Appendices

Most of the material included here is needed, in one way or another, for the development in the main text. Results have been deferred to an appendix for various reasons: They may not fit naturally into the basic exposition, or their proofs may be too technical or complicated to include in the regular text. In fact, some of the longer proofs are omitted altogether and replaced by references to the standard literature.

We begin, in Appendix A1, with a review of some basic results about extremal or ergodic decompositions. Though referred to explicitly only in the later chapters, the subject is clearly of fundamental importance for the entire book. Unfortunately, some key results in the area are quite deep, and their proofs often require methods outside the scope of the present exposition. Some more specialized results in this area appear in Section 9.2.

Appendix A2 contains some technical results about convergence in distribution, especially for random measures, needed in Chapter 3. In Appendix A3 we review some results about multiple stochastic integrals, required in Chapters 8 and 9, where the underlying processes may be either Gaussian or Poisson. In particular, we give a short proof of Nelson's hypercontraction theorem for multiple Wiener–Itô integrals. Next, to fill some needs in Sections 1.6 and 1.7, we list in Appendix A4 some classical results about completely monotone and positive definite functions, including the Hausdorff–Bernstein characterizations and the celebrated theorems of Bochner and Schoenberg. Finally, Appendix A5 reviews the basic theory of Palm measures and Papangelou kernels, required for our discussion in Section 2.7.

A1. Decomposition and Selection

Given a family $\mathcal{T}$ of measurable transformations on a probability space $(S, \mathcal{S}, \mu)$, we say that a set $I \in \mathcal{S}$ is *(strictly) $\mathcal{T}$-invariant* if $T^{-1}I = I$ for every $T \in \mathcal{T}$ and *a.s. $\mathcal{T}$-invariant* if $\mu(I \,\Delta\, T^{-1}I) = 0$ for every $T \in \mathcal{T}$, where $A \,\Delta\, B$ denotes the symmetric difference of A and B. Furthermore, we say that μ is *ergodic* if $\mu I = 0$ or 1 for every a.s. $\mathcal{T}$-invariant set $I \in \mathcal{S}$, and *weakly ergodic* if the same condition holds for every strictly invariant set I. A random element ξ of S is said to be *$\mathcal{T}$-symmetric* if its distribution μ is *$\mathcal{T}$-invariant*, in the sense that $\mu \circ T^{-1} = \mu$ for every $T \in \mathcal{T}$, and we say that ξ is *$\mathcal{T}$-ergodic* if the corresponding property holds for μ.

Let us begin with a common situation where the symmetry or ergodicity of a random element is preserved by a measurable mapping.

Lemma A1.1 *(preservation laws) Let $\mathcal{T}$ and $\mathcal{T}'$ be families of measurable transformations on some measurable spaces S and S', and fix a measurable mapping $f\colon S \to S'$ such that*

$$\{f \circ T;\, T \in \mathcal{T}\} = \{T' \circ f;\, T' \in \mathcal{T}'\}.$$

Then for any $\mathcal{T}$-symmetric or $\mathcal{T}$-ergodic random element ξ in S, the random element $\eta = f(\xi)$ in S is $\mathcal{T}'$-symmetric or $\mathcal{T}'$-ergodic, respectively.

Proof: Suppose that ξ is $\mathcal{T}$-symmetric. Letting $T' \in \mathcal{T}'$ be arbitrary and choosing $T \in \mathcal{T}$ with $f \circ T = T' \circ f$, we get

$$T' \circ \eta = T' \circ f \circ \xi = f \circ T \circ \xi \stackrel{d}{=} f \circ \xi = \eta,$$

which shows that η is $\mathcal{T}'$-symmetric.

Next we note that the invariant σ-fields $\mathcal{I}$ in S and $\mathcal{J}$ in S' are related by $f^{-1}\mathcal{J} \subset \mathcal{I}$. In fact, letting $J \in \mathcal{J}$ and $T \in \mathcal{T}$ be arbitrary and choosing $T' \in \mathcal{T}'$ with $T' \circ f = f \circ T$, we get a.s. $\mathcal{L}(\xi)$

$$T^{-1} f^{-1} J = f^{-1} T'^{-1} J = f^{-1} J.$$

Hence, if ξ is $\mathcal{T}$-ergodic,

$$P\{\eta \in J\} = P\{\xi \in f^{-1}J\} = 0 \text{ or } 1, \quad J \in \mathcal{J},$$

which means that η is $\mathcal{T}'$-ergodic. □

For any class $\mathcal{T}$ of measurable transformations on a measurable space S, the set of $\mathcal{T}$-invariant probability measures μ on S is clearly convex. A $\mathcal{T}$-invariant distribution μ is said to be *extreme* if it has no non-trivial representation as a convex combination of invariant measures. We examine the relationship between the notions of ergodicity and extremality.

Lemma A1.2 *(ergodicity and extremality) Let $\mathcal{T}$ be a family of measurable transformations on a measurable space S, and consider a $\mathcal{T}$-invariant distribution μ on S. If μ is extreme, it is even ergodic, and the two notions are equivalent when $\mathcal{T}$ is a group. For countable groups $\mathcal{T}$, it is also equivalent that μ be weakly ergodic.*

Proof: Suppose that μ is not ergodic. Then there exists an a.s. $\mathcal{T}$-invariant set $I \in \mathcal{S}$ such that $0 < \mu I < 1$, and we get a decomposition

$$\mu = \mu(I)\, \mu[\,\cdot\,|I] + \mu(I^c)\, \mu[\,\cdot\,|I^c].$$

Here $\mu[\cdot|I] \neq \mu[\cdot|I^c]$, since $\mu[I|I] = 1$ and $\mu[I|I^c] = 0$. Furthermore, the invariance of I and μ implies that, for any $T \in \mathcal{T}$ and $B \in \mathcal{S}$,

$$\begin{aligned}\mu(I)\,\mu[T^{-1}B|I] &= \mu(T^{-1}B \cap I) = \mu \circ T^{-1}(B \cap I) \\ &= \mu(B \cap I) = \mu(I)\,\mu[B|I],\end{aligned}$$

which shows that $\mu[\cdot|I]$ is again invariant. Hence, μ is not extreme either.

Now let $\mathcal{T}$ be a group, and assume that μ is ergodic. Consider any convex combination

$$\mu = c\mu_1 + (1-c)\mu_2,$$

where μ_1 and μ_2 are $\mathcal{T}$-invariant distributions on S and $c \in (0,1)$. Introduce the Radon–Nikodým density $f = d\mu_1/d\mu$. Letting $T \in \mathcal{T}$ and $B \in \mathcal{S}$ be arbitrary and using the invariance of μ_1 and μ, we get

$$\mu_1 B = \mu_1(TB) = \int_{TB} f d\mu = \int_{TB} f d(\mu \circ T^{-1}) = \int_B (f \circ T) d\mu.$$

The uniqueness of f yields $f \circ T = f$ a.s., and so for any $a \geq 0$

$$T^{-1}\{f \geq a\} = \{f \circ T \geq a\} = \{f \geq a\} \text{ a.s. } \mu,$$

which means that the set $\{f \geq a\}$ is a.s. invariant. Since μ is ergodic, we get $\mu\{f \geq a\} = 0$ or 1 for every a, and it follows that f is a.s. a constant. But then $\mu_1 = \mu_2 = \mu$, and the stated decomposition is trivial. This shows that μ is extreme.

If $\mathcal{T}$ is a countable group and I is a.s. invariant, then the set $I' = \bigcap_{T \in \mathcal{T}} T^{-1} I$ is strictly invariant with $\mu(I \Delta I') = 0$. Assuming μ to be weakly ergodic, we get $\mu I = \mu I' = 0$ or 1, which shows that μ is even ergodic. □

The following result identifies two cases where every invariant distribution has an integral representation in terms of extreme points. This decomposition may not be unique in general.

Theorem A1.3 *(extremal decomposition, Choquet, Kallenberg) Let $\mathcal{T}$ be a class of measurable transformations on a measurable space S, and assume one of these conditions:*

(i) *$S = B^\infty$ for some Borel space B, and $\mathcal{T}$ is induced by a class of transformations on $\mathbb{N}$;*

(ii) *S is Polish, and the set of $\mathcal{T}$-invariant distributions on S is weakly closed.*

Then every $\mathcal{T}$-invariant probability measure on S is a mixture of extreme, $\mathcal{T}$-invariant distributions.

Proof: (i) Embedding B as a Borel set in $[0,1]$, we may regard μ as an invariant distribution on the compact space $J = [0,1]^\infty$. The space $\mathcal{M}_1(J)$ is again compact and metrizable (cf. Rogers and Williams (1994), Theorem

81.3), and the sub-set M of $\mathcal{T}$-invariant distributions on J is convex and closed, hence compact. By a standard form of Choquet's theorem (cf. Alfsen (1971)), any measure $\mu \in M$ has then an integral representation

$$\mu A = \int m(A)\,\nu(dm), \quad A \in \mathcal{B}(J), \tag{1}$$

where ν is a probability measure on the set $\mathrm{ex}(M)$ of extreme elements in M. In particular, we note that $\nu\{m;\, mB^\infty = 1\} = 1$. Writing $R(m)$ for the restriction of m to B^∞ and putting $\pi_A m = m(A)$, we get for any Borel set $A \subset B^\infty$

$$\begin{aligned} \mu A &= \mu(A \cap B^\infty) = \int m(A \cap B^\infty)\,\nu(dm) \\ &= \int (\pi_A \circ R)\,d\nu = \int \pi_A\, d(\nu \circ R^{-1}) \\ &= \int m(A)\,(\nu \circ R^{-1})(dm). \end{aligned} \tag{2}$$

It remains to note that if $m \in \mathrm{ex}(M)$ with $mB^\infty = 1$, then m is an extreme, $\mathcal{T}$-invariant distribution on B^∞.

(ii) Here we may embed S as a Borel sub-set of a compact metric space J (cf. Rogers and Williams (1994), Theorem 82.5). The space $\mathcal{M}_1(J)$ is again compact and metrizable, and $\mathcal{M}_1(S)$ can be identified with the sub-set $\{\mu \in \mathcal{M}_1(J);\, \mu S = 1\}$ (op. cit., Theorem 83.7). Now the set M of all $\mathcal{T}$-invariant distributions on S remains convex as a sub-set of $\mathcal{M}_1(J)$, and its closure $\overline{M}$ in $\mathcal{M}_1(J)$ is both convex and compact. Thus, Choquet's theorem yields an integral representation as in (1), in terms of a probability measure ν on $\mathrm{ex}(\overline{M})$.

Since μ is restricted to S, we may proceed as before to obtain a representation of type (2), where $R(m)$ now denotes the restriction of m to S. It remains to show that $\mathrm{ex}(\overline{M}) \cap \mathcal{M}_1(S) \subset \mathrm{ex}(M)$. But this follows easily from the relation $\overline{M} \cap \mathcal{M}_1(S) = M$, which holds since M is closed in $\mathcal{M}_1(S)$. □

Under suitable regularity conditions, the ergodic decomposition is unique and can be obtained by conditioning on the invariant σ-field.

Theorem A1.4 *(decomposition by conditioning, Farrel, Varadarajan) Let $\mathcal{T}$ be a countable group of measurable transformations on a Borel space S, and consider a $\mathcal{T}$-symmetric random element ξ in S with distribution μ and invariant σ-field $\mathcal{I}_\xi$. Then the conditional distributions $P[\xi \in \cdot|\mathcal{I}_\xi]$ are a.s. ergodic and $\mathcal{T}$-invariant, and μ has the unique ergodic decomposition*

$$\mu = \int m\,\nu(dm), \qquad \nu = \mathcal{L}(P[\xi \in \cdot|\mathcal{I}_\xi]).$$

Proof: See Dynkin (1978) or Maitra (1977). □

The following result allows us to extend a representation from the ergodic to the general case through a suitable randomization.

Lemma A1.5 *(randomization) Given some Borel spaces S, T, and U, a measurable mapping $f: S \times T \to U$, and some random elements ξ in S and η in U, define $m_t = \mathcal{L}(f(\xi,t))$, $t \in T$. Then $\mathcal{L}(\eta)$ is a mixture over the class $\mathcal{C} = \{m_t\}$ iff $\eta = f(\tilde{\xi},\tau)$ a.s. for some random elements $\tilde{\xi} \overset{d}{=} \xi$ in S and $\tau \perp\!\!\!\perp \tilde{\xi}$ in T. If even $P[\eta \in \cdot|\mathcal{I}] \in \mathcal{C}$ a.s. for some σ-field $\mathcal{I}$, we can choose τ to be $\mathcal{I}$-measurable.*

Proof: Put $M = \mathcal{M}_1(U)$, and note that M is again Borel. In $M \times T$ we may introduce the product measurable set

$$A = \{(\mu,t) \in M \times T;\ \mu = \mathcal{L}(f(\xi,t))\}.$$

Then, by hypothesis, $\mathcal{L}(\eta) = \int \mu\,\nu(d\mu)$ for some probability measure ν on M satisfying $\nu(\pi A) = \nu(\mathcal{C}) = 1$, where πA denotes the projection of A onto M. Hence, the general section theorem (FMP A1.4) yields a measurable mapping $g: M \to T$ satisfying

$$\nu\{\mu = \mathcal{L}(f(\xi, g(\mu)))\} = \nu\{(\mu, g(\mu)) \in A\} = 1.$$

Letting $\alpha \perp\!\!\!\perp \xi$ be a random element in M with distribution ν, we get $\eta \overset{d}{=} f(\xi, g(\alpha))$ by Fubini's theorem. Finally, the transfer theorem (FMP 6.10) yields a random pair $(\tilde{\xi},\tau) \overset{d}{=} (\xi, g(\alpha))$ in $S \times T$ satisfying $\eta = f(\tilde{\xi},\tau)$ a.s.

If $P[\eta \in \cdot|\mathcal{I}] \in \mathcal{C}$ a.s., we may instead apply the section theorem to the product-measurable set

$$A' = \{(\omega,t) \in \Omega \times T;\ P[\eta \in \cdot|\mathcal{I}](\omega) = \mathcal{L}(f(\xi,t))\},$$

to obtain an $\mathcal{I}$-measurable random element τ in T satisfying

$$P[\eta \in \cdot|\mathcal{I}] = \mathcal{L}(f(\xi,t))|_{t=\tau} \ \text{ a.s.}$$

Letting $\zeta \overset{d}{=} \xi$ in S with $\zeta \perp\!\!\!\perp \tau$ and using Fubini's theorem, we get a.s.

$$P[f(\zeta,\tau) \in \cdot|\tau] = P\{f(\zeta,t) \in \cdot\}|_{t=\tau} = \mu_\tau = P[\eta \in \cdot|\tau],$$

which implies $(f(\zeta,\tau),\tau) \overset{d}{=} (\eta,\tau)$. By FMP 6.11 we may then choose a random pair $(\tilde{\xi},\tilde{\tau}) \overset{d}{=} (\zeta,\tau)$ in $S \times T$ such that

$$(f(\tilde{\xi},\tilde{\tau}),\tilde{\tau}) = (\eta,\tau) \ \text{ a.s.}$$

In particular, we get $\tilde{\tau} = \tau$ a.s., and so $\eta = f(\tilde{\xi},\tau)$ a.s. Finally, we note that $\tilde{\xi} \overset{d}{=} \zeta \overset{d}{=} \xi$, and also that $\tilde{\xi} \perp\!\!\!\perp \tau$ since $\zeta \perp\!\!\!\perp \tau$ by construction. □

We proceed with a result on measurable selections.

Lemma A1.6 *(measurable selection) Let ξ and η be random elements in some measurable spaces S and T, where T is Borel, and let f be a measurable function on $S \times T$ such that $f(\xi,\eta) = 0$ a.s. Then there exists a ξ-measurable random element $\hat{\eta}$ in T satisfying $f(\xi,\hat{\eta}) = 0$ a.s.*

Proof: Define a measurable set $A \subset S \times T$ and its S-projection πA by

$$\begin{aligned} A &= \{(s,t) \in S \times T;\ f(s,t) = 0\}, \\ \pi A &= \textstyle\bigcup_{t \in T} \{s \in S;\ f(s,t) = 0\}. \end{aligned}$$

By the general section theorem (FMP A1.4), there exists a measurable function $g\colon S \to T$ satisfying

$$(\xi, g(\xi)) \in A \ \text{ a.s. on } \ \{\xi \in \pi A\}.$$

Since $(\xi, \eta) \in A$ implies $\xi \in \pi A$, we also have

$$P\{\xi \in \pi A\} \geq P\{(\xi,\eta) \in A\} = P\{f(\xi,\eta) = 0\} = 1.$$

This proves the assertion with $\hat{\eta} = g(\xi)$. □

A2. Weak Convergence

For any metric or metrizable space S, let $\hat{\mathcal{M}}(S)$ denote the space of bounded measures on S, write $\hat{\mathcal{M}}_c(S)$ for the sub-set of measures bounded by c, and let $\mathcal{M}_c(S)$ be the further sub-class of measures μ with $\mu S = c$. For any $\mu \in \hat{\mathcal{M}}(S)$, we define $\hat{\mu} = \mu/(\mu S \vee 1)$. On $\hat{\mathcal{M}}(S)$ we introduce the *weak topology*, induced by the mappings $\mu \mapsto \mu f = \int f d\mu$ for any f belonging to the space $\hat{C}_+(S)$ of bounded, continuous, non-negative functions on S. Then the weak convergence $\mu_n \xrightarrow{w} \mu$ means that $\mu_n f \to \mu f$ for every $f \in \hat{C}_+(S)$. A set $M \subset \hat{\mathcal{M}}(S)$ is said to be *tight* if

$$\inf_{K \in \mathcal{K}} \sup_{\mu \in M} \mu K^c = 0,$$

where $\mathcal{K} = \mathcal{K}(S)$ denotes the class of compact sets in S. We begin with a simple extension of Prohorov's theorem (FMP 16.3).

Lemma A2.1 *(weak compactness, Prohorov) For any Polish space S, a set $M \subset \hat{\mathcal{M}}(S)$ is weakly, relatively compact iff*

(i) $\sup_{\mu \in M} \mu S < \infty$,
(ii) *M is tight.*

Proof: First suppose that M is weakly, relatively compact. Since the mapping $\mu \mapsto \mu S$ is weakly continuous, we conclude that the set $\{\mu S;\ \mu \in M\}$ is relatively compact in $\mathbb{R}_+$, and (i) follows by the Heine–Borel theorem. To prove (ii), we note that the set $\hat{M} = \{\hat{\mu};\ \mu \in M\}$ is again weakly, relatively compact by the weak continuity of the mapping $\mu \mapsto \hat{\mu}$. By (i) we may then assume that $M \subset \hat{\mathcal{M}}_1(S)$. For fixed $s \in S$, define $\mu' = \mu + (1 - \mu S)\delta_s$. The mapping $\mu \mapsto \mu'$ is weakly continuous on $\hat{\mathcal{M}}_1(S)$, and so the set $M' = \{\mu';\ \mu \in M\}$ in $\mathcal{M}_1(S)$ is again relatively compact. But then Prohorov's theorem yields (ii) for the set M', and the same condition follows for M.

Conversely, assume conditions (i) and (ii), and fix any $\mu_1, \mu_2, \ldots \in M$. If $\liminf_n \mu_n S = 0$, then $\mu_n \xrightarrow{w} 0$ along a sub-sequence. If instead $\liminf_n \mu_n S > 0$, then by (i) we have $\mu_n S \to c \in (0,\infty)$ along a sub-sequence $N' \subset \mathbb{N}$. But then by (ii) the sequence $\tilde{\mu}_n = \mu_n/\mu_n S$, $n \in N'$, is tight, and so by Prohorov's theorem we have $\tilde{\mu}_n \xrightarrow{w} \mu$ along a further sub-sequence N''. Hence, $\mu_n = (\mu_n S)\tilde{\mu}_n \xrightarrow{w} c\mu$ along N'', which shows that M is weakly, relatively compact. □

Next we prove a tightness criterion for a.s. bounded random measures.

Lemma A2.2 *(weak tightness, Prohorov, Aldous)* *Let $\xi_1, \xi_2, \ldots$ be a.s. bounded random measures on a Polish space S. Then (ξ_n) is tight for the weak topology on $\hat{\mathcal{M}}(S)$ iff*

(i) *$(\xi_n S)$ is tight in $\mathbb{R}_+$,*

(ii) *$(E\hat{\xi}_n)$ is weakly tight in $\hat{\mathcal{M}}(S)$.*

Proof: First suppose that (ξ_n) is weakly tight. Since the mapping $\mu \mapsto \mu S$ is weakly continuous on $\hat{\mathcal{M}}(S)$, condition (i) follows by FMP 16.4. Using the continuity of the mapping $\mu \mapsto \hat{\mu}$, we see from the same result that $(\hat{\xi}_n)$ is weakly tight. Hence, there exist some weakly compact sets $M_k \subset \hat{\mathcal{M}}_1(S)$ such that

$$P\{\hat{\xi}_n \notin M_k\} \le 2^{-k-1}, \quad k, n \in \mathbb{N}.$$

By Lemma A2.1 we may next choose some compact sets $K_k \subset S$ such that

$$M_k \subset \{\mu;\, \mu K_k^c \le 2^{-k-1}\}, \quad k \in \mathbb{N}.$$

Then for any n and k

$$E\hat{\xi}_n K_k^c \le P\{\hat{\xi}_n \notin M_k\} + E[\hat{\xi}_n K_k^c;\, \xi_n \in M_k] \le 2^{-k},$$

and (ii) follows.

Conversely, assume conditions (i) and (ii). We may then choose some constants $c_k > 0$ and some compact sets $K_k \subset S$ such that

$$P\{\xi_n S > c_k\} \le 2^{-k}, \quad E\hat{\xi}_n K_k^c \le 2^{-2k}, \qquad k, n \in \mathbb{N}.$$

Now introduce in $\hat{\mathcal{M}}(S)$ the sets

$$M_m = \{\mu;\, \mu S \le c_m\} \cap \bigcap_{k>m} \{\mu;\, \hat{\mu} K_k^c \le 2^{-k}\}, \quad m \in \mathbb{N},$$

which are weakly, relatively compact by Lemma A2.1. Using the countable sub-additivity of P and Chebyshev's inequality, we get

$$\begin{aligned} P\{\xi_n \notin M_m\} &= P\Big(\{\xi_n S > c_m\} \cup \textstyle\bigcup_{k>m} \{\hat{\xi}_n K_k^c > 2^{-k}\}\Big) \\ &\le P\{\xi_n S > c_m\} + \textstyle\sum_{k>m} P\{\hat{\xi}_n K_k^c > 2^{-k}\} \\ &\le 2^{-m} + \textstyle\sum_{k>m} 2^k E\hat{\xi}_n K_k^c \\ &\le 2^{-m} + \textstyle\sum_{k>m} 2^{-k} = 2^{-m+1} \to 0, \end{aligned}$$

which shows that (ξ_n) is weakly tight in $\hat{\mathcal{M}}(S)$. □

The last result leads to some useful criteria for convergence in distribution of a.s. bounded random measures on a metrizable space S. When S is Polish, we write $\xi_n \xrightarrow{wd} \xi$ for such convergence with respect to the weak topology on $\hat{\mathcal{M}}(S)$. In the special case where S is locally compact, we may also consider the corresponding convergence with respect to the vague topology, here denoted by $\xi_n \xrightarrow{vd} \xi$. In the latter case, write $\overline{S}$ for the one-point compactification of S. With a slight abuse of notation, we say that $(\hat{E}\xi_n)$ is *tight* if

$$\inf_{K\in\mathcal{K}} \limsup_{n\to\infty} E[\xi_n K^c \wedge 1] = 0.$$

Theorem A2.3 *(convergence in distribution) Let $\xi, \xi_1, \xi_2, \ldots$ be a.s. bounded random measures on a Polish space S. Then these two conditions are equivalent:*

(i) $\xi_n \xrightarrow{wd} \xi$,

(ii) $\xi_n f \xrightarrow{d} \xi f$ *for all* $f \in \hat{C}_+(S)$.

If S is locally compact, then (i) *is also equivalent to each of the following conditions:*

(iii) $\xi_n \xrightarrow{vd} \xi$ *and* $\xi_n S \xrightarrow{d} \xi S$,

(iv) $\xi_n \xrightarrow{vd} \xi$ *and* $(\hat{E}\xi_n)$ *is tight,*

(v) $\xi_n \xrightarrow{vd} \xi$ *on* $\overline{S}$.

Proof: If $\xi_n \xrightarrow{wd} \xi$, then $\xi_n f \xrightarrow{d} \xi f$ for all $f \in \hat{C}_+(S)$ by continuous mapping, and $(\hat{E}\xi_n)$ is tight by Prohorov's theorem and Lemma A2.2. Hence, (i) implies conditions (ii)–(v).

Now assume instead condition (ii). Then the Cramér–Wold theorem (FMP 5.5) yields $(\xi_n f, \xi_n S) \xrightarrow{d} (\xi f, \xi S)$ for all $f \in \hat{C}_+$, and so $\hat{\xi}_n f \xrightarrow{d} \hat{\xi} f$ for the same functions f, where $\hat{\xi} = \xi/(\xi S \vee 1)$ and $\hat{\xi}_n = \xi_n/(\xi_n S \vee 1)$, as before. Since $\hat{\xi} S \leq 1$ and $\hat{\xi}_n S \leq 1$ for all n, it follows that $E\hat{\xi}_n \xrightarrow{w} E\hat{\xi}$, and so by Lemma A2.1 the sequence $(E\hat{\xi}_n)$ is weakly tight. Since also $\xi_n S \xrightarrow{d} \xi S$, Lemma A2.2 shows that the random sequence (ξ_n) is tight with respect to the weak topology. Then Prohorov's theorem (FMP 16.3) shows that (ξ_n) is also relatively compact in distribution. In other words, any sub-sequence $N' \subset \mathbb{N}$ has a further sub-sequence N'', such that $\xi_n \xrightarrow{wd} \eta$ along N'' for some random measure η on S with $\eta S < \infty$ a.s. Thus, for any $f \in \hat{C}_+$, we have both $\xi_n f \xrightarrow{d} \xi f$ and $\xi_n f \xrightarrow{d} \eta f$, and then also $\xi f \stackrel{d}{=} \eta f$. Applying the Cramér–Wold theorem once again, we obtain

$$(\xi f_1, \ldots, \xi f_n) \stackrel{d}{=} (\eta f_1, \ldots, \eta f_n), \quad f_1, \ldots, f_n \in \hat{C}_+.$$

By a monotone-class argument we conclude that $\xi \stackrel{d}{=} \eta$, on the σ-field $\mathcal{C}$ generated by the mappings $\mu \mapsto \mu f$ for arbitrary $f \in \hat{C}_+$. Noting that $\mathcal{C}$ agrees with the Borel σ-field on $\hat{\mathcal{M}}(S)$ since the latter space is separable, we obtain $\xi \stackrel{d}{=} \eta$. This shows that $\xi_n \xrightarrow{wd} \xi$ along N'', and (i) follows since N' was arbitrary.

Now suppose that S is locally compact. In each of the cases (iii)–(v) it suffices to show that (ξ_n) is weakly tight, since the required weak convergence will then follow as before, by means of Prohorov's theorem. First we assume condition (iii). For any compact set $K \subset S$, continuous function f on S with $0 \le f \le 1_K$, and measure $\mu \in \hat{\mathcal{M}}(S)$, we note that

$$1 - e^{-\mu K^c} \le 1 - e^{-\mu(1-f)} \le e^{\mu S}(e^{-\mu f} - e^{-\mu S}).$$

Letting $c > 0$ with $\xi S \neq c$ a.s., we obtain

$$\begin{aligned} E(1 - e^{-\xi_n K^c}) &\le E[1 - e^{-\xi_n K^c}; \xi_n S \le c] + P\{\xi_n S > c\} \\ &\le e^c E(e^{-\xi_n f} - e^{-\xi_n S}) + P\{\xi_n S > c\} \\ &\to e^c E(e^{-\xi f} - e^{-\xi S}) + P\{\xi S > c\}. \end{aligned}$$

Here the right-hand side tends to 0 as $f \uparrow 1$ and then $c \to \infty$, which shows that $(\hat{E}\xi_n)$ is tight. The required tightness of (ξ_n) now follows by Lemma A2.2.

Now assume condition (iv). By the tightness of $(\hat{E}\xi_n)$, we may choose some compact sets $K_k \subset S$ such that

$$E[\xi_n K_k^c \wedge 1] \le 2^{-k}, \quad k, n \in \mathbb{N}.$$

Next, for every $k \in \mathbb{N}$, we may choose some $f_k \in C_K(S)$ with $f_k \ge 1_{K_k}$. Since $\xi_n K_k \le \xi_n f_k \xrightarrow{d} \xi f_k$, the sequences $(\xi_n K_k)$ are tight in $\mathbb{R}_+$, and we may choose some constants $c_k > 0$ such that

$$P\{\xi_n K_k > c_k\} \le 2^{-k}, \quad k, n \in \mathbb{N}.$$

Letting $r_k = c_k + 1$, we get by combination

$$\begin{aligned} P\{\xi_n S > r_k\} &\le P\{\xi_n K_k > c_k\} + P\{\xi_n K_k^c > 1\} \\ &\le 2^{-k} + 2^{-k} = 2^{-k+1}, \end{aligned}$$

which shows that the sequence $(\xi_n S)$ is tight in $\mathbb{R}_+$. By Lemma A2.2 it follows that (ξ_n) is weakly tight in $\hat{\mathcal{M}}(S)$.

Finally, we assume condition (v). If S is compact, then $\overline{S} = S$, and the weak and vague topologies coincide. Otherwise, we have $\infty \in K^c$ in $\overline{S}$ for $K \subset \mathcal{K}(S)$, and every $f \in C_K(S)$ may be extended to a continuous function $\bar{f}$ on $\overline{S}$ satisfying $\bar{f}(\infty) = 0$. By continuous mapping,

$$\xi_n f = \xi_n \bar{f} \xrightarrow{d} \xi \bar{f} = \xi f, \quad f \in C_K(S),$$

and so $\xi_n \xrightarrow{vd} \xi$ on S by FMP 16.16. Since also $1 \in C_K(\overline{S})$, we get

$$\xi_n S = \xi_n \overline{S} \xrightarrow{d} \xi \overline{S} = \xi S.$$

This reduces the proof to the case of (iii). □

The following lemma is often useful to extend a convergence criterion from the ergodic to the general case.

Lemma A2.4 *(randomization) Let $\mu, \mu_1, \mu_2, \dots$ be probability kernels between two metric spaces S and T such that $s_n \to s$ in S implies $\mu_n(s_n, \cdot) \xrightarrow{w} \mu(s, \cdot)$ in $\mathcal{M}_1(T)$. Then for any random elements $\xi, \xi_1, \xi_2, \dots$ in S with $\xi_n \xrightarrow{d} \xi$, we have $E\mu_n(\xi_n, \cdot) \xrightarrow{w} E\mu(\xi, \cdot)$. This remains true when the constants $s, s_1, s_2, \dots$ and random elements $\xi, \xi_1, \xi_2, \dots$ are restricted to some measurable subsets $S_0, S_1, \dots \subset S$.*

Proof: For any bounded, continuous function f on T, the integrals μf and $\mu_n f$ are bounded, measurable functions on S such that $s_n \to s$ implies $\mu_n f(s_n) \to \mu f(s)$. Hence, the continuous-mapping theorem in FMP 4.27 yields $\mu_n f(\xi_n) \xrightarrow{d} \mu f(\xi)$, and so $E\mu_n f(\xi_n) \to E\mu f(\xi)$. The assertion follows since f was arbitrary. The extended version follows by the same argument from the corresponding extension of FMP 4.27. □

We also need the following elementary tightness criterion.

Lemma A2.5 *(hyper-contraction and tightness) Consider some random variables $\xi_1, \xi_2, \dots \geq 0$, σ-fields $\mathcal{F}_1, \mathcal{F}_2, \dots$, and constant $c \in (0, \infty)$ such that a.s.*

$$E[\xi_n^2 | \mathcal{F}_n] \leq c(E[\xi_n | \mathcal{F}_n])^2 < \infty, \quad n \in \mathbb{N}.$$

Then, assuming (ξ_n) to be tight, so is the sequence $\eta_n = E[\xi_n | \mathcal{F}_n]$, $n \in \mathbb{N}$.

Proof: Fix any r, ε, and $p_1, p_2, \dots \in (0,1)$ with $p_n \to 0$. Using the Paley–Zygmund inequality in FMP 4.1, we have a.s. on $\{\eta_n > 0\}$

$$\begin{aligned} 0 \; &< \; \frac{(1-r)^2}{c} \leq (1-r)^2 \frac{(E[\xi_n | \mathcal{F}_n])^2}{E[\xi_n^2 | \mathcal{F}_n]} \\ &\leq \; P[\xi_n > r\eta_n | \, \mathcal{F}_n] \\ &\leq \; P[p_n \xi_n > r\varepsilon | \, \mathcal{F}_n] + 1\{p_n \eta_n < \varepsilon\}, \end{aligned} \tag{1}$$

which is also trivially true when $\eta_n = 0$. The tightness of (ξ_n) yields $p_n \xi_n \xrightarrow{P} 0$ by the criterion in FMP 4.9, and so the first term on the right of (1) tends to 0 in L^1, hence also in probability. Since the sum is bounded from below, we obtain $1\{p_n \eta_n < \varepsilon\} \xrightarrow{P} 1$, which shows that $p_n \eta_n \xrightarrow{P} 0$. Using FMP 4.9 in the opposite direction, we conclude that even (η_n) is tight. □

A3. Multiple Stochastic Integrals

Multiple Gaussian and Poisson integrals are needed to represent processes with higher-dimensional symmetries. The former are defined, most naturally, on tensor products $\bigotimes_i H_i = H_1 \otimes \cdots \otimes H_d$ of Hilbert spaces, which are understood to be infinite-dimensional and separable, unless otherwise specified. Given an ortho-normal basis (ONB) $h_{i1}, h_{i2}, \ldots$ in each H_i, we recall that the tensor products $\bigotimes_i h_{i,j_i} = h_{1,j_1} \otimes \cdots \otimes h_{d,j_d}$ form an ONB in $\bigotimes_i H_i$. Since the H_i are isomorphic, it is often convenient to take $H_i = H$ for all i, in which case we may write $\bigotimes_i H_i = H^{\otimes d}$.

For any $f \in H_1 \otimes \cdots \otimes H_d$, we define the *supporting spaces* $M_1, \ldots, M_d$ of f to be the smallest closed, linear sub-spaces $M_i \subset H_i$, $i \le d$, such that $f \in M_1 \otimes \cdots \otimes M_d$. More precisely, the M_i are the smallest closed, linear sub-spaces of H_i such that f has an orthogonal expansion in terms of tensor products $h_1 \otimes \cdots \otimes h_d$ with $h_i \in M_i$ for all i. (Such an expansion is said to be *minimal* if it is based on the supporting spaces M_i.) For a basis-free description of the M_i, write $H_i' = \bigotimes_{j \ne i} H_j$, and define some bounded, linear operators $A_i \colon H_i' \to H_i$ by

$$\langle g, A_i h\rangle = \langle f, g \otimes h\rangle, \quad g \in H_i,\ h \in H_i',\ i \le d.$$

Then M_i equals $\overline{R(A_i)}$, the closed range of A_i.

The orthogonal representation of an element $f \in H_1 \otimes \cdots \otimes H_d$ clearly depends on the choice of ortho-normal bases in the d spaces. In the two-dimensional case, however, there is a simple diagonal version, which is essentially unique.

Lemma A3.1 *(diagonalization) For any $f \in H_1 \otimes H_2$, there exist a finite or infinite sequence $\lambda_1, \lambda_2, \ldots > 0$ with $\sum_j \lambda_j^2 < \infty$ and some ortho-normal sequences $\varphi_1, \varphi_2, \ldots$ in H_1 and $\psi_1, \psi_2, \ldots$ in H_2, such that*

$$f = \sum\nolimits_j \lambda_j\, (\varphi_j \otimes \psi_j). \tag{1}$$

This representation is unique, apart from joint, orthogonal transformations of the elements φ_j and ψ_j, within sets of indices j with a common value $\lambda_j > 0$. Furthermore, the expansion in (1) *is minimal.*

Proof: Define a bounded linear operator $A : H_2 \to H_1$ and its adjoint $A^* \colon H_1 \to H_2$ by

$$\langle g, Ah\rangle = \langle A^* g, h\rangle = \langle f, g \otimes h\rangle, \quad g \in H_1,\ h \in H_2. \tag{2}$$

Then A^*A is a positive, self-adjoint, and compact operator on H_2, and so it has a finite or infinite sequence of eigen-values $\lambda_1^2 \ge \lambda_2^2 \ge \cdots > 0$ with associated ortho-normal eigen-vectors $\psi_1, \psi_2, \ldots \in H_2$. For definiteness, we may choose $\lambda_j > 0$ for all j. It is easy to check that the elements $\varphi_j = \lambda_j^{-1} A\psi_j \in H_1$ are ortho-normal eigen-vectors of the operator AA^* on H_1 with the same eigen-values, and that

$$A\psi_j = \lambda_j \varphi_j, \qquad A^*\varphi_j = \lambda_j \psi_j. \tag{3}$$

This remains true for any extension of the sequences (φ_i) and (ψ_i) to ONBs of H_1 and H_2. Since the tensor products $\varphi_i \otimes \psi_j$ form an ONB in $H_1 \otimes H_2$, we have a representation

$$f = \sum\nolimits_{i,j} c_{ij} \, (\varphi_i \otimes \psi_j),$$

for some constants c_{ij} with $\sum_{i,j} c_{ij}^2 < \infty$. Using (2) and (3), we obtain

$$c_{ij} = \langle f, \varphi_i \otimes \psi_j \rangle = \langle A\psi_j, \varphi_i \rangle = \lambda_j \langle \varphi_j, \varphi_i \rangle = \lambda_j \delta_{ij},$$

and (1) follows.

Conversely, suppose that (1) holds for some constants $\lambda_j > 0$ and ortho-normal elements $\varphi_j \in H_1$ and $\psi_j \in H_2$. Combining this with (2) yields (3), and so the λ_j^2 are eigen-values of A^*A with associated ortho-normal eigen-vectors ψ_j. This implies the asserted uniqueness of the λ_j and ψ_j. Since any other sets of eigen-vectors ψ_j' and φ_j' must satisfy (3), the two sets are related by a common set of orthogonal transformations. Finally, the minimality of the representation (1) is clear from (3). □

A process η on a Hilbert space H is said to be *iso-normal* if $\eta h \in L^2$ for every $h \in H$ and the mapping η preserves inner products, so that $E(\eta h \, \eta k) = \langle h, k \rangle$. By a *G-process* on H we mean an iso-normal, centered Gaussian process on H. More generally, we define a *continuous, linear, random functional (CLRF)* on H as a process η on H such that

$$\begin{aligned} &\eta(ah + bk) = a\eta h + b\eta k \ \text{ a.s.}, \quad h, k \in H, \ a, b \in \mathbb{R}, \\ &\eta h_n \xrightarrow{P} 0, \qquad \|h_n\| \to 0. \end{aligned}$$

We need the following basic existence and uniqueness result, valid for multiple stochastic integrals based on independent or identical G-processes.

Theorem A3.2 *(multiple Gaussian integrals, Wiener, Itô) Let $\eta_1, \ldots, \eta_d$ be independent G-processes on some Hilbert spaces $H_1, \ldots, H_d$, and fix any $k_1, \ldots, k_d \in \mathbb{N}$. Then there exists an a.s. unique CLRF $\eta = \bigotimes_i \eta^{\otimes k_i}$ on $H = \bigotimes_i H_i^{\otimes k_i}$ such that*

$$\Big(\bigotimes\nolimits_i \eta_i^{\otimes k_i}\Big) \Big(\bigotimes\nolimits_{i,j} f_{ij}\Big) = \prod\nolimits_{i,j} \eta_i f_{ij},$$

whenever the elements $f_{ij} \in H_i$ are orthogonal in j for fixed i. The functional η is L^2-bounded with mean 0.

Proof: The result for $d = 1$ is classical (cf. FMP 13.21). In general, we may introduce the G-process $\zeta = \bigoplus_i \eta_i$ on the Hilbert space $H = \bigoplus_i H_i$ and put $\chi = \zeta^{\otimes k}$, where $k = \sum_i k_i$. The restriction η of χ to $\bigotimes_i H_i^{\otimes k_i}$ has clearly the desired properties. □

We also need some norm estimates for multiple Wiener–Itô integrals.

Lemma A3.3 *(hyper-contraction, Nelson) Let $\eta_1, \ldots, \eta_d$ be independent G-processes on some Hilbert spaces $H_1, \ldots, H_d$, and fix any $p > 0$ and $k_1, \ldots, k_d \in \mathbb{N}$. Then we have, uniformly in f,*

$$\Big\| \Big(\bigotimes_i \eta_i^{\otimes k_i} \Big) f \Big\|_p \lesssim \|f\|, \quad f \in \bigotimes_i H_i^{\otimes k_i}.$$

Proof: Considering the G-process $\eta = \eta_1 + \cdots + \eta_d$ in $H = \bigoplus_i H_i$, we may write the multiple integral $(\bigotimes_i \eta_i^{\otimes k_i}) f$ in the form $\eta^{\otimes k} g$ for a suitable element $g \in H^{\otimes k}$ with $\|g\| = \|f\|$, where $k = \sum_i k_i$. It is then enough to prove that, for any $p > 0$ and $n \in \mathbb{N}$,

$$\|\eta^{\otimes n} f\|_p \lesssim \|f\|, \quad f \in H^{\otimes n}. \tag{4}$$

Here we may take $H = L^2(\mathbb{R}_+, \lambda)$ and let η be generated by a standard Brownian motion B on $\mathbb{R}_+$, in which case f may be regarded as an element of $L^2(\lambda^n)$. Considering separately each of the $n!$ tetrahedral regions in $\mathbb{R}_+^n$, we may further reduce to the case where f is supported by the set

$$\Delta_n = \{(t_1, \ldots, t_n) \in \mathbb{R}_+^n;\ t_1 < \cdots < t_n\}.$$

Then $\eta^{\otimes n} f$ can be written as an iterated Itô integral (FMP 18.13)

$$\begin{aligned} \eta^{\otimes n} f &= \int dB_{t_n} \int dB_{t_{n-1}} \cdots \int f(t_1, \ldots, t_n)\, dB_{t_1} \\ &= \int (\eta^{\otimes(n-1)} \hat f_t)\, dB_t, \end{aligned} \tag{5}$$

where $\hat f_{t_n}(t_1, \ldots, t_{n-1}) = f(t_1, \ldots, t_n)$.

For $n = 1$ we note that ηf is $N(0, \|f\|_2^2)$, and therefore

$$\|\eta f\|_p = \|\eta(f/\|f\|_2)\|_p \, \|f\|_2 \lesssim \|f\|_2, \quad f \in L^2(\lambda),$$

as required. Now suppose that (4) holds for all multiple integrals up to order $n - 1$. Using the representation (5), a BDG-inequality from FMP 17.7, the extended Minkowski inequality in FMP 1.30, the induction hypothesis, and Fubini's theorem, we get for any $p \geq 2$

$$\begin{aligned} \|\eta^{\otimes n} f\|_p &= \Big\| \int (\eta^{\otimes(n-1)} \hat f_t)\, dB_t \Big\|_p \\ &\lesssim \Big\| \int (\eta^{\otimes(n-1)} \hat f_t)^2\, dt \Big\|_{p/2}^{1/2} \\ &\leq \Big(\int \|\eta^{\otimes(n-1)} \hat f_t\|_p^2\, dt \Big)^{1/2} \\ &\lesssim \Big(\int \|\hat f_t\|_2^2\, dt \Big)^{1/2} = \|f\|_2. \end{aligned}$$

Taking $p = 2$ and using Jensen's inequality, we get for any $p \leq 2$

$$\|\eta^{\otimes n} f\|_p \leq \|\eta^{\otimes n} f\|_2 \lesssim \|f\|_2, \quad f \in L^2(\lambda^n).$$

This completes the induction and proves (4). □

Next we construct the tensor product $X \otimes h$ of a CLRF X on H with a fixed element $h \in H$.

Lemma A3.4 *(mixed multiple integrals) For any CLRF X on H and element $h \in H$, there exists an a.s. unique CLRF $X \otimes h$ on $H^{\otimes 2}$ such that*

$$(X \otimes h)(f \otimes g) = (Xf)\,\langle h, g\rangle \ \ a.s., \quad f, g \in H. \tag{6}$$

Proof: For any $h \in H$, define a linear operator A_h from H to $H^{\otimes 2}$ by $A_h f = f \otimes h$, $f \in H$, and note that A_h is bounded since

$$\|A_h f\| = \|f \otimes h\| = \|f\|\,\|h\|, \quad f \in H.$$

The adjoint A_h^* is a bounded, linear operator from $H^{\otimes 2}$ to H, and we may define a CLRF $X \otimes h$ on $H^{\otimes 2}$ by

$$(X \otimes h)f = XA_h^* f, \quad f \in H^{\otimes 2}. \tag{7}$$

For any $f, g, k \in H$, we have

$$\langle k, A_h^*(f \otimes g)\rangle = \langle A_h k,\, f \otimes g\rangle = \langle k \otimes h,\, f \otimes g\rangle = \langle k, f\rangle\,\langle h, g\rangle,$$

which implies

$$A_h^*(f \otimes g) = \langle h, g\rangle\, f, \quad f, g \in H.$$

Combining with (7) and using the linearity of X, we obtain (6). To prove the asserted uniqueness, fix any ONB $h_1, h_2, \ldots$ of H and apply (6) to the the tensor products $h_i \otimes h_j$, $i, j \in \mathbb{N}$, which form an ONB in $H^{\otimes 2}$. The required uniqueness now follows by the linearity and continuity of X. □

We turn our attention to double Poisson integrals of the form $\xi\eta f$ or $\xi^2 f$, where ξ and η are independent Poisson processes on some measurable spaces S and T. We assume the underlying intensity measures $E\xi$ and $E\eta$ to be σ-finite, and to simplify the writing we may take $S = T = \mathbb{R}_+$ and $E\xi = E\eta = \lambda$. The existence poses no problem, since the mentioned integrals can be defined as path-wise, Lebesgue-type integrals with respect to the product measures $\xi \otimes \eta$ and $\xi^2 = \xi \otimes \xi$, respectively. It is less obvious when these integrals converge.

The following result gives necessary and sufficient conditions for the a.s. convergence of the stochastic integrals ξf, $\xi^2 f$, or $\xi\eta f$, where ξ and η are independent, unit rate Poisson processes on $\mathbb{R}_+$ and $f \geq 0$ is a measurable function on $\mathbb{R}_+$ or $\mathbb{R}_+^2$, respectively. Given a measurable function $f \geq 0$ on $\mathbb{R}_+^2$, we define $f_1 = \lambda_2 \hat{f}$ and $f_2 = \lambda_1 \hat{f}$, where $\hat{f} = f \wedge 1$, and $\lambda_i f$ denotes the Lebesgue integral of f in the i-th coordinate. Also put $f^* = \sup_s |f(s)|$, and write λ_D for normalized Lebesgue measure along the diagonal $D = \{(x, y) \in \mathbb{R}_+^2;\, x = y\}$.

Theorem A3.5 *(convergence of Poisson integrals, Kallenberg and Szulga) Let ξ and η be independent, unit rate Poisson processes on $\mathbb{R}_+$. Then for any measurable function $f \geq 0$ on $\mathbb{R}_+$, we have $\xi f < \infty$ a.s. iff $\lambda \hat{f} < \infty$. If $f \geq 0$ is instead measurable on $\mathbb{R}_+^2$ and $f_i = \lambda_j \hat{f}$ for $i \neq j$, we have $\xi\eta f < \infty$ a.s. iff these three conditions are fulfilled:*

(i) $\lambda\{f_i = \infty\} = 0$ *for* $i = 1, 2$,

(ii) $\lambda\{f_i > 1\} < \infty$ *for* $i = 1, 2$,

(iii) $\lambda^2[\hat{f};\, f_1 \vee f_2 \leq 1] < \infty$.

Finally, $\xi^2 f < \infty$ a.s. iff (i)–(iii) *hold and $\lambda_D \hat{f} < \infty$.*

Our proof is based on a sequence of lemmas. We begin with some elementary moment formulas, where we write $\psi(t) = 1 - e^{-t}$ for $t \geq 0$.

Lemma A3.6 *(moment identities) Let ξ and η be independent, unit rate Poisson processes on $\mathbb{R}_+$. Then for any measurable set $B \subset \mathbb{R}_+$ or function $f \geq 0$ on $\mathbb{R}_+$ or $\mathbb{R}_+^2$, we have*

(i) $E\psi(\xi f) = \psi(\lambda(\psi \circ f))$,

(ii) $P\{\xi B > 0\} = \psi(\lambda B)$,

(iii) $E(\xi f)^2 = \lambda f^2 + (\lambda f)^2,\ E\xi\eta f = \lambda^2 f$,

(iv) $E(\xi\eta f)^2 = \lambda^2 f^2 + \lambda(\lambda_1 f)^2 + \lambda(\lambda_2 f)^2 + (\lambda^2 f)^2$.

Proof: Statements (i) and (ii) appear in FMP 12.2. For claim (iii), note that $E(\xi B)^2 = \lambda B + (\lambda B)^2$, and extend by linearity, independence, and monotone convergence. To prove (iv), conclude from (iii) and Fubini's theorem that

$$\begin{aligned} E(\xi\eta f)^2 &= E(\xi(\eta f))^2 = E\lambda_1(\eta f)^2 + E(\lambda_1 \eta f)^2 \\ &= \lambda_1 E(\eta f)^2 + E(\eta(\lambda_1 f))^2 \\ &= \lambda^2 f^2 + \lambda_1(\lambda_2 f)^2 + \lambda_2(\lambda_1 f)^2 + (\lambda^2 f)^2. \qquad \square \end{aligned}$$

We proceed to estimate the tails in the distribution of $\xi\eta f$. For simplicity we let $f \leq 1$, so that $f_i = \lambda_j f$ for $i \neq j$.

Lemma A3.7 *(tail estimate) Let ξ be a unit rate Poisson process on $\mathbb{R}_+$, and consider a measurable function $f\colon \mathbb{R}_+^2 \to [0,1]$ with $\lambda^2 f < \infty$. Then*

$$P\Big\{\xi\eta f > \tfrac{1}{2}\lambda^2 f\Big\} \gtrsim \psi\left(\frac{\lambda^2 f}{1 + f_1^* \vee f_2^*}\right).$$

Proof: We may clearly assume that $\lambda^2 f > 0$. By Lemma A3.6 we have $E\xi\eta f = \lambda^2 f$ and

$$\begin{aligned} E(\xi\eta f)^2 &\leq (\lambda^2 f)^2 + \lambda^2 f + f_1^* \lambda f_1 + f_2^* \lambda f_2 \\ &= (\lambda^2 f)^2 + (1 + f_1^* + f_2^*)\lambda^2 f, \end{aligned}$$

and so the Paley–Zygmund inequality in FMP 4.1 yields

$$\begin{aligned} P\left\{\xi\eta f > \tfrac{1}{2}\lambda^2 f\right\} &\geq \frac{(1-\frac{1}{2})^2(\lambda^2 f)^2}{(\lambda^2 f)^2 + (1 + f_1^* \vee f_2^*)\lambda^2 f} \\ &\gtrsim \left(1 + \frac{1 + f_1^* \vee f_2^*}{\lambda^2 f}\right)^{-1} \\ &\gtrsim \psi\left(\frac{\lambda^2 f}{1 + f_1^* \vee f_2^*}\right). \end{aligned}$$

□

Lemma A3.8 *(decoupling) Let ξ and η be independent, unit rate Poisson processes on $\mathbb{R}_+$. Then for any measurable functions $f \geq 0$ on $\mathbb{R}_+^2$ with $f = 0$ on D, we have*

$$E\psi(\xi^2 f) \asymp E\psi(\xi\eta f).$$

Proof: We may clearly assume that f is supported by the wedge $W = \{(s,t);\ 0 \leq s < t\}$. It is equivalent to show that

$$E[(V \cdot \xi)_\infty \wedge 1] \asymp E[(V \cdot \eta)_\infty \wedge 1],$$

where $V_t = \xi f(\cdot, t) \wedge 1$. Here V is clearly predictable with respect to the right-continuous and complete filtration induced by ξ and η (FMP 25.23). The random time

$$\tau = \inf\{t \geq 0;\ (V \cdot \eta)_t > 1\}$$

is then optional (FMP 7.6), and so the process $1\{\tau \geq t\}$ is again predictable (FMP 25.1). Noting that ξ and η are both compensated by λ (FMP 25.25), we get (FMP 25.22)

$$\begin{aligned} E[(V \cdot \xi)_\infty;\ \tau = \infty] &\leq E(V \cdot \xi)_\tau = E(V \cdot \lambda)_\tau = E(V \cdot \eta)_\tau \\ &\leq E[(V \cdot \eta)_\infty \wedge 1] + 2P\{\tau < \infty\}, \end{aligned}$$

and therefore

$$\begin{aligned} E[(V \cdot \xi)_\infty \wedge 1] &\leq E[(V \cdot \xi)_\infty;\ \tau = \infty] + P\{\tau < \infty\} \\ &\leq E[(V \cdot \eta)_\infty \wedge 1] + 3P\{\tau < \infty\} \\ &\leq 4E[(V \cdot \eta)_\infty \wedge 1]. \end{aligned}$$

The same argument applies with the roles of ξ and η interchanged. □

Proof of Theorem A3.5: To prove the first assertion, let $f \geq 0$ on $\mathbb{R}_+$. Then Lemma A3.6 (i) shows that $\xi f = \infty$ a.s. iff $\lambda\hat{f} \geq \lambda(\psi \circ f) = \infty$, and so by Kolmogorov's zero-one law we have $\xi f < \infty$ a.s. iff $\lambda\hat{f} < \infty$.

Turning to the second assertion, let $f \geq 0$ on $\mathbb{R}_+^2$. Since $\xi \otimes \eta$ is a.s. simple, we have $\xi\eta f < \infty$ iff $\xi\eta\hat{f} < \infty$ a.s., which allows us to take $f \leq 1$. First assume conditions (i)–(iii). Here (i) yields

$$E\xi\eta[f;\ f_1 \vee f_2 = \infty] \leq \sum\nolimits_i \lambda^2[f;\ f_i = \infty] = 0,$$

and so we may assume that $f_1, f_2 < \infty$. Then the first assertion yields $\eta f(s,\cdot) < \infty$ and $\xi f(\cdot, s) < \infty$ a.s. for every $s \geq 0$. Furthermore, (ii) implies $\xi\{f_1 > 1\} < \infty$ and $\eta\{f_2 > 1\} < \infty$ a.s. By Fubini's theorem we get a.s.

$$\xi\eta[f;\, f_1 \vee f_2 > 1] \leq \xi[\eta f;\, f_1 > 1] + \eta[\xi f;\, f_2 > 1] < \infty,$$

which allows us to assume that even $f_1, f_2 \leq 1$. Then (iii) yields $E\xi\eta f = \lambda^2 f < \infty$, which implies $\xi\eta f < \infty$ a.s.

Conversely, suppose that $\xi\eta f < \infty$ a.s. for some function f into $[0,1]$. By Lemma A3.6 (i) and Fubini's theorem we have

$$E\psi(\lambda\psi(t\eta f)) = E\psi(t\xi\eta f) \to 0, \quad t \downarrow 0,$$

which implies $\lambda\psi(t\eta f) \to 0$ a.s., and hence $\eta f < \infty$ a.e. $\lambda \otimes P$. By the first assertion and Fubini's theorem we get $f_1 = \lambda_2 f < \infty$ a.e. λ, and the symmetric argument yields $f_2 = \lambda_1 f < \infty$ a.e. This proves (i).

Next, Lemma A3.6 (i) yields, on the set $\{f_1 > 1\}$,

$$\begin{aligned} E\psi(\eta f) &= \psi(\lambda_2(\psi \circ f)) \geq \psi((1 - e^{-1})f_1) \\ &\geq \psi(1 - e^{-1}) \equiv c > 0. \end{aligned}$$

Hence, for any measurable set $B \subset \{f_1 > 1\}$,

$$E\lambda[1 - \psi(\eta f); B] \leq (1 - c)\lambda B,$$

and so, by Chebyshev's inequality,

$$\begin{aligned} P\{\lambda\psi(\eta f) < \tfrac{1}{2}c\lambda B\} &\leq P\{\lambda[1 - \psi(\eta f); B] > (1 - \tfrac{1}{2}c)\lambda B\} \\ &\leq \frac{E\lambda[1 - \psi(\eta f); B]}{(1 - \frac{1}{2}c)\lambda B} \leq \frac{1 - c}{1 - \frac{1}{2}c}. \end{aligned}$$

Since B was arbitrary, we conclude that

$$P\{\lambda\psi(\eta f) \geq \tfrac{1}{2}c\lambda\{f_1 > 1\}\} \geq 1 - \frac{1 - c}{1 - \frac{1}{2}c} = \frac{c}{2 - c} > 0.$$

Noting that $\lambda\psi(\eta f) < \infty$ a.s. by the one-dimensional result and Fubini's theorem, we obtain $\lambda\{f_1 > 1\} < \infty$. This, together with the corresponding result for f_2, proves (ii).

Finally, we may apply Lemma A3.7 to the function $f1\{f_1 \vee f_2 \leq 1\}$ to obtain

$$P\Big\{\xi\eta f > \tfrac{1}{2}\lambda^2[f;\, f_1 \vee f_2 \leq 1]\Big\} \gtrsim \psi\Big(\tfrac{1}{2}\lambda^2[f;\, f_1 \vee f_2 \leq 1]\Big).$$

This implies (iii), since the opposite statement would yield the contradiction $P\{\xi\eta f = \infty\} > 0$.

Now turn to the last assertion. Since $1_D \cdot \xi^2$ has coordinate projections ξ, the one-dimensional result shows that $\xi^2 f 1_D < \infty$ a.s. iff $\lambda_D \hat{f} < \infty$. Thus, we may henceforth assume that $f = 0$ on D. Then Lemma A3.8 yields

$$E\psi(t\xi^2 f) \asymp E\psi(t\xi\eta f), \quad t > 0,$$

and so, as $t \to 0$,

$$P\{\xi^2 f = \infty\} \asymp P\{\xi\eta f = \infty\},$$

which shows that $\xi^2 f < \infty$ a.s. iff $\xi\eta f < \infty$ a.s. □

A4. Complete Monotonicity

For any infinite sequence $c_0, c_1, \ldots \in \mathbb{R}$, put $\Delta c_k = c_{k+1} - c_k$, and define recursively the higher order differences by

$$\Delta^0 c_k = c_k, \quad \Delta^{n+1} c_k = \Delta(\Delta^n c_k), \qquad k, n \geq 0,$$

where all differences are with respect to k. We say that (c_k) is *completely monotone* if

$$(-1)^n \Delta^n c_k \geq 0, \quad k, n \geq 0. \tag{1}$$

The definition for finite sequences is the same, apart from the obvious restrictions on the parameters k and n.

Next we say that a function $f: \mathbb{R}_+ \to \mathbb{R}$ is *completely monotone*, if the sequence $f(nh)$, $n \in \mathbb{Z}_+$, is completely monotone in the discrete-time sense for every $h > 0$. Thus, we require

$$(-1)^n \Delta_h^n f(t) \geq 0, \quad t, h, n \geq 0, \tag{2}$$

where $\Delta_h f(t) = f(t+h) - f(t)$, and the higher order differences are defined recursively by

$$\Delta_h^0 f(t) = f(t), \quad \Delta_h^{n+1} f(t) = \Delta_h(\Delta_h^n f(t)), \qquad t, h, n \geq 0,$$

where all differences are now with respect to t. For functions on $[0,1]$ the definitions are the same, apart from the obvious restrictions on t, h, and n. When $f \in C^k(\mathbb{R}_+)$ or $C^k[0,1]$, it is easy to verify by induction that

$$\Delta_h^n f(t) = \int_0^h \cdots \int_0^h f^{(n)}(t + s_1 + \cdots + s_n)\, ds_1 \cdots ds_n,$$

for appropriate values of t, h, and n, where $f^{(n)}$ denotes the nth derivative of f. In this case, (2) is clearly equivalent to the condition

$$(-1)^n f^{(n)}(t) \geq 0, \quad t \in I^\circ,\ n \geq 0. \tag{3}$$

Sometimes it is more convenient to consider the related notion of *absolute monotonicity*, defined as in (1), (2), or (3), respectively, except that the factor $(-1)^n$ is now omitted. Note that a sequence (c_k) or function $f(t)$ is absolutely monotone on a discrete or continuous interval I iff (c_{-k}) or $f(-t)$ is completely monotone on the reflected interval $-I = \{t;\, -t \in I\}$.

We may now state the basic characterizations of completely monotone sequences and functions. For completeness, we include the elementary case of finite sequences, using the notation

$$n^{(k)} = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1), \quad 0 \le k \le n.$$

Theorem A4.1 *(complete monotonicity, Hausdorff, Bernstein)*

(i) *A finite sequence* $c_0, \ldots, c_n \in \mathbb{R}$ *is completely monotone with* $c_0 = 1$ *iff there exists a random variable* κ *in* $\{0, \ldots, n\}$ *such that*

$$c_k = E\kappa^{(k)}/n^{(k)}, \quad k = 0, \ldots, n.$$

(ii) *An infinite sequence* $c_0, c_1, \ldots \in \mathbb{R}$ *is completely monotone with* $c_0 = 1$ *iff there exists a random variable* α *in* $[0,1]$ *such that*

$$c_k = E\alpha^k, \quad k \in \mathbb{Z}_+.$$

(iii) *A function* $f\colon [0,1] \to \mathbb{R}$ *is completely monotone with* $f(0) = f(0+) = 1$ *iff there exists a random variable* κ *in* $\mathbb{Z}_+$ *such that*

$$f(t) = E(1-t)^\kappa, \quad t \in [0,1].$$

(iv) *A function* $f\colon \mathbb{R}_+ \to \mathbb{R}$ *is completely monotone with* $f(0) = f(0+) = 1$ *iff there exists a random variable* $\rho \ge 0$ *such that*

$$f(t) = Ee^{-\rho t}, \quad t \ge 0.$$

In each case, the associated distribution is unique.

Proof: See Feller (1971), pp. 223, 225, 439. □

Next we say that a function $f\colon \mathbb{R}^d \to \mathbb{C}$ is *non-negative definite* if

$$\sum\nolimits_{h,k} c_h \bar{c}_k f(x_h - x_k) \ge 0, \quad c_1, \ldots, c_d \in \mathbb{C}, \;\; x_1, \ldots, x_d \in \mathbb{R}^d.$$

The following result characterizes non-negative definite functions in terms of characteristic functions.

Theorem A4.2 *(non-negative definite functions, Bochner) A function* $f\colon \mathbb{R}^d \to \mathbb{C}$ *is continuous and non-negative definite with* $f(0) = 1$ *iff there exists a random vector* ξ *in* $\mathbb{R}^d$ *such that*

$$f(t) = Ee^{it\xi}, \quad t \in \mathbb{R}^d.$$

Proof: See Feller (1971), p. 622, for the case $d = 1$. The proof for $d > 1$ is similar. □

The next result gives a remarkable connection between non-negative definite and completely monotone functions. Given a function f on $\mathbb{R}_+$, we define the functions f_n on $\mathbb{R}^n$ by

$$f_n(x_1, \ldots, x_n) = f(x_1^2 + \cdots + x_n^2), \quad x_1, \ldots, x_n \in \mathbb{R}. \tag{4}$$

Theorem A4.3 *(Fourier and Laplace transforms, Schoenberg) A continuous function $f : \mathbb{R}_+ \to \mathbb{R}$ with $f(0) = 1$ is completely monotone iff the function f_n in* (4) *is non-negative definite on $\mathbb{R}^n$ for every $n \in \mathbb{N}$.*

Proof: See Schoenberg (1938a), Theorem 2, Donoghue (1969), pp. 201–206, or Berg et al. (1984), pp. 144–148. □

A5. Palm and Papangelou Kernels

Given a random measure ξ on a measurable space $(S, \mathcal{S})$, we define the associated *Campbell measure* C on $S \times \mathcal{M}(S)$ by

$$Cf = E \int f(s, \xi)\, \xi(ds), \quad f \in (S \times \mathcal{M}(S))_+,$$

where $(S \times \mathcal{M}(S))_+$ denotes the class of measurable functions $f \geq 0$ on $S \times \mathcal{M}(S)$. Since ξ is assumed to be a.s. σ-finite, the same thing is true for C, and we may choose f to be strictly positive with $Cf < \infty$, in which case the projection $\nu = (f \cdot C)(\cdot \times \mathcal{M})$ is bounded and satisfies $\nu B = 0$ iff $\xi B = 0$ a.s. for every $B \in \mathcal{S}$. Any σ-finite measure ν with the latter property is called a *supporting measure* for ξ, and we note that ν is unique up to an equivalence, in the sense of mutual absolute continuity. In particular, we may choose $\nu = E\xi$ when the latter measure is σ-finite.

If S is Borel, then so is $\mathcal{M}(S)$, and there exists a kernel $Q = (Q_s)$ from S to $\mathcal{M}(S)$ satisfying the disintegration formula

$$Cf = \int \nu(ds) \int f(s, \mu)\, Q_s(d\mu), \quad f \in (S \times \mathcal{M}(S))_+, \tag{1}$$

or simply $C = \nu \otimes Q$. This can be proved in the same way as the existence of regular conditional distributions (FMP 6.3–4). The *Palm measures* Q_s of ξ are a.e. unique up to a normalization, and they can be chosen to be probability measures—the *Palm distributions* of ξ—iff $\nu = E\xi$ is σ-finite.

When ξ is a point process (i.e. integer-valued), we can also introduce the *reduced Campbell measure* C' on $S \times \mathcal{N}(S)$, given by

$$C'f = E \int f(s, \xi - \delta_s)\, \xi(ds), \quad f \in (S \times \mathcal{N}(S))_+,$$

and consider the disintegration of C' into *reduced Palm measures*,

$$C'f = \int \nu(ds) \int f(s,\mu)\, Q'_s(d\mu), \quad f \in (S \times \mathcal{N}(S))_+,$$

or $C' = \nu \otimes Q'$. Comparing with (1), we see that

$$\begin{aligned} \int_B \nu(ds) \int f(\mu)\, Q'_s(d\mu) &= E \int_B f(\xi - \delta_s)\, \xi(ds) \\ &= \int_B \nu(ds) \int f(\mu - \delta_s)\, Q_s(d\mu), \end{aligned}$$

which implies $Q'_s = Q_s \circ (\mu - \delta_s)^{-1}$ a.e. ν, in the sense that

$$\int f(\mu)\, Q'_s(d\mu) = \int f(\mu - \delta_s)\, Q_s(d\mu), \quad s \in S \text{ a.e. } \nu.$$

If $E\xi$ is σ-finite, then again we may choose $\nu = E\xi$, in which case Q'_s may be thought of as the conditional distribution of $\xi - \delta_s$, given that $\xi\{s\} > 0$.

We may also consider a disintegration of C' in the other variable. This is especially useful when $C'(S \times \cdot) \ll \mathcal{L}(\xi)$, in the sense that $P\{\xi \in A\} = 0$ implies $C'(S \times A) = 0$, since we can then choose the supporting measure on $\mathcal{N}(S)$ to be equal to $\mathcal{L}(\xi)$. Unfortunately, the stated absolute continuity may fail in general, which makes the present theory more complicated. In the following we shall often write $1_B \cdot \xi = 1_B \xi$, for convenience.

Theorem A5.1 *(disintegration kernel, Papangelou, Kallenberg) Let C' be the reduced Campbell measure of a point process ξ on a Borel space S. Then there exists a maximal kernel R from $\mathcal{N}(S)$ to S such that*

$$E \int f(s,\xi)\, R(\xi, ds) \le C'f, \quad f \in (S \times \mathcal{N}(S))_+, \tag{2}$$

where the random measure $\eta = R(\xi, \cdot)$ is given on $(\operatorname{supp} \xi)^c$ by

$$1_B \eta = \frac{E[1_B \xi;\, \xi B = 1 | 1_{B^c} \xi]}{P[\xi B = 0 | 1_{B^c} \xi]} \quad \text{a.s. on } \{\xi B = 0\}, \quad B \in \mathcal{S}. \tag{3}$$

If ξ is simple, we have $\eta = 0$ a.s. on $\operatorname{supp} \xi$, and equality holds in (2) iff

$$P[\xi B = 0 | 1_{B^c} \xi] > 0 \quad \text{a.s. on } \{\xi B = 1\}, \quad B \in \mathcal{S}, \tag{4}$$

which then remains true on $\{\xi B < \infty\}$.

By the *maximality* of R we mean that, if R' is any other kernel satisfying (2), then $R'(\xi, \cdot) \le R(\xi, \cdot)$ a.s. The maximal solution R is called the *Papangelou kernel* of ξ, and we may refer to the associated random measure $\eta = R(\xi, \cdot)$ on S as the *Papangelou measure*. The requirement (4) for absolute continuity is often referred to as condition (Σ).

Proof: Since C' is σ-finite, we may fix a measurable function $g > 0$ on $S \times \mathcal{N}(S)$ with $C'g < \infty$ and introduce the projection $\nu = (g \cdot C')(S \times \cdot)$ on $\mathcal{N}(S)$. Then ν is a supporting measure of C' in the second variable, in the sense that $C'(S \times M) = 0$ iff $\nu M = 0$ for any measurable set $M \subset \mathcal{N}(S)$. Now consider the Lebesgue decomposition $\nu = \nu_a + \nu_s$ with respect to $\mathcal{L}(\xi)$ (FMP 2.10), and fix a measurable subset $A \subset \mathcal{N}(S)$ with $\nu_s A = 0$ and $P\{\xi \in A^c\} = 0$, so that $\nu_a = 1_A \cdot \nu$. Then $C'(S \times (A \cap \cdot)) \ll \mathcal{L}(\xi)$, in the sense that the left-hand side vanishes for any set M with $P\{\xi \in M\} = 0$, and we may introduce the associated disintegration

$$E \int f(s,\xi)\, R(\xi, ds) = (1_{S\times A} \cdot C')f \le C'f, \quad f \in (S \times \mathcal{N}(S))_+.$$

If R' is any other kernel satisfying (2), then for any f as above,

$$\begin{aligned} E \int f(s,\xi)\, R'(\xi, ds) &= E\Big[\int f(s,\xi)\, R'(\xi, ds);\ \xi \in A\Big] \\ &\le C'(1_{S\times A} f) = (1_{S\times A} \cdot C')f \\ &= E \int f(s,\xi)\, R(\xi, ds). \end{aligned}$$

Choosing $g > 0$ as before with $C'g < \infty$, we get a similar relationship between the kernels

$$\tilde{R}(\xi, ds) = g(s,\xi) R(\xi, ds), \qquad \tilde{R}'(\xi, ds) = g(s,\xi) R'(\xi, ds).$$

In particular, for any $B \in \mathcal{S}$, we may take

$$f(s,\xi) = 1_B(s)\, 1\{\tilde{R}(\xi, B) < \tilde{R}'(\xi, B)\}$$

to obtain

$$E[\tilde{R}(\xi, B) - \tilde{R}'(\xi, B);\ \tilde{R}(\xi, B) < \tilde{R}'(\xi, B)] \ge 0,$$

which implies $\tilde{R}(\xi, B) \ge \tilde{R}'(\xi, B)$ a.s. Starting from countably many such relations and extending by a monotone-class argument, we conclude that $\tilde{R}(\xi, \cdot) \ge \tilde{R}'(\xi, \cdot)$ and hence $R(\xi, \cdot) \ge R'(\xi, \cdot)$, outside a fixed null set, which establishes the required maximality.

Next we show that

$$1\{\xi B = 0\} \ll P[\xi B = 0 | 1_{B^c}\xi] \text{ a.s.}, \quad B \in \mathcal{S}, \tag{5}$$

where a relation $a \ll b$ between two quantities $a, b \ge 0$ means that $b = 0$ implies $a = 0$. Formula (5) follows from the fact that

$$\begin{aligned} &P\{P[\xi B = 0 | 1_{B^c}\xi] = 0,\ \xi B = 0\} \\ &\qquad = E[P[\xi B = 0 | 1_{B^c}\xi];\ P[\xi B = 0 | 1_{B^c}\xi] = 0] = 0. \end{aligned}$$

Fixing any $B \in \mathcal{S}$ and letting $M \subset \mathcal{N}(S)$ be measurable with

$$M \subset M_0 \equiv \{\mu;\ \mu B = 0,\ P[\xi B = 0 | 1_{B^c}\xi \in d\mu] > 0\},$$

we obtain

$$\begin{aligned} C'(B\times M) &= P\{1_{B^c}\xi\in M,\, \xi B=1,\, P[\xi B=0|1_{B^c}\xi]>0\} \\ &\ll E[P[\xi B=0|1_{B^c}\xi];\, 1_{B^c}\xi\in M] \\ &= P\{\xi B=0,\, 1_{B^c}\xi\in M\}\le P\{\xi\in M\}, \end{aligned}$$

which shows that $C'(B\times\cdot)\ll\mathcal{L}(\xi)$ on M_0. Combining (5) with the maximality of R, we get for any $A\in\mathcal{S}\cap B$

$$\begin{aligned} &E[E[\xi A;\, \xi B=1|1_{B^c}\xi];\, 1_{B^c}\xi\in M,\, P[\xi B=0|1_{B^c}\xi]>0] \\ &\quad= E[\xi A;\, 1_{B^c}\xi\in M,\, \xi B=1,\, P[\xi B=0|1_{B^c}\xi]>0] \\ &\quad= C'(A\times(M\cap M_0))=E[\eta A;\, \xi\in(M\cap M_0)] \\ &\quad= E[\eta A;\, 1_{B^c}\xi\in M,\, \xi B=0,\, P[\xi B=0|1_{B^c}\xi]>0] \\ &\quad= E[R(1_{B^c}\xi,A);\, 1_{B^c}\xi\in M,\, \xi B=0] \\ &\quad= E[R(1_{B^c}\xi,A)\, P[\xi B=0|1_{B^c}\xi];\, 1_{B^c}\xi\in M]. \end{aligned}$$

Since M was arbitrary, it follows that

$$\begin{aligned} &E[\xi A;\, \xi B=1|1_{B^c}\xi]\, 1\{P[\xi B=0|1_{B^c}\xi]>0\} \\ &\qquad= R(1_{B^c}\xi,A)\, P[\xi B=0|1_{B^c}\xi], \end{aligned}$$

and so by (5) we have

$$\eta A=\frac{E[\xi A;\, \xi B=1|1_{B^c}\xi]}{P[\xi B=0|1_{B^c}\xi]}\ \text{ a.s. on } \{\xi B=0\},$$

which extends to (3) since both sides are random measures on B.

From this point on, we assume that ξ is simple. Applying (2) to the product-measurable function

$$f(s,\mu)=\mu\{s\},\quad s\in S,\ \ \mu\in\mathcal{N}(S),$$

we obtain

$$\begin{aligned} E\int\eta\{s\}\,\xi(ds) &= E\int\xi\{s\}\,\eta(ds)\le C'f \\ &= E\int(\xi-\delta_s)\{s\}\,\xi(ds) \\ &= E\int(\xi\{s\}-1)\,\xi(ds)=0. \end{aligned}$$

Hence, $\int\eta\{s\}\xi(ds)=0$ a.s., which implies $\eta=0$ a.s. on $\operatorname{supp}\xi$.

Now suppose that (4) is fulfilled. Fix any $B\in\mathcal{S}$ with $\xi B<\infty$ a.s., and consider any measurable subset $M\subset\mathcal{N}(S)$ with $P\{\xi\in M\}=0$. Assuming first that $M\subset\{\mu B=0\}$, we get by (4)

$$\begin{aligned} C'(B\times M) &= E\int_B 1\{(\xi-\delta_s)B=0,\, 1_{B^c}\xi\in M\}\,\xi(ds) \\ &= P\{\xi B=1,\, 1_{B^c}\xi\in M\} \\ &\ll E[P[\xi B=0|1_{B^c}\xi];\, 1_{B^c}\xi\in M] \\ &= P\{\xi B=0,\, 1_{B^c}\xi\in M\} \\ &\le P\{\xi\in M\}=0, \end{aligned}$$

which shows that $C'(B\times M)=0$.

Next let $M \subset \{\mu B < m\}$ for some $m \in \mathbb{N}$. Fix a nested sequence of countable partitions of B into measurable subsets B_{nj}, such that any two points $s \neq t$ of B eventually lie in different sets B_{nj}. By the result for $m = 1$ and dominated convergence, we obtain

$$\begin{aligned} C'(B \times M) &= \sum_j C'(B_{nj} \times M) \\ &= \sum_j C'(B_{nj} \times (M \cap \{\mu B_{nj} > 0\})) \\ &= \sum_j E \int_{B_{nj}} 1\{\xi - \delta_s \in M,\, \xi B_{nj} > 1\}\, \xi(ds) \\ &\leq m \sum_j P\{\xi B \leq m,\, \xi B_{nj} > 1\} \\ &= mE\Big[\sum_j 1\{\xi B_{nj} > 1\};\, \xi B \leq m\Big] \to 0, \end{aligned}$$

which shows that again $C'(B \times M) = 0$. The result extends by monotone convergence, first to the case $M \subset \{\mu B < \infty\}$, and then to $B = S$ for general M. Thus, (4) implies $C'(S \times \cdot) \ll \mathcal{L}(\xi)$.

Conversely, suppose that $C'(S \times \cdot) \ll \mathcal{L}(\xi)$. Then for any $n \in \mathbb{N}$, $B \in \mathcal{S}$, and measurable $M \subset \mathcal{N}(S)$, we have

$$\begin{aligned} nP\{\xi B = n,\, 1_{B^c}\xi \in M\} &= E[\xi B;\, \xi B = n,\, 1_{B^c}\xi \in M] \\ &= C'(B \times \{\mu B = n-1,\, 1_{B^c}\mu \in M\}) \\ &\ll P\{\xi B = n-1,\, 1_{B^c}\xi \in M\}. \end{aligned}$$

Iterating this relation and then summing over n, we obtain

$$P\{\xi B < \infty,\, 1_{B^c}\xi \in M\} \ll P\{\xi B = 0,\, 1_{B^c}\xi \in M\},$$

which together with (5) yields

$$\begin{aligned} &P\{\xi B < \infty,\, P[\xi B = 0 | 1_{B^c}\xi] = 0\} \\ &\qquad \ll P\{\xi B = 0,\, P[\xi B = 0 | 1_{B^c}\xi] = 0\} = 0. \end{aligned}$$

This shows that $P[\xi B = 0 | 1_{B^c}\xi] > 0$, a.s. on $\{\xi B < \infty\}$. □

Historical and Bibliographical Notes

Only publications closely related to topics in the main text are mentioned. No completeness is claimed, and I apologize in advance for inevitable errors and omissions. References to my own papers are indicated by K(·).

1. The Basic Symmetries

The notion of exchangeability was introduced by Haag (1928), who derived some formulas for finite sequences of exchangeable events, some of which are implicit already in the work of de Moivre (1718–56). Further information about the early development of the subject appears in Dale (1985).

The characterization of infinite, exchangeable sequences as mixtures of i.i.d. sequences was established by de Finetti, first (1929, 1930) for random events, and then (1937) for general random variables. The result was extended by Hewitt and Savage (1955) to random elements in a compact Hausdorff space. Dubins and Freedman (1979) and Freedman (1980) showed by examples that de Finetti's theorem fails, even in its weaker mixing form, without some regularity conditions on the underlying space.

De Finetti's study of exchangeable sequences was continued by many people, including Khinchin (1932, 1952), de Finetti (1933a,b), Dynkin (1953), Loève (1960–63), Aldous (1982a,b), and Ressel (1985). Olshen (1971, 1973) noted the equivalence of the various σ-fields occurring in the conditional form of the result. The paper by Hewitt and Savage (1955) also contains the celebrated zero-one law named after these authors, extensions of which are given by Aldous and Pitman (1979).

The connection between finite, exchangeable sequences and sampling from a finite population has been noted by many authors. A vast literature deals with comparative results for sampling with or without replacement, translating into comparisons of finite and infinite, exchangeable sequences. A notable result in this area is the inequality of Hoeffding (1963), which has been generalized by many authors, including Rosén (1967) and Pathak (1974). Our continuous-time version in Proposition 3.19 may be new. Error estimates in the approximation of finite, exchangeable sequences by infinite ones are given by Diaconis and Freedman (1980b).

For infinite sequences of random variables, Ryll-Nardzewski (1957) noted that the properties of exchangeability and contractability are equivalent. The result fails for finite sequences, as noted by Kingman (1978a).

In that case, the relationship between the two notions was investigated in K(2000). Part (i) of Theorem 1.13 was conjectured by IVANOFF and WEBER (personal communication) and proved by the author.

Processes with exchangeable increments were first studied by BÜHLMANN (1960), who proved a version of Theorem 1.19. Alternative approaches appear in ACCARDI and LU (1993) and in FREEDMAN (1996). Exchangeable random measures on $[0,1]$ were first characterized in K(1973b, 1975a); the corresponding characterizations on the product spaces $S \times \mathbb{R}_+$ and $S \times [0,1]$ appeared in K(1990a). The relationship between contractable sequences and processes, stated in Theorem 1.23, is quoted from K(2000). General symmetry properties, such as those in Theorems 1.17 and 1.18, were first noted for random measures in K(1975–86). The fact that exchangeability is preserved under composition of independent processes was noted in K(1982).

FELLER (1966–71) noted the connection between exchangeable sequences and HAUSDORFF's (1921) characterization of absolutely monotone sequences. Further discussion on the subject appears in KIMBERLING (1973). The corresponding relationship between exchangeable processes and BERNSTEIN's (1928) characterization of completely monotone functions was explored in K(1972, 1975–86), and independently by DABONI (1975, 1982). Related remarks appear in FREEDMAN (1996). The fact, from K(1973a), that any exchangeable, simple point process on $[0,1]$ is a mixed binomial process was also noted by both DAVIDSON (1974) and MATTHES, KERSTAN, and MECKE (1974–82). The characterization of mixed Poisson processes by the order-statistics property in Corollary 1.28 (iii) goes back to NAWROTZKI (1962) (see also FEIGIN (1979)). The result in Exercise 12 was explored by RÉNYI (1953), in the context of order statistics. The description of the linear birth (or YULE) process in Exercise 13 is due to KENDALL (1966), and alternative proofs appear in WAUGH (1970), NEUTS and RESNICK (1971), and ATHREYA and NEY (1972).

MAXWELL (1875, 1878) derived the normal distribution for the velocities of the molecules in a gas, assuming spherical symmetry and independence in orthogonal directions. The Gaussian approximation of spherically symmetric distributions on a high-dimensional sphere seems to be due to MAXWELL (op. cit.) and BOREL (1914), though the result is often attributed to POINCARÉ (1912). (See the historical remarks in EVERITT (1974), p. 134, and DIACONIS and FREEDMAN (1987).) Related discussions and error estimates appear in MCKEAN (1973), STAM (1982), GALLARDO (1983), YOR (1985), and DIACONIS and FREEDMAN (1987).

Random sequences and processes with more general symmetries than exchangeability have been studied by many authors, beginning with DE FINETTI himself (1938, 1959). In particular, FREEDMAN (1962–63) obtained the discrete- and continuous-time versions of Theorem 1.31. Alternative proofs and further discussion appear in papers by KELKER (1970), KINGMAN (1972b), EATON (1981), LETAC (1981), and SMITH (1981). The result was later recognized as equivalent to the celebrated theorem of SCHOENBERG

(1938a,b) in classical analysis.

Related symmetries, leading to mixtures of stable distributions, have been studied by many authors, including Bretagnolle, Dacunha-Castelle, and Krivine (1966), Dacunha-Castelle (1975), Berman (1980), Ressel (1985, 1988), and Diaconis and Freedman (1987, 1990). The underlying characterization of stable distributions goes back to Lévy (1924, 1925). The general operator version in Theorem 1.37 is quoted from K(1993). Our L^p-invariance is related to a notion of "pseudo-isotropy," studied by Misiewicz (1990). For $p \neq 2$, every linear isometry on $L^p[0,1]$ is essentially of the type employed in the proof of Lemma 1.39, according to a characterization of Lamperti, quoted by Royden (1988). In the Hilbert-space case of $p = 2$, a much larger class of isometries is clearly available.

The only full-length survey of exchangeability theory, previously available, is the set of lecture notes by Aldous (1985), which also contain an extensive bibliography. The reader may also enjoy the short but beautifully crafted survey article by Kingman (1978a). Brief introductions to exchangeability theory appear in Chow and Teicher (1997) and in K(2002).

De Finetti himself, one of the founders of Bayesian statistics, turned gradually (1972, 1974–75) away from mathematics to become a philosopher of science, developing theories of subjective probability of great originality, where his celebrated representation theorem plays a central role among the theoretical underpinnings. An enthusiastic gathering of converts payed tribute to de Finetti at a 1981 Rome conference, held on the occasion of his 75th birthday. The ensuing proceedings (eds. Koch and Spizzichino (1982)) exhibit a curious mix of mathematics and philosophy, ranging from abstract probability theory to the subtle art of assigning subjective probabilities to the possible outcomes of a soccer game.

2. Conditioning and Martingales

The first use of martingale methods in exchangeability theory may be credited to Loève (1960–63), who used the reverse martingale property of the empirical distributions to give a short proof of de Finetti's theorem. A related martingale argument had previously been employed by Doob (1953) to give a simple proof of the strong law of large numbers for integrable, i.i.d. random variables. Though Loève himself (1978) eventually abandoned his martingale approach to exchangeability, the method has subsequently been adopted by many text-book authors. The present characterizations of exchangeability in terms of reverse martingales, stated in Theorems 2.4, 2.12, and 2.20, appear to be new.

Martingale characterizations of special processes go back to Lévy (1937–54), with his celebrated characterization of Brownian motion. A similar characterization of the Poisson process was dicovered by Watanabe (1964). Local characteristics of semi-martingales were introduced, independently, by Jacod (1975) and Grigelionis (1975, 1977), both of whom used them to characterize processes with independent increments. The criteria for mixed Lévy processes in Theorem 2.14 (i) were obtained by GRIGELIONIS (1975), and a wide range of further extensions and applications may be found in JACOD and SHIRYAEV (1987).

The relationship between exchangeability and martingale theory was explored more systematically in K(1982), which contains the basic discrete- and continuous-time characterizations of exchangeability in terms of strong stationarity and reflection invariance, as well as a primitive version of Theorem 2.15. For finite exchangeable sequences, the relation $\xi_\tau \stackrel{d}{=} \xi_1$ was also noted, independently, by BLOM (1985). Further semi-martingale characterizations in discrete and continuous time, including some early versions of Theorems 2.13 and 2.15, were established in K(1988a). The paper K(2000) provides improved and extended versions of the same results, gives the basic norm relations for contractable processes in Theorem 2.23, and contains the regularization properties in Theorems 2.13 and 2.25. Martingale characterizations of exchangeable and contractable arrays were recently obtained by IVANOFF and WEBER (1996, 2003, 2004). A property strictly weaker than strong stationarity has been studied by BERTI, PRATELLI, and RIGO (2004).

The general representation of exchangeable processes in Theorem 2.18, originating with K(1972), was first published in K(1973a). The obvious point-wise convergence of the series of compensated jumps was strengthened in K(1974a) to a.s. uniform convergence, after a similar result for Lévy processes had been obtained by FERGUSON and KLASS (1972). Hyper-contraction methods were first applied to exchangeability theory in K(2002).

The discrete super-martingale in Proposition 2.28 was discovered and explored in K(1975c), where it was used, along with some special continuous-time versions, to study stochastic integration with respect to Lévy processes. The special case of $f(x) = x^2$ had been previously considered by DUBINS and FREEDMAN (1965), and a related maximum inequality appears in DUBINS and SAVAGE (1965). The proof of the general result relies on an elementary estimate, due to ESSEEN and VON BAHR (1965). The growth rates in Theorem 2.32, originally derived for Lévy processes by FRISTEDT (1967) and MILLAR (1971, 1972), were extended to more general exchangeable processes in K(1974b). Finally, a version of Proposition 2.33 appears in K(1989b).

Pure and mixed Poisson processes were characterized by SLIVNYAK (1962) and PAPANGELOU (1974b), through the invariance of their reduced Palm distributions. The general characterization of mixed Poisson and binomial processes appeared in K(1972, 1973c), and the version for general random measures was obtained in K(1975a). Motivated by problems in stochastic geometry, PAPANGELOU (1976) also derived related characterizations of suitable Cox processes, in terms of invariance properties of the associated Papangelou kernels. Various extensions and asymptotic results were derived in K(1978a,b; 1983–86).

3. Convergence and Approximation

Central-limit type theorems for sampling from a finite population and for finite or infinite sequences of exchangeable random variables have been established by many authors, including BLUM, CHERNOFF, ROSENBLATT, and TEICHER (1958), BÜHLMANN (1958, 1960), CHERNOFF and TEICHER (1958), ERDÖS and RÉNYI (1959), TEICHER (1960), BIKELIS (1969), MORAN (1973), and KLASS and TEICHER (1987). Asymptotically invariant sampling from stationary and related processes was studied in K(1999a).

Criteria for Poisson convergence have been given by KENDALL (1967), RIDLER-ROWE (1967), and BENCZUR (1968), and convergence to more general limits was considered by HÁJEK (1960) and ROSÉN (1964). More recent developments along those lines include some martingale-type limit theorems of EAGLESON and WEBER (1978), WEBER (1980), EAGLESON (1982), and BROWN (1982), and some extensions to Banach-space valued random variables, due to DAFFER (1984) and TAYLOR, DAFFER, and PATTERSON (1985). Limit theorems for exchangeable arrays were obtained by IVANOFF and WEBER (1992, 1995).

Functional limit theorems for general random walks and Lévy processes go back to the seminal work of SKOROHOD (1957). The first genuine finite-interval results are due to ROSÉN (1964), who derived criteria for convergence to a Brownian bridge, in the context of sampling from a finite population. His results were extended by BILLINGSLEY (1968) to summation processes based on more general exchangeable sequences. HAGBERG (1973), still working in the special context of sampling theory, derived necessary and sufficient conditions for convergence to more general processes. The general convergence criteria for exchangeable sequences and processes, here presented in Sections 3.1 and 3.3, were first developed in K(1973a), along with the general representation theorem for exchangeable processes on $[0,1]$. (The latter result appeared first in K(1972), written independently of Hagberg's work.)

The restriction and extension results for exchangeable processes, here exhibited in Section 3.5, also originated with K(1973a). In K(1982), the basic convergence criterion for exchangeable processes on $[0,1]$ was strengthened to a uniform version, in the spirit of SKOROHOD (1957); the even stronger coupling result in Theorem 3.25 (iv) is new. The remaining coupling methods of Section 3.6 were first explored in K(1974b), along with applications to a wide range of path properties for exchangeable processes. The underlying results for Lévy processes, quoted in the text, were obtained by KHINCHIN (1939), FRISTEDT (1967), and MILLAR (1971, 1972).

Though a general convergence theory for exchangeable random measures on $[0,1]$ or $\mathbb{R}_+$ was developed already in K(1975a), the more general results in Section 3.2, involving random measures on the product spaces $S \times [0,1]$ and $S \times \mathbb{R}_+$, appear to be new. The one-dimensional convergence criteria of Section 3.5 were first obtained in K(1988c), along with additional results of a similar nature. Finally, an extensive theory for contractable processes was developed in K(2000), including the distributional coupling theorem and related convergence and tightness criteria.

The sub-sequence principles have a rich history, going back to some early attempts to extend the classical limit theorems of probability theory to the context of orthogonal functions and lacunary series. (See Gaposhkin (1966) for a survey and extensive bibliography.) For general sequences of random variables, sub-sequence principles associated with the law of large numbers were obtained by Révész (1965) and Komlós (1967). Chatterji (1972) stated the sub-sequence property for arbitrary limit theorems as an heuristic principle, and the special cases of the central limit theorem and the law of the iterated logarithm were settled by Gaposhkin (1972), Berkes (1974), and Chatterji (1974a,b).

In independent developments, Rényi and Révész (1963) proved that exchangeable sequences converge in the stable sense, a mode of convergence previously introduced by Rényi (1963). Motivated by problems in functional analysis, Dacunha-Castelle (1975) proved that every tight sequence of random variables contains an asymptotically exchangeable sub-sequence. A related Banach-space result was obtained, independently, by Figiel and Sucheston (1976). The stronger version in Theorem 3.32 is essentially due to Aldous (1977), and the present proof follows the approach in Aldous (1985).

Kingman (unpublished) noted the subtle connection between the two sub-sequence problems, and Aldous (1977), in a deep analysis, proved a general sub-sequence principle for broad classes of weak and strong limit theorems. Berkes and Péter (1986) proved Theorem 3.35, along with some more refined approximation results, using sequential coupling techniques akin to our Lemma 3.36, previously developed by Berkes and Philipp (1979). A further discussion of strong approximation by exchangeable sequences appears in Berkes and Rosenthal (1985).

4. Predictable Sampling and Mapping

Doob (1936) proved the optional skipping property for i.i.d. sequences of random variables, thereby explaining the futility of the gambler's attempts to beat the odds. His paper is historically significant for being possibly the first one to employ general optional times. Modernized versions of the same result are given by Doob (1953), Theorem III.5.2, and Billingsley (1986), Theorem 7.1, and some pertinent historical remarks appear in Halmos (1985), pp. 74–76.

The mentioned result was extended in K(1982) to any finite or infinite, exchangeable sequence, and in K(1988a) the order restriction on the sampling times $\tau_1, \ldots, \tau_m$ was eliminated. Furthermore, the optional skipping property was extended in K(2000) to arbitrary contractable sequences. The cited papers K(1982, 1988a, 2000) also contain continuous-time versions of the same results. The optional skipping and predictable sampling theorems were extended by Ivanoff and Weber (2003, 2004) to finite or infinite, contractable or exchangeable arrays.

The fluctuation identity of Sparre-Andersen (1953–54) was "a sensation greeted with incredulity, and the original proof was of an extraordinary intricacy and complexity," according to Feller (1971). The argument was later simplified by Feller (op. cit.) and others. The present proof, based on the predictable sampling theorem, is quoted from K(2002), where the result is used to give simple proofs of the arcsine laws for Brownian motion and symmetric Lévy processes.

The time-change reduction of a continuous local martingale to a Brownian motion was discovered, independently, by Dambis (1965) and Dubins and Schwarz (1965), and the corresponding multi-variate result was proved by Knight (1970, 1971). The similar reduction of a quasi-left-continuous, simple point process to Poisson was proved, independently, by Meyer (1971) and Papangelou (1972). Cocozza and Yor (1980) derived some more general reduction theorems of the same type. The general Gauss–Poisson reduction in Theorem 4.5 is taken from K(1990b).

The invariance criteria for the Brownian motion and bridge are quoted from K(1989b). Integrability criteria for strictly stable Lévy processes were first established in K(1975b), and a more careful discussion, covering even the weakly stable case, appears in K(1992a). The time-change representations in Theorem 4.24 were obtained for symmetric processes by Rosiński and Woyczyński (1986), and independently, in the general case, in K(1992a). (Those more general results were first announced in an invited plenary talk of 1984.) The general invariance theorem for stable processes appears to be new.

Apart from the publications already mentioned, there is an extensive literature dealing with Poisson and related reduction and approximation results for point processes, going back to the seminal papers of Watanabe (1964) and Grigelionis (1971). Let us only mention the subsequent papers by Karoui and Lepeltier (1977), Aalen and Hoem (1978), Kurtz (1980), Brown (1982, 1983), Merzbach and Nualart (1986), Pitman and Yor (1986), and Brown and Nair (1988a,b), as well as the monograph of Barbour, Holst, and Janson (1992). Our present development in Section 4.5 is based on results in K(1990b).

5. Decoupling Identities

For a random walk $S_n = \xi_1 + \cdots + \xi_n$ on $\mathbb{R}$ and for suitable optional times $\tau < \infty$, the elementary relations $ES_\tau = E\tau\, E\xi$ and $ES_\tau^2 = E\tau\, E\xi^2$ were first noted and explored by Wald (1944, 1945), in connection with his development of sequential analysis. The formulas soon became standard tools in renewal and fluctuation theory. Elementary accounts of the two equations and their numerous applications appear in many many texts, including Feller (1971) and Chow and Teicher (1997).

Many authors have improved on Wald's original statements by relaxing the underlying moment conditions. Thus, Blackwell (1946) established the first order relation under the minimal conditions $E\tau < \infty$ and $E|\xi| < \infty$. When $E\xi = 0$ and $E\xi^2 < \infty$, the first order Wald equation $ES_\tau = 0$ remains true under the weaker condition $E\tau^{1/2} < \infty$, as noted by Burkholder and Gundy (1970), and independently by Gordon (as reported in Chung (1974), p. 343). The latter result was extended by Chow, Robbins, and Siegmund (1971), who showed that if $E\xi = 0$ and $E|\xi|^p < \infty$ for some $p \in [1, 2]$, then the condition $E\tau^{1/p} < \infty$ suffices for the validity of $ES_\tau = 0$. The ultimate result in this direction was obtained by Klass (1988), who showed that if $E\xi = 0$, then $ES_\tau = 0$ holds already under the minimal condition $Ea_\tau < \infty$, where $a_n = E|S_n|$.

We should also mention an extension of the first order Wald identity to dependent random variables, established by Franken and Lisek (1982) in the context of Palm distributions. The result also appears in Franken, König, Arndt, and Schmidt (1981).

The continuous-time Wald identities $EB_\tau = 0$ and $EB_\tau^2 = E\tau$, where B is a standard, one-dimensional Brownian motion, have been used extensively in the literature, and detailed discussions appear in, e.g., Loève (1978) and Karatzas and Shreve (1991). Here the first order relation holds when $E\tau^{1/2} < \infty$ and the second order formula is valid for $E\tau < \infty$. Those equations, along with some more general, first and second order moment identities for the Itô integral $V \cdot B$, follow from the basic martingale properties of stochastic L^2-integrals, first noted and explored by Doob (1953).

The general decoupling identities for exchangeable sums and integrals on bounded or unbounded index sets were originally obtained in K(1989b), where the connection with predictable sampling was also noted. The tetrahedral moment identities for contractable sums and integrals were first derived, under more restrictive boundedness conditions, in K(2000).

Predictable integration with respect to Lévy processes was essentially covered already by the elementary discussion of the stochastic L^2-integral in Doob (1953). The theory is subsumed by the more general, but also more sophisticated, semi-martingale theory, as exhibited in Chapter 26 of K(2002). A detailed study of Lévy integrals was undertaken in Millar (1972) and K(1975b). The quoted norm conditions, presented here without any claims to optimality, were originally derived in K(1989b), along with similar estimates in the general, exchangeable case.

6. Homogeneity and Reflections

The notions of local homogeneity and reflection invariance were first introduced and studied in K(1982), where versions can be found for random sets of the basic representation theorems. The same paper includes a discussion of the multi-state case and its connections with the Markov property, and it also contains a simple version of the uniform sampling Theorem 6.17.

Excursion theory and the associated notion of local time can be traced back to a seminal paper of LÉVY (1939). Building on Lévy's ideas, ITÔ (1972) showed how the excursion structure of a Markov process can be represented in terms of a Poisson process on the local time scale. Regenerative sets with positive Lebesgue measure were studied extensively by KINGMAN (1972a). An introduction to regenerative processes and excursion theory appears in Chapter 22 of K(2002), and a more detailed exposition is given by DELLACHERIE, MAISONNEUVE, and MEYER (1987, 1992), Chapters 15 and 20.

Exchangeable partitions of finite intervals arise naturally in both theory and applications. In particular, they appear to play a significant role in game theory. Their Hausdorff measure was studied in K(1974b), where results for regenerative sets were extended to the finite-interval case. The associated distributions also arose naturally in K(1981), as Palm measures of regenerative sets. The latter paper further contains discussions, in the regenerative case, of continuity properties for the density of the local time intensity measure $E\xi$. Similar results on a finite interval were obtained in K(1983), where a slightly weaker version of Theorem 6.16 appears.

Both CARLESON and KESTEN, in independent work, proved that a subordinator with zero drift and infinite Lévy measure will a.s. avoid any fixed point in $(0, \infty)$, a result originally conjectured by CHUNG. Their entirely different approaches are summarized in ASSOUAD (1971) and BERTOIN (1996), Theorem III.4. The corresponding result for exchangeable partitions of $[0, 1]$ was established by BERBEE (1981). Our simple proof, restricted to the regular case, is adapted from K(1983). The asymptotic probability for an exchangeable random set in $\mathbb{R}_+$ or $[0, 1]$ to hit a short interval was studied extensively in K(1999c, 2001, 2003), along with the associated conditional distribution. Some extensions of those results to higher dimensions were derived by ELALAOUI-TALIBI (1999).

BLUMENTHAL and GETOOR (1968) noted that the strong Markov property follows from a condition of global homogeneity, in a version for optional times that may take infinite values. (To obtain a true Markov process, in the sense of the usual axiomatic definition, one needs to go on and construct an associated transition kernel, a technical problem addressed by WALSH (1972).) Connections with exchangeability were noted in K(1982), and some related but less elementary results along the same lines, though with homogeneity defined in terms of finite optional times, were established in K(1987) (cf. Theorem 8.23 in K(2002) for a short proof in a special case). In K(1998), the homogeneity and independence components of the strong Markov property were shown, under suitable regularity conditions, to be essentially equivalent. A totally unrelated characterization of mixed Markov chains was noted by DIACONIS and FREEDMAN (1980a) and subsequently studied by ZAMAN (1984, 1986).

7. Symmetries on Arrays

The notion of *partial exchangeability* of a sequence of random variables—the invariance in distribution under a sub-group of permutations—goes back to DE FINETTI (1938, 1959, 1972), and has later been explored by many authors. Separately (or *row-column*) exchangeable arrays $X = (X_{ij})$, first introduced by DAWID (1972), arise naturally in connection with a Bayesian approach to the analysis of variance. Jointly exchangeable, set-indexed arrays $X = (X_{\{i,j\}})$ (often referred to as *weakly exchangeable*) arise naturally in the context of U-statistics, a notion first introduced by HOEFFDING (1948). Dissociated and exchangeable arrays were studied in some early work of MCGINLEY and SIBSON (1975), SILVERMAN (1976), and EAGLESON and WEBER (1978).

A crucial breakthrough in the development of the subject came with the first representation theorems for row-column and weakly exchangeable arrays, established independently by ALDOUS (1981) and HOOVER (1982a), using entirely different methods. While the proof appearing in ALDOUS (1981), also outlined in ALDOUS (1985) and attributed by the author to KINGMAN, is purely probabilistic, the proof of HOOVER (1982a), also outlined in HOOVER (1982b), is based on profound ideas in symbolic logic and non-standard analysis. The underlying symmetry properties were considered much earlier by logicians, such as GAIFMAN (1961) and KRAUSS (1969), who obtained de Finetti-type results in the two cases, from which the functional representations can be derived.

HOOVER (1979), in a formidable, unpublished manuscript, went on to prove the general representation theorems for separately or jointly exchangeable arrays on $\mathbb{N}^d$, $d \geq 1$, using similar techniques from mathematical logic. His paper also provides criteria for the equivalence of two representing functions, corresponding to our condition (iii) in Theorems 7.28 and 7.29. (Condition (ii) was later added in K(1989a, 1992b, 1995).) A probabilistic approach to Hoover's main results, some of them in a slightly extended form, was provided in K(1988b, 1989a, 1992b, 1995). Contractable arrays on $\widetilde{\mathbb{N}}$ were studied in K(1992b), where the corresponding functional representation and equivalence criteria were established.

A different type of representation was obtained by DOVBYSH and SUDAKOV (1982), in the special case of positive definite, symmetric, jointly exchangeable arrays of dimension two. HESTIR (1986) shows how their result can be deduced from Hoover's representation.

Conditional properties of exchangeable arrays have been studied by many authors. In particular, extensions and alternative proofs of the basic Lemma 7.6 have been provided by LYNCH (1984), K(1989a), and HOOVER (1989). The conditional relations presented here were developed in K(1995). Convergence criteria and martingale-type properties for exchangeable arrays have been studied extensively by IVANOFF and WEBER (1992, 1995, 1996, 2003). The relevance of such arrays and their representations, in the contexts of the analysis of variance and Bayesian statistics, has been discussed by Kingman (1979), Speed (1987), and others.

An application to the study of visual perception was considered by Diaconis and Freedman (1981), who used simulations of exchangeable arrays of zeros and ones to disprove some conjectures in the area, posed by the psychologist Julesz and his followers. Such considerations lead naturally to the statistical problem of estimating (a version of) the representation function for an exchangeable array, given a single realization of the process. The problem was solved in K(1999b), to the extent that a solution seems at all possible. A non-technical introduction to exchangeable arrays appears in Aldous (1985), where more discussion and further references can be found.

Motivated by some problems in population genetics, Kingman (1978b, 1982b) studied exchangeable partitions of $\mathbb{N}$ and proved the corresponding special case of the so-called paint-box representation in Theorem 7.38. (The term comes from the mental picture, also suggested by Kingman, of an infinite sequence of objects, painted in randomly selected colors chosen from a possibly infinite paint box, where one considers the partition into classes of objects with the same color.) The present result for general symmetries is new; its proof was inspired by the first of two alternative approaches to Kingman's result suggested by Aldous (1985). A continuity theorem for exchangeable partitions appears in Kingman (1978c, 1980). Some algebraic and combinatorial aspects of exchangeable partitions have been studied extensively by Pitman (1995, 2002) and Gnedin (1997).

8. Multi-variate Rotations

The early developments in this area were motivated by some limit theorems for U-statistics, going back to Hoeffding (1948). Here the classical theory is summarized in Serfling (1980), and some more recent results are given by Dynkin and Mandelbaum (1983) and Mandelbaum and Taqqu (1984). Related results for random matrices are considered in Hayakawa (1966) and Wachter (1974).

Jointly rotatable matrices may have been mentioned for the first time in Olson and Uppuluri (1970, 1973), where a related characterization is proved. Infinite, two-dimensional, separately and jointly rotatable arrays were introduced by Dawid (1977, 1978), who also conjectured the general representation formula in the separate case. The result was subsequently proved by Aldous (1981), for dissociated arrays and under a moment condition, restrictions that were later removed in K(1988b). The latter paper also characterizes two-dimensional, jointly rotatable arrays, as well as separately or jointly exchangeable random sheets. It further contains some related uniqueness and continuity criteria.

Independently of K(1988b), Hestir (1986) derived the representation formula for separately exchangeable random sheets on $\mathbb{R}_+^2$, in the special case of vanishing drift components and under a moment condition, using a characterization of jointly exchangeable, positive definite arrays from Dovbysh and Sudakov (1982). Other, apparently independent, developments include a paper by Olshanski and Vershik (1996), where a related representation is established for jointly rotatable, Hermitian arrays.

The general theory of separately or jointly rotatable random arrays and functionals of arbitrary dimension was originally developed in K(1994, 1995), with some crucial ideas adopted from Aldous (1981), as indicated in the main text. The paper K(1995) also provides characterizations of exchangeable and contractable random sheets of arbitrary dimension. Brownian sheets and related processes have previously been studied by many authors, including Orey and Pruitt (1973) and Adler (1981).

9. Symmetric Measures in the Plane

The topic of exchangeable point processes in the plane first arose in discussions with Aldous (personal communication, 1979), and a general representation in the ergodic, separately exchangeable case was conjectured in Aldous (1985). The five basic representation theorems for exchangeable random measures in the plane were subsequently established in K(1990a). The extension theorem for contractable random measures was proved by Casukhela (1997).

Appendices

The general theory of integral representations over extreme points was developed by Choquet (1960), and modern expositions of his deep results may be found in Alfsen (1971) and in Dellacherie and Meyer (1983), Chapter 10. The special case of integral representations of invariant distributions has been studied extensively by many authors, including Farrell (1962), Varadarajan (1963), Phelps (1966), Maitra (1977), and Kerstan and Wakolbinger (1981). The connection with sufficient statistics was explored by Dynkin (1978), Diaconis and Freedman (1984), and Lauritzen (1989). Our general extremal decompositions are adapted from K(2000). Ergodicity is usually defined in terms of strictly invariant sets, which may require an extra condition on the family of transformations in Lemma A1.2. The randomization and selection Lemmas A1.5 and A1.6 are quoted from K(1988b) and K(2000), respectively, and the simple conservation law in Lemma A1.1 is adapted from K(1990a).

The equivalence of tightness and weak relative compactness, for probability measures on a Polish space, was established by Prohorov (1956). The general criteria were applied in Prohorov (1961) to random measures on a compact metric space. Tightness criteria for the vague topology on a Polish space were developed by Debes, Kerstan, Liemant, and Matthes (1970–71) and Harris (1971), and the corresponding criteria in the locally compact case were noted by Jagers (1974). The relationship between the weak and vague topologies for bounded random measures on a locally compact space was examined in K(1975–86). The present criterion for weak tightness of bounded random measures on a Polish space is adapted from ALDOUS (1985). The associated convergence criterion may be new.

The theory of multiple integrals with respect to Brownian motion was developed in a seminal paper of ITÔ (1951). The underlying idea is implicit in WIENER's (1938) discussion of chaos expansions of Gaussian functionals. For a modern account, see Chapters 13 and 18 of K(2002). The associated hyper-contraction property was discovered by NELSON (1973), and related developments are surveyed in DELLACHERIE et al. (1992) and KWAPIEŃ and WOYCZYŃSKI (1992). Convergence criteria for Poisson and related integrals were derived by KALLENBERG and SZULGA (1989).

The characterization of completely monotone sequences is due to HAUSDORFF (1921), and the corresponding results for completely monotone functions on $\mathbb{R}_+$ and $[0,1]$ were obtained by BERNSTEIN (1928). Simple, probabilistic proofs of these results were given by FELLER (1971), pp. 223–225 and 439–440. BOCHNER (1932) proved his famous characterization of positive definite functions, after a corresponding discrete-parameter result had been noted by HERGLOTZ (1911). Simple proofs of both results appear in FELLER (1971), pp. 622 and 634. The fundamental relationship between completely monotone and positive definite functions was established by SCHOENBERG (1938a,b). Modern proofs and further discussion appear in DONOGHUE (1969) and BERG, CHRISTENSEN, and RESSEL (1984).

The idea of Palm probabilities goes back to PALM (1943), and the modern definition of Palm distributions, via disintegration of the Campbell measure, is due to RYLL-NARDZEWSKI (1961). The Papangelou kernel of a simple point process was introduced in a special case by PAPANGELOU (1974b), and then in general in K(1978a). NGUYEN and ZESSIN (1979) noted the profound connection between point process theory and statistical mechanics, and MATTHES, WARMUTH, and MECKE (1979) showed how the Papangelou kernel can be obtained, under the regularity condition (Σ), through disintegration of the reduced Campbell measure. The present approach, covering even the general case, is based on K(1983–86).

Bibliography

Here I include only publications related, directly or indirectly, to material in the main text, including many papers that are not mentioned in the historical notes. No completeness is claimed, and any omissions are unintentional. No effort has been made to list the huge number of papers dealing with statistical applications. There are also numerous papers devoted to symmetries other than those considered in this book.

Aalen, O.O., Hoem, J.M. (1978). Random time changes for multivariate counting processes. *Scand. Actuarial J.*, 81–101.

Accardi, L., Lu, Y.G. (1993). A continuous version of de Finetti's theorem. *Ann. Probab.* **21**, 1478–1493.

Adler, R.J. (1981). *The Geometry of Random Fields.* Wiley, Chichester.

Aldous, D.J. (1977). Limit theorems for subsequences of arbitrarily-dependent sequences of random variables. *Z. Wahrsch. verw. Geb.* **40**, 59–82.

— (1981). Representations for partially exchangeable arrays of random variables. *J. Multivar. Anal.* **11**, 581–598.

— (1982a). On exchangeability and conditional independence. In Koch and Spizzichino (eds.), pp. 165–170.

— (1982b). Partial exchangeability and $\bar{d}$-topologies. *Ibid.*, pp. 23–38.

— (1985). Exchangeability and related topics. In: *École d'Été de Probabilités de Saint-Flour XIII—1983. Lect. Notes in Math.* **1117**, pp. 1–198. Springer, Berlin.

Aldous, D.J., Pitman, J.W. (1979). On the zero-one law for exchangeable events. *Ann. Probab.* **7**, 704–723.

Alfsen, E.M. (1971). *Compact Convex Sets and Boundary Integrals.* Springer, Berlin.

Anis, A.A., Gharib, M. (1980). On the variance of the maximum of partial sums of n-exchangeable random variables with applications. *J. Appl. Probab.* **17**, 432–439.

Assouad, P. (1971). Démonstration de la "Conjecture de Chung" par Carleson. *Séminaire de Probabilités V. Lect. Notes in Math.* **191**, 17–20. Springer, Berlin.

Athreya, K.B, Ney, P.E. (1972). *Branching Processes.* Springer, Berlin.

Balasubramanian, K, Balakrishnan, N. (1994). Equivalence of relations for order statistics for exchangeable and arbitrary cases. *Statist. Probab. Lett.* **21**, 405–407.

Ballerini, R. (1994). Archimedean copulas, exchangeability, and max-stability. *J. Appl. Probab.* **31**, 383–390.

Barbour, A.D., Eagleson, G.K. (1983). Poisson approximations for some statistics based on exchangeable trials. *Adv. Appl. Probab.* **15**, 585–600.

Barbour, A.D., Holst, L., Janson, S. (1992). *Poisson Approximation.* Clarendon Press, Oxford.

Bartfai, P. (1980). Remarks on exchangeable random variables. *Publ. Math. Debrecen* **27**, 143–148.

Benczur, A. (1968). On sequences of equivalent events and the compound Poisson process. *Studia Sci. Math. Hungar.* **3**, 451–458.

Berbee, H. (1981). On covering single points by randomly ordered intervals. *Ann. Probab.* **9**, 520–528.

Berg, C., Christensen, J.P.R., Ressel, P. (1984). *Harmonic Analysis on Semigroups.* Springer, New York.

Berkes, I. (1974). The law of the iterated logarithm for subsequences of random variables. *Z. Wahrsch. verw. Geb.* **30**, 209–215.

— (1984). Exchangeability and limit theorems for subsequences of random variables. In: *Limit Theorems in Probability and Statistics* (ed. P. Révész), pp. 109–152. North-Holland/Elsevier, Amsterdam.

Berkes, I., Péter, E. (1986). Exchangeable random variables and the subsequence principle. *Probab. Th. Rel. Fields* **73**, 395–413.

Berkes, I., Philipp, W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.* **7**, 29–54.

Berkes, I, Rosenthal, H.P. (1985). Almost exchangeable sequences of random variables. *Z. Wahrsch. verw. Geb.* **70**, 473–507.

Berman, S. (1980). Stationarity, isotropy, and sphericity in l^p. *Z. Wahrsch. verw. Geb.* **54**, 21–23.

Bernstein, S. (1928). Sur les fonctions absolument monotones. *Acta Math.* **52**, 1–66.

Berti, P., Pratelli, L., Rigo, P. (2004). Limit theorems for a class of identically distributed random variables. *Ann. Probab.* **32**, 2029–2052.

Berti, P., Rigo, P. (1997). A Glivenko–Cantelli theorem for exchangeable random variables. *Statist. Probab. Lett.* **32**, 385–391.

Bertoin, J. (1996). *Lévy Processes.* Cambridge University Press, Cambridge.

— (2001a). Eternal additive coalescents and certain bridges with exchangeable increments. *Ann. Probab.* **29**, 344–360.

— (2001b). Homogeneous fragmentation processes. *Probab. Theor. Rel. Fields* **121**, 301–318.

— (2002). Self-similar fragmentations. *Ann. Inst. H. Poincaré, Sec. B* **38**, 319–340.

Bikelis, A. (1969). On the estimation of the remainder term in the central limit theorem for samples from finite populations (Russian). *Studia Sci. Math. Hungar.* **4**, 345–354.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

— (1986). *Probability and Measure*, 2nd ed. Wiley, New York.

Blackwell, D. (1946). On an equation of Wald. *Ann. Math. Statist.* **17**, 84–87.

Blom, G. (1985). A simple property of exchangeable random variables. *Amer. Math. Monthly* **92**, 491–492.

Blum, J.R. (1982). Exchangeability and quasi-exchangeability. In Koch and Spizzichino (eds.), pp. 171–176.

Blum, J.R., Chernoff, H., Rosenblatt, M., Teicher, H. (1958). Central limit theorems for interchangeable processes. *Canad. J. Math.* **10**, 222–229.

Blumenthal, R.M., Getoor, R.K. (1968). *Markov Processes and Potential Theory.* Academic Press, New York.

Bochner, S. (1932). *Vorlesungen über Fouriersche Integrale.* Reprint ed., Chelsea, New York 1948.

Boes, D.C., Salas, L.C.J.D. (1973). On the expected range and expected adjusted range of partial sums of exchangeable random variables. *J. Appl. Probab.* (1973), 671–677.

Borel, E. (1914). *Introduction géométrique à quelques théories physiques.* Gauthier-Villars, Paris.

Bretagnolle, J, Dacunha-Castelle, D., Krivine, J.L. (1966): Lois stables et espaces L^p. *Ann. Inst. H. Poincaré B* **2**, 231–259.

Bretagnolle, J., Klopotowski, A. (1995). Sur l'existence des suites de variables aléatoires s à s indépendantes échangeables ou stationnaires. *Ann. Inst. H. Poincaré B* **31**, 325–350.

Brown, T.C. (1982). Poisson approximations and exchangeable random variables. In Koch and Spizzichino (eds.), pp. 177–183.

— (1983). Some Poisson approximations using compensators. *Ann. Probab.* **11**, 726–744.

Brown, T.C., Ivanoff, B.G., Weber, N.C. (1986). Poisson convergence in two dimensions with application to row and column exchangeable arrays. *Stoch. Proc. Appl.* **23**, 307–318.

Brown, T.C., Nair, M.G. (1988a). A simple proof of the multivariate random time change theorem for point processes. *J. Appl. Probab.* **25**, 210–214.

— (1988b). Poisson approximations for time-changed point processes. *Stoch. Proc. Appl.* **29**, 247–256.

Bru, B., Heinich, H., Lootgitier, J.C. (1981). Lois de grands nombres pour les variables échangeables. *C.R. Acad. Sci. Paris* **293**, 485–488.

Bühlmann, H. (1958). Le problème "limit centrale" pour les variables aléatoires échangeables. *C.R. Acad. Sci. Paris* **246**, 534–536.

— (1960). Austauschbare stochastische Variabeln und ihre Grenzwertsätze. *Univ. Calif. Publ. Statist.* **3**, 1–35.

Burkholder, D.L., Gundy, R.F. (1970). Extrapolation and interpolation of quasi-linear operators on martingales. *Acta Math.* **124**, 249-304.

Casukhela, K.S. (1997). Symmetric distributions of random measures in higher dimensions. *J. Theor. Probab.* **10**, 759–771.

Chanda, C. (1971). Asymptotic distribution of sample quantiles for exchangeable random variables. *Calcutta Statist. Assoc. Bull.* **20**, 135–142.

Chatterji, S.D. (1972). Un principe de sous-suites dans la théorie des probabilités. *Séminaire de probabilités VI. Lect. Notes in Math.* **258**, 72–89. Springer, Berlin.

— (1974a). A principle of subsequences in probability theory: The central limit theorem. *Adv. Math.* **13**, 31–54.

— (1974b). A subsequence principle in probability theory II. The law of the iterated logarithm. *Invent. Math.* **25**, 241–251.

Chernoff, H., Teicher, H. (1958). A central limit theorem for sums of interchangeable random variables. *Ann. Math. Statist.* **29**, 118–130.

Choquet, G. (1960). Le théorème de représentation intégrale dans les ensembles convexes compacts. *Ann. Inst. Fourier* **10**, 333–344.

Chow, Y.S., Robbins, H., Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping.* Houghton Mifflin, New York.

Chow, Y.S., Teicher, H. (1997). *Probability Theory: Independence, Interchangeability, Martingales*, 3rd ed. Springer, New York (1st ed. 1978).

Chung, K.L. (1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

Cocozza, C., Yor, M. (1980). Démonstration d'un théorème de F. Knight à l'aide de martingales exponentielles. *Lect. Notes in Math.* **784**, 496–499. Springer, Berlin.

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.

— (1982). Exchangeability and completely monotone functions. In Koch and Spizzichino (eds.), pp. 39–45.

Dabrowski, A.R. (1988). Strassen-type invariance principles for exchangeable sequences. *Statist. Probab. Lett.* **7**, 23–26.

Dacunha-Castelle, D. (1975). Indiscernability and exchangeability in L^p spaces. *Aarhus Proceedings* **24**, 50–56.

— (1982). A survey on exchangeable random variables in normed spaces. In Koch and Spizzichino (eds.), pp. 47–60.

Daffer, P.Z. (1984). Central limit theorems for weighted sums of exchangeable random elements in Banach spaces. *Stoch. Anal. Appl.* **2**, 229–244.

Dale, A.I. (1985). A study of some early investigations into exchangeability. *Hist. Math.* **12**, 323–336.

Dambis, K.E. (1965). On the decomposition of continuous submartingales. *Theory Probab. Appl.* **10**, 401–410.

Daras, T. (1997). Large and moderate deviations for the empirical measures of an exchangeable sequence. *Statist. Probab. Lett.* **36**, 91–100.

— (1998). Trajectories of exchangeable sequences: Large and moderate deviations results. *Statist. Probab. Lett.* **39**, 289–304.

Davidson, R. (1974). Stochastic processes of flats and exchangeability. In: *Stochastic Geometry* (eds. E.F. Harding, D.G. Kendall), pp. 13–45. Wiley, London.

Dawid, A.P. (1972). Contribution to discussion of Lindley and Smith (1972), *J. Roy. Statist. Soc. Ser. B* **34**, 29–30.

— (1977). Spherical matrix distributions and a multivariate model. *J. Roy. Statist. Soc.* (B) **39**, 254–261.

— (1978). Extendibility of spherical matrix distributions. *J. Multivar. Anal.* **8**, 559–566.

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., Maisonneuve, B., Meyer, P.A. (1983, 1987, 1992). *Probabilités et Potentiel*, Vols. **3–5**, Hermann, Paris.

Diaconis, P. (1977). Finite forms of de Finetti's theorem on exchangeability. *Synthese* **36**, 271–281.

— (1988). Recent progress on de Finetti's notions of exchangeability. In: *Bayesian Statistics* (eds. J.M. Bernardo, M.H. Degroot, D.V. Lindley, A.F.M. Smith), **3**, 111–125. Oxford Univ. Press.

Diaconis, P., Eaton, M.L., Lauritzen, S.L. (1992). Finite de Finetti theorems in linear models and multivariate analysis. *Scand. J. Statist.* **19**, 289–315.

Diaconis, P., Freedman, D.A. (1980a). De Finetti's theorem for Markov chains. *Ann. Probab.* **8**, 115–130.

— (1980b). Finite exchangeable sequences. *Ann. Probab.* **8**, 745–764.

— (1980c). De Finetti's generalizations of exchangeability. In: *Studies in Inductive Logic and Probability*, pp. 233–249. Univ. Calif. Press.

— (1981). On the statistics of vision: the Julesz conjecture. *J. Math. Psychol.* **24**, 112–138.

— (1982). De Finetti's theorem for symmetric location families. *Ann. Statist.* **10**, 184–189.

— (1984). Partial exchangeability and sufficiency. In: *Statistics: Applications and New Directions* (eds. J.K. Ghosh, J. Roy), pp. 205–236. Indian Statistical Institute, Calcutta.

— (1987). A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré* **23**, 397–423.

— (1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. *J. Theor. Probab.* **1**, 381–410.

— (1990). Cauchy's equation and de Finetti's theorem. *Scand. J. Statist.* **17**, 235–250.

Dinwoodie, I.H., Zabell, S.L. (1992). Large deviations for exchangeable random vectors. *Ann. Probab.* **20**, 1147–1166.

Donoghue, W.F. (1969). *Distributions and Fourier Transforms.* Academic Press, New York.

DOOB, J.L. (1936). Note on probability. *Ann. Math.* (2) **37**, 363–367.

— (1953). *Stochastic Processes.* Wiley, New York.

DOVBYSH, L.N., SUDAKOV, V.N. (1982). Gram–de Finetti matrices. *J. Soviet Math.* **24**, 3047–3054.

DUBINS, L.E. (1982). Towards characterizing the set of ergodic probabilities. In KOCH and SPIZZICHINO (eds.), pp. 61–74.

— (1983). Some exchangeable probabilities are singular with respect to all presentable probabilities. *Z. Wahrsch. verw. Geb.* **64**, 1–5.

DUBINS, L.E., FREEDMAN, D.A. (1965). A sharper form of the Borel–Cantelli lemma and the strong law. *Ann. Math. Statist.* **36**, 800–807.

— (1979). Exchangeable processes need not be mixtures of independent, identically distributed random variables. *Z. Wahrsch. verw. Geb.* **48**, 115–132.

DUBINS, L.E., SAVAGE, L.J. (1965), A Tchebycheff-like inequality for stochastic processes. *Proc. Nat. Acad. Sci. USA* **53**, 274–275.

DUBINS, L.E., SCHWARZ, G. (1965). On continuous martingales. *Proc. Natl. Acad. Sci. USA* **53**, 913–916.

DYNKIN, E.B. (1953). Classes of equivalent random variables (in Russian). *Uspehi Mat. Nauk (8)* **54**, 125–134.

— (1978). Sufficient statistics and extreme points. *Ann. Probab.* **6**, 705–730.

DYNKIN, E.B., MANDELBAUM, A. (1983). Symmetric statistics, Poisson point processes and multiple Wiener integrals. *Ann. Statist.* **11**, 739–745.

EAGLESON, G.K. (1979). A Poisson limit theorem for weakly exchangeable events. *J. Appl. Probab.* **16**, 794–802.

— (1982). Weak limit theorems for exchangeable random variables. In KOCH and SPIZZICHINO (eds.), pp. 251–268.

EAGLESON, G.K., WEBER, N.C. (1978). Limit theorems for weakly exchangeable arrays. *Math. Proc. Cambridge Phil. Soc.* **84**, 123–130.

EATON, M. (1981). On the projections of isotropic distributions. *Ann. Statist.* **9**, 391–400.

ELALAOUI-TALIBI, H. (1999). Multivariate Palm measure duality for exchangeable interval partitions. *Stoch. Stoch. Reports* **66**, 311–328.

ELALAOUI-TALIBI, H., CASUKHELA, K.S. (2000). A note on the multivariate local time intensity of exchangeable interval partitions. *Statist. Probab. Lett.* **48**, 269–273.

ERDÖS, P., RÉNYI, A. (1959). On the central limit theorem for samples from a finite population. *Publ. Math. Inst. Hungar. Acad. Sci. A* **4**, 49–61.

ESSEEN, C.G., VON BAHR, B. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \le r \le 2$. *Ann. Math. Statist.* **36**, 299–303.

ETEMADI, N., KAMINSKI, M. (1996). Strong law of large numbers for 2-exchangeable random variables. *Statist. Probab. Letters* **28**, 245–250.

EVANS, S.N., ZHOU, X. (1999). Identifiability of exchangeable sequences with identically distributed partial sums. *Elec. Comm. Probab.* **4**, 9–13.

EVERITT, C.W.F. (1974). *James Clerk Maxwell.* Scribner's, New York.

FARRELL, R.H. (1962). Representation of invariant measures. *Ill. J. Math.* **6**, 447–467.

FEIGIN, P. (1979). On the characterization of point processes with the order statistic property. *J. Appl. Probab.* **16**, 297–304.

FELLER, W. (1966, 1971). *An Introduction of Probability Theory and its Applications*, Vol. **2**, 1st and 2nd eds. Wiley, New York.

FERGUSON, T.S., KLASS, M.J. (1972). A representation of independent increment processes without Gaussian component. *Ann. Math. Statist.* **43**, 1634–1643.

FIGIEL, T, SUCHESTON, L. (1976). An application of Ramsey sets in analysis. *Adv. Math.* **20**, 103–105.

FINETTI, B. DE (1929). Fuzione caratteristica di un fenomeno aleatorio. In: *Atti Congr. Int. Mat., Bologna, 1928* (ed. ZANICHELLI) **6**, 179–190.

— (1930). Fuzione caratteristica di un fenomeno aleatorio. *Mem. R. Acc. Lincei* (6) **4**, 86–133.

— (1933a). Classi di numeri aleatori equivalenti. *Rend. Accad. Naz. Lincei* **18**, 107–110.

— (1933b). La legge dei grandi numeri nel caso dei numeri aleatori equivalenti. *Rend. Accad. Naz. Lincei* **18**, 279–284.

— (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7**, 1–68. Engl. transl.: *Studies in Subjective Probability* (eds. H.E. KYBURG, H.E. SMOKLER), pp. 99–158. Wiley, New York, 2nd ed. 1980.

— (1938). Sur la condition d'équivalence partielle. *Act. Sci. Ind.* **739**, 5–18. Engl. transl.: *Studies in Inductive Logic and Probability, II* (ed. R.C. JEFFREY). Univ. California Press, Berkeley.

— (1959). La probabilità e la statistica nei rapporti con l'induzione, secondo i diversi punti di vista. In: *Atti Corso CIME su Induzione e Statistica*, Varenna, pp. 1–115.

— (1972). *Probability, Induction and Statistics* (collected papers in English translation). Wiley, New York.

— (1974–75). *Theory of Probability, I–II.* Wiley, New York.

FORTINI, S., LADELLI, L., REGAZZINI, E. (1996). A central limit problem for partially exchangeable random variables. *Theor. Probab. Appl.* **41**, 224–246.

FRANKEN, P., KÖNIG, D., ARNDT, U., SCHMIDT, V. (1981). *Queues and Point Processes.* Akademie-Verlag, Berlin.

FRANKEN, P., LISEK, B. (1982). On Wald's identity for dependent variables. *Z. Wahrsch. verw. Geb.* **60**, 134–150.

FRÉCHET, M. (1943). Les probabilités associées à un système d'événements compatibles et dépendants, II. *Actual. Scient. Indust.* **942**. Hermann, Paris.

FREEDMAN, D.A. (1962–63). Invariants under mixing which generalize de Finetti's theorem. *Ann. Math. Statist.* **33**, 916–923; **34**, 1194–1216.

— (1980). A mixture of independent identically distributed random variables need not admit a regular conditional probability given the exchangeable σ-field. *Z. Wahrsch. verw. Geb.* **51**, 239–248.

— (1996). De Finetti's theorem in continuous time. In: *Statistics, Probability and Game Theory* (eds. T.S. Ferguson, L.S. Shapley, J.B. MacQueen), pp. 83–98. Inst. Math. Statist., Hayward, CA.

Fristedt, B.E. (1967). Sample function behavior of increasing processes with stationary, independent increments. *Pacific J. Math.* **21**, 21–33.

Gaifman, H. (1961). Concerning measures in first order calculi. *Israel J. Math.* **2**, 1–18.

Galambos, J. (1982). The role of exchangeability in the theory of order statistics. In Koch and Spizzichino (eds.), pp. 75–86.

Gallardo, L. (1983). Au sujet du contenu probabiliste d'un lemma d'Henri Poincaré. *Ann. Univ. Clermont* **69**, 192–197.

Gaposhkin, V.F. (1966). Lacunary series and independent functions. *Russian Math. Surveys* **21/6**, 3–82.

— (1972). Convergence and limit theorems for sequences of random variables. *Theory Probab. Appl.* **17**, 379–399.

Gnedin, A.V. (1995). On a class of exchangeable sequences. *Statist. Probab. Lett.* **25**, 351–355.

— (1996). On a class of exchangeable sequences. *Statist. Probab. Lett.* **28**, 159–164.

— (1997). The representation of composition structures. *Ann. Probab.* **25**, 1437–1450.

Griffiths, R.C., Milne, R.K. (1986). Structure of exchangeable infinitely divisible sequences of Poisson random vectors. *Stoch. Proc. Appl.* **22**, 145–160.

Grigelionis, B (1971). On representation of integer-valued random measures by means of stochastic integrals with respect to the Poisson measure (in Russian). *Litov. Mat. Sb.* **11**, 93–108.

— (1975). The characterization of stochastic processes with conditionally independent increments. *Litovsk. Mat. Sb.* **15**, 53–60.

— (1977). Martingale characterization of stochastic processes with independent increments. *Litovsk. Mat. Sb.* **17**, 75–86.

Guerre, S. (1986). Sur les suites presque échangeables dans L^q, $1 \le q < 2$. *Israel J. Math.* **56**, 361–380.

Guglielmi, A., Melilli, E. (2000). Approximating de Finetti's measures for partially exchangeable sequences. *Statist. Probab. Lett.* **48**, 309–315.

Haag, J. (1928). Sur un problème général de probabilités et ses diverses applications. In: *Proc. Int. Congr. Math., Toronto 1924*, pp. 659–674.

Hagberg, J. (1973). Approximation of the summation process obtained by sampling from a finite population. *Theory Probab. Appl.* **18**, 790–803.

Hájek, J. (1960). Limiting distributions in simple random sampling from a finite population. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5**, 361–374.

Halmos, P.R. (1985). *I want to be a Mathematician.* Springer, New York.

Harris, T.E. (1971). Random measures and motions of point processes. *Z. Wahrsch. verw. Geb.* **18**, 85–115.

HAUSDORFF, F. (1921). Summationsmethoden und Momentfolgen. *Math. Z.* **9**, 280–299.

HAYAKAWA, T. (1966). On the distribution of a quadratic form in a multivariate normal sample. *Ann. Inst. Statist. Math.* **18**, 191–201.

HAYAKAWA, Y. (2000). A new characterization property of mixed Poisson processes via Berman's theorem. *J. Appl. Probab.* **37**, 261–268.

HEATH, D., SUDDERTH, W. (1976). De Finetti's theorem for exchangeable random variables. *Amer. Statistician* **30**, 188–189.

HERGLOTZ, G. (1911). Über Potenzreihen mit positivem, reelem Teil im Einheitskreis. *Ber. verh. Sächs. Ges. Wiss. Leipzig, Math.-Phys. Kl.* **63**, 501–511.

HESTIR, K. (1986). The Aldous representation theorem and weakly exchangeable non-negative definite arrays. Ph.D. dissertation, Statistics Dept., Univ. of California, Berkeley.

HEWITT, E., SAVAGE, L.J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80**, 470–501.

HILL, B.M., LANE, D., SUDDERTH, W. (1987). Exchangeable urn processes. *Ann. Probab.* **15**, 1586–1592.

HIRTH, U., RESSEL, P. (1999). Random partitions by semigroup methods. *Semigr. For.* **59**, 126–140.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distributions. *Ann. Math. Statist.* **19**, 293–325.

— (1963). Probability estimates for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30.

HOOVER, D.N. (1979). Relations on probability spaces and arrays of random variables. Preprint, Institute of Advanced Study, Princeton.

— (1982a). A normal form theorem for the probability logic $L_{\omega_1 P}$, with applications. *J. Symbolic Logic* **47**.

— (1982b). Row-column exchangeability and a generalized model for probability. In KOCH and SPIZZICHINO (eds.), pp. 281–291.

— (1989). Tail fields of partially exchangeable arrays. *J. Multivar. Anal.* **31**, 160–163.

HSU, Y.S. (1979). A note on exchangeable events. *J. Appl. Probab.* **16**, 662–664.

HU, T.C. (1997). On pairwise independent and independent exchangeable random variables. *Stoch. Anal. Appl.* **15**, 51–57.

HU, Y.S. (1979). A note on exchangeable events. *J. Appl. Probab.* **16**, 662–664.

HUANG, W.J., SU, J.C. (1999). Reverse submartingale property arising from exchangeable random variables. *Metrika* **49**, 257–262.

HUDSON, R.L., MOODY, G.R. (1976). Locally normal symmetric states and an analogue of de Finetti's theorem. *Z. Wahrsch. verw. Geb.* **33**, 343–351.

ITÔ, K. (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3**, 157–169.

— (1972). Poisson point processes attached to Markov processes. *Proc. 6th Berkeley Symp. Math. Statist. Probab.* **3**, 225–239. Univ. of California Press, Berkeley.

Ivanoff, B.G., Weber, N.C. (1992). Weak convergence of row and column exchangeable arrays. *Stoch. Stoch. Rep.* **40**, 1–22.

— (1995). Functional limit theorems for row and column exchangeable arrays. *J. Multivar. Anal.* **55**, 133–148.

— (1996). Some characterizations of partial exchangeability. *J. Austral. Math. Soc.* (A) **61**, 345–359.

— (2003). Spreadable arrays and martingale structures. *J. Austral. Math. Soc.*, to appear.

— (2004). Predictable sampling for partially exchangeable arrays. *Statist. Probab. Lett.* **70**, 95–108.

Jacod, J. (1975). Multivariate point processes: Predictable projection, Radon–Nikodym derivative, representation of martingales. *Z. Wahrsch. verw. Geb.* **31**, 235–253.

Jacod, J., Shiryaev, A.N. (1987). *Limit Theorems for Stochastic Processes.* Springer, Berlin.

Jagers, P. (1974). Aspects of random measures and point processes. *Adv. Probab. Rel. Topics* **3**, 179–239. Marcel Dekker, New York.

Jiang, X., Hahn, M.G. (2002). Empirical central limit theorems for exchangeable random variables. *Statist. Probab. Lett.* **59**, 75–81.

Johnson, W.B., Schechtman, G., Zinn, J. (1985). Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probab.* **13**, 234–253.

Kallenberg, O. (1972). Ph.D. dissertation, Math. Dept., Chalmers Univ. of Technology, Gothenburg, Sweden.

— (1973a). Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrsch. verw. Geb.* **27**, 23–36.

— (1973b). A canonical representation of symmetrically distributed random measures. In: *Mathematics and Statistics, Essays in Honour of Harald Bergström* (eds. P. Jagers, L. Råde), pp. 41–48. Chalmers Univ. of Technology, Gothenburg.

— (1973c). Characterization and convergence of random measures and point processes. *Z. Wahrsch. verw. Geb.* **27**, 9–21.

— (1974a). Series of random processes without discontinuities of the second kind. *Ann. Probab.* **2**, 729–737.

— (1974b). Path properties of processes with independent and interchangeable increments. *Z. Wahrsch. verw. Geb.* **28**, 257–271.

— (1975a). On symmetrically distributed random measures. *Trans. Amer. Math. Soc.* **202**, 105–121.

— (1975b). On the existence and path properties of stochastic integrals. *Ann. Probab.* **3**, 262–280.

— (1975c). Infinitely divisible processes with interchangeable increments and random measures under convolution. *Z. Wahrsch. verw. Geb.* **32**, 309–321.

— (1975–86). *Random Measures*, 1st to 4th eds. Akademie-Verlag and Academic Press, Berlin and London.

— (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.

— (1981). Splitting at backward times in regenerative sets. *Ann. Probab.* **9**, 781–799.

— (1982). Characterizations and embedding properties in exchangeability. *Z. Wahrsch. verw. Geb.* **60**, 249–281.

— (1983). The local time intensity of an exchangeable interval partition. In: *Probability and Statistics, Essays in Honour of Carl-Gustav Esseen* (eds. A. Gut, L. Holst), pp. 85–94. Uppsala University.

— (1983–86). *Random Measures*, 3rd and 4th eds. Akademie-Verlag and Academic Press, Berlin and London.

— (1987). Homogeneity and the strong Markov property. *Ann. Probab.* **15**, 213–240.

— (1988a). Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16**, 508–534.

— (1988b). Some new representations in bivariate exchangeability. *Probab. Th. Rel. Fields* **77**, 415–455.

— (1988c). One-dimensional uniqueness and convergence criteria for exchangeable processes. *Stoch. Proc. Appl.* **28**, 159–183.

— (1989a). On the representation theorem for exchangeable arrays. *J. Multivar. Anal.* **30**, 137–154.

— (1989b). General Wald-type identities for exchangeable sequences and processes. *Probab. Th. Rel. Fields* **83**, 447–487.

— (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.

— (1992a). Some time change representations of stable integrals, via predictable transformations of local martingales. *Stoch. Proc. Appl.* **40**, 199–223.

— (1992b). Symmetries on random arrays and set-indexed processes. *J. Theor. Probab.* **5**, 727–765.

— (1993). Some linear random functionals characterized by L^p-symmetries. In: *Stochastic Processes, a Festschrift in Honour of Gopinath Kallianpur* (eds. S. Cambanis, J.K. Ghosh, R.L. Karandikar, P.K. Sen), pp. 171–180. Springer, Berlin.

— (1994). Multiple Wiener–Itô integrals and a multivariate version of Schoenberg's theorem. In: *Chaos Expansions, Multiple Wiener–Itô Integrals and their Applications* (eds. C. Houdré, V. Pérez-Abreu), pp. 73–86. CRC Press, Boca Raton.

— (1995). Random arrays and functionals with multivariate rotational symmetries. *Probab. Th. Rel. Fields* **103**, 91–141.

— (1998). Components of the strong Markov property. In: *Stochastic Processes and Related Topics: In Memory of Stamatis Cambanis, 1943–1995* (eds. I. Karatzas, B.S. Rajput, M.S. Taqqu), pp. 219–230. Birkhäuser, Boston.

— (1999a). Asymptotically invariant sampling and averaging from stationary-like processes. *Stoch. Proc. Appl.* **82**, 195–204.

— (1999b). Multivariate sampling and the estimation problem for exchangeable arrays. *J. Theor. Probab.* **12**, 859–883.

— (1999c). Palm measure duality and conditioning in regenerative sets. *Ann. Probab.* **27**, 945–969.

— (2000). Spreading-invariant sequences and processes on bounded index sets. *Probab. Th. Rel. Fields* **118**, 211–250.

— (2001). Local hitting and conditioning in symmetric interval partitions. *Stoch. Proc. Appl.* **94**, 241–270.

— (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York (1st ed. 1997).

— (2003). Palm distributions and local approximation of regenerative processes. *Probab. Th. Rel. Fields* **125**, 1–41.

Kallenberg, O., Szulga, J. (1989). Multiple integration with respect to Poisson and Lévy processes. *Probab. Th. Rel. Fields* **83**, 101–134.

Karatzas, I., Shreve, S. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, New York.

Karoui, N., Lepeltier, J.P. (1977). Représentation des processus ponctuels multivariés à l'aide d'un processus de Poisson. *Z. Wahrsch. verw. Geb.* **39**, 111–133.

Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā* A **32**, 419–430.

Kendall, D.G. (1966). Branching processes since 1873. *J. London Math. Soc.* **41**, 385–406.

— (1967). On finite and infinite sequences of exchangeable events. *Studia Sci. Math. Hung.* **2**, 319–327.

Kerstan, J., Wakolbinger, A. (1981). Ergodic decomposition of probability laws. *Z. Wahrsch. verw. Geb.* **56**, 399–414.

Khinchin, A.Y. (1932). Sur les classes d'événements équivalents. *Mat. Sbornik* **33**, 40–43.

— (1939). Sur la croissance locale des processus stochastiques homogènes à accroissements indépendants (in Russian with French summary). *Izv. Akad. Nauk SSSR, Ser. Mat.* **3**, 487–508.

— (1952). On classes of equivalent events (in Russian). *Dokl. Akad. Nuak SSSR* **85**, 713–714.

Kimberling, C.H. (1973). Exchangeable events and completely monotonic sequences. *Rocky Mtn. J.* **3**, 565–574.

KINGMAN, J.F.C. (1972a). *Regenerative Phenomena.* Wiley, London.

— (1972b). On random sequences with spherical symmetry. *Biometrika* **59**, 492–493.

— (1978a). Uses of exchangeability. *Ann. Probab.* **6**, 183–197.

— (1978b). The representation of partition structures. *J. London Math. Soc.* **18**, 374–380.

— (1978c). Random partitions in population genetics. *Proc. R. Soc. Lond. A* **361**, 1–20.

— (1979). Contribution to discussion on: The reconciliation of probability assessments. *J. Roy. Statist. Soc. Ser. A* **142**, 171.

— (1980). *The Mathematics of Genetic Diversity.* SIAM, Philadelphia.

— (1982a). The coalescent. *Stoch. Proc. Appl.* **13**, 235–248.

— (1982b). Exchangeability and the evolution of large populations. In KOCH and SPIZZICHINO (eds.), pp. 97–112.

KLASS, M.J. (1988). A best possible improvement of Wald's equation. *Ann. Probab.* **16**, 840–853.

KLASS, M.J., TEICHER, H. (1987). The central limit theorem for exchangeable random variables without moments. *Ann. Probab.* **15**, 138–153.

KNIGHT, F.B. (1970). An infinitesimal decomposition for a class of Markov processes. *Ann. Math. Statist.* **41**, 1510–1529.

— (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.

KOCH, G, SPIZZICHINO, F. (eds.) (1982). *Exchangeability in Probability and Statistics.* North-Holland, Amsterdam.

KOMLÓS, J. (1967). A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.* **18**, 217–229.

KRAUSS, P.H. (1969). Representations of symmetric probability models. *J. Symbolic Logic* **34**, 183–193.

KRICKEBERG, K. (1974). Moments of point processes. In: *Stochastic Geometry* (eds. E.F. HARDING, D.G. KENDALL), pp. 89–113. Wiley, London.

KURITSYN, Y.G. (1984). On monotonicity in the law of large numbers for exchangeable random variables. *Theor. Probab. Appl.* **29**, 150–153.

— (1987). On strong monotonicity of arithmetic means of exchangeable random variables. *Theor. Probab. Appl.* **32**, 165–166.

KURTZ, T.G. (1980). Representations of Markov processes as multiparameter time changes. *Ann. Probab.* **8**, 682–715.

KWAPIEŃ, S., WOYCZYŃSKI, W.A. (1992). *Random Series and Stochastic Integrals: Single and Multiple.* Birkhäuser, Boston.

LAURITZEN, S.L. (1989). *Extremal Families and Systems of Sufficient Statistics. Lect. Notes in Statist.* **49**. Springer, Berlin.

Lefévre, C., Utev, S. (1996). Comparing sums of exchangeable Bernoulli random variables. *J. Appl. Probab.* **33**, 285–310.

Letac, G. (1981). Isotropy and sphericity: some characterizations of the normal distribution. *Ann. Statist.* **9**, 408–417.

Lévy, P. (1924). Théorie des erreurs. La lois de Gauss et les lois exceptionelles. *Bull. Soc. Math. France* **52**, 49–85.

— (1925). *Calcul des Probabilités.* Gauthier-Villars, Paris.

— (1937–54). *Théorie de l'Addition des Variables Aléatoires*, 1st and 2nd eds. Gauthier-Villars, Paris.

— (1939). Sur certain processus stochastiques homogènes. *Comp. Math,* **7**, 283–339.

Loève, M. (1960–63). *Probability Theory*, 2nd and 3rd eds. Van Nostrand, Princeton, NJ.

— (1978). *Probability Theory*, Vol. **2**, 4th ed. Springer, New York.

Lynch, J. (1984). Canonical row-column-exchangeable arrays. *J. Multivar. Anal.* **15**, 135–140.

Maitra, A. (1977). Integral representations of invariant measures. *Trans. Amer. Math. Soc.* **229**, 209–225.

Mandelbaum, A., Taqqu, M.S. (1983). Invariance principle for symmetric statistics. *Ann. Statist.* **12**, 483–496.

Mandrekar, V., Patterson, R.F. (1993). Limit theorems for symmetric statistics of exchangeable random variables. *Statist. Probab. Lett.* **17**, 157–161.

Matthes, K., Kerstan, J., Mecke, J. (1974–82). *Infinitely Divisible Point Processes.* Wiley, Chichester 1978 (German ed., Akademie-Verlag, Berlin 1974; Russian ed., Nauka, Moscow 1982).

Matthes, K., Warmuth, W., Mecke, J. (1979). Bemerkungen zu einer Arbeit von Nguyen Xuan Xanh and Hans Zessin. *Math. Nachr.* **88**, 117–127.

Maxwell, J.C. (1875). *Theory of Heat*, 4th ed. Longmans, London.

— (1878). On Boltzmann's theorem on the average distribution of energy in a system of material points. *Trans. Cambridge Phil. Soc.* **12**, 547.

McGinley, W.G., Sibson, R. (1975). Dissociated random variables. *Math. Proc. Cambridge Phil. Soc.* **77**, 185–188.

McKean, H.P. (1973). Geometry of differential space. *Ann. Probab.* **1**, 197–206.

Merzbach, E., Nualart, D. (1986). A characterization of the spatial Poisson process and changing time. *Ann. Probab.* **14**, 1380–1390.

Meyer, P.A. (1971). Démonstration simplifiée d'un théorème de Knight. *Lect. Notes in Math.* **191**, 191–195. Springer, Berlin.

Millar, P.W. (1971). Path behavior of processes with stationary independent increments. *Z. Wahrsch. verw. Geb.* **17**, 53–73.

— (1972). Stochastic integrals and processes with stationary independent increments. *Proc. 6th Berkeley Symp. Math. Statist. Probab.* **3**, 307–331. Univ. of California Press, Berkeley.

MISIEWICZ, J. (1990). Pseudo isotropic measures. *Nieuw Arch. Wisk.* **8**, 111–152.

MOIVRE, A. DE (1718–56). *The Doctrine of Chances; or, a Method of Calculating the Probability of Events in Play*, 3rd ed. (post.) Reprint, F. Case and Chelsea, London, NY 1967.

MORAN, P.A.P. (1973). A central limit theorem for exchangeable variates with geometrical applications. *J. Appl. Probab.* **10**, 837–846.

MOSCOVICI, E., POPESCU, O. (1973). A theorem on exchangeable random variables. *Stud. Cerc. Mat.* **25**, 379–383.

NAWROTZKI, K. (1962). Ein Grenzwertsatz für homogene zufällige Punktfolgen (Verallgemeinerung eines Satzes von A. Rényi). *Math. Nachr.* **24**, 201–217.

NEKKACHI, M.J. (1994). Weak convergence for weakly exchangeable arrays. *C.R. Acad. Sci. Ser. Math.* **319**, 717–722.

NELSON, E. (1973). The free Markov field. *J. Funct. Anal.* **12**, 211–227.

NEUTS, M., RESNICK, S. (1971). On the times of birth in a linear birth process. *J. Austral. Math. Soc.* **12**, 473–475.

NGUYEN X.X., ZESSIN, H. (1979). Integral and differential characterizations of the Gibbs process. *Math. Nachr.* **88**, 105–115.

OLSHANSKY, G.I., VERSHIK, A.M. (1996). Ergodic unitarily invariant measures on the space of infinite Hermitian matrices. In: *Contemporary Mathematical Physics*, 137–175. *AMS Transl. Ser.* 2, **175**, Providence, RI.

OLSHEN, R.A. (1971). The coincidence of measure algebras under an exchangeable probability. *Z. Wahrsch. verw. Geb.* **18**, 153–158.

— (1973). A note on exchangeable sequences. *Z. Wahrsch. verw. Geb.* **28**, 317–321.

OLSON, W.H., UPPULURI, V.R.R. (1970). Characterization of the distribution of a random matrix by rotational invariance. *Sankhyā Ser. A* **32**, 325–328.

— (1973). Asymptotic distribution of eigenvalues of random matrices. *Proc. 6th Berkeley Symp. Math. Statist. Probab.* **3**, 615–644.

OREY, S., PRUITT, W.E. (1973). Sample functions of the N-parameter Wiener process. *Ann. Probab.* **1**, 138–163.

PALM, C. (1943). Intensitätsschwankungen in Fernsprechverkehr. *Ericsson Technics* **44**, 1–189. Engl. trans., *North-Holland Studies in Telecommunication* **10**, Elsevier 1988.

PAPANGELOU, F. (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* (eds. E.F. HARDING, D.G. KENDALL), 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.

Pathak, P.K. (1974). An extension of an inequality of Hoeffding. *Bull. Inst. Math. Statist.* **3**, 156.

Patterson, R.F. (1989). Strong convergence for U-statistics in arrays of rowwise exchangeable random variables. *Stoch. Anal. Appl.* **7**, 89–102.

Patterson, R.F., Taylor, R.L., Inoue, H. (1989). Strong convergence for sums of randomly weighted, rowwise exchangeable random variables. *Stoch. Anal. Appl.* **7**, 309–323.

Phelps, R.R. (1966). *Lectures on Choquet's Theorem.* Van Nostrand, Princeton, NJ.

Pitman, J.W. (1978). An extension of de Finetti's theorem. *Adv. Appl. Probab.* **10**, 268–270.

— (1995). Exchangeable and partially exchangeable random partitions. *Probab. Th. Rel. Fields* **102**, 145–158.

— (2002). Combinatorial stochastic processes. In: *École d'Été de Probabilités de Saint-Flour. Lect. Notes in Math.* Springer, Berlin (to appear).

Pitman, J.W., Yor, M. (1986). Asymptotic laws of planar Brownian motion. *Ann. Probab.* **14**, 733–779.

Poincaré, H. (1912). *Calcul des Probabilités.* Gauthier-Villars, Paris.

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.

Pruss, A.R. (1998). A maximal inequality for partial sums of finite exchangeable sequences of random variables. *Proc. Amer. Math. Soc.* **126**, 1811–1819.

Rao, C.R., Shanbhag, D.N. (2001). Exchangeability, functional equations, and characterizations. In: *Handbook of Statistics* **19** (eds. D.N. Shanbhag, C.R. Rao), 733–763. Elsevier/North-Holland, New York.

Regazzini, E., Sazonov, V.V. (1997). On the central limit problem for partially exchangeable random variables with values in a Hilbert space. *Theor. Probab. Appl.* **42**, 656–670.

Rényi, A. (1953). On the theory of order statistics. *Acta Math. Acad. Sci. Hung.* **4**, 191–231.

— (1963). On stable sequences of events. *Sankhyā, Ser. A* **25**, 293–302.

Rényi, A., Révész, P. (1963). A study of sequences of equivalent events as special stable sequences. *Publ. Math. Debrecen* **10**, 319–325.

Ressel, P. (1985). De Finetti-type theorems: an analytical approach. *Ann. Probab.* **13**, 898–922.

— (1988). Integral representations for distributions of symmetric stochastic processes. *Probab. Th. Rel. Fields* **79**, 451–467.

— (1994). Non-homogeneous de Finetti-type theorems. *J. Theor. Probab.* **7**, 469–482.

Révész, P. (1965). On a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.* **16**, 310–318.

Ridler-Rowe, C.J. (1967). On two problems on exchangeable events. *Studia Sci. Math. Hung.* **2**, 415–418.

Robinson, J., Chen, K.H. (1989). Limit theorems for standardized partial sums of weakly exchangeable arrays. *Austral. J. Statist.* **31**, 200–214.

Rogers, L.C.G., Williams, D. (1994). *Diffusions, Markov Processes, and Martingales*, **I**, 2nd ed. Wiley, Chichester.

Romanowska, M. (1983). Poisson theorems for triangle arrays of exchangeable events. *Bull. Acad. Pol. Sci. Ser. Math.* **31**, 211–214.

Rosalsky, A. (1987). A strong law for a set-indexed partial sum process with applications to exchangeable and stationary sequences. *Stoch. Proc. Appl.* **26**, 277–287.

Rosén, B (1964). Limit theorems for sampling from a finite population. *Ark. Mat.* **5**, 383–424.

— (1967). On an inequality of Hoeffding. *Ann. Math. Statist.* **38**, 382–392.

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.

Royden, H.L. (1988). *Real Analysis*, 3rd ed. Macmillan, New York.

Ryll-Nardzewski, C. (1957). On stationary sequences of random variables and the de Finetti's [sic] equivalence. *Colloq. Math.* **4**, 149–156.

— (1961). Remarks on processes of calls. In: *Proc. 4th Berkeley Symp. Math. Statist. Probab.* **2**, 455–465.

Saunders, R. (1976). On joint exchangeability and conservative processes with stochastic rates. *J. Appl. Probab.* **13**, 584–590.

Scarsini, M. (1985). Lower bounds for the distribution function of a k-dimensional n-extendible exchangeable process. *Statist. Probab. Lett.* **3**, 57–62.

Scarsini, M., Verdicchio, L. (1993). On the extendibility of partially exchangeable random vectors. *Statist. Probab. Lett.* **16**, 43–46.

Schoenberg, I.J. (1938a). Metric spaces and completely monotone functions. *Ann. Math.* **39**, 811–841.

— (1938b). Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.* **44**, 522–536.

Scott, D.J., Huggins, R.M. (1985). A law of the iterated logarithm for weakly exchangeable arrays. *Math. Proc. Camb. Phil. Soc.* **98**, 541–545.

Seneta, E. (1987). Chuprov on finite exchangeability, expectation of ratios, and measures of association. *Hist. Math.* **14**, 249–257.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Shaked, M. (1977). A concept of positive dependence for exchangeable random variables. *Ann. Statist.* **5**, 505–515.

Silverman, B.W. (1976). Limit theorems for dissociated random variables. *Adv. Appl. Probab.* **8**, 806–819.

Skorohod, A.V. (1957). Limit theorems for stochastic processes with independent increments. *Theory Probab. Appl.* **2**, 138–171.

Slivnyak, I.M. (1962). Some properties of stationary flows of homogeneous random events. *Th. Probab. Appl.* **7**, 336–341; **9**, 168.

Slud, E.V. (1978). A note on exchangeable sequences of events. *Rocky Mtn. J. Math.* **8**, 439–442.

Smith, A.M.F. (1981). On random sequences with centered spherical symmetry. *J. Roy. Statist. Soc. Ser. B* **43**, 208–209.

Sparre-Andersen, E. (1953–54). On the fluctuations of sums of random variables, I–II. *Math. Scand.* **1**, 263–285; **2**, 193–194, 195–223.

Speed, T. (1987). What is an analysis of variance? *Ann. Statist.* **15**, 885–910.

Spizzichino, F. (1982). Extendibility of symmetric probability distributions and related bounds. In Koch and Spizzichino (eds.), pp. 313–320.

Stam, A.J. (1982). Limit theorems for uniform distributions on high dimensional Euclidean spaces. *J. Appl. Probab.* **19**, 221–228.

Sun, Y. (1998). The almost equivalence of pairwise and mutual independence and the duality with exchangeability. *Probab. Theor. Rel. Fields* **112**, 425–456.

Szczotka, W. (1980). A characterization of the distribution of an exchangeable random vector. *Bull. Acad. Polon., Ser. Sci. Math.* **28**, 411–414.

Takács, L. (1967). *Combinatorial Methods in the Theory of Stochastic Processes.* Wiley, New York.

Taylor, R.L. (1986). Limit theorems for sums of exchangeable random variables. *Int. Statist. Symp.* **1986**, 785–805.

Taylor, R.L., Daffer, P.Z., Patterson, R.F. (1985). *Limit Theorems for Sums of Exchangeable Random Variables.* Rowman & Allanheld, Totowa, NJ.

Taylor, R.L., Hu, T.C. (1987). On laws of large numbers for exchangeable random variables. *Stoch. Anal. Appl.* **5**, 323–334.

Teicher, H. (1960). On the mixture of distributions. *Ann. Math. Statistist.* **31**, 55–73.

— (1971). On interchangeable random variables. In: *Studi di Probabilita Statistica e Ricerca in Onore di Giuseppe Pompilj*, pp. 141–148.

— (1982). Renewal theory for interchangeable random variables. In Koch and Spizzichino (eds.), pp. 113–121.

Trashorras, J. (2002). Large deviations for a triangular array of exchangeable random variables. *Ann. Inst. H. Poincaré, Sec. B* **38**, 649–680.

Troutman, B.M. (1983). Weak convergence of the adjusted range of cumulative sums of exchangeable random variables. *J. Appl. Probab.* **20**, 297–304.

Vaillancourt, J. (1988). On the existence of random McKean–Vlasov limits for triangular arrays of exchangeable diffusions. *Stoch. Anal. Appl.* **6**, 431–446.

Varadarajan, V.S. (1963). Groups of automorphisms of Borel spaces. *Trans. Amer. Math. Soc.* **109**, 191–220.

WACHTER, K.W. (1974). Exchangeability and asymptotic random matrix spectra. In: *Progress in Statistics* (eds. J. GANI, K. SARKADI, I. VINCE), pp. 895–908. North-Holland, Amsterdam.

WALD, A. (1944). On cumulative sums of random variables. *Ann. Math. Statist.* **15**, 283–296.

— (1945). Sequential tests of statistical hypotheses. *Ann. Math. Statist.* **16**, 117–186.

WALSH, J.B. (1972). Transition functions of Markov processes. *Séminaire de Probabilités VI. Lect. Notes in Math.* **258**, 215–232. Springer, Berlin.

WATANABE, S. (1964). On discontinuous additive functionals and Lévy measures of a Markov process. *Japan. J. Math.* **34**, 53–79.

WATSON, G.S. (1986). An ordering inequality for exchangeable random variables. *Adv. Appl. Probab.* **18**, 274–276.

WAUGH, W.A.O'N. (1970). Transformation of a birth process into a Poisson process. *J. Roy. Statist. Soc. B* **32**, 418–431.

WAYMIRE, E.C., WILLIAMS, S.C. (1996). A cascade decomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math. Soc.* **348**, 585–632.

WEBER, N.C. (1980). A martingale approach to central limit theorems for exchangeable random variables. *J. Appl. Probab.* **17**, 662–673.

WIENER, N. (1938). The homogeneous chaos. *Amer. J. Math.* **60**, 897–936.

YOR, M. (1985). Inégalités de martingales continues arrétées à un temps quelconque, I. *Lect. Notes in Math.* **1118**, 110–171. Springer, Berlin.

ZABELL, S.L. (1995). Characterizing Markov exchangeable sequences. *J. Theor. Probab.* **8**, 175–178.

ZAMAN, A. (1984). Urn models for Markov exchangeability. *Ann. Probab.* **12**, 223–229.

— (1986). A finite form of de Finetti's theorem for stationary Markov exchangeability. *Ann. Probab.* **14**, 1418–1427.

Author Index

Subject Index

Symbol Index