references. Prohorov (1956) noted that the space of probability measures on a Polish space is again Polish under the weak topology, which implies the Borel property in Theorem 1.5.

Monotone-class theorems were first established by Sierpiński (1928). Their use in probability theory goes back to Halmos (1950), Loève (1955), and Dynkin (1961). The usefulness of s-finite measures was recognized by Sharpe (1988) and Getoor (1990). Dissection and covering systems of various kinds have been employed by many authors, including Leadbetter (1972), RM_1, p. 11, and DVJ_1, pp. 282, 382.

1.2. Factorial measures and associated factorial moment measures were formally introduced in RM_2, pp. 109f, though related ideas have been noted by earlier authors, as detailed in DVJ_1, Section 5.4. Our first factorial decomposition is implicit in previous work, whereas the second one may be new.

1.3. Kernels are constantly used in all areas of modern probability theory. Their basic properties were listed in FMP 1.40–41, and some further discussion appears in Sharpe (1988). The use of Laplace transforms to verify the kernel property goes back to RM_1 1.7.

1.4. Disintegration of probability measures on a product space is a classical subject, arising naturally in the context of conditioning. Here Doob

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7

(1938) proved the existence of regular conditional distributions. The general disintegration formula, constantly needed throughout probability theory, may have been stated explicitly for the first time in FMP 6.4. It is often used without proof or established repeatedly in special cases. A version for Palm measures was proved in RM_1, and the general disintegration theorem was given in RM_2 15.3.3. In the context of stationary point processes, it appears in MKM_2, pp. 228, 311.

Doob (1953) noted the martingale approach to the Radon-Nikodym theorem, which yields the existence of product-measurable kernel densities. Meyer & Yor (1976) showed that the density result fails without some regularity condition on T. The associated construction of product-measurable disintegration densities was given in FMP 7.26, and the general kernel disintegration appears in K(10). Partial disintegrations go back to RM_2 13.1. Universal densities and disintegrations were first obtained in K(14).

1.5. Differentiation of measures is a classical subject, going back to Lebesgue (1904). More general differentiation results are usually derived from Vitali's (1908) covering theorem. The differentiation property of general dissection systems was established by de Possel (1936), and the simple martingale proof of his result was noted by Doob (1953). The lifting theorem was proved by von Neumann (1931). Extensive surveys of classical differentiation theory appear in Bogachev (2007), Bruchner et al. (1997), and Shilov & Gurevich (1966). The local comparison and decoupling results were first obtained, with different proofs, in K(09).

2. Distributions and Local Structure

2.1. Foundations of random measure and point process theory have been developed by many authors, including Harris (1968), $\mathrm{MKM}_{1,2}$, Jagers (1974), RM_1, and DVJ_2, mostly assuming the underlying space S to be Polish or even locally compact. An exception is Moyal (1962), and more recently Last & Penrose (2017), who studied the atomic and counting properties of point processes on an abstract measurable space. The Borel assumption on S was used systematically in FMP.

Laplace and characteristic functionals, first introduced by Kolmogorov (1935), were applied to point processes by Kendall (1949) and Bartlett & Kendall (1951), and to general random measures by Mecke (1968) and von Waldenfels (1968). Elementary cases of Theorem 2.8, going back to Korolyuk and Dobrushin, as quoted by Khinchin (1955), have been discussed by many authors, including Volkonsky (1960), Belyaev (1969), Leadbetter (1972), and Daley (1974). Related results for processes in $D[0,1]$ appear in Section 4.5 of Gikhman & Skorokhod (1965). The present extended version was given in Ch. 2 of RM_1.

Further remarks on the early history of point processes appear in DVJ_1

and Guttorp & Thorarinsdottir (2012). General accounts of point process theory have been given by $\mathrm{MKM}_{1,2}$, $\mathrm{RM}_{1,2}$, $\mathrm{DVJ}_{1,2}$, Neveu (1977), Cox & Isham (1980), Brémaud (1981), Jabobsen (2006), and others, and special subareas have been surveyed by many authors, often as part of a broader subject. Statistical applications have been considered by Liptser & Shiryaev (1977/78) and Krickeberg (1982). Apart from the pioneering work of Jagers (1974) and some discussion in DVJ_2, my earlier book $\mathrm{RM}_{1,2}$ has remained, until now, the only comprehensive account of general random measure theory.

2.2. The surprising fact that the distribution of an integral ξX is not determined by the joint distribution of ξ and X was noted in K(14), which also contains the associated uniqueness theorem, as well as the extension property of the absolute continuity $\xi \ll \eta$. The notion of L^p-intensity was introduced in K(79), which contains the given characterization and representation.

2.3. Doob (1938) proved the existence of regular conditional distributions, and Harris (1968/71) extended the result to any additive processes. Avoidance functions of simple point processes were characterized by Karbe (1973) and Kurtz (1974). Alternative approaches to both results appear in MKM_2, pp. 19ff, DVJ_2, pp. 30ff, and $\mathrm{RM}_{1,2}$, pp. 41ff.

Maxitive processes have been studied by Kendall (1974), Norberg (1984), and Vervaat (1997). Choquet (1953–54) explored the deep connection between random sets and alternating capacities, using methods from functional analysis. The equivalent version for additive processes may be new, though a related but weaker result appears in Norberg (1984). General expositions of random set theory are given by Matheron (1975), Molchanov (2005), and Schneider & Weil (2008).

3. Poisson and Related Processes

3.1. The classical Poisson approximation of binomial probabilities was first obtained by de Moivre (1711–12), and Poisson (1837) noted how the limiting probabilities can be combined into distributions on Z_+. Homogeneous Poisson processes on R_+ were constructed already by Ellis (1844) as renewal processes with exponential holding times. He also calculated the distributions of the n-th events, and derived an associated central limit theorem. However, he never deduced the Poisson distributions of the counting variables, and it is unclear whether he realized that the constructed process has independent increments.

This was later done by Bateman (1910), consultant of Rutherford and Geiger in their studies of radioactive decay, who derived the Poisson property by solving a system of differential equations. Further Poisson studies at about the same time were motivated by applications to risk theory and telecommunication. Thus, Lundberg (1903) set up a difference/differential equation, which upon solution yields the desired Poisson distributions for the increments. His obscure writing was partly clarified by CRAMÉR (1969) (see the colorful passage from an earlier review of Cramér translated in MARTIN-LÖF (1995)). ERLANG (1909) derived (semi-heuristically) the Poisson distribution of the counting variables from the classical binomial approximation. (He admits that he is only reviewing some unpublished work of a certain Mr. F. Johannsen, director of the Copenhagen telephone company.) The same method was used by LÉVY (1934), who also used a time-change argument to extend the result to the inhomogeneous case, also treated by COPELAND & REGAN (1936). Lévy went on to characterize general stochastically continuous processes with independent increments, and in the homogeneous case, ITÔ (1942a) recognized the jump point process as a stationary Poisson process in the plane.

Poisson processes on abstract spaces were first constructed by WIENER (1938). PALM (1943) studied homogeneous Poisson processes on the line as special renewal processes. DOOB (1953) proves that the distribution of a homogeneous Poisson process is preserved under independent i.i.d. displacements of its points. The preservation property for general Poisson processes and randomization kernels was obtained by PRÉKOPA (1958).

Independent thinnings of a point process were first considered by RÉNYI (1956). The mixed binomial representation of a general Poisson process, implicit in WIENER (1938), was stated more explicitly by MOYAL (1962). Cox processes, first introduced by COX (1955), were studied extensively by KINGMAN (1964), KRICKEBERG (1972), and GRANDELL (1976). The binomial extension in Theorem 3.7 first appeared in RM_1 9.1.

RÉNYI (1967) noted that the distribution of a simple Poisson process is determined by its one-dimensional distributions. His result was extended to arbitrary simple point processes by MÖNCH (1971), and then to any diffuse random measures by GRANDELL (1976), who pioneered the systematic use of the Cox transform. Both results were noted independently in K(73a/c), and Theorem 3.8 (ii) was added in RM_1 3.4. RIPLEY (1976) extended the point process result to more general classes of sets, and XIA (2004) noted that a single exceptional point is allowed. The potential for further extensions is limited by counter-examples in GOLDMAN (1967) (attributed to L. SHEPP) and SZÁSZ (1970).

Poisson and related processes form a recurrent theme in the books $\mathrm{MKM}_{1,2}$, $\mathrm{RM}_{1,2}$, $\mathrm{DVJ}_{1,2}$, and FMP, each of which provides most of their basic properties. More specialized treatments of the subject have been given by KINGMAN (1993) and LAST & PENROSE (2017).

3.2. General birth & death processes were introduced by YULE (1924), who studied in particular the pure birth process named after him. Part (i) of Theorem 3.12 goes back at least to KOLMOGOROV (1931), whereas statements (ii)−(iii) were noted by RÉNYI (1953) and KENDALL (1966), respectively. Alternative approaches to Kendall's theorem appear in WAUGH

(1970), NEUTS & RESNICK (1971), and ATHREYA & NEY (1972). The remaining results of this section were developed for the needs in K(13).

3.3. The connection between infinitely divisible distributions and processes with independent increments was noted by DE FINETTI (1929), and the corresponding characterization is due to KOLMOGOROV (1932) and LÉVY (1934–35), leading up to the celebrated *Lévy–Khinchin formula* and associated *Lévy–Itô representation*. The random-variable case of Theorem 3.23 (iii) was proved by DOEBLIN (1939).

The independent-increment characterization of Poisson processes goes back to ERLANG (1909) in the stationary case, to LÉVY (1934) in the inhomogeneous case, and to WIENER (1938) in higher dimensions. The corresponding result for marked point processes is essentially due to ITÔ (1942a) in the stationary case and to KINGMAN (1967) in general. Random measures with independent increments have been studied by PRÉKOPA (1958), MOYAL (1962), and KINGMAN (1967).

MATTHES (1963a) established the characterization of infinitely divisible point processes. The extension to general random measures is due to JIŘINA (1964) in the a.s. bounded case, and to LEE (1967) in general. The finite-dimensional characterization of infinitely divisible random measures was noted in RM_1.

3.4. Positive, symmetric, and compensated Poisson integrals first appeared in ITÔ's (1942a) pathwise representation of Lévy processes and infinitely divisible laws. General existence and convergence criteria for positive or symmetric Poisson integrals were given in K & SZULGA (1989), whereas the more difficult case of compensated Poisson integrals was studied in K(91). The present extensions to marked point processes with independent increments are new.

3.5. DE FINETTI (1930/37) proved that any infinite sequence of exchangeable random variables is mixed i.i.d. The result was extended by RYLL-NARDZEWSKI (1957) to *contractable* sequences[1], where all subsequences are assumed to have the same distribution. BÜHLMANN (1960) proved a continuous-time version of de Finetti's theorem. Combining those results, we obtain the mixed-Poisson characterization of simple, contractable point processes on R_+. The thinning characterization of Cox processes was first noted, with a different proof, by MECKE (1968).

The mixed binomial characterization of exchangeable point processes on $[0,1]$ was proved independently in DAVIDSON (1974b), $\mathrm{MKM}_{1,2}$, and K(73a). The result was extended to general random measures in K(73b/75a) and RM_1, and the present version is adapted from K(05). The connection between exchangeable point processes and completely monotone functions was noted in K(73a), RM_1, p.75, and DABONI (1975). A comprehensive account of exchangeability and related symmetries is given in K(05).

[1]also called *spreading invariant* or *spreadable*

4. Convergence and Approximation

4.1. Weak convergence of probability measures on topological spaces was first studied by ALEXANDROV (1940/41/43), who proved a version of the "portmanteau" theorem, and derived compactness criteria when S is locally compact. PROHOROV (1956) introduced his metric for probability measures on a complete, separable metric space S, extending the *Lévy metric* on the real line, and proved the fundamental Theorem 4.2 in this setting. An alternative metrization is given by DUDLEY (1989). The extension to locally finite measures on S, first announced in DEBES et al. (1970), was developed in $\mathrm{MKM}_{1,2}$ and DVJ_1.

Comprehensive accounts of general weak convergence theory are given by PARTHASARATHY (1967), BILLINGSLEY (1968), and ETHIER & KURTZ (1986), and a more compact treatment appears in Ch. 16 of FMP.

4.2. Theorem 4.11 was established by PROHOROV (1961) for compact spaces, by VON WALDENFELS (1968) for locally compact spaces, and by HARRIS (1971) for any separable and complete metric space. In the latter setting, convergence and tightness criteria for point processes were also announced in DEBES et al. (1970) and proved in $\mathrm{MKM}_{1,2}$. For locally compact spaces, the relationship between weak and vague convergence in distribution was clarified in RM_1 4.9.

One-dimensional convergence criteria were obtained in K(73a) and RM_1, pp. 35f, including versions of Theorem 4.18 (often misunderstood) for locally compact spaces. The stronger versions in Theorems 4.15 and 4.16, first established for locally compact spaces in K(96), were extended to Polish spaces by PETERSON (2001). The power of the Cox transform in this context was explored by GRANDELL (1976). The one-dimensional criteria provide a connection to random sets, with the associated local topology of FELL (1962), also discussed in FMP A2.5.

SKOROHOD (1956) introduced four topologies on $D[0,1]$, of which the J_1-topology is the most important. The latter was extended to $D(\mathsf{R}_+)$ by STONE (1963) and LINDVALL (1973), and detailed discussions appear in ETHIER & KURTZ (1986) and JACOD & SHIRYAEV (1987). The relationship between the Skorohod and vague topologies is often misunderstood.

4.3. The general limit problem for null arrays of random variables, originally posed by KOLMOGOROV, was subsequently studied by many authors. The infinite divisibility of the limit was proved by FELLER (1937) and KHINCHIN (1937), and general convergence criteria were obtained by DOEBLIN (1939) and GNEDENKO (1939). Convergence criteria for infinitely divisible distributions on R_+^d were given by JIŘINA (1966) and NAWROTZKI (1968). Our modern exposition, in Ch. 15 of FMP, may be compared with the traditional accounts in LOÈVE (1977) and PETROV (1995).

Criteria for Poisson convergence, for superpositions of point processes on the line, were established in increasing generality by PALM (1943), KHIN-

CHIN (1955), OSOSKOV (1956), and GRIGELIONIS (1963). Versions for more general spaces were given by GOLDMAN (1967) and JAGERS (1972). The general limit theorem for null arrays of point processes was proved by KERSTAN & MATTHES (1964a) and $\mathrm{MKM}_{1,2}$, and the extension to general random measures (often misunderstood) appeared in K(73a). The one-dimensional criterion, for limits with independent increments, was obtained in RM_1 7.2, and criteria for simple or diffuse limits were explored in K(73a/96) and RM_1 7.6–10. The relationship between the supports of an infinitely divisible distribution and its Lévy measure was clarified in RM_1 6.8.

4.4. Strong convergence of point processes was studied extensively by MATTHES et al., forming the foundation of their theory of infinitely divisible point processes. A cornerstone in their development is the deep continuity Theorem 4.38 from MKM_2, here included with a simplified proof. Some of the quoted results on Poisson and Cox approximation may be new. Precise error estimates for the Poisson approximation are developed by BARBOUR et al. (1992), and an elementary discussion appears in DVJ_2, p. 162.

4.5. RÉNYI (1956) established the Poisson convergence, for a sequence of independent thinnings and reciprocal contractions of a fixed renewal process. Related statements and further discussion appear in NAWROTZKI (1962), BELYAEV (1963), GOLDMAN (1967), and WESTCOTT (1976). Those results were all subsumed by the general criteria in Corollary 4.41, first established for locally compact spaces in K(75b), along with some more general versions for compound point processes. The extension to variable thinning rates appeared as an exercise in RM_1. MKM_2 extended the result to general Polish spaces, and DVJ_2 noted the extension to certain random thinning rates. The present version for dissipative transition kernels may be new.

Functional limit theorems for random walks and Lévy processes go back to DONSKER (1951–52) and SKOROHOD (1957), and related results for sampling from a finite population were obtained by ROSÉN (1964), BILLINGSLEY (1968), and HAGBERG (1973). General convergence criteria for exchangeable random sequences, processes, and measures were established in K(73b) and RM_1. The present versions, for random measures on a product space, extend some results for compact spaces in K(05).

5. Stationarity in Euclidean Spaces

5.1. The idea of viewing a stationary stream of incoming phone calls from the vantage point of a typical arrival time was conceived by the Swedish engineer and applied mathematician PALM (1943), whose seminal thesis may be regarded as the beginning of modern point process theory. A more rigorous approach to the resulting Palm probabilities was provided by KHINCHIN (1955), whose simple convexity arguments were later extended by SLIVNYAK (1962) to the entire distribution.

Independently of Palm's work, and motivated by Doob's (1948) ideas in renewal theory, Kaplan (1955) introduced the modified Palm distribution of a stationary point process on the line, and proved the fundamental relationship between discrete and continuous-time stationarity. A discrete-time version of his result had been previously noted by Kac (1947). Kaplan's theorem was rediscovered by both Ryll-Nardzewski (1961) and Slivnyak (1962). Neveu (1968) and Papangelou (1970) noted the relations to some celebrated results in ergodic theory, due to Ambrose (1941) and Ambrose & Kakutani (1942).

The simple skew factorization approach to Palm distributions, first noted by Matthes (1963b), provides an easy access to the basic uniqueness and inversion theorems. For sub-intervals of the real line, part (i) of Theorem 5.5 goes back to Korolyuk, as cited by Khinchin (1955), whereas part (iii) was obtained by Ryll-Nardwewski (1961). The general versions are due to König & Matthes (1963) and Matthes (1963b) for $d = 1$, and to $\mathrm{MKM}_{1,2}$ for $d > 1$. The original proofs were simplified in FMP_2.

Palm distributions of stationary point processes on the line have been studied extensively by many authors, including Fieger (1965), Nawrotzki (1978), Glynn & Sigman (1992), Nieuwenhuis (1989/94), and Thorisson (1995, 2000). Applications to queuing theory have been developed by König et al. (1974), Franken et al. (1981), and Bachelli & Bremaud (2003).

5.2. For point processes on the real line, Kaplan's (1955) theorem may be stated, most naturally, in terms of spacing measures, which are then essentially equivalent to the Palm measures. The notion was extended in K(00) to any stationary random measure on the line, leading to the present symmetric version of Kaplan's theorem. Further results involving spacing measures appear in K(02a).

5.3. The notion of weak asymptotic invariance was introduced by Dobrushin (1956), to provide necessary and sufficient conditions for a sequence of particle systems to converge to Poisson, under independent random shifts. A more careful discussion was given by Stone (1968), and further results appear in Debes et al. (1970), Fichtner (1974), and $\mathrm{MKM}_{1,2}$. The fact that the successive convolution powers of a non-lattice distribution on R^d are weakly asymptotically invariant goes back to Maruyama (1955) and H. Watanabe (1956), and further discussion appears in Stam (1967). Local central limit theorems, related to our Lemma 5.19, were surveyed by Stone (1967).

5.4. The sample intensity of a stationary point process on the line is implicit in Slivnyak (1962), and the corresponding notion in R^d was explored in $\mathrm{MKM}_{1,2}$. The random measure version of Wiener's (1939) ergodic theorem on R^d was established by Nguyen & Zessin (1979). The mean ergodic theorems in the text were given in FMP, whereas the weak and strong smoothing versions may be new, leading to short proofs of the basic limit theorems for ν-transforms of stationary point processes. The weak version of the latter is due to DOBRUSHIN (1956), whereas the strong version may be new. The criterion for preservation of independence under a μ-transformation was obtained in K(78c), in the context of particle systems. The present argument is a simplification of the original proof.

5.5. Limit theorems relating any stationary point process to its Palm version have been discussed extensively by many authors, including NIEUWENHUIS (1989), THORISSON (2000), and K(02). For stationary point processes on the line, the relationship between ordinary and modified Palm distributions, along with the role of the sample intensity, are implicit in SLIVNYAK (1962), and further discussion appears in $\mathrm{MKM}_{1,2}$ and THORISSON (2000). SLIVNYAK (1962) also showed how the distribution of a point process can be recovered from its Palm version by a suitable averaging. His original version in terms of weak convergence was strengthened by ZÄHLE (1980) to convergence in total variation. Here, as elsewhere, the powerful coupling theorem of THORISSON (1996) provides short proofs. The simple symmetric versions for spacing measures were discovered and explored in K(02).

5.6. The fact that absolute continuity implies local invariance was noted by STONE (1968). A more systematic study of local invariance was initiated in K(78b), which also contains a discussion of local L^p-invariance for $p = 1, 2$. The extension to general $p \geq 1$ is new, though the characterization in Theorem 5.47 was claimed without proof, in a footnote of K(78b).

5.7. The ballot theorem of BERTRAND (1887), one of the earliest rigorous results in probability theory, states that if two candidates in an election, A and B, are getting the proportions p and $q = 1 - p$ of the votes, then the probability that A will lead throughout the ballot count equals $(p - q)_+$. Extensions and new approaches have been noted by many authors, beginning with ANDRÉ (1887) and BARBIER (1887). A modern discussion, based on combinatorial arguments, is given by FELLER (1968), and a simple martingale proof appears in CHOW & TEICHER (1997). A version for cyclically stationary sequences and processes was obtained by TAKÁCS (1967). All earlier versions are subsumed by the present statement from K(99b), whose proof relies heavily on Takács' ideas.

The three arcsine laws for Brownian motion were discovered by LÉVY (1939), along with the related uniform laws for the Brownian bridge. Various extensions to random walks and Lévy processes are included in FMP. The uniform laws were extended by FITZSIMMONS & GETOOR (1995) and KNIGHT (1996) to Lévy bridges and other exchangeable processes on the unit interval. The present versions are again adapted from K(99b).

6. Palm and Related Kernels

6.1. The Campbell measure of a stationary Poisson process on the line is implicit in CAMPBELL's (1909) discussion of shot-noise processes. For general point processes, RYLL-NARDZEWSKI (1961) constructed the Palm distributions by disintegration of the associated Campbell measures. His approach was adopted by JAGERS (1973), PAPANGELOU (1974a), RM_1, and others. By contrast, KUMMER & MATTHES (1970) and $\mathrm{MKM}_{1,2}$ avoid the disintegration by expressing properties of the Palm measures directly in terms of the Campbell measure.

Multi-variate Palm distributions were first mentioned by JAGERS (1973). Campbell measures are also important in martingale theory, where they were introduced and studied by DOLÉANS (1967).

6.2. First-order, reduced Campbell measures, implicit in RM_1, were formally introduced by MATTHES et al. (1979). Higher order reduced and compound Campbell measures were studied, along with their disintegrations, in Ch. 12 of RM_2. The conditioning approach to Palm distributions was discussed in K(03/10/11b).

6.3. For a stationary point process ξ on the line, SLIVNYAK (1962) proved that the reduced, modified Palm measures agree with the underlying distribution iff ξ is mixed Poisson. The corresponding characterization of Poisson processes is due to MECKE (1967). The Palm measures of a general mixed Poisson or binomial process were identified in K(73a). The characterization of infinitely divisible random measures in terms of the factorization $\mathcal{L}(\xi) \prec \mathcal{L}(\xi \,\|\, \xi)_s$ was obtained, in the stationary case, by KERSTAN & MATTHES (1964b), AMBARTZUMJAN (1966), MECKE (1967), and TORTRAT (1969), and then in general by KUMMER & MATTHES (1970). The weaker conditions in Theorem 6.17 (i) and (iv) were obtained in RM_1. The related criterion in Theorem 6.20 was proved for point processes by MATTHES (1969), and in general in RM_1.

6.4. The recursive construction of higher order Palm and reduced Palm measures goes back to RM_2. The general theory of iterated conditioning and Palm disintegration was developed in K(10/11b).

6.5. Higher order moment measures of Poisson and Cox processes were derived by KRICKEBERG (1974b). The Palm distributions of an exchangeable random measure were studied in RM_1 and K(05). The classical formulas, for the Palm distributions of Poisson and Cox processes, were extended in K(11b) to the multi-variate case. Poisson and Cox cluster processes, occurring frequently in applications, such as in NEYMAN & SCOTT (1958/72), are fundamental for the understanding of infinitely divisible random measures (cf. $\mathrm{MKM}_{1,2}$), as well as for various branching and super-processes (cf. DAWSON (1993)). In particular, the multi-variate Slivnyak formula plays a crucial role in K(13).

6.6. For point processes on a suitable topological space, elementary approximations of Palm distributions were derived, under various continuity assumptions, by JAGERS (1973) and in RM_2. Most results in this section originated in K(09), though the idea behind the ratio limit theorem appears already in RM_2 12.9. The decoupling properties were suggested by similar results for regenerative sets in K(03), and for super-processes in K(13).

6.7. The duality between Palm measures and moment densities was developed in K(99a), for applications to the local time random measure of a regenerative set. It has also been applied in K(01) to symmetric interval partitions, and in K(13) to super-processes.

7. Group Stationarity and Invariance

7.1. Invariant measures on special groups were constructed by HURWITZ (1897) and others by direct computation. HAAR (1933) proved the existence of invariant measures on a general lcscH group. The uniqueness was later established by WEIL (1936/40), who also extended the theory of invariant measures to homogeneous spaces. A comprehensive treatment of the classical theory appears in HEWITT & ROSS (1979), whereas SCHINDLER (2003) contains some recent developments and further references. Invariant measures are usually discussed in a topological setting, and the subject is often regarded as part of harmonic analysis. The present non-topological treatment is based on K(11a). The usefulness of s-finite measures was noted by SHARPE (1988) and GETOOR (1990).

Invariant disintegrations were first used to construct Palm distributions of stationary point processes. In this context, the regularization approach was developed by RYLL-NARDZEWSKI (1961) and PAPANGELOU (1974a), for point processes on Euclidean and more general homogeneous spaces. The simple skew factorization, first noted by MATTHES (1963b), was extended by MECKE (1967) and TORTRAT (1969) to random measures on a locally compact group.

In the non-transitive case, invariant disintegrations with respect to a fixed measure were discussed by KRICKEBERG (1974a/b), in connection with problems in stochastic geometry. The general integral representation of invariant measures was established in K(07a/11a). For the disintegration of invariant measures on a product space, the general factorization and regularization results were obtained in K(07a).

7.2. Stationary random measures are often discussed in terms of flows of measure-preserving transformations on the underlying probability space, as in NEVEU (1977), which makes the invariant representation part of the definition. An exception is the treatment of GETOOR (1990), who uses perfection techniques to derive invariant representations in special cases. The present more general representations were given in K(07a/11a). The construction of invariant Palm kernels goes back to RYLL-NARDZEWSKI (1961), SLIVNYAK (1962), and MATTHES (1963), though the multi-variate case was not considered until K(07a).

Theorem 7.10 was proved for $S = G = \mathsf{R}$ and $T = \Omega$ by GETOOR (1990), and the general version appears in K(07a). Similarly, Theorem 7.12, quoted from K(07a), extends a result of THORISSON (1995/2000).

7.3. The idea behind the inversion kernel goes back to ROTHER & ZÄHLE (1990), who extended the method of skew factorization to general homogeneous spaces. Our construction of orbit selectors and inversion kernels, from K(11a), led to a further extension of the skew factorization approach to the non-transitive case. An alternative approach to invariant disintegration was independently developed by GENTNER & LAST (2011).

The inversion method was used in K(11a) to construct invariant representations and disintegrations, based on disintegrations in the more elementary non-invariant case. A probabilistic interpretation of the measure inversion and related kernels was noted by LAST (2010), and further developed in K(11a).

7.4. MECKE (1967) characterized the Palm measures of stationary random measures on a locally compact, Abelian group, and established the associated inversion formula. His results were extended to homogeneous spaces by ROTHER & ZÄHLE (1990). Versions of the duality criterion, in the non-transitive case, were derived independently by GENTNER & LAST (2011) and in K(11a).

Invariant transformations of stationary random measures have been considered by many authors, and versions of the corresponding Palm transformations were noted by HARRIS (1971), PORT & STONE (1973), GEMAN & HOROWITZ (1975), MECKE (1975), and HOLROYD & PERES (2005). Two-sided versions and extensions to homogeneous spaces were obtained by LAST & THORISSON (2009) and LAST (2009/10), and the present version for non-transitive group actions was given in K(11a). The classical subject of mass transport is covered by an extensive literature, including the treatise of RACHEV & RÜSCHENDORF (1998). An associated mass-transport principle was considered by BENJAMINI et al. (1999) and ALDOUS & LYONS (2007).

NEVEU (1976/77) proved his exchange formula for closed subgroups of R^d, and explored its use as a basic tool in Palm measure theory. His result was extended by LAST & THORISSON (2009) and LAST (2009/10) to more general groups and homogeneous spaces. The present non-transitive criterion appears in K(11a).

7.5. The general theory of invariant disintegration was developed in K(14). A more elementary result appears in Theorem 7.6 above.

7.6. Special cases of stationary disintegration were identified by PAPANGELOU (1976) and in K(76/80), in connection with problems in stochastic geometry. The general theory is adapted from K(14), which also clarifies the need for an extra condition, to ensure that the stationarity will be preserved by integration. Ergodic decompositions is a classical subject, covered by a vast literature. Different aspects and approaches are considered in Theorem 7.3 above, in FMP 10.26, and in Dynkin (1978).

8. Exterior Conditioning

8.1. Exterior conditioning plays a basic role in statistical mechanics, in the study of equilibrium states, corresponding to a given interaction potential, determined by physical considerations. The resulting theory of *Gibbs measures*, named after Gibbs (1902), has been explored by Preston (1976), Georgii (1988), and others. The *DLR-equation* in Theorem 8.13 (i) goes back to Dobrushin (1970) and Lanford & Ruelle (1969). The connections to modern point process theory were noted by Nguyen & Zessin (1979), and further explored by Matthes et al. (1979), Glötzl (1980a/b), Rauchenschwandtner (1980), and R & Wakolbinger (1981).

Motivated by problems in stochastic geometry, Papangelou (1974b) used the basic consistency relation in Lemma 8.10 to construct the first order kernel named after him, under some simplifying assumptions, which were later removed in K(78a). The present partial disintegration approach was noted by Matthes et al. (1979). The Gibbs kernel was introduced in Ch. 13 of RM_2, as part of a general theory of conditioning in point processes, also clarifying the connections with Palm measures.

8.2. The results in this section all go back to Ch. 13 in RM_2.

8.3. Condition (Σ) was introduced, independently and with different motivations, by Papangelou (1974b) in stochastic geometry, and by Kozlov (1976) in statistical mechanics. In Matthes et al. (1979), it was shown to imply the absolute continuity $C_1(\cdot \times S) \ll \mathcal{L}(\xi)$, and a version of the DLR-equation. The condition itself, and its various extensions, were studied in Ch. 13 of RM_2, with minor improvements added in K(05)

Matthes et al. (1979) also noted the physical significance of absolute continuity, where the density of the first order Papangelou kernel is equivalent to the local energy function in statistical mechanics. In particular, this led to a version of the Poisson domination principle. Further connections between Papangelou kernels and local energies were noted by Glötzl (1980b) and Rauchenschwandtner (1980).

8.4. The equivalence between the symmetry of a point process and the invariance of its first order Papangelou kernel was proved in increasing generality by Papangelou (1974b/76) and in K(78a,b/80). A similar asymptotic result appears in K(78b).

8.5. Local limit theorems, involving the random measure π, were obtained, for simple point processes, in K(78a). More general versions were given in Ch. 14 of RM_2, which also contains the associated regularity decomposition.

8.6. A version of the global limit theorem was obtained, under simplifying assumptions, in the seminal paper of PAPANGELOU (1974b). More general versions were developed in K(78a), and in Ch. 14 of RM_2. VAN DER HOEVEN (1982/83) introduced the notion of external measurability, for random measures and processes. Mimicking the construction of compensators as dual predictable projections, he obtained the external intensity $\hat{\xi}$ as the dual, externally measurable projection of ξ. He also extended the global limit theorem to non-simple point processes. The version for more general random measures was given in RM_2 14.12, which provides the present, simplified approach to van der Hoeven's results.

9. Compensation and Time Change

9.1. Predictable and totally inaccessible times appear implicitly, along with quasi-leftcontinuous processes, in the work of BLUMENTHAL (1957) and HUNT (1957/58). A systematic study of optional times and their associated σ-fields was initiated by CHUNG & DOOB (1965), MEYER (1966), and DOLÉANS (1967). Their ideas were further developed by DELLACHERIE (1972), DELLACHERIE & MEYER (1975/87), and many others into a *"general theory of processes*," revolutionizing many areas of probability theory. A concise introduction appears in Ch.'s 25–26 of FMP.

9.2. The Doob–Meyer decomposition, suggested by DOOB's elementary decomposition of discrete-time sub-martingales, was first established by MEYER (1962/63), in the preliminary form of Lemma 9.15. DOLÉANS (1967) proved the equivalence of natural and predictable increasing processes, leading to the ultimate version of the decomposition. The original proofs, appearing in DELLACHERIE (1972) and DELLACHERIE & MEYER (1975/87), were based on some deep results in capacity theory. The present shorter and more elementary argument, adapted from FMP 25.5, combines RAO's (1969) simple proof of the preliminary version, based on DUNFORD's (1939) weak compactness criterion, with DOOB's (1984) ingenious approximation of totally inaccessible times.

Induced compensators of optional times first arose, as *hazard functions*, in reliability theory. More general compensators were later studied in the Markovian context by WATANABE (1964) and others, under the name of *Lévy systems*. Compensators of general random measures, on product spaces $\mathsf{R}_+ \times S$, were constructed and employed by GRIGELIONIS (1971) and JACOD (1975). Detailed discussions of this and related material appear in JACOD (1979), JACOD & SHIRYAEV (1987), and Ch.'s 25–26 of FMP. More elementary accounts of the point process case have been given, under various simplifying assumptions, by many authors, including LAST & BRANDT (1995) and JACOBSEN (2006).

9.3. The martingale characterization of Brownian motion was discovered by LÉVY (1937/54), whereas the corresponding characterization of ql-continuous Poisson processes, in terms of compensators, was obtained by WATANABE (1964). The latter result was extended by JACOD (1975) to marked point processes with independent increments, and by GRIGELIONIS (1977) to general semi-martingales. See also JACOD & SHIRYAEV (1987).

DOEBLIN (1940), in a long forgotten deposition to the Academy of Sciences in Paris, proved that any continuous (local) martingale can be time-changed into a Brownian motion. The result was rediscovered by both DAMBIS (1965) and DUBINS & SCHWARZ (1965). The corresponding reduction to Poisson, of a simple, ql-continuous point process, was proved independently by MEYER (1971) and PAPANGELOU (1972), and multi-variate versions of both results were noted by KNIGHT (1971). Further time-change reductions were given in COCOZZA & YOR (1980) and K(90b), and some point-process applications appear in AALEN & HOEM (1978) and BROWN & NAIR (1988).

The predictable sampling theorem, along with its continuous-time counterpart, were obtained in K(88). The special case of optional skipping in i.i.d. sequences goes back to DOOB (1936), and related ideas in fluctuation theory were used by SPARRE-ANDERSEN (1953–54). Time-change reductions of stable integrals were obtained, first in special cases by ROSIŃSKI & WOYCZYŃSKI (1986), and then more generally in K(92). A comprehensive account of the subject appears in K(05).

9.4. The formula for the natural compensator of a random pair (τ, χ) is classical, and can be traced back to DELLACHERIE (1972) and JACOD (1975). The stated solution to the stochastic integral equation $Z = 1 - Z_- \cdot \bar{\eta}$ is a special case of the *exponential process*, introduced by DOLÉANS (1970). Our fundamental martingale was discovered in K(90b), along with the observation that $E|U_{\tau,\chi}| < \infty$ implies $EV_{\tau,\chi} = 0$. The same paper explores the various predictable mapping and time-change reductions, obtainable from this fact. BROWN (1978/79) used a more direct approach, in deriving some asymptotic versions of Watanabe's Poisson and Cox characterizations, in terms of compensators.

9.5. One-point processes have been studied by many authors, including DELLACHERIE (1972). The integral representation of an adapted pair (τ, χ) with associated compensator was obtained in K(90b), by an intricate time-change argument. The simplified proof given here was anticipated by a concluding remark in the same paper.

9.6. Tangent processes, first mentioned by ITÔ (1942), were used explicitly by JACOD (1984) to extend the notion of semi-martingales. A discrete-time version of the basic existence theorem was noted by KWAPIEŃ & WOYCZYŃSKI (1992), pp. 104 f, and a careful proof appears in DE LA PEÑA & GINÉ (1999), pp. 293 f. A discrete-time version of the tangential preservation lemma appears in Kwapień & Woyczyński (1992), p. 134. Continuous-time versions of those results are given in K(16).

9.7. Discrete-time versions of the basic tangential comparison theorem were first obtained by Zinn (1986) and Hitchenko (1988), and a detailed proof appears in Section 5.2 of Kwapień & Woyczyński (1992), which also contains, on p. 134, a discrete-time version of the one-sided comparison theorem. The idea of using tangential sequences to extend results for independent random variables to the dependent case, often referred to as the *principle of conditioning*, was developed systematically by Jakubowski (1986) and others. Related results, detailed proofs, and numerous applications appear in Kwapień & Woyczyński (1987/91/92) and de la Peña & Giné (1999). The present continuous-time comparison theorems, and related criteria for boundedness and convergence, appear in K(16).

Stochastic integration with respect to general semi-martingales was developed by Courrège (1962–63), Kunita & Watanabe (1967), Doléans & Meyer (1970), and Meyer (1967/76), and a summary appears in Ch. 26 or FMP. The BDG inequalities go back to Burkholder et al. (1966/72), Millar (1968), and others, and short proofs appear in FMP 17.7 and 26.12.

10. Multiple Integration

10.1. Multiple Poisson integrals were first mentioned explicitly by Itô (1956), though some underlying ideas can be traced back to the work of Wiener (1938) and Wiener & Wintner (1943) on *discrete chaos.* An L^2-theory for Poisson integrals, analogous to that for multiple Wiener–Itô integrals with respect to Brownian motion, has been developed by many authors, including Ogura (1972), Kabanov (1975), Segall & Kailath (1976), Engel (1982), and Surgailis (1984). Moment formulas for Poisson and related integrals appear in Krickeberg (1974b). Decoupling methods, explored by many authors, are surveyed by Kwapień & Woyczyński (1992) and de la Peña & Giné (1999).

The convergence problem for multiple Poisson integrals was originally motivated by the representations in K(90a). Most results in this section are adapted from K & Szulga (1989).

10.2. Here again, the main results are adapted from K & Szulga (1989). The hyper-contraction property of multi-linear forms was proved by Krakowiak & Szulga (1986). An analogous result for multiple Wiener–Itô integrals is due to Nelson (1973), cf. K(05). The basic hyper-contraction lemma goes back to Paley & Zygmund (1932) and Marcinkiewicz & Zygmund (1937).

10.3. This material is again adapted from K & Szulga (1989).

10.4. Many authors have studied multiple integrals, based on symmetric stable and more general Lévy processes, using different methods of construction, valid under more or less restrictive assumptions. Thus, Surgailis (1985) employs an interpolation theorem in Lorentz–Zygmund spaces, whereas Krakowiak & Szulga (1988) use vector methods and ideas from harmonic analysis. Rosiński & Woyczyński (1986) obtained necessary and sufficient conditions for the existence of certain double stable integrals, using methods involving geometry and probability in Banach spaces. Existence criteria for double integrals, based on symmetric Lévy processes, were found by Kwapień & Woyczyński (1987), through a deep analysis involving random series in Orlicz spaces. Their conditions are accordingly complicated.

Though the present approach, adapted from K & Szulga (1989), is more elementary, it provides general criteria for convergence of multiple integrals of arbitrary order, based on positive or symmetric Lévy processes. In overlapping cases, the present conditions, though formally simpler, can be shown to be equivalent to those obtained by previous authors.

10.5. This material was developed in K(91), which also contains results for multiple integrals based on non-symmetric Lévy processes.

10.6. This material is adapted from K(16).

11. Line and Flat Processes

11.1. Poisson processes of flats were studied extensively by Miles (1969/71/74) and others, and more general line processes in the plane were explored by Davidson (1974a). Assuming stationarity under rigid motions, local finiteness of second moments, and no pairs of parallel lines, he obtained his representation of the covariance measure, and proved the existence of a Cox process with the same first and second moments. A general disintegration approach, developed by Krickeberg (1974a/b), allowed some extensions to higher dimensions. The underlying assumptions were weakened in K(76) to stationary under translations, and local finiteness of the first moments. The idea of using Palm measures and conditional intensities, to establish the Cox property of a stationary line or flat process, goes back to some ingenious work of Papangelou (1974a,b/76). General accounts of stochastic geometry are given by Stoyan et al. (1995) and Schneider & Weil (2008).

11.2. Based on the indicated Cox connection, and the apparent difficulty to construct more general examples, Davidson (1974c) conjectured that all line processes with the stated properties are indeed Cox. Despite his untimely death, in a climbing accident at age 25, his work inspired an intense activity. The stated conjecture turned out to be false, and a counter-example was given in K(77b), based on the stationary lattice point process. Alternative constructions of the latter were provided by Kingman, as quoted in

K(77b), and by MECKE (1979).

11.3. The mentioned moment and invariance properties of a stationary line process in the plane were extended in K(76) to processes of k-flats in R^d, for arbitrary $k \geq d/2$, under the same assumptions of stationarity under translations and local finiteness of first moments. For $k = d - 1$, the basic covariance formula was extended by KRICKEBERG (1974b) to moments of order $\leq d$.

11.4. The a.s. invariance of a stationary random measure η on F_k^d was established in PAPANGELOU (1976) and K(76/80), under suitable assumptions of absolute continuity $\eta \circ \pi^{-1} \ll \mu$, where μ was initially taken to be the homogenous measure on Φ_k^d. The regularity conditions were relaxed in K(78b/81) to a version of local invariance, which was shown in K(81) to be nearly sharp.

11.5. BREIMAN (1963) noted that, if the points of a stationary and ergodic point process are put in motion, with independent constant velocities, chosen from some absolutely continuous distribution, the resulting particle system converges in distribution to a Poisson process. STONE (1968) noted that the result follows from a general claim of DOBRUSHIN (1956). Extensions and alternative approaches were provided by THEDÉEN (1964), GOLDMAN (1967), CLIFFORD & SUDBURY (1974), $\mathrm{MKM}_{1,2}$, and JACOBS (1977). The result also applies to one-dimensional particle systems with elastic interaction, since two colliding particles exchange their velocities, so that the space-time diagram remains the same. In this setting, the path of a single particle was studied by HARRIS (1965), SPITZER (1969), and SZATZSCHNEIDER (1975).

The independence property in Breiman's theorem was shown in K(78c) to be destroyed immediately, unless the configuration is Poisson to begin with, so that the entire space-time process becomes stationary. In K(78b) we showed that a line process on R^{d+1}, stationary in d directions, and subject to appropriate regularity conditions, is asymptotically Cox under translations in the perpendicular direction. In terms of particle systems, this means essentially that, under suitable regularity conditions, the independence assumption in Breiman's theorem can be weakened to joint stationarity.

11.6. In his studies of line processes in the plane, DAVIDSON (1974a/c) found the need to exclude the possibility of parallel lines. For more general flat processes, some counter-examples of PAPANGELOU (1976) show that we must also exclude the possibility of outer degeneracies. A systematic study of inner and outer degeneracies was initiated in K(80), which is also the source of the $0 - \infty$ law in Theorem 11.24. The latter paper further contains some related results for stationary flat processes.

11.7. This section is again based on K(80), which contains versions of the crucial Lemma 11.25, and all subsequent theorems. The deep Corollary 11.28

was anticipated by remarks in K(80), though the result was never stated explicitly. Further invariance criteria, often of a more technical nature, were given in K(78b/80/81).

12. Regeneration and Local Time

12.1. Though excursions of Markov chains are implicit already in some early work of KOLMOGOROV and LÉVY, general renewal processes may not have been introduced explicitly until PALM (1943), followed by subsequent studies of DOOB (1948) and others. KAPLAN (1955) thought of his construction of Palm distributions as an extension of renewal theory.

The first renewal theorem was obtained by ERDÖS et al. (1949), for random walks in Z_+, though the statement in that case follows immediately from KOLMOGOROV's (1936a/b) ergodic theorem for Markov chains, as pointed out by CHUNG. BLACKWELL (1948/53) extended the result to general random walks in R_+. The two-sided version, for transient random walks, is due to FELLER & OREY (1961). The present density version is essentially a special case of some general results of STONE (1966).

The coupling method, first devised by DOEBLIN (1938), was explored by LINDVALL (1977) to give a probabilistic proof of the renewal theorem. The present approach, adapted from FMP 9.20, is a modification of an argument of ATHREYA et al. (1978), which originally did not cover all cases.

12.2. Local time of Brownian motion at a fixed point was discovered and explored by LÉVY (1939), who devised several explicit constructions, based on approximations from the excursions. General regenerative sets were studied by HOROWITZ (1972), who noted their connection with subordinators. An alternative approach to local time, via additive functionals, was pioneered by DYNKIN (1965) and others, and further developed by BLUMENTHAL & GETOOR (1964/68), as outlined in FMP 22.24.

Assuming the existence of local time, ITÔ (1972) showed how the excursion structure of a Markov process can be described by a Poisson process on the local time scale. The present approach is adapted from FMP 22.11, where the local time was constructed, jointly with Itô's excursion point process. A two-sided regenerative process was recognized, in K(02a/03), as the Palm version of a stationary regenerative process.

Comprehensive accounts of regenerative processes and excursion theory are given in BLUMENTHAL (1992) and DELLACHERIE et al. (1992). Many basic properties carry over to the analogous, but more difficult, case of *exchangeable interval partitions* of $[0,1]$. In particular, some extensions of the subsequent material are discussed in K(01/05).

12.3. The semi-martingale approach to local time, first suggested by MEYER (1976), is based on some earlier work of TANAKA (1963) and SKOROHOD (1961/62). The joint continuity of the space-time local time of Brownian motion, leading to its interpretation as an occupation density, was established by Trotter (1958). The extension to arbitrary continuous semi-martingales is due to Yor (1978). The general Itô–Tanaka formula was also obtained, independently, by Wang (1977).

12.4. The basic role of Palm measures in the context of regenerative processes was first recognized in K(81a). The material in this section is adapted from K(03).

12.5. Regenerative sets Ξ with positive Lebesgue measure were studied extensively by Kingman (1972), under the name of *regenerative phenomena*. In particular, he proved in this case the existence of a continuous renewal density p, called the *p-function* of Ξ. The condition $\hat{\mu}_s \in L^1$ was introduced in K(81a), along with the regularity indices m_r, c, and c', and the approximate description of the continuity set of p was obtained in K(81a/99a). Related regularity indices, for general Lévy processes, were considered by Blumenthal & Getoor (1961) and others. A general criterion for absolute continuity of an infinitely divisible distribution was given by Tucker (1965).

12.6. The fundamental identity, first established in K(99a), may be regarded as a special case of a basic relation, for "good" potential densities of a Markov process. The existence of such a density is a classical problem, discussed in Blumenthal & Getoor (1968). The duality approach to Palm distributions was first devised in K(99a), for application to regenerative processes. The present approach, via the conditioning kernel, is adapted from K(03), which also provides the associated continuity and symmetry properties.

12.7. Though regular versions of the Palm distributions of X were considered already in K(81a/99a), the present approach was first developed in K(03). Of the asymptotic factorizations of Palm distributions in Theorem 12.44, part (iv) was established, with a different proof, in K(81a), whereas parts (i)−(iii) were added in K(03). The latter paper also explored the basic properties of shift invariance and time reversal.

12.8. One-dimensional versions of Theorems 12.45 and 12.50 were derived, by different methods, in K(99a). All multi-variate results in this section are adapted from K(03). Without our regularity assumptions, the theory becomes much more difficult. Here Chung posed the problem of finding the hitting probabilities of individual points. The intricate solutions, independently devised by Carleson and Kesten, are recorded in Assouad (1971) and Bertoin (1996), respectively. An extension to exchangeable interval partitions was given by Berbee (1981).

13. Branching Systems and Super-processes

13.1. The classical, discrete-time branching process was first introduced and studied by Bienaymé (1845), who in particular found the extinction probability to be the smallest root of the equation $f(s) = s$. His results were partly rediscovered by Watson & Galton[2] (1974), after whom the process has traditionally been named. See Kendall (1966/75) for further comments on the early history. In the critical case, the asymptotic survival probability and associated distribution were found by Kolmogorov (1938) and Yaglom (1947), respectively, and the present comparison argument is due to Spitzer (unpublished). The transition probabilities of a birth & death process with time-dependent rates $n\mu$ and $n\lambda$ were given by Kendall (1948), though the result for constant μ and λ had been noted already by Palm (1943, unpublished). Our exact description of the associated ancestral process may be new.

The diffusion limit of a Bienaymé process was obtained by Feller (1951), and a measure-valued branching process was considered by Jiřina (1966). The DW-process, first constructed by Watanabe (1968), was studied extensively by Dawson (1977/93) and others, along with various extensions. The discovery of historical super-processes, explored in profound papers by Dynkin (1991), Dawson & Perkins (1991), and Le Gall (1991), paved the way for some powerful probabilistic techniques, leading to a further vigorous development of the subject.

Comprehensive accounts of the theory of classical branching processes have been given by Harris (1963), Athreya & Ney (1972), Jagers (1975), and Asmussen & Hering (1983). Discrete-time cluster processes in Euclidean spaces were studied extensively in Debes et al. (1970) and $\mathrm{MKM}_{1,2}$. The vast literature on super-processes has been surveyed by Dawson (1993), Dynkin (1994), Le Gall (1999), Etheridge (2000), Perkins (2002), and Li (2011).

For the uninitiated reader, Etheridge's book may provide the easiest access to DW-processes. In particular, it contains elementary discussions of scaling limits, the log-Laplace equation, martingale characterizations, the Brownian snake, and the genealogical and conditional structure. A detailed proof of the second martingale characterization appears in Section II.5 of Perkin's book.

13.2. The genealogy of a DW-process, first described by Dawson & Perkins (1991), is also implicit in Le Gall's (1991) ingenious identification of the branching structure of a DW-process with some structural properties of a reflecting Brownian motion. Equivalent results for Brownian excursion have been uncovered by Williams (1974) and others. The criteria for lo-

---

[2]This is the same Galton who got infamous as the founder of *eugenics*, the "science" of breeding a human master race, later adopted and put into action by Hitler and others. I feel relieved to be justified in dropping his name from any association with the subject.

cal finiteness were obtained in K(08), as was the equivalence between the two versions of local extinction, whereas the explicit criteria in the measure-valued case were obtained by Bramson et al. (1993) and Klenke (1998). The scaling properties of a DW-process and its canonical cluster, mostly implicit in previous work, were summarized in K(13), which is also the source of the strong temporal continuity of the process.

13.3. The higher order moment measures of a general super-process were derived by Dynkin (1991), who also described the resulting formulas in terms of a combinatorial tree structure. The cluster decomposition and forward recursion are implicit in Dynkin's work, whereas the backward recursion and Markov property first appeared in K(13). Etheridge (2000) noted the connection with the uniform Brownian tree, as constructed by her method of sideways recursion.

The relationship between moment measures and Brownian trees, first noted in K(13) as a formal coincidence, has now been clarified by the tree representation of multi-variate Campbell measures, based on ideas that were partially implicit in K(13). The resulting representation of Palm trees, here stated as Corollary 13.22, was previously known, in the univariate case, from Dawson & Perkins (1991), with some further clarifications in Dawson (1993).

The study of super-processes was revolutionized by the profound and powerful work of Le Gall (1991/99), whose *Brownian snake* represents the entire DW-process, along with its genealogical structure. Discrete-time versions of some underlying ideas had earlier been noted by Neveu (1986), and related properties of the Brownian excursion had previously been obtained by Williams (1974). The extended Brownian excursion and snake were constructed in K(13).

13.4. The results in this section were developed in K(08/13).

13.5. This material is adapted from K(13).

13.6. The basic hitting estimates for $d \geq 3$ were obtained by Dawson et al. (1989), through a deep analysis of the non-linear PDE in Lemma 13.6. The more difficult case of $d = 2$ was settled by Le Gall (1994), using estimates based on the Brownian snake. Our versions in Theorem 13.38 are quoted from K(08). The remaining estimates and approximations in this section were obtained in K(08/13).

13.7. A version of the Lebesgue approximation was obtained for $d \geq 3$ by Tribe (1994), whereas the result for $d = 2$ was proved in K(08). Our unified approach is adapted from the latter paper.

13.8. The space-time stationary version $\tilde{\xi}$ of the DW-process was constructed in K(08), along with the associated strong approximation. A related but weaker result is mentioned without proof in Dawson & Perkins (1991), with reference to some unpublished joint work with Iscoe. The analogous approximation by the canonical cluster $\tilde{\eta}$ was obtained in K(13), which is also our source of the remaining continuity and invariance properties of $\tilde{\xi}$ and $\tilde{\eta}$.

13.9. The results in this section, partly anticipated in K(07b), were developed in K(13).

13.10. The persistence-extinction dichotomy was noted by Debes et al. (1970), who also established the stated truncation criteria. This led to the basic problem of finding necessary and sufficient conditions for stability, where partial progress was made by Liemant & Matthes (1974/75/77), and others. Our method of Palm or *backward trees* was developed in K(77a), which is the source of the remaining results in this section.

An introduction to cluster processes is given in $\mathrm{MKM}_{1,2}$. The method of backward trees has been adapted to more general processes by U. Zähle (1988a/b) and others. Further developments are documented in Fleischmann et al. (1981/82), Liemant et al. (1988), and Gorostiza et al. (1990/91).

Appendices

A1. In an excusable oversight[3], Lebesgue (1905) claimed that the one-dimensional projections of a Borel set in the plane are again Borel. Suslin (1917) found an error in the proof, which led him to introduce *analytic sets*. Then Lusin (1917) proved that every analytic set in R is Lebesgue measurable, thereby establishing a version of the projection theorem. The general theory of analytic sets was developed by the Russian and Polish schools, culminating with the work of Kuratowski (1933). The projection and section theorems are often derived from capacity theory, as developed by Choquet (1953–54), Dellacherie (1972), and Dellacherie & Meyer (1975). A direct approach, with detailed proofs, appears in Dudley (1989).

The weak compactness criterion, due to Dunford (1939), is known also to be necessary. Some basic properties of conditional independence are given in FMP 6.6–8, and the present statement appeared in K(03). Concentration functions were first introduced by Lévy (1937), and the basic estimates for random sums were obtained by Kolmogorov (1958) and Rogozin (1961). Short proofs based on characteristic functions were devised by Esseen (1968), and comprehensive surveys are given by Hengartner & Theodorescu (1973) and Petrov (1995).

A2. The process in Lemma A2.1 is a special case of Doléans' (1970) *exponential process*, reviewed in FMP 26.8. Basic facts about Brownian excursions are summarized in FMP 22.15. The associated path decomposition was obtained by Williams (1974), with alternative proofs given by Le Gall

[3]It is comforting to know that even the giants can make mistakes.

(1986). Sub-martingales of class (D) were introduced by Meyer (1962/63), for his version of the Doob–Meyer decomposition.

Wiener (1939) extended Birkhoff's pointwise ergodic theorem to averages over concentric balls in R^d. A modern proof appears in FMP_2 10.14, along with various extensions and ramifications. Comprehensive surveys of ergodic theorems are given by Tempel'man (1972) and Krengel (1985). The powerful shift coupling theorem is due to Thorisson (1996/2000), and a short proof appears in FMP_2 10.28. The closure property in Lemma A2.7 is quoted from K(03).

A3. Lemma A3.3 appeared in K(08), and the remaining results are taken from K(13).

A4. Lemmas A4.1 and A4.2 are quoted from K(13), and Lemma A4.3 appears in FMP 10.15. The stated covering theorem and its corollary go back to Besicovitch (1945), and careful proofs appear in Bogachev (2007).

A5. Boothby (1986) gives a concise and accessible introduction to Lie groups and Riemannian manifolds. Projective limits of Lie groups, along with the representation theorem for locally compact groups, are carefully explained, with detailed proofs, in Montgomery & Zippin (1955).

A6. Absolutely and completely monotone sequences and functions were characterized by Hausdorff (1921) and Bernstein (1928), respectively. Feller (1971) gives short probabilistic proofs, and shows how de Finetti's (1930) theorem for exchangeable events follows from Hausdorff's result. For further discussion with relations to exchangeability, see Sections 1.6 and A4 in K(05). The elementary Lemma A6.2 is quoted from K(08).

References

Here I am only listing publications mentioned in the main text or historical notes. No completeness is claimed, and the knowledgable reader will notice that many papers of historical significance are missing. Again I apologize for any unintentional errors or omissions.

AALEN, O.O., HOEM, J.M. (1978). Random time changes for multivariate counting processes. *Scand. Actuar. J.*, 81–101.

ALDOUS, D., LYONS, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.* **12**, 1454–1508.

ALEXANDROV, A.D. (1940/41/43). Additive set functions in abstract spaces. *Mat. Sb.* **8**, 307–348; **9**, 563–628; **13**, 169–238.

ALEXANDROV, P. (1916). Sur la puissance des ensembles mesurables B. *C.R. Acad. Sci. Paris* **162**, 323–325.

AMBARTZUMJAN, R.V. (1966). On an equation for stationary point processes (in Russian). *Dokl. Akad. Nauk Armjanskoi SSR* **42**, 141–147.

AMBROSE, W. (1941). Representation of ergodic flows. *Ann. Math.* **42**, 723–739.

AMBROSE, W., KAKUTANI, S. (1942). Structure and continuity of measurable flows. *Duke Math. J.* **9**, 25–42.

ANDRÉ, D. (1887). Solution directe du problème résolu par M. Bertrand. *C.R. Acad. Sci. Paris* **105**, 436–437.

ASMUSSEN, S., HERING, H. (1983). *Branching Processes.* Birkhäuser, Boston.

ASSOUAD, P. (1971). Démonstration de la “Conjecture de Chung” par Carleson. In: *Lect. Notes in Math.* **191**, 17–20. Springer, Berlin.

ATHREYA, K.B., MCDONALD, D., NEY, P. (1978). Coupling and the renewal theorem. *Amer. Math. Monthly* **85**, 809–814.

ATHREYA, K.B., NEY, P.E. (1972). *Branching Processes.* Springer, Berlin.

BACCELLI, F., BRÉMAUD, P. (2003). *Elements of Queuing Theory*, 2nd ed. Springer, Berlin.

BARBIER, É. (1887). Généralisation du problème résolu par M. J. Bertrand. *C.R. Acad. Sci. Paris* **105**, 407, 440.

BARBOUR, A.D., HOLST, L., JANSON, S. (1992). *Poisson Approximation.* Clarendon Press, Oxford.

BARTLETT, M.S., KENDALL, D.G. (1951). On the use of the characteristic functional in the analysis of some stochastic processes in physics and biology. *Proc. Cambridge Phil. Soc.* **47**, 65–76.

BATEMAN, H. (1910). Note on the probability distribution of α-particles. *Philos. Mag.* **20** (6), 704–707.

BELYAEV, Y.K. (1963). Limit theorems for dissipative flows. *Th. Probab. Appl.* **8**, 165–173.

— (1969). Elements of the general theory of random streams (in Russian). In: H. Cramér and M.R. Leadbetter: *Stationary and Related Stochastic Processes*, Russian ed. MIR, Moscow.

BENJAMINI, I., LYONS, R., PERES, Y., SCHRAMM, O. (1999). Group-invariant percolation on graphs. *Geom. Func. Anal.* **9**, 29–66.

BERBEE, H. (1981). On covering single points by randomly ordered intervals. *Ann. Probab.* **9**, 520–528.

BERNSTEIN, S. (1928). Sur les fonctions absolument monotones. *Acta Math.* **52**, 1–66.

BERTOIN, J. (1996). *Lévy Processes.* Cambridge Univ. Press.

BERTRAND, J. (1887). Solution d’un problème. *C.R. Acad. Sci. Paris* **105**, 369.

BESICOVITCH, A.S. (1945). A general form of the covering principle and relative differentiation of additive functions. *Proc. Cambridge Philos. Soc.* **41**, 103–110.

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7

Bienaymé, I.J. (1845). De la loi de multiplication et de la durée des familles. *Soc. Philomat. Paris Extraits, Sér.* **5**, 37–39.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, NY.

Blackwell, D. (1948). A renewal theorem. *Duke Math. J.* **15**, 145–150.

— (1953). Extension of a renewal theorem. *Pacific J. Math.* **3**, 315–320.

Blumenthal, R.M. (1957). An extended Markov property. *Trans. Amer. Math. Soc.* **82**, 52–72.

— (1992). *Excursions of Markov Processes.* Birkhäuser, Boston.

Blumenthal, R.M., Getoor, R.K. (1961). Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.* **10**, 493–516.

— (1964). Local times for Markov processes. *Z. Wahrsch. verw. Geb.* **3**, 50–74.

— (1968). *Markov Processes and Potential Theory.* Academic Press, NY.

Bogachev, V.I. (2007). *Measure Theory I–II.* Springer, Berlin.

Boothby, W.M. (1986). *An Introduction to Differentiable Manifolds and Riemannian Geometry*, 2nd ed. Academic Press, New York.

Bramson, M., Cox, J.T., Greven, A. (1993). Ergodicity of critical branching processes in low dimensions. *Ann. Probab.* **21**, 1946–1957.

Breiman, L. (1963). The Poisson tendency in traffic distribution. *Ann. Math. Statist.* **34**, 308–311.

Brémaud, P. (1981). *Point Processes and Queues.* Springer, NY.

Brown, T.C. (1978). A martingale approach to the Poisson convergence of simple point processes. *Ann. Probab.* **6**, 615–628.

— (1979). Compensators and Cox convergence. *Math. Proc. Cambridge Phil. Soc.* **90**, 305–319.

Brown, T.C., Nair, M.G. (1988). Poisson approximation for time-changed point processes. *Stoch. Proc. Appl.* **29**, 247–256.

Bruchner, A.M., Bruchner, J.B., Thomson, B.S. (1997). *Real Analysis.* Prentice-Hall, Upper Saddle River.

Bühlmann, H. (1960). Austauschbare stochastische Variabeln und ihre Grenzwertsätze. *Univ. California Publ. Statist.* **3**, 1–36.

Burkholder, D.L. (1966). Martingale transforms. *Ann. Math. Statist.* **37**, 1494–1504.

Burkholder, D.L., Davis, B.J., Gundy, R.F. (1972). Integral inequalities for convex functions of operators on martingales. *Proc. 6th Berkeley Symp. Math. Statist. Probab.* **2**, 223–240.

Campbell, N.R. (1909). The study of discontinuous phenomena. *Proc. Cambridge Phil. Soc.* **15**, 117–136.

Choquet, G. (1953–54). Theory of capacities. *Ann. Inst. Fourier Grenoble* **5**, 131–295.

Chow, Y.S., Teicher, H. (1997). *Probability Theory: Independence, Interchangeability, Martingales*, 3rd ed. Springer, NY.

Chung, K.L., Doob, J.L. (1965). Fields, optionality and measurability. *Amer. J. Math.* **87**, 397–424.

Clifford, P., Sudbury, A. (1974). Some results on the limiting behaviour of infinite particle systems. *Z. Wahrsch. verw. Gebiete* **30**, 273–278.

Cocozza, C, Yor, M. (1980). Démonstration d'un théorème de F. Knight a l'aide de martingales exponentielles. *Lect. Notes in Math.* **784**, 496–499. Springer, Berlin.

Copeland, A.H., Regan, F. (1936). A postulational treatment of the Poisson law. *Ann. Math.* **37**, 357–362.

Courrège, P. (1962–63). Intégrales stochastiques et martingales de carré intégrable. *Sem. Brelot–Choquet–Deny* **7**. Publ. Inst. H. Poincaré.

Cox, D.R. (1955). Some statistical methods connected with series of events. *J. Roy. Statist. Soc., Ser. B*, **17**, 129–164.

Cox, D.R., Isham, V. (1980). *Point Processes.* Chapman & Hall, London.

Cramér, H. (1969). Historical review of Filip Lundberg's work on risk theory. On streams of random events. *Scandin. Actuarial J.* **1969** 3-4 Supplement, 6–23.

Daboni, L. (1975). Caratterizzazione delle successioni (funzioni) completamente monotone in termini di rappresentabilità delle funzioni di sopravvivenza di particolari intervalli scambiabili tra successi (arrivi) contigui. *Rend. Math.* **8**, 399–412.

Daley, D.J. (1974). Various concepts of orderliness for point processes. In: *Stochastic Geometry* (E.F. Harding & D.G. Kendall, eds.), pp. 148–161. Wiley, London.

Daley, D.J., Vere-Jones, D. (2003/08). *An Introduction to the Theory of Point Processes, I–II* ($\mathrm{DVJ}_{1,2}$), 2nd ed. Springer, NY.

Dambis, K.E. (1965). On the decomposition of continuous submartingales. *Theor. Probab. Appl.* **10**, 401–410.

Davidson, R. (1974a). Stochastic processes of flats and exchangeability. In: *Stochastic Geometry* (E.F. Harding & D.G. Kendall, eds.), pp. 13–45. Wiley, London.

— (1974b). Exchangeable point-processes. *Ibid.*, 46–51.

— (1974c). Construction of line-processes: Second-order properties. *Ibid.*, 55–75.

Dawson, D.A. (1977). The critical measure diffusion. *Z. Wahrsch. verw. Geb.* **40**, 125–145.

— (1993). Measure-valued Markov processes. In: *École d'été de probabilités de Saint-Flour XXI–1991, Lect. Notes in Math.* **1541**, pp. 1–260. Springer, Berlin.

Dawson, D.A., Iscoe, I., Perkins, E.A. (1989). Super-Brownian motion: path properties and hitting probabilities. *Probab. Th. Rel. Fields* **83**, 135–205.

Dawson, D.A., Perkins, E.A. (1991). Historical processes. *Mem. Amer. Math. Soc.* **93**, #454.

Debes, H., Kerstan, J, Liemant, A., Matthes, K. (1970/71). Verallgemeinerung eines Satzes von Dobrushin I, III. *Math. Nachr.* **47**, 183–244; **50**, 99–139.

Dellacherie, C. (1972). *Capacités et Processus Stochastiques.* Springer, Berlin.

Dellacherie, C., (Maisonneuve, B.,) Meyer, P.A. (1975/80/83/87/92). *Probabilités et Potentiel, I–V.* Hermann, Paris.

Dobrushin, R.L. (1956). On Poisson laws for distributions of particles in space (in Russian). *Ukrain. Mat. Z.* **8**, 127–134.

— (1970). Gibbsian random fields without hard core conditions. *Theor. Math. Phys.* **4**, 705–719.

Doeblin, W. (1938). Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables. *Bull. Soc. Math. France* **66**, 210–220.

— (1939). Sur les sommes d'un grand nombre de variables aléatoires indépendantes. *Bull. Sci. Math.* **63**, 23–64.

— (1940). Sur l'équation de Kolmogoroff. *Pli cacheté 11,668. C.R. Acad. Sci. Paris* **31** (2000), 1059–1102.

Doléans(-Dade), C. (1967). Processus croissants naturel et processus croissants très bien mesurable. *C.R. Acad. Sci. Paris* **264**, 874–876.

— (1970). Quelques applications de la formule de changement de variables pour les semimartingales. *Z. Wahrsch. verw. Geb.* **16**, 181–194.

Doléans-Dade, C., Meyer, P.A. (1970). Intégrales stochastiques par rapport aux martingales locales. *Lect. Notes in Math.* **124**, 77–107. Springer, Berlin.

Donsker, M.D. (1951–52). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6**.

Doob, J.L. (1938). Stochastic processes with an integral-valued parameter. *Trans. Amer. Math. Soc.* **44**, 87–150.

— (1948). Renewal theory from the point of view of the theory of probability. *Trans. Amer. Math. Soc.* **63**, 422–438.

— (1953). *Stochastic Processes.* Wiley, NY.

— (1984). *Classical Potential Theory and its Probabilistic Counterpart.* Springer, NY.

Dubins, L.E., Schwarz, G. (1965). On continuous martingales. *Proc. Natl. Acad. Sci. USA* **53**, 913–916.

Dudley, R.M. (1989). *Real Analysis and Probability.* Wadsworth & Brooks/Cole, Pacific Grove, CA.

Dunford, N. (1939). A mean ergodic theorem. *Duke Math. J.* **5**, 635–646.

Dynkin, E.B. (1961). *Theory of Markov Processes.* Engl. trans., Prentice Hall & Pergamon Press, Englewood Cliffs, NJ, and Oxford (Russian orig. 1959).

— (1965). *Markov Processes, I–II.* Engl. trans., Springer, Berlin (Russian orig. 1963).

— (1978). Sufficient statistics and extreme points. *Ann. Probab.* **6**, 705–730.

— (1991). Branching particle systems and superprocesses. *Ann. Probab.* **19**, 1157–1194.

— (1994). *An Introduction to Branching Measure-Valued Processes. CRM Monograph Ser.* **6**. AMS, Providence, RI.

Ellis, R.L. (1844). On a question in the theory of probabilities. *Cambridge Math. J.* **4** (21), 127–133. Reprinted in W. Walton (ed.), *The Mathematical and Other Writings of Robert Leslie Ellis*, pp. 173–179, Deighton Bell, Cambridge 1863.

Engel, D.D. (1982). The multiple stochastic integral. *Mem. Amer. Math. Soc.* **38**, #265.

Erdös, P., Feller, W., Pollard, H. (1949). A theorem on power series. *Bull. Amer. Math. Soc.* **55**, 201–204.

Erlang, A.K. (1909). The theory of probabilities and telephone conversations. *Nyt. Tidskr. Mat. B*, **20**, 33–41. Reprinted in E. Brockmeyer, H.L. Halstrom, A. Jensen, *The Life and Works of A.K. Erlang*, pp. 131–137, Copenhagen Telephone Company, Copenhagen 1948.

Esseen, C.G. (1968). On the concentration function of a sum of independent random variables. *Z. Wahrsch. verw. Geb.* **9**, 290–308.

Etheridge, A.M. (2000). *An Introduction to Superprocesses. Univ. Lect. Ser.* **20**. AMS, Providence, RI.

Ethier, S.N., Kurtz, T.G. (1986). *Markov Processes, Characterization and Convergence.* Wiley, NY.

Fell, J.M.G. (1962). A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff space. *Proc. Amer. Math. Soc.* **13**, 472–476.

Feller, W. (1936). Zur Theorie der stochastischen Prozesse (Existenz und Eindeutigkeitssätze). *Math. Ann.* **113**, 113–160.

— (1937). On the Kolmogoroff—P. Lévy formula for infinitely divisible distribution functions. *Proc. Yugoslav. Acad. Sci.* **82**, 95–112.

— (1951). Diffusion processes in genetics. *Proc. 2nd Berkeley Symp. Math. Statist. Probab.*, 227–246. Univ. of Calif. Press, Berkeley.

— (1968/71). *An Introduction to Probability Theory and its Applications*, I, 3rd ed.; II, 2nd. ed. Wiley, NY.

Feller, W., Orey, S. (1961). A renewal theorem. *J. Math. Mech.* **10**, 619–624.

Fichtner, K.H. (1974). Gleichverteilungseigenschaften substochastischer Kerne und zufällige Punktfolgen. *Math. Nachr.* **62**, 251–260.

Fieger, W. (1965). Eine für beliebige Call-Prozesse geltende Verallgemeinerung der Palmschen Formeln. *Math. Scand.* **16**, 121–147.

de Finetti, B. (1929). Sulle funzioni ad incremento aleatorio. *Rend. Acc. Naz. Lincei* **10**, 163–168.

— (1930). Fuzione caratteristica di un fenomeno aleatorio. *Mem. R. Acc. Lincei* (6) **4**, 86–133.

— (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7**, 1–68.

Fitzsimmons, P.J., Getoor, R.K. (1995). Occupation time distributions for Lévy bridges and excursions. *Stoch. Proc. Appl.* **58**, 73–89.

Fleischmann, K., (Hermann, K., Liemant, A.,) Matthes, K. (1981/82). Kritische Verzweigungsprozesse mit allgemeinem Phasenraum III, VI, VIII. *Math. Nachr.* **102**, 63–107; **105**, 307–316; **109**, 239–247.

Franken, P, König, D., Arndt, U., Schmidt, V. (1981). *Queues and Point Processes.* Akademie-Verlag, Berlin.

Geman, D., Horowitz, J. (1975). Random shifts which preserve measure. *Proc. Amer. Math. Soc.* **49**, 143–150.

Gentner, D., Last, G. (2011). Palm pairs and the general mass transport principle. *Math. Z.* **267**, 695–716.

Georgii, H.O. (1988). *Gibbs Measures and Phase Transitions.* W. de Gruyter, Berlin.

Getoor, R.K. (1990). *Excessive Measures.* Birkhäuser, Boston.

Gibbs, J.W. (1902). *Elementary Principles in Statistical Mechanics.* Yale Univ. Press.

Gikhman, I.I., Skorokhod, A.V. (1965). Introduction to the Theory of Random Processes (in Russian). Engl. transl., Saunders, Philadelphia 1969.

Glötzl, E. (1980a). Bemerkungen zu einer Arbeit von O.K. Kozlov. *Math. Nachr.* **94**, 277–289.

— (1980b). Lokale Energien und Potentiale für Punktprozesse. *Math. Nachr.* **96**, 195–206.

Glynn, P., Sigman, K. (1992). Uniform Cesàro limit theorems for synchronous processes with applications to queues. *Stoch. Proc. Appl.* **40**, 29–43.

Gnedenko, B.V. (1939). On the theory of limit theorems for sums of independent random variables (in Russian). *Izv. Akad. Nauk SSSR Ser. Mat.* 181–232, 643–647.

Goldman, J.R. (1967). Stochastic point processes: limit theorems. *Ann. Math. Statist.* **38**, 771–779.

Gorostiza, L.G., Roelly-Coppoletta, S., Wakolbinger, A. (1990). Sur la persistance du processus de Dawson–Watanabe stable. L'intervertion de la limite en temps et de la renormalisation. In: *Lecture Notes in Math.* **1426**, 275–281. Springer, Berlin.

Gorostiza, L.G., Wakolbinger, A. (1991). Persistence criteria for a class of critical branching systems in continuous time. *Ann. Probab.* **19**, 266–288.

Grandell, J. (1976). *Doubly stochastic Poisson processes. Lect. Notes in Math.* **529**. Springer, Berlin.

Grigelionis, B. (1963). On the convergence of sums of random step processes to a Poisson process. *Th. Probab. Appl.* **8**, 172–182.

— (1971). On the representation of integer-valued measures by means of stochastic integrals with respect to Poisson measure. *Litovsk. Mat. Sb.* **11**, 93–108.

— (1977). Martingale characterization of stochastic processes with independent increments. *Litovsk. Mat. Sb.* **17**, 75–86.

Guttorp, P., Thorarinsdottir, T.I. (2012). What happened to discrete chaos, the Quenouille process, and the sharp Markov property? Some history of stochastic point processes. *Intern. Statist. Revue* **80**, 253–268.

Haar, A. (1933). Der Maßbegriff in der Theorie der kontinuerlichen Gruppen. *Ann. Math.* **34**, 147–169.

Hagberg, J. (1973). Approximation of the summation process obtained by sampling from a finite population. *Th. Probab. Appl.* **18**, 790–803.

Halmos, P.R. (1950). *Measure Theory.* Van Nostrand, NY. Reprint ed., Springer, NY 1974.

Harris, T.E. (1963). *The Theory of Branching Processes.* Springer, Berlin. Reprint ed., Dover, NY 1989.

— (1965). Diffusions with collisions between particles. *J. Appl. Probab.* **2**, 323–338.

— (1968). Counting measures, monotone random set functions. *Z. Wahrsch. verw. Geb.* **10**, 102–119.

— (1971). Random measures and motions of point processes. *Z. Wahrsch. verw. Geb.* **18**, 85–115.

Hausdorff, F. (1916). Die Mächtigkeit der Borelschen Mengen. *Math. Ann.* **77**, 430–437.

— (1921). Summationsmethoden und Momentfolgen. *Math. Z.* **9**, 280–299.

Heinonen, J. (2001). *Lectures on Analysis on Metric Spaces.* Springer, New York.

Hengartner, W., Theodorescu, R. (1973). *Concentration functions.* Academic Press, New York.

Hewitt, E., Ross, K.A. (1979). *Abstract Harmonic Analysis I*, 2nd ed. Springer, NY.

Hitczenko, P. (1988). Comparison of moments for tangent sequences of random variables. *Probab. Th. Rel. Fields* **78**, 223–230.

van der Hoeven, P.C.T. (1982). Une projection de processus ponctuels. *Z. Wahrsch. verw. Geb.* **61**, 483–499.

— (1983). *On Point Processes.* Dissertation, Math. Centrum, Amsterdam.

Holroyd, A.E., Peres, Y. (2005). Extra heads and invariant allocations. *Ann. Probab.* **33**, 31–52.

Horowitz, J. (1972). Semilinear Markov processes, subordinators and renewal theory. *Z. Wahrsch. verw. Geb.* **24**, 167–193.

Hunt, G.A. (1957/58). Markoff processes and potentials, I–III. *Illinois J. Math.* **1**, 44–93, 316–369; **2**, 151–213.

Hurwitz, A. (1897). Über die Erzeugung der Invarianten durch Integration. *Nachr. Ges. Göttingen, Math.-Phys. KL* 71–90.

Itô, K. (1942a). On stochastic processes I: Infinitely divisible laws of probability. *Jap. J. Math.* **18**, 261–301.

— (1942b). Differential equations determining a Markov process. *J. Pan-Japan Math. Coll.* 1077.

— (1956). Spectral type of the shift transformation of differential processes with stationary increments. *Trans. Amer. Math. Soc.* **81**, 253–263.

— (1972). Poisson point processes attached to Markov processes. In: *Proc. 6th Berkeley Symp. Math. Statist. Probab.* **3**, 225–239. Univ. California Press, Berkeley.

Jacobs, P.A. (1977). A Poisson convergence theorem for a particle system with dependent constant velocities. *Stoch. Proc. Appl.* **6**, 41–52.

Jacobsen, M. (2006). *Point Process Theory and Applications.* Birkhäuser, Boston.

Jacod, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Z. Wahrsch. verw. Geb.* **31**, 235–253.

— (1979). *Calcul Stochastique et Problèmes de Martingales. Lect. Notes in Math.* **714**. Springer, Berlin.

— (1984). Une généralisation des semimartingales: les processus admettant un processus à accroissements indépendants tangent. *Lect. Notes in Math.* **1059**, 91–118. Springer, Berlin.

Jacod, J., Shiryaev, A.N. (1987). *Limit Theorems for Stochastic Processes.* Springer, Berlin.

Jagers, P. (1972). On the weak convergence of superpositions of point processes. *Z. Wahrsch. verw. Geb.* **22**, 1–7.

— (1973). On Palm probabilities. *Z. Wahrsch. verw. Geb.* **26**, 17–32.

— (1974). Aspects of random measures and point processes. In: *Advances in Probability and Related Topics* **3**, 179–239. Marcel Dekker, NY.

— (1975). *Branching Processes with Biological Applications.* Wiley, London.

Jakubowski, A. (1986). Principle of conditioning in limit theorems for sums of random variables. *Ann. Probab.* **14**, 902–915.

Jiřina, M. (1964). Branching processes with measure-valued states. In: *Trans. 3rd Prague Conf. Inf. Th. Statist. Dec. Func. Rand. Proc.*, pp. 333–357.

— (1966). Asymptotic behavior of measure-valued branching processes. *Rozpravy Československé Akad. Věd. Řada Mat. Přirod. Věd* **76**:3, Praha.

Kabanov, Y.M. (1975). On extended stochastic integrals. *Th. Probab. Appl.* **20**, 710–722.

Kac, M. (1947). On the notion of recurrence in discrete stochastic processes. *Bull. Amer. Math. Soc.* **53**, 1002–1010.

Kallenberg, O. (1973a). Characterization and convergence of random measures and point processes. *Z. Wahrsch. verw. Geb.* **27**, 9–21.

— (1973b). Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrsch. verw. Geb.* **27**, 23–36.

— (1973c). Characterization of continuous random processes and signed measures. *Studia Sci. Math. Hungar.* **8**, 473–477.

— (1975a). On symmetrically distributed random measures. *Trans. Amer. Math. Soc.* **202**, 105–121.

— (1975b). Limits of compound and thinned point processes. *J. Appl. Probab.* **12**, 269–278.

— (1975/76, 1983/86). *Random Measures*, 1st–2nd eds. (RM_1); 3rd–4th eds. (RM_2). Akademie-Verlag and Academic Press, Berlin and London.

— (1976/80/81). On the structure of stationary flat processes, I–III. *Z. Wahrsch. verw. Geb.* **37**, 157–174; **52**, 127–147; **56**, 239–253.

— (1977a). Stability of critical cluster fields. *Math. Nachr.* **77**, 1–43.

— (1977b). A counterexample to R. Davidson's conjecture on line processes. *Math. Proc. Cambridge Philos. Soc.* **82**, 301–307.

— (1978a). On conditional intensities of point processes. *Z. Wahrsch. verw. Geb.* **41**, 205–220.

— (1978b). On the asymptotic behavior of line processes and systems of non-interacting particles. *Z. Wahrsch. verw. Geb.* **43**, 65–95.

— (1978c). On the independence of velocities in a system of non-interacting particles. *Ann. Probab.* **6**, 885–890.

— (1979). L_p-intensities of random measures. *Stoch. Proc. Appl.* **9**, 155–161.

— (1981a). Splitting at backward times in regenerative sets. *Ann. Probab.* **9**, 781–799.

— (1988). Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16**, 508–534.

— (1990a). Exchangeable random measures in the plane. *J. Theor. Probab.* **3**, 81–136.

— (1990b). Random time change and an integral representation for marked stopping times. *Probab. Th. Rel. Fields* **86**, 167–202.

— (1991). Some uses of point processes in multiple stochastic integration. *Math. Nachr.* **151**, 7–31.

— (1992). Some time change representations of stable integrals, via predictable transformations of local martingales. *Stoch. Proc. Appl.* **40**, 199–223.

— (1996). Improved criteria for distributional convergence of point processes. *Stoch. Proc. Appl.* **64**, 93–102.

— (1997/2002). *Foundations of Modern Probability* (FMP), 1st–2nd eds. Springer, NY.

— (1999a). Palm measure duality and conditioning in regenerative sets. *Ann. Probab.* **27**, 945–969.

— (1999b). Ballot theorems and sojourn laws for stationary processes. *Ann. Probab.* **27**, 2011–2019.

— (2000). An extension of the basic Palm measure correspondence. *Probab. Th. Rel. Fields* **117**, 118–131.

— (2001). Local hitting and conditioning in symmetric interval partitions. *Stoch. Proc. Appl.* **94**, 241–270.

— (2002a). Ergodic properties of Palm and spacing measures. *J. Theor. Probab.* **15**, 77–96.

— (2003). Palm distributions and local approximation of regenerative processes. *Probab. Th. Rel. Fields* **125**, 1–41.

— (2005). *Probabilistic Symmetries and Invariance Principles.* Springer, NY.

— (2007a). Invariant measures and disintegrations with applications to Palm and related kernels. *Probab. Th. Rel. Fields* **139**, 285–310, 311.

— (2007b). Some problems of local hitting, scaling, and conditioning. *Acta Appl. Math.* **96**, 271–282.

— (2008). Some local approximations of Dawson–Watanabe superprocesses. *Ann. Probab.* **36**, 2176–2214.

— (2009). Some local approximation properties of simple point processes. *Probab. Th. Rel. Fields* **143**, 73–96.

— (2010). Commutativity properties of conditional distributions and Palm measures. *Comm. Stoch. Anal.* **4**, 21–34.

— (2011a). Invariant Palm and related disintegrations via skew factorization. *Probab. Th. Rel. Fields* **149**, 279–301.

— (2011b). Iterated Palm conditioning and some Slivnyak-type theorems for Cox and cluster processes. *J. Theor. Probab.* **24**, 875–893.

— (2013). Local conditioning in Dawson–Watanabe superprocesses. *Ann. Probab.* **41**, 385–443.

— (2014). Stationary and invariant densities and disintegration kernels. *Probab. Theor. Rel. Fields* **160**, 567–592.

— (2016). Tangential existence and comparison, with applications to single and multiple integration. *Probab. Math. Statist.* (to appear)

Kallenberg, O., Szulga, J. (1989). Multiple integration with respect to Poisson and Lévy processes. *Probab. Th. Rel. Fields* **83**, 101–134.

Kaplan, E.L. (1955). Transformations of stationary random sequences. *Math. Scand.* **3**, 127–149.

Karbe, W. (1973). Konstruktion einfacher zufälliger Punktfolgen. Diplomarbeit, Friedrich–Schiller–Universität Jena, Sektion Mathematik.

Kendall, D.G. (1948). On the generalized “birth-and-death” process. *Ann. Math. Statist.* **19**, 1–15.

— (1949). Stochastic processes and population growth. *J. Roy. Statist. Soc. Ser. B* **11**, 230–264.

— (1966). Branching processes since 1873. *J. London Math. Soc.* **41**, 385–406.

— (1974). Foundations of a theory of random sets. In: *Stochastic Geometry* (E.F. Harding & D.G. Kendall, eds.), pp. 322–376. Wiley, London.

— (1975). The genealogy of genealogy: Branching processes before (and after) 1873. *Bull. London Math. Soc.* **7**, 225–253.

Kerstan, J., Matthes, K. (1964a). Stationäre zufällige Punktfolgen II. *Jahresber. Deutsch. Mat.-Verein.* **66**, 106–118.

— (1964b). Verallgemeinerung eines Satzes von Sliwnjak. *Rev. Roumaine Math. Pures Appl.* **9**, 811–829.

Kerstan, J., Matthes, K.. Mecke, J. (1974/82). *Unbegrenzt teilbare Punktprozesse.* See Matthes et al. (1974/78/82).

Khinchin, A.Y. (1937). Zur Theorie der unbeschränkt teilbaren Verteilungsgesetze. *Mat. Sb.* **2**, 79–119.

— (1955). *Mathematical Methods in the Theory of Queuing* (Russian). Engl. transl., Griffin, London 1960.

Kingman, J.F.C. (1964). On doubly stochastic Poisson processes. *Proc. Cambridge Phil. Soc.* **60**, 923–930.

— (1967). Completely random measures. *Pac. J. Math.* **21**, 59–78.

— (1972). *Regenerative Phenomena.* Wiley, London.

— (1973). An intrinsic description of local time. *J. London Math. Soc.* **6** (2), 725–731.

— (1993). *Poisson Processes.* Clarendon Press, Oxford.

Klenke, A. (1998). Clustering and invariant measures for spatial branching models with infinite variance. *Ann. Probab.* **26**, 1057–1087.

Knight, F. (1971). A reduction of continuous, square-integrable martingales to Brownian motion. *Lect. Notes in Math.* **190**, 19–31. Springer, Berlin.

— (1996). The uniform law for exchangeable and Lévy process bridges. In: *Hommage à P.A. Meyer et J. Neveu. Astérisque* **236**, 171–188.

Kolmogorov, A.N. (1931). Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* **104**, 415–458.

— (1932). Sulla forma generale di un processo stocastico omogeneo (un problema di B. de Finetti). *Atti Accad. Naz. Lincei Rend.* (6) **15**, 805–808, 866–869.

— (1935). La transformation de Laplace dans les espaces linéaires. *C.R. Acad. Sci. Paris* **200**, 1717–1718.

— (1936a). Anfangsgründe der Markoffschen Ketten mit unendlich vielen möglichen Zuständen. *Mat. Sb.* **1**, 607–610.

— (1936b). Zur Theorie der Markoffschen Ketten. *Math. Ann.* **112**, 155–160.

— (1938). Zur Lösung einer biologischen Aufgabe. *Izvest. Nauchno—Issled. Inst. Mat. Mech. Tomskom Gosud. Univ.* **2**, 1–6.

— (1958). Sur les propriétés des fonctions de concentration de M. P. Lévy. *Ann. Inst. Henri Poincaré* **16**, 27–34.

König, D., Matthes, K. (1963). Verallgemeinerung der Erlangschen Formeln I. *Math. Nachr.* **26**, 45–56.

König, D., Matthes, K., Nawrotzki, K. (1974). Unempfindlichkeitseigenschaften von Bedienungsprozessen. In: B.W. Gnedenko & I.N. Kowalenko: *Einführung in die Bedienuingstheorie*, pp. 357–450. Akademie-Verlag, Berlin.

Kozlov, O.K. (1976). Gibbsian description of point random fields. *Th. Probab. Appl.* **21**, 339–355.

Krakowiak, W., Szulga, J. (1986). Random multilinear forms. *Ann. Probab.* **14**, 955–973.

— (1988). A multiple stochastic integral with respect to a strictly p-stable random measure. *Ann. Probab.* **16**, 764–777.

Krengel, U. (1985). *Ergodic Theorems.* de Gruyter, Berlin.

Krickeberg, K. (1972). The Cox process. *Symp. Math.* **9**, 151–167.

— (1974a). Invariance properties of the correlation measure of line-processes. In: *Stochastic Geometry* (E.F. Harding & D.G. Kendall, eds.), pp. 76–88. Wiley, London.

— (1974b). Moments of point processes. *Ibid.*, pp. 89–113.

— (1982). Processus ponctuels en statistique. In: *École d'été de probabilités de Saint-Flour X–1980. Lect. Notes in Math.* **929**, 205–313.

Kummer, G., Matthes, K. (1970). Verallgemeinerung eines Satzes von Sliwnjak II–III. *Rev. Roumaine Math. Pures Appl.* **15**, 845–870, 1631–1642.

Kunita, H., Watanabe, S. (1967). On square integrable martingales. *Nagoya Math. J.* **30**, 209–245.

Kuratowski, K. (1933). *Topology, I.* English transl., Academic Press, NY 1966.

Kurtz, T.G. (1974). Point processes and completely monotone set functions. *Z. Wahrsch. verw. Geb.* **31**, 57–67.

Kwapień, S., Woyczyński, W.A. (1987). Double stochastic integrals, random quadratic forms and random series in Orlicz spaces. *Ann. Probab.* **15**, 1072–1096.

— (1991). Semimartingale integrals via decoupling inequalities and tangent processes. *Probab. Math. Statist.* **12**, 165–200.

— (1992). *Random Series and Stochastic Integrals: Single and Multiple.* Birkhäuser, Boston.

Lanford, O.E., Ruelle, D. (1969). Observables at infinity and states with short range correlations in statistical mechanics. *Comm. Math. Phys.* **13**, 194–215.

Last, G. (2009). Modern random measures: Palm theory and related models. In: *New Perspectives in Stochastic Geometry* (W. Kendall & I. Molchanov, eds.). Oxford Univ. Press.

— (2010). Stationary random measures on homogeneous spaces. *J. Theor. Probab.* **23**, 478–497.

Last, G., Brandt, A. (1995). *Marked Point Processes on the Real Line: The Dynamic Approach.* Springer, NY.

Last, G., Penrose, M.D. (2017). *Lectures on the Poisson Process.* Cambridge Univ. Press, to appear.

Last, G., Thorisson, H. (2009). Invariant transports of stationary random measures and mass-stationarity. *Ann. Probab.* **37**, 790–813.

Leadbetter, M.R. (1972). On basic results of point process theory. *Proc. 6th Berkeley Symp. Math. Statist. Probab.* **3**, 449–462. Univ. California Press, Berkeley.

Lebesgue, H. (1904). Leçons sur l'intégration et la recherche des fonctions primitives. *Oeuvres Scientifiques* **2**, 111–154.

— (1905). Sur les fonctions représentables analytiquement. *J. Math.* (Ser 6) **1**, 139–216.

Lee, P.M. (1967). Infinitely divisible stochastic processes. *Z. Wahrsch. verw. Geb.* **7**, 147–160.

Le Gall, J.F. (1986). Une approche élémentaire des théorèmes de décomposition de Williams. *Lect. Notes in Math.* **1204**, 447–464. Springer, Berlin.

— (1991). Brownian excursions, trees and measure-valued branching processes. *Ann. Probab.* **19**, 1399–1439.

— (1994). A lemma on super-Brownian motion with some applications. In: *The Dynkin Festschrift*, pp. 237–251. Birkhäuser, Boston.

— (1999). *Spatial Branching Processes, Random Snakes and Partial Differential Equations. Lect. in Math., ETH Zürich.* Birkhäuser, Basel.

Lévy, P. (1934–35). Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Ann. Scuola Norm. Sup. Pisa (2)*, **3**, 337–366; **4**, 217–218.

— (1937/54). *Théorie de l'addition des variables aléatoires*, eds. 1–2. Gautier-Villars, Paris.

— (1939). Sur certain processus stochastiques homogènes. *Comp. Math.* **7**, 283–339.

Li, Z. (2011). *Measure-valued Branching Markov Processes.* Springer, Heidelberg.

Liemant, A., Matthes, K. (1974/75/77). Verallgemeinerungen eines Satzes von Dobrushin IV, VI. *Math. Nachr.* **59**, 311–317; **70**, 387–390; **80**, 7–18.

Liemant, A., Matthes, K., Wakolbinger, A. (1988). *Equilibrium Distributions of Branching Processes.* Akademie-Verlag, Berlin.

Lindvall, T. (1973). Weak convergence of probability measures and random functions in the function space $D[0,\infty)$. *J. Appl. Probab.* **10**, 109–121.

— (1977). A probabilistic proof of Blackwell's renewal theorem. *Ann. Probab.* **5**, 482–485.

Liptser, R.S., Shiryayev, A.N. (1977/78). *Statistics of Random Processes, I–II.* Springer, NY (2nd ed. 2000).

Loève, M. (1955/77–78). *Probability Theory, I–II*, eds. 1–4. Springer, NY.

Lundberg, F. (1903). *Approximate representation of the probability function. Reinsurance of collective risks* (in Swedish). Ph.D. thesis, Almqvist & Wiksell, Uppsala.

Lusin, N.N. (1917). Sur la classification de M. Baire. *C.R. Acad. Sci. Paris* **164**, 91–94.

Marcinkiewicz, J., Zygmund, A. (1937). Sur les fonctions independantes. *Fund. Math.* **29**, 60–90.

Martin-Löf, A. (1995). Harald Cramér and insurance mathematics. *Appl. Stoch. Models and Data Anal.* **11**, 271–276.

Maruyama, G. (1955). On the Poisson distribution derived from independent random walks. *Nat. Sci. Rep. Ochanomizu Univ.* **6**, 1–6.

Matheron, G. (1975). *Random Sets and Integral Geometry.* Wiley, London.

Matthes, K. (1963a). Unbeschränkt teilbare Verteilungsgesetze stationärer zufälliger Punktfolgen. *Wiss. Z. Hochsch. Elektrotechnik Ilmenau* **9**, 235–238.

— (1963b). Stationäre zufällige Punktfolgen I. *Jahresber. Deutsch. Math.-Verein.* **66**, 66–79.

— (1969). Eine Characterisierung der kontinuerlichen unbegrenzt teilbaren Verteilungsgesetze zufälliger Punktfolgen. *Rev. Roumaine Math. Pures Appl.* **14**, 1121–1127.

Matthes, K., Kerstan, J., Mecke, J. (1974/78/82). *Infinitely Divisible Point Processes.* 1st German ed. (MKM_1), Akademie-Verlag, Berlin; 2nd English ed. (MKM_2), Wiley, Chicester; 3rd Russian ed. (MKM_3), Nauka, Moscow.

Matthes, K., Warmuth, W., Mecke, J. (1979). Bemerkungen zu einer Arbeit von Nguyen Xuan Xanh and Hans Zessin. *Math. Nachr.* **88**, 117–127.

Mazurkiewicz, S. (1916). Über Borelsche Mengen. *Bull. Acad. Cracovie* **1916**, 490–494.

Mecke, J. (1967). Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen. *Z. Wahrsch. verw. Geb.* **9**, 36–58.

— (1968). Eine characteristische Eigenschaft der doppelt stochastischen Poisson-Prozesse. *Z. Wahrsch. verw. Geb.* **11**, 74–81.

— (1975). Invarianzeigenschaften allgemeiner Palmscher Maße. *Math. Nachr.* **65**, 335–344.

— (1979). An explicit description of Kallenberg's lattice-type point process. *Math. Nachr.* **89**, 185–195.

Meyer, P.A. (1962). A decomposition theorem for supermartingales. *Illinois J. Math.* **6**, 193–205.

— (1963). Decomposition of supermartingales: the uniqueness theorem. *Illinois J. Math.* **7**, 1–17.

— (1966). *Probability and Potentials.* Engl. trans., Blaisdell, Waltham.

— (1967). Intégrales stochastiques, I–IV. *Lect. Notes in Math.* **39**, 72–162. Springer, Berlin.

— (1971). Démonstration simplifiée d'un théorème de Knight. *Lect. Notes in Math.* **191**, 191–195. Springer, Berlin.

— (1976). Un cours sur les intégrales stochastiques. *Lect. Notes in Math.* **511**, 245–398. Springer, Berlin.

Meyer, P.A., Yor, M. (1976). Sur la théorie de la prédiction, et le problème de décomposition des tribus $\mathcal{F}^o_{t+}$. In: *Séminaire de Probabilités X*, 104–117. *Lect. Notes in Math.* **511**. Springer, Berlin.

Miles, R.E. (1969/71). Poisson flats in Euclidean spaces, I–II. *Adv. Appl. Probab.* **1**, 211–237; **3**, 1–43.

— (1974). A synopsis of 'Poisson flats in Euclidean spaces'. In: *Stochastic Geometry* (E.F. Harding & D.G. Kendall, eds.), pp. 202–227. Wiley, London.

Millar, P.W. (1968). Martingale integrals. *Trans. Amer. Math. Soc.* **133**, 145–166.

de Moivre, A. (1711–12). On the measurement of chance. Engl. trans.: *Int. Statist. Review* **52** (1984), 229–262.

Molchanov, I. (2005). *Theory of Random sets.* Springer, London.

Mönch, G. (1971). Verallgemeinerung eines Satzes von A. Rényi. *Studia Sci. Math. Hungar.* **6**, 81–90.

Montgomery, D., Zippin, L. (1955). *Topological Transformation Groups.* Interscience Publishers, New York.

Moyal, J.E. (1962). The general theory of stochastic population processes. *Acta Math.* **108**, 1–31.

Nawrotzki, K. (1962). Ein Grenzwertsatz für homogene zufällige Punktfolgen (Verallgemeinerung eines Satzes von A. Rényi). *Math. Nachr.* **24**, 201–217.

— (1968). Mischungseigenschaften stationärer unbegrenzt teilbarer zufälliger Maße. *Math. Nachr.* **38**, 97–114.

— (1978). Einige Bemerkungen zur Verwendung der Palmschen Verteilung in der Bedienungstheorie. *Math. Oper.-forsch. Statist., Ser. Optim.* **9**, 241–253.

Nelson, E. (1973). The free Markov field. *J. Funct. Anal.* **12**, 211–227.

von Neumann, J. (1931). Algebraische Repräsentanten der Funktionen "bis auf eine Menge vom Masse Null". *J. Crelle* **165**, 109–115.

Neuts, M., Resnick, S. (1971). On the times of birth in a linear birth process. *J. Austral. Math. Soc.* **12**, 473–475.

Neveu, J. (1968). Sur la structure des processus ponctuels stationnaires. *C.R. Acad. Sc. Paris* **267A**, 561–564.

— (1976). Sur les mesures de Palm de deux processus ponctuels stationnaires. *Z. Wahrsch. verw. Geb.* **34**, 199–203.

— (1977). Processus ponctuels. In: *École d'été de probabilités de Saint-Flour VI–1976. Lect. Notes in Math.* **598**, 249–445. Springer, Berlin.

— (1986). Arbres et processus de Galton–Watson. *Ann. Inst. Henri Poincaré* **22**, 199–207.

Neyman, J., Scott, E.L. (1958). Statistical approach to problems of cosmology (with discussion). *J. Roy. Statist. Soc. Ser. B* **20**, 1–43.

— (1972). Processes of clustering and applications. In: *Stochastic Point Processes* (P.A.W. Lewis, ed.), pp. 646–681. Wiley, NY.

Nguyen, X.X., Zessin, H. (1979). Ergodic theorems for spatial processes. *Z. Wahrsch. verw. Geb.* **48**, 133–158.

Nieuwenhuis, G. (1989). Asymptotics of functional limit theorems for stationary point processes and their Palm distributions. *Probab. Th. Rel. Fields* **81**, 593–608.

— (1994). Bridging the gap between a stationary point process and its Palm distribution. *Statist. Nederl.* **48**, 37–62.

Norberg, T. (1984). Convergence and existence of random set distributions. *Ann. Probab.* **12**, 726–732.

Ogura, H. (1972). Orthogonal functionals of the Poisson process. *IEEE Trans. Inf. Theory* **18**, 473–481.

Ososkov, G.A. (1956). A limit theorem for flows of homogeneous events. *Th. Probab. Appl.* **1**, 248–255.

Paley, R.E.A.C., Zygmund, A. (1932). A note on analytic functions on the unit circle. *Proc. Cambridge Phil. Soc.* **28**, 266–272.

Palm, C. (1943). Intensity variations in telephone traffic (in German). *Eriksson Technics* **44**, 1–189. Engl. transl., *North-Holland Studies in Telecommunication* **10**, Elsevier 1988.

Papangelou, F. (1970). The Ambrose–Kakutani theorem and the Poisson process. In: *Lect. Notes in Math.* **160**, 234–240. Springer, Berlin.

— (1972). Integrability of expected increments of point processes and a related random change of scale. *Trans. Amer. Math. Soc.* **165**, 486–506.

— (1974a). On the Palm probabilities of processes of points and processes of lines. In: *Stochastic Geometry* (E.F. Harding & D.G. Kendall, eds.), pp. 114–147. Wiley, London.

— (1974b). The conditional intensity of general point processes and an application to line processes. *Z. Wahrsch. verw. Geb.* **28**, 207–226.

— (1976). Point processes on spaces of flats and other homogeneous spaces. *Math. Proc. Cambridge Phil. Soc.* **80**, 297–314.

Parthasarathy, K.R. (1967). *Probability Measures on Metric Spaces.* Academic Press, NY.

de la Peña, V.H., Giné, E. (1999). *Decoupling: From Dependence to Independence.* Springer, NY.

Perkins, E.A. (2002). Dawson–Watanabe superprocesses and measure-valued diffusions. In: *École d'été de probabilités de Saint-Flour XXIX–1999. Lect. Notes in Math.* **1781**, 125–329. Springer, Berlin.

Peterson, L.D. (2001). Convergence of random measures on Polish spaces. Dissertation, Auburn University.

Petrov, V.V. (1995). *Limit theorems of Probability Theory: Sequences of Independent Random Variables.* Clarendon Press, Oxford.

Poisson, S.D. (1837). *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile, Précédées des Règles Générales du Calcul des Probabilités.* Bachelier, Paris.

Port, S.C., Stone, C.J. (1973). Infinite particle systems. *Trans. Amer. Math. Soc.* **178**, 307–340.

de Possel, R. (1936). Sur la dérivation abstraite des fonctions d'ensembles. *J. Math. Pures Appl.* **15**, 391–409.

Prékopa, A. (1958). On secondary processes generated by a random point distribution of Poisson type. *Ann. Univ. Sci. Budapest Sect. Math.* **1**, 153–170.

Preston, C. (1976). *Random Fields. Lect. Notes in Math.* **534**. Springer, Berlin.

Prohorov, Y.V. (1956). Convergence of random processes and limit theorems in probability theory. *Th. Probab. Appl.* **1**, 157–214.

— (1961). Random measures on a compactum. *Soviet Math. Dokl.* **2**, 539–541.

Rachev, S.T., Rüschendorf, L. (1998). *Mass Transportation Problems, I: Theory.* Springer, NY.

Rao, K.M. (1969). On decomposition theorems of Meyer. *Math. Scand.* **24**, 66–78.

Rauchenschwandtner, B. (1980). *Gibbsprozesse und Papangeloukerne.* Verb. wissensch. Gesellsch. Österreichs, Vienna.

Rauchenschwandtner, B., Wakolbinger, A. (1981). Some aspects of the Papangelou kernel. In: *Point Processes and Queuing Problems*, pp. 325–336. North-Holland, Amsterdam.

Rényi, A. (1953). On the theory of order statistics. *Acta Math. Acad. Sci. Hung.* **4**, 191–231.

— (1956). A characterization of Poisson processes. *Magyar Tud. Akad. Mat. Kutato Int. Közl.* **1**, 519–527.

— (1967). Remarks on the Poisson process. *Studia Sci. Math. Hungar.* **2**, 119–123.

Ripley, B.D. (1976). Locally finite random sets: foundations for point process theory. *Ann. Probab.* **4**, 983–994.

Rogozin, B.A. (1961). On an estimate of the concentration function. *Th. Probab. Appl.* **6**, 94–97.

Rosén, B. (1964). Limit theorems for sampling from a finite population. *Ark. Mat.* **5**, 383–424.

Rosiński, J., Woyczyński, W.A. (1986). On Itô stochastic integration with respect to p-stable motion: Inner clock, integrability of sample paths, double and multiple integrals. *Ann. Probab.* **14**, 271–286.

Rother, W., Zähle, M. (1990). Palm distributions in homogeneous spaces. *Math. Nachr.* **149**, 255–263.

Ryll-Nardzewski, C. (1957). On stationary sequences of random variables and the de Finetti's equivalence. *Colloq. Math.* **4**, 149–156.

— (1961). Remarks on processes of calls. In: *Proc. 4th Berkeley Symp. Math. Statist. Probab.* **2**, 455–465. Univ. California Press, Berkeley.

Schindler, W. (2003). *Measures with Symmetry Properties. Lect. Notes in Math.* **1808**. Springer, Berlin.

Schneider, R., Weil, W. (2008). *Stochastic and Integral Geometry.* Springer, Berlin.

Segall, A., Kailath, T. (1976). Orthogonal functionals of independent-increment processes. *IEEE Trans. Inf. Theory* **22**, 287–298.

Sharpe, M. (1988). *General Theory of Markov Processes.* Academic Press, Boston.

Shilov, G.E., Gurevich, B.L. (1966). *Integral, Measure and Derivative: A Unified Approach.* Prentice-Hall, Englewood Cliffs, NJ.

Sierpiński, W. (1928). Une théorème générale sur les familles d'ensemble. *Fund. Math.* **12**, 206–210.

Skorohod, A.V. (1956). Limit theorems for stochastic processes. *Th. Probab. Appl.* **1**, 261–290.

— (1957). Limit theorems for stochastic processes with independent increments. *Th. Probab. Appl.* **2**, 122–142.

— (1961/62). Stochastic equations for diffusion processes in a bounded region, I–II. *Th. Probab. Appl.* **6**, 264–274; **7**, 3–23.

Slivnyak, I.M. (1962). Some properties of stationary flows of homogeneous random events. *Th. Probab. Appl.* **7**, 336–341; **9**, 168.

Sparre-Andersen, E. (1953/54). On the fluctuations of sums of random variables, I–II. *Math. Scand.* **1**, 263–285; **2**, 195–223.

Spitzer, F. (1969). Uniform motion with elastic collision of an infinite particle system. *J. Math. Mech.* **18**, 973–989.

Stam, A.J. (1967). On shifting iterated convolutions I–II. *Compositio Math.* **17**, 268–280; **18**, 201–228.

Stone, C.J. (1963). Weak convergence of stochastic processes defined on a semi-infinite time interval. *Proc. Amer. Math. Soc.* **14**, 694–696.

— (1966). On absolutely continuous components and renewal theory. *Ann. Math. Statist.* **37**, 271–275.

— (1967). On local and ratio limit theorems. *Proc. 5th Berkeley Symp. Math. Statist. Probab.* **2**, 217–224.

— (1968). On a theorem by Dobrushin. *Ann. Math. Statist.* **39**, 1391–1401.

Stoyan, D., Kendall, W.S., Mecke, J. (1995). *Stochastic Geometry and its Applications*, 2nd ed. Wiley, Chichester.

Surgailis, D. (1984). On multiple Poisson stochastic integrals and associated Markov semigroups. *Probab. Math. Statist.* **3**, 217–239.

— (1985). On the multiple stable integral. *Z. Wahrsch. verw. Geb.* **70**, 621–632.

Suslin, M.Y. (1917). Sur une définition des ensembles mesurables B sans nombres transfinis. *C.R. Acad. Sci. Paris* **164**, 88–91.

Szász, D.O.H. (1970). Once more on the Poisson process. *Studia Sci. Math. Hungar.* **5**, 441–444.

Szatzschneider, W. (1975). A more deterministic version of Harris–Spitzer's "random constant velocity" model for infinite systems of particles. *Lect. Notes in Math.* **472**, 157–168. Springer, Berlin.

Takács, L. (1967). *Combinatorial Methods in the Theory of Stochastic Processes.* Wiley, NY.

Tanaka, H. (1963). Note on continuous additive functionals of the 1-dimensional Brownian path. *Z. Wahrsch. verw. Geb.* **1**, 251–257.

Tempel'man, A.A. (1972). Ergodic theorems for general dynamical systems. *Trans. Moscow Math. Soc.* **26**, 94–132.

Thedéen, T. (1964). A note on the Poisson tendency in traffic distribution. *Ann. Math. Statist.* **35**, 1823–1824.

Thorisson, H. (1995). On time and cycle stationarity. *Stoch. Proc. Appl.* **55**, 183–209.

— (1996). Transforming random elements and shifting random fields. *Ann. Probab.* **24**, 2057–2064.

— (2000). *Coupling, Stationarity, and Regeneration.* Springer, NY.

Tortrat, A. (1969). Sur les mesures aléatoires dans les groupes non abéliens. *Ann. Inst. H. Poincaré, Sec. B*, **5**, 31–47.

Tribe, R. (1994). A representation for super Brownian motion. *Stoch. Proc. Appl.* **51**, 207–219.

Trotter, H.F. (1958). A property of Brownian motion paths. *Illinois J. Math.* **2**, 425–433.

Tucker, H. (1965). On a necessary and sufficient condition that an infinitely divisible distribution be absolutely continuous. *Trans. Amer. Math. Soc.* **118**, 316–330.

Vervaat, W. (1997). Random upper semicontinuous functions and extremal processes. In: *Probability and Lattices* (W. Vervaat & H. Holwerda, eds.), *CWI Tracts* **110**, 1–56. CWI, Amsterdam.

Vitali, G. (1908). Sui gruppi di punti e sulle funzioni di variabili reali. *Atti. R. Accad. Sci. Torino* **43**, 75–92.

Volkonsky, V.A. (1960). Additive functionals of Markov processes. *Trudy Mosk. Mat. Obshc.* **9**, 143–189.

von Waldenfels, W. (1968). Characteristische Funktionale zufälliger Maße. *Z. Wahrsch. verw. Geb.* **10**, 279–283.

Wang, A.T. (1977). Generalized Itô's formula and additive functionals of Brownian motion. *Z. Wahrsch. verw. Geb.* **41**, 153–159.

Watanabe, H. (1956). On the Poisson distribution. *J. Math. Soc. Japan* **8**, 127–134.

Watanabe, S. (1964). On discontinuous additive functionals and Lévy measures of a Markov process. *Japan. J. Math.* **34**, 53–79.

— (1968). A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.* **8**, 141–167.

Watson, H.W., Galton, F. (1874). On the probability of the extinction of families. *J. Anthropol. Inst. Gr. Britain and Ireland* **4**, 138–144.

Waugh, W.A.O'N. (1970). Transformation of a birth process into a Poisson process. *J. Roy. Statist. Soc. B* **32**, 418–431.

Weil, A. (1936). La mesure invariante dans les espaces de groupes et les espaces homogènes. *Enseignement Math.* **35**, 241.

— (1940). L'intégration dans les groupes topologique et ses applications. *Actualités Sci. et Ind.* **869**, 1145. Hermann et Cie, Paris.

Westcott, M. (1976). Simple proof of a result on thinned point processes. *Ann. Probab.* **4**, 89–90.

Wiener, N. (1938). The homogeneous chaos. *Amer. J. Math.* **60**, 897–936.

— (1939). The ergodic theorem. *Duke Math. J.* **5**, 1–18.

Wiener, N., Wintner, A. (1943). The discrete chaos. *Amer. J. Math.* **65**, 279–298.

Williams, D. (1974). Path decomposition and continuity of local time for one-dimensional diffusions, I. *Proc. London Math. Soc.* (3) **28**, 738–768.

Xia, A. (2004). Point processes characterized by their one dimensional distributions. *Probab. Th. Rel. Fields* **128**, 467–474; **138**, 323.

Yaglom, A.M. (1947). Certain limit theorems in the theory of branching processes. *Dokl. Acad. Nauk SSSR* **56**, 795–798.

Yor, M. (1978). Sur la continuité des temps locaux associée à certaines semimartingales. *Astérisque* **52–53**, 23–36.

Yule, G.U. (1924). A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, F.R.S. *Phil. Trans. Roy. Soc. London B* **213**, 21–87.

Zähle, M. (1980). Ergodic properties of general Palm measures. *Math. Nachr.* **95**, 93–106.

Zähle, U. (1988a). Self-similar random measures I. Notion, carrying Hausdorff dimension, and hyperbolic distribution. *Probab. Th. Rel. Fields* **80**, 79–100.

— (1988b). The fractal character of localizable measure-valued processes II. Localizable processes and backward trees. *Math. Nachr.* **137**, 35–48.

Zinn, J. (1986). Comparison of martingale difference sequences. In: *Probability in Banach Spaces. Lect. Notes in Math.* **1153**, 453–457. Springer, Berlin.

Indices

Authors

O. Kallenberg, *Random Measures, Theory and Applications*, Probability Theory and Stochastic Modelling 77, DOI 10.1007/978-3-319-41598-7

Subjects

Symbols