Some Contributions to the Stochastic Analysis of the Dirichlet Process

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Babette Josephine Picker, M.Sc.

Tag der mündlichen Prüfung:	22. Oktober 2025
Referent:	Prof. Dr. Günter Last
Korreferent:	Prof. Dr. Daniel Hug

Danksagung

Mein besonderer Dank gilt meinem Betreuer Prof. Dr. Günter Last, der mir die Promotion ermöglicht hat. Während dieser Zeit konnte ich nicht nur fachlich mein Wissen ausbauen, sondern auch persönlich sehr viel lernen. Seine Haltung, sich in erster Linie als Mathematiker zu sehen, prägt seine Arbeit und hat auch meine eigene Sicht beeinflusst. Ich konnte sehr von seinem umfangreichen Fachwissen, das weit über die Stochastik hinausgeht, profitieren. Gleichzeitig weiß ich die große Freiheit und das mir entgegengebrachte Vertrauen sehr zu schätzen und empfinde es nicht als selbstverständlich.

Des Weiteren möchte ich mich bei meinem Korreferenten Prof. Dr. Daniel Hug bedanken. Ich freue mich sehr über sein großes Interesse an meiner Arbeit bereits während meiner Zeit als Doktorandin. Seine zahlreichen Ideen, Anregungen und umfangreiches Feedback haben sehr zum Gelingen dieser Arbeit beigetragen.

Außerdem schulde ich PD Dr. Steffen Winter großen Dank. Sein Glaube an meine Fähigkeiten, seine Unterstützung und offene Tür, an der ich stets mit allen meinen Anliegen willkommen war, haben mich immer wieder bestärkt, motiviert und aufgebaut. Insbesondere bin ich dankbar für das große Vertrauen in unseren gemeinsamen Lehrveranstaltungen und die Möglichkeit, speziell in Mathe 3, meine verrückten Ideen umzusetzen.

Auch allen anderen Dozenten, die ich bei der Betreuung von Lehrveranstaltungen unterstützen durfte, möchte ich Dank aussprechen. An dieser Stelle geht auch ein großes Dankeschön an alle Studierenden, die ich in Lehrveranstaltungen oder als Tutorinnen und Tutoren am Institut betreuen durfte. Sie sind der Grund, warum ich auch an schlechten Tagen mit einem Lächeln ins Büro gekommen bin.

Darüber hinaus möchte ich meine aufrichtige Wertschätzung gegenüber allen aktuellen und ehemaligen Mitgliedern des Instituts für Stochastik für das stets angenehme Arbeitsumfeld zum Ausdruck bringen.

Von Herzen danke ich meinen Freunden und meiner Familie für ihre bedingungslose Unterstützung.

Contents

Chapter 1

Introduction

Imagine throwing a dice. In a perfect setting, each side of the dice has an equal chance of appearing. However, due to manufacturing imperfections, different dice can have different biases. Now imagine we have a bag full of dice and we randomly select one to throw. This scenario illustrates a probability distribution (selecting a dice) over a set of probability distributions (different dice) and illustrates the idea underlying the Dirichlet process, which defines a probability distribution over probability measures. The Dirichlet process is widely used in statistical modelling, particularly for tasks such as clustering, density estimation and modelling distributions where the number of components is not fixed in advance (for example, clustering a dataset into groups without specifying the number of groups beforehand). Examples are Ferreira da Silva (2007), which uses a model based on a Dirichlet process to segment brain tissue in MRI images, Rodríguez-Álvarez, Inácio and Klein (2025) for a recent density regression model and Topkaya, Erdogan and Porikli (2013), where a method to automatically detect and track a previously unknown number of objects in videos is presented. A comprehensive statistical perspective on Dirichlet processes can be found in Teh (2010). Moreover, Shiga (1990) shows that the distribution of a Dirichlet process is the unique stationary reversible distribution of a Fleming–Viot process with parent-independent mutation, a measure-valued Markov process in population genetics. The Dirichlet process is well-known in number theory and combinatorics (cf. Billingsley (1972) showing that the sequence consisting of the atoms of a Dirichlet process arises as limiting distribution in prime factorisation or Pitman (2006) for random partitions). A survey connecting population genetics, statistics and number theory is Crane (2016).

Returning to our dice example, consider the task of defining a probability distribution over a set of probability distributions on a finite set with $n \geq 1$ elements (in case of a six-sided dice, $n = 6$). This can be approached as follows. Any probability distribution on a set with n elements is characterised by a vector $(p_1, \dots, p_n)$ whose entries are non-negative and sum to one. In other words, every probability distribution on this set corresponds to a point in the simplex

$$\Delta_n \coloneqq \{(p_1, \dots, p_n) \in [0,1]^n : p_1 + \ldots + p_n = 1\}.$$

The Dirichlet distribution is a probability distribution on this simplex. For $n \geq 2$ and a parameter vector $(\alpha_1, \dots, \alpha_n)$ with $\alpha_1, \dots, \alpha_n \in (0, \infty)$ the Dirichlet distribution $\mathrm{Dir}(\alpha_1, \dots, \alpha_n)$ assigns to each Borel subset S of Δ_n the value

$$\frac{\Gamma(\alpha_1 + \ldots + \alpha_n)}{\Gamma(\alpha_1) \ldots \Gamma(\alpha_n)} \int_0^1 \cdots \int_0^1 \mathbb{1}_S(x_1, \dots, x_n) x_1^{\alpha_1 - 1} \cdots x_n^{\alpha_n - 1} \, \mathrm{d}x_1 \ldots \mathrm{d}x_n,$$

where Γ denotes the Gamma function $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, \mathrm{d}t$, $z > 0$. As is customary, we extend the definition to allow for parameter vectors in which some entries vanish. Let $n \geq 2$ and suppose $\alpha_1 = \ldots = \alpha_k = 0$ for some $k < n$ while $\alpha_{k+1}, \dots, \alpha_n > 0$. Then we define $\mathrm{Dir}(\alpha_1, \dots, \alpha_n) = \delta_0^{\otimes k} \otimes \mathrm{Dir}(\alpha_{k+1}, \dots, \alpha_n)$, the product measure consisting of k times the Dirac measure in 0, denoted by δ_0, and the Dirichlet distribution $\mathrm{Dir}(\alpha_{k+1}, \dots, \alpha_n)$ on the lower-dimensional simplex Δ_{n-k}. This reflects the fact that components corresponding to zero parameters are almost surely equal to zero. A similar definition applies when parameter vectors $(\alpha_1, \dots, \alpha_n)$ with $\alpha_1, \dots, \alpha_n \in [0, \infty)$ and $\alpha_1 + \ldots + \alpha_n > 0$ are considered. Furthermore, if $n = 1$ and $\alpha_1 > 0$, we set $\mathrm{Dir}(\alpha_1) = \delta_1$, the Dirac measure in 1. The Dirichlet distribution is widely applied in various settings, including machine learning, Bayesian statistics and population genetics. A thorough treatment can be found for example in Ng, Tian and Tang (2011).

The task of defining a probability distribution over distributions on arbitrary spaces proves to be more intricate if the probability distribution is expected to place positive mass on a rich class of measures while remaining analytically and computationally tractable. The fundamental idea introduced by Ferguson was to extend the concept of the Dirichlet distribution to a general setting. The Dirichlet process is a random probability measure such that, for any finite partition of the underlying space, the vector of probabilities assigned to the partition elements follows a Dirichlet distribution governed by some parameter measure. Formally, let $(\mathbb{X}, \mathcal{X})$ be a measurable space carrying a finite measure $\rho \neq 0$ and let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A random probability measure is a measurable mapping from Ω to the set of all finite measures, equipped with the standard σ-field. Ferguson's definition now reads as follows.

Definition 1.1 (Ferguson (1973))**.** A random probability measure ζ on $\mathbb{X}$ is called a Dirichlet process with parameter measure ρ if

$$\mathbb{P}((\zeta(B_1), \ldots, \zeta(B_n)) \in \cdot) = \mathrm{Dir}(\rho(B_1), \ldots, \rho(B_n)) \tag{1.1}$$

for every measurable partition $B_1, \ldots, B_n$, $n \in \mathbb{N}$, of $\mathbb{X}$.

One method of studying random processes is stochastic analysis. Stochastic analysis or stochastic calculus is an extension of calculus to stochastic processes. In calculus, an elementary identity is the integration by parts formula. Generalised to an open subset D of $\mathbb{R}^n$ with sufficiently smooth boundary, it states

$$\int_D \langle \nabla f(x), h(x) \rangle \,\mathrm{d}x = -\int_D f(x) \operatorname{div} h(x) \,\mathrm{d}x$$

for continuously differentiable $f\colon D \to \mathbb{R}$ and $h\colon D \to \mathbb{R}^n$ with compact support in D. (Of course, the assumptions can be weakened. Since we are here primarily interested in using the intuition from elementary calculus as a guiding star, we impose stronger assumptions than necessary.) Here, $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in $\mathbb{R}^n$, the gradient ∇f is, as usual, a vector containing the partial derivatives of f and the divergence $\operatorname{div} h$ is defined by

$$\operatorname{div} h \coloneqq \partial_1 h_1 + \ldots + \partial_n h_n,$$

where h_i is for $i \in \{1, \ldots, n\}$ the ith component of h. Introducing the Laplace operator Δ, which is for $f \in C^2(D, \mathbb{R})$ given by

$$\Delta f \coloneqq \partial_{11} f + \ldots + \partial_{nn} f,$$

the identity

$$\operatorname{div} \nabla f = \Delta f$$

follows for $f \in C^2(D, \mathbb{R})$. Moreover, for $f, g \in C^2(D, \mathbb{R})$ with compact support, we obtain

$$\int_D \Delta f(x) g(x) \,\mathrm{d}x = \int_D f(x) \Delta g(x) \,\mathrm{d}x = -\int_D \langle \nabla f(x), \nabla g(x) \rangle \,\mathrm{d}x.$$

In stochastic analysis, the Lebesgue measure is replaced by the distribution of a random process and the considered functions are random variables which are square integrable with respect to this distribution. The construction of derivatives becomes more intricate in this infinite-dimensional setting. One possible approach is to consider random variables of a specific form for which a gradient can be defined explicitly. Gradients for a broader class of random variables can subsequently be defined through approximation techniques. Another approach is to use an orthogonal decomposition of the space of all square integrable random variables, often referred to as chaos expansion. With the help of this decomposition, every square-integrable random variable can be represented as a (possibly infinite) sum of integrals with respect to the process. Using this representation of a random variable, it is possible to define a gradient in terms of the chaos expansion. Once a gradient is available, one can define an adjoint operator that satisfies an integration by parts formula, analogous to the divergence in $\mathbb{R}^n$. Within the framework of random processes, the adjoint operator to the gradient in a suitable space, commonly also referred to as the divergence operator, is denoted by δ. The composition of the divergence δ and the gradient ∇ results, as in $\mathbb{R}^n$, in yet another operator. This operator is usually denoted by L and is the generator of an associated Markov process. We conclude the motivating example with a connection to stochastic analysis. Let

$$P_t f(x) \coloneqq \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} f(y) \,\mathrm{d}y, \quad t > 0,\ x \in \mathbb{R}^n,$$

for bounded and continuous functions $f\colon \mathbb{R}^n \to \mathbb{R}$. It can be shown that $\{P_t, t \geq 0\}$, where P_0 is the identity mapping, forms a semigroup of operators and that the function $u\colon (0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ with $u(t, x) = P_t f(x)$ for a bounded and continuous function f is a solution of the heat equation $\partial_t = \Delta_x$ with $\lim_{t \to 0} u(t, x) = f(x)$ for every $x \in \mathbb{R}^n$. Moreover, it holds

$$P_t f(x) = \mathbb{E}[f(x + B_{2t})], \quad t > 0,\ x \in \mathbb{R}^n,$$

for a bounded and continuous function f and a Brownian motion $(B_t)_{t \geq 0}$ in $\mathbb{R}^n$. This sparks the idea that the Brownian motion is the associated Markov process. (For rigorous statements and proofs concerning the heat semigroup, see e.g. Bakry, Gentil and Ledoux (2014), Ethier and Kurtz (2005) or Jost (2013).)

Regarding the Dirichlet process, preliminary steps in stochastic analysis have already been taken. In Peccati (2008), a chaos expansion for square integrable functions is established and in Flint and Torrisi (2023) a partial integration formula with what Dello Schiavo and Lytvynov (2023) refer to as "the discrete gradient" is derived. The exact results will be recalled in Chapter 2. In the context of Fleming–Viot processes, also a gradient and a generator, defined with the help of a class of smooth cylindrical functions, are studied. We recall the Fleming–Viot process with parent-independent mutation in Chapter 2. There are several other measure-valued processes closely related, e.g. the process on the space of all probability measures on the unit interval $[0, 1]$ constructed in Shao (2011), which is reversible with respect to the distribution of a Dirichlet process, the Wasserstein diffusion on the same state space from von Renesse and Sturm (2009) or the process considered in Dello Schiavo (2022), where also a comparison to the Wasserstein diffusion as well as the Fleming–Viot process with parent-independent mutation can be found. More generally, processes related to the Poisson–Dirichlet distribution or the Pitman–Yor process are explored for instance in Petrov (2009) or Feng and Sun (2010). The study of Markov processes also encompasses functional inequalities, such as the Poincaré inequality. A Poincaré inequality for the Dirichlet form of the Fleming–Viot process with parent-independent mutation is proven in Stannat (2000). Functional inequalities for the process from von Renesse and Sturm (2009) and the Poisson–Dirichlet distribution are derived in Döring and Stannat (2009) and Feng, Sun, Wang and Xu (2011), respectively. The spectral gap of another Dirichlet form related to the Dirichlet and the Gamma process is estimated in Ren and Wang (2020). A Poincaré inequality for the Dirichlet distribution is established in Feng, Miclo and Wang (2017).

Despite the various works concerning different facets of stochastic analysis of the Dirichlet process, a thorough and unified examination remains absent. The present thesis proposes an approach based on explicit formulas for the chaos expansion. These explicit formulas are derived with the help of a multivariate integral equation for the Dirichlet process. Before we can formulate the equation, we introduce some notation. Let $(\mathbb{X}, \mathcal{X})$ be a measurable space. Given a measure μ on $\mathbb{X}$ and a natural number $n \in \mathbb{N}$, we define a measure $\mu^{[n]}$ by

$$\mu^{[n]}(B) := \int_{\mathbb{X}} \dots \int_{\mathbb{X}} \mathbb{1}_B(x_1, \dots, x_n)\,(\mu + \delta_{x_1} + \dots + \delta_{x_{n-1}})(\mathrm{d}x_n) \dots (\mu + \delta_{x_1})(\mathrm{d}x_2)\,\mu(\mathrm{d}x_1), \quad B \in \mathcal{X}^{\otimes n}.$$

If the distribution of a Dirichlet process ζ on $\mathbb{X}$ with parameter measure ρ is denoted by $\mathrm{Dir}(\rho)$ and the set of all finite measures on $\mathbb{X}$ by $\mathbf{M}(\mathbb{X})$, the multivariate Mecke-type equation from Chapter 3 states

$$\begin{aligned}\int_{\mathbf{M}(\mathbb{X})} \int_{\mathbb{X}^n} & f(\mu, x_1, \dots, x_n)\,\mu^n(\mathrm{d}(x_1, \dots, x_n))\ \mathrm{Dir}(\rho, \mathrm{d}\mu) \\ &= \frac{1}{\rho^{[n]}(\mathbb{X}^n)} \int_{\mathbb{X}^n} \int_{\mathbf{M}(\mathbb{X})} f(\mu, x_1, \dots, x_n)\ \mathrm{Dir}(\rho + \delta_{x_1} + \dots + \delta_{x_n}, \mathrm{d}\mu)\,\rho^{[n]}(\mathrm{d}(x_1, \dots, x_n))\end{aligned}$$

for all measurable $f\colon \mathbf{M}(\mathbb{X}) \times \mathbb{X}^n \to [0, \infty)$. Using this equation, in Chapter 4, we are able to show that every measurable mapping $F\colon \mathbf{M}(\mathbb{X}) \to \mathbb{R}$ such that $\mathbb{E}[F(\zeta)^2] < \infty$ can be written as an orthogonal series

$$F(\zeta) = \mathbb{E}[F] + \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} f_n(x)\,\zeta^n(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.}$$

where the convergence is in $L^2(\mathbb{P})$. Moreover, we can explicitly calculate the kernel functions f_n, $n \in \mathbb{N}$. In terms of the chaos expansion, we specify a dense subset $\mathrm{dom}(\nabla)$ of all square-integrable mappings F for which we then define a gradient ∇F in Chapter 5. As a next step, we introduce a divergence operator δ acting as an adjoint of ∇ in the sense that

$$\mathbb{E}\left[\int_{\mathbb{X}} H(x) \nabla_x F\,\zeta(\mathrm{d}x)\right] = \mathbb{E}\left[\delta(H) F\right]$$

for mappings F such that ∇F is defined and suitable $H\colon \Omega \times \mathbb{X} \to \mathbb{R}$. Note that the integration on the left-hand side is taken with respect to the random measure ζ, in contrast to the Gaussian and Poisson case, where integration with respect to a deterministic measure is considered. Finally, we construct an operator L satisfying

$$-\delta(\nabla F) = LF, \quad \mathbb{P}\text{-a.s.}$$

as well as

$$\mathbb{E}[(LF)G] = \mathbb{E}[F(LG)] = -\mathcal{E}(\nabla F, \nabla G)$$

for all mappings F, G in its domain. Here, $\mathcal{E}$ is the Dirichlet form associated with the Fleming–Viot process with parent-independent mutation. We show that L is the generator of this process. An application of the theory developed in the present thesis is a sharp Poincaré inequality and a reverse Poincaré inequality for the Dirichlet process, deduced in Chapter 7.

This thesis is structured as follows. Chapter 2 reviews concepts and results from the literature relevant for this thesis and explains the notation used. Chapter 3 is devoted to the study of a multivariate version of the Mecke-type equation, which is stated in Theorem 3.7, along with an analysis of the properties of the associated moment measures. As a corollary, an explicit formula for the moment measures of a Dirichlet process is obtained in (3.4). The multivariate Mecke-type equation is the basis for Chapter 4, in which the chaos expansion, stated in Theorem 4.18, is derived. An explicit formula for the involved kernel functions is given in (4.16). Chapter 5 makes use of the chaos expansion to define the operators ∇, δ and L in dedicated sections. The integration by parts formula connecting these three operators is established in Theorem 5.20. The connection to the Fleming–Viot process with parent-independent mutation is drawn in Chapter 6. In particular, in Theorem 6.5 the associated Dirichlet form is obtained (in terms of the chaos decomposition). Chapter 7 contains the variance bounds in Theorem 7.1 (sharp Poincaré inequality) and Theorem 7.6 (reverse Poincaré inequality). For the sake of readability, two elementary summation formulas used in various parts of this thesis have been collected in the appendix.

CHAPTER 2

Dirichlet processes

This chapter provides an overview of key concepts from the existing literature. The first section collects some constructions of Dirichlet processes to give the reader a brief impression of the process's versatility. To situate these constructions within a broader conceptual landscape, this section includes blue boxes that point to related concepts or provide background information. The second section is devoted to the Mecke-type equation. The third section reviews relevant concepts from the literature. The final section explains the notation used in this thesis.

2.1. Construction

Since Ferguson's seminal work in Ferguson (1973), in which he defined and proved the existence of Dirichlet processes through the verification of Kolmogorov's consistency conditions (conditions that ensure that a collection of finite-dimensional distributions gives rise to a stochastic process; Ferguson shows that it is possible to define a random probability measure by specifying the distribution on partitions under certain conditions, which, in particular, guarantee consistency under refinements of the partitions) and gave a construction using Gamma processes (§ 4 in Ferguson (1973)), a variety of alternative constructions have been proposed in the subsequent literature. In this section, we provide a brief overview of different constructions.

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space with a non-zero finite measure ρ.

Construction with a Gamma random measure

Analogous to the construction of a random vector with a Dirichlet distribution via independent Gamma random variables, a Dirichlet process can be constructed. This approach is described, for instance, in Appendix 2 in Kingman (1975), in § 9.2 in Kingman (1993) or in Example 15.6 in combination with Exercise 15.1 in Chapter 15 in Last and Penrose (2017).

Construction of a Dirichlet-distributed random vector using Gamma distributed random variables (cf. e.g. Example 3.3.6 in Hogg, McKean and Craig (2019))

Let $k \in \mathbb{N}$ and let $\theta_1, \dots, \theta_k > 0$. Moreover, let $Y_1, \dots, Y_k$ be independent random variables such that Y_i follows a Gamma distribution with shape parameter θ_i and rate parameter 1 for each $i \in \{1, \dots, k\}$. By standard properties of the Gamma distribution, $Y = Y_1 + \dots + Y_k$ has a Gamma distribution with shape parameter $\theta_1 + \dots + \theta_k$ and rate parameter 1.
A change of variable argument for the corresponding densities shows that the vector $(Y_1/Y, \dots, Y_k/Y)$ is Dirichlet-distributed with parameter $(\theta_1, \dots, \theta_k)$ and independent of Y.

Let η be a Poisson process on $\mathbb{X} \times (0, \infty)$ whose intensity measure is the product of ρ and the measure on $(0, \infty)$ that has Lebesgue-density $r \mapsto r^{-1}e^{-r}$. Then,

$$\xi(B) := \int_{B \times (0,\infty)} r\, \eta(\mathrm{d}(x, r)), \quad B \in \mathcal{X},$$

defines a random measure ξ on $\mathbb{X}$ (cf. Example 15.6 in Last and Penrose (2017)), which is completely independent, i.e. for pairwise disjoint measurable sets $B_1, \dots, B_m$ the random variables $\xi(B_1), \dots, \xi(B_m)$ are independent ($m \in \mathbb{N}$). Furthermore, $\xi(B)$ follows a Gamma distribution with shape parameter $\rho(B)$ and scale parameter 1 for $B \in \mathcal{X}$ (cf. Example 15.6 in Last and Penrose (2017)). Thus, ξ is called *Gamma random measure.*

As ρ is assumed to be finite, $\xi(\mathbb{X})$ is finite with probability one and the random probability measure ζ, defined by

$$\zeta(B) := \frac{\xi(B)}{\xi(\mathbb{X})}, \quad B \in \mathcal{X}, \tag{2.1}$$

is a Dirichlet process with parameter measure ρ (cf. e.g. Appendix 2 in Kingman (1975)), independent of $\xi(\mathbb{X})$ (cf. e.g. § 4 in Ferguson (1973) or Lemma 1 in Tsilevich and Vershik (1999)). We note that the distribution of a Dirichlet process is determined by its parameter measure (cf. Proposition 13.2 in Last and Penrose (2017)).

Construction with the Poisson–Dirichlet distribution

Closely related to the construction with a Gamma random measure is the already mentioned construction from § 4 in Ferguson (1973).

Let $\Gamma_1 > \Gamma_2 > \ldots$ be the points of a Poisson process $\widetilde{\eta}$ on $(0,\infty)$ whose intensity measure has the density $r \mapsto \rho(\mathbb{X})r^{-1}e^{-r}$ with respect to the Lebesgue measure on $(0,\infty)$. (We can also view $\widetilde{\eta}$ as the restriction $\eta(\mathbb{X}\times\cdot)$ of the Poisson process from the last section). Since

$$\mathbb{E}\left[\sum_{i=1}^{\infty}\Gamma_i\right] = \mathbb{E}\left[\int_0^{\infty} x\,\widetilde{\eta}(\mathrm{d}x)\right] = \rho(\mathbb{X})\int_0^{\infty} x(x^{-1}e^{-x})\,\mathrm{d}x = \rho(\mathbb{X}),$$

the sum $\Gamma = \sum_{i=1}^{\infty}\Gamma_i$ is almost surely finite and we can set

$$P_i = \frac{\Gamma_i}{\Gamma}, \quad i \in \mathbb{N}.$$

Let $Y_1, Y_2, \ldots$ be independent random variables in $\mathbb{X}$ with distribution $\frac{\rho}{\rho(\mathbb{X})}$, independent of $\Gamma_1, \Gamma_2, \ldots$.

Theorem 2.1 (Theorem 2 in § 4 in Ferguson (1973)). *The random probability measure $\sum_{i=1}^{\infty} P_i\delta_{Y_i}$ is a Dirichlet process with parameter measure ρ.*

The law of the sequence $(P_i)_{i\in\mathbb{N}}$ on the set

$$\nabla_\infty = \left\{(a_n)_{n\in\mathbb{N}} : 1 \geq a_1 \geq a_2 \geq \ldots \geq 0, \sum_{n=1}^{\infty} a_n = 1\right\}$$

is called *Poisson–Dirichlet distribution with parameter 0 and* $\rho(\mathbb{X})$, abbreviated $\mathrm{PD}(0,\rho(\mathbb{X}))$. This distribution, sometimes also referred to as $\mathrm{PD}(\rho(\mathbb{X}))$, was introduced in Kingman (1975) and later generalised to the two-parameter setting in Pitman and Yor (1997). The component "Dirichlet" in the name Poisson–Dirichlet distribution is explained by the following box.

Relation between the one-parameter Poisson–Dirichlet distribution and the Dirichlet distribution (cf. Kingman (1975) and § 9.3 in Kingman (1993)))

Let $Q^{(n)} = (Q_1^{(n)}, \ldots, Q_n^{(n)})$ for $n \in \mathbb{N}$ be a Dirichlet-distributed vector with parameter $(\theta_1^{(n)}, \ldots, \theta_n^{(n)})$ where $\theta_1^{(n)}, \ldots, \theta_n^{(n)} > 0$. Furthermore, assume

$$\max\left(\theta_1^{(n)}, \ldots, \theta_n^{(n)}\right) \to 0 \qquad \text{and} \qquad \sum_{i=1}^{n} \theta_i^{(n)} \to \lambda, \quad n \to \infty,$$

for a constant $\lambda > 0$. Then, for any $k \in \mathbb{N}$, the decreasing order statistic $(Q_{(1)}^{(n)}, \ldots, Q_{(k)}^{(n)})$, i.e. $Q_{(1)}^{(n)} \geq \ldots \geq Q_{(k)}^{(n)}$ are the k largest values in $(Q_{(1)}^{(n)}, \ldots, Q_{(n)}^{(n)})$, converges as $n \to \infty$ to $(P_1, \ldots, P_k)$, where $(P_i)_{i\in\mathbb{N}}$ follows a Poisson–Dirichlet distribution with parameter λ.

We note that, according to § 9.6 in Kingman (1993), the Poisson–Dirichlet distribution is "rather less than user-friendly". However, in the context of population genetics, where this distribution arises as the stationary distribution of a studied Markov process (the infinitely-many neutral alleles diffusion model, cf. e.g. Ethier and Kurtz (1981) for a rigorous derivation of the model and a proof of stationarity in Theorem 4.3), Watterson (1976) establishes results on the distribution of the vector $(P_1, \dots, P_k)$, $k \in \mathbb{N}$, for a sequence $(P_i)_{i\in\mathbb{N}}$ with Poisson–Dirichlet distribution. A detailed treatment of the Poisson–Dirichlet distribution is given in Feng (2010).

Construction with a Polyá urn scheme

Urn models can be used to describe many phenomena in stochastic (cf. Mahmoud (2008) for a general overview and applications to informatics and population genetics). An urn scheme that can be used to describe Dirichlet processes was given by Blackwell and MacQueen (1973). It is now widely known as the Blackwell–MacQueen urn scheme in their honour. In their paper, Blackwell and MacQueen assumed that $(\mathbb{X}, \mathcal{X})$ is a complete, separable metric space with Borel-σ-field. This assumption can be weakened, as can be seen, for example, in Pitman (1996).

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables such that

$$\mathbb{P}(X_1 \in \cdot\,) = \frac{\rho(\cdot)}{\rho(\mathbb{X})} \qquad \text{and} \qquad \mathbb{P}(X_{n+1} \in \cdot\,| X_1, \dots, X_n) = \frac{(\rho + \delta_{X_1} + \ldots + \delta_{X_n})(\cdot)}{\rho(\mathbb{X}) + n}. \tag{2.2}$$

The sequence $(X_n)_{n\in\mathbb{N}}$ is called a *Pólya sequence with parameter* ρ.

Theorem 2.2 (Blackwell and MacQueen (1973), cf. Theorem 2 in Pitman (1996))**.** *Let* $(X_n)_{n\in\mathbb{N}}$ *be a sequence of random variables with distribution given by* (2.2)*. Moreover, let the random measures* ρ_n*,* $n \in \mathbb{N}$*, be constructed by*

$$\rho_n = \rho + \sum_{i=1}^{n} \delta_{X_i}, \quad n \in \mathbb{N}.$$

Then the measures $\frac{\rho_n}{\rho_n(\mathbb{X})}$ *converge in total variation norm for* $n \to \infty$ *almost surely to a discrete random measure* ζ *which is a Dirichlet process with parameter measure* ρ*. Furthermore, given* ζ*, the random variables* $X_1, X_2, \dots$ *are independent with distribution* ζ*.*

For every fixed $N \in \mathbb{N}$, the vector $(X_1, \dots, X_N)$ constructed as in (2.2) induces a random partition of the set $\{1, \dots, N\}$ by grouping together integers $i, j \in \{1, \dots, N\}$ for which $X_i = X_j$. That is, the partition corresponds to the equivalence classes of the relation $i \sim j$ if and only if $X_i = X_j$ for $i, j \in \{1, \dots, N\}$. If ρ is assumed to be diffuse, i.e., it has no atoms, the resulting sequence of partitions corresponds to a realisation of the so-called Chinese restaurant process.

Chinese restaurant process (cf. Chapter 3 in Pitman (2006) or (11.19) in Aldous (1985))

Let $\theta > 0$. Consider a restaurant with infinitely many tables, each with infinite seating capacity. The first customer enters and sits at the first table. The second customer either joins the first at the same table with probability $\frac{1}{\theta+1}$ or chooses to sit at a new table with probability $\frac{\theta}{\theta+1}$. More generally, suppose $n \in \mathbb{N}$ customers are currently seated at $k \in \{1, \dots, n\}$ tables, with $n_1, \dots, n_k$ customers at each table, where $\sum_{j=1}^{k} n_j = n$. Then, customer $n+1$ decides to sit at table $j \in \{1, \dots, k\}$ with probability $\frac{n_j}{\theta+n}$ or opens a new table with probability $\frac{\theta}{\theta+n}$.
This procedure yields an exchangeable random partition of $\mathbb{N}$, that is, a sequence $(\Pi_n)_{n\in\mathbb{N}}$ of random partitions satisfying two properties: First, for each $n \in \mathbb{N}$, the partition Π_n is a random partition of $\{1, \dots, n\}$ whose distribution is invariant under permutations of $\{1, \dots, n\}$. Second, the sequence $(\Pi_n)_{n\in\mathbb{N}}$ is consistent in the sense that, for $n > 1$, the partition Π_{n-1} is almost surely obtained from Π_n by removing the element n.

Construction with Stick-breaking

In general, stick-breaking is a method for constructing a discrete random probability measure by distributing the total mass of one to different points. The idea is as follows. We identify the mass 1 with a stick of length 1. In the first step, we break the stick at the random variable V_1 taking values in $[0, 1]$. The amount V_1 is then assigned to the first point. Next, the remainder of the stick with length $1 - V_1$ is broken into two pieces with lengths relative to V_2 and $1 - V_2$ for the realisation of a second random variable V_2 in $[0, 1]$. We allocate the mass $(1-V_1)V_2$ to the second point and are left with a stick of length $(1-V_1)(1-V_2)$. By iteratively introducing random variables $V_3,\dots,V_n$ with values in $[0,1]$, the mass of the n-th point $(n\in\mathbb{N})$ is $V_n\prod_{k=1}^{n-1}(1-V_k)$. In the statistics literature, a stick-breaking construction for Dirichlet processes was introduced by Sethuraman (1994). However, in the context of population genetics, it was already discovered in McCloskey (1965).

Theorem 2.3 (Theorem 3.4 in Sethuraman (1994))**.** *Let $Z_1,Z_2,\dots$ and $V_1,V_2,\dots$ be independent random variables such that Z_n is distributed according to $\frac{\rho}{\rho(\mathbb{X})}$ and V_n follows a* $\mathrm{Beta}(1,\rho(\mathbb{X}))$*-distribution, $n\in\mathbb{N}$. Then*

$$\zeta=\sum_{n=1}^{\infty}V_n\prod_{k=1}^{n-1}(1-V_k)\delta_{Z_n}$$

is a Dirichlet process with parameter measure ρ.

To prove the theorem, Sethuraman (1994) uses a distributional equation for Dirichlet processes that is interesting in its own right. Let $V\sim\mathrm{Beta}(1,\rho(\mathbb{X}))$ and $X\sim\frac{\rho}{\rho(\mathbb{X})}$ be independent. We consider the equation

$$P\overset{d}{=}V\delta_X+(1-V)P \tag{2.3}$$

where $\overset{d}{=}$ denotes equality in distribution and P is a random measure independent of (V,X). It can be shown (cf. proof of Theorem 3.4 in Sethuraman (1994)) that a Dirichlet process with parameter measure ρ is the unique solution to this equation.

Let

$$W_1=V_1 \qquad\text{and}\qquad W_n=V_n\prod_{k=1}^{n-1}(1-V_k),\quad n\in\mathbb{N},n>1,$$

for the random variables $V_1,V_2,\dots$ from the preceding Theorem. Then the distribution of the sequence $(W_n)_{n\in\mathbb{N}}$ in the set

$$\Delta_\infty=\left\{(a_n)_{n\in\mathbb{N}}:0\le a_1,a_2,\dots\le 1,\sum_{n=1}^{\infty}a_n=1\right\}$$

is called $\mathrm{GEM}(\rho(\mathbb{X}))$*-distribution* (cf. p. 321 in Ewens (2004)). The distribution is named after Griffiths (1980), Engen (1975) and McCloskey (1965), who introduced it in the context of population genetics and studied its properties.

One may inquire about the connection between the stick-breaking representation of a Dirichlet process and its representation via the Poisson–Dirichlet distribution. The relationship between a sequence $(W_n)_{n\in\mathbb{N}}$ following a GEM-distribution and a sequence $(P_i)_{i\in\mathbb{N}}$ distributed according to the Poisson–Dirichlet law is addressed by Patil and Taillie (1977), who show that $(W_n)_{n\in\mathbb{N}}$ is a sized-bias permutation of $(P_i)_{i\in\mathbb{N}}$. A random sequence $(\widetilde{P}_i)_{i\in\mathbb{N}}$ is called a size-biased permutation of $P=(P_i)_{i\in\mathbb{N}}$ if, conditionally given P, the first element $\widetilde{P}_1$ equals P_i with probability P_i, $i\in\mathbb{N}$, and, given P, the element $\widetilde{P}_2$ is equal to P_j, $j\neq i$, with probability $\frac{P_j}{1-P_i}$. This process continues iteratively, selecting each subsequent element with probability proportional to its weight among the remaining unchosen components.

If the parameter measure ρ is diffuse, the locations of the weights in the above theorem (and in the construction with a Poisson–Dirichlet distribution) are almost surely distinct. In this case, both constructions yield almost surely the same process (cf. Corollary 9 and Corollary 10 in Pitman (1996)). By sorting the weights from the stick-breaking construction in decreasing order, one obtains a sequence that follows the Poisson–Dirichlet distribution. The associated locations are reordered accordingly, maintaining the correspondence between each weight and its location. Conversely, starting from a Poisson–Dirichlet distributed sequence of weights, one can recover the GEM-representation by applying a size-biased permutation to the weights and reordering the associated locations in the same manner.

By modifying the parameters in the Beta-distribution that determine the weights in the stick-breaking representation, we obtain the two-parameter Poisson–Dirichlet distribution and a generalisation of Dirichlet processes: Pitman–Yor processes.

Pitman–Yor processes (cf. Pitman and Yor (1997))

Let $0 \leq \alpha < 1$ and $\theta > -\alpha$. Let $V_1, V_2, \ldots$ be independent random variables with $V_n \sim \text{Beta}(1-\alpha, \theta + n\alpha)$, $n \in \mathbb{N}$. We define $\widetilde{Y}_1 = V_1$ and $\widetilde{Y}_n = \prod_{i=1}^{n-1}(1-V_i)V_n$ for $n \in \mathbb{N}$, $n > 1$. Further, we denote the decreasing order statistics of $(\widetilde{Y}_n)_{n\in\mathbb{N}}$ by $(Y_n)_{n\in\mathbb{N}}$, i.e. $Y_1 \geq Y_2 \geq \ldots$. Then the law of the sequence $(Y_n)_{n\in\mathbb{N}}$ is called two-parameter Poisson–Dirichlet distribution with parameter α and θ or $\text{PD}(\alpha,\theta)$-distribution for short. Moreover, let $X_1, X_2, \ldots$ be independent and identically distributed random variables in $\mathbb{X}$ such that the sequences $(X_n)_{n\in\mathbb{N}}$ and $(V_n)_{n\in\mathbb{N}}$ are independent. Let ν be the distribution of X_n, $n \in \mathbb{N}$. The random measure

$$\sum_{n=1}^{\infty} Y_n \delta_{X_n}$$

is called Pitman–Yor process with discount parameter α, strength parameter θ and base distribution ν. We note that the case $\alpha = 0$ corresponds to a Dirichlet process with parameter measure $\theta\nu$.

2.2. Mecke-type equation

A deeper structural understanding of random measures can be gained through integral identities. In the context of Poisson processes, a fundamental result in this regard is the Mecke equation, originally established by Mecke (1967). Let $(\mathbb{X}, \mathcal{X})$ be a measurable space with a s-finite measure λ, i.e. λ is a countable sum of finite measures and denote by $\mathbf{N}(\mathbb{X})$ the set of all measures on $\mathbb{X}$ that can be written as a countable sum of finite counting measures, i.e. measures that take only values in $\mathbb{N}_0$. Equipped with the smallest σ-field such that all mappings $\mathbf{N}(\mathbb{X}) \ni \mu \mapsto \mu(B)$, $B \in \mathcal{X}$, become measurable, this space becomes a measurable space. The Mecke equation (cf. e.g. Theorem 4.1 in Last and Penrose (2017)) states that a point process, i.e. a measurable mapping $\eta \colon \Omega \to \mathbf{N}(\mathbb{X})$ is a Poisson process with intensity measure λ if and only if

$$\mathbb{E}\left[\int_{\mathbb{X}} f(\eta, x)\, \eta(\mathrm{d}x)\right] = \int_{\mathbb{X}} \mathbb{E}\left[f(\eta + \delta_x, x)\right]\, \lambda(\mathrm{d}x)$$

for all measurable functions $f \colon \mathbf{N}(\mathbb{X}) \times \mathbb{X} \to [0, \infty]$.

This identity for Poisson processes has inspired analogous identities for other random measures, commonly referred to as Mecke-type equations. In the case of Dirichlet processes, the fixed point equation (2.3) from Sethuraman (1994) shows that a random probability measure ζ on a measurable space $(\mathbb{X}, \mathcal{X})$ which satisfies

$$\mathbb{E}\left[f(\zeta)\right] = \mathbb{E}\left[\int_{\mathbb{X}} \int_0^1 f((1-u)\zeta + u\delta_x)\, \text{Beta}(1,\theta)(\mathrm{d}u)\, \nu(\mathrm{d}x)\right]$$

for all measurable $f \colon \mathbf{M}(\mathbb{X}) \to [0, \infty)$ as well as a constant $\theta > 0$ and a probability measure ν is a Dirichlet process. Here, $\mathbf{M}(\mathbb{X})$ denotes the set of all s-finite measures on $\mathbb{X}$, which we equip with the smallest σ-field that makes the mappings $\mathbf{M}(\mathbb{X}) \ni \mu \mapsto \mu(B)$, $B \in \mathcal{X}$, measurable. A characterisation in terms of a Mecke-type equation in the case of a diffuse parameter measure is proven in Dello Schiavo and Lytvynov (2023) using a Mecke-type equation for a Gamma random measure (derived in Lemma 2.2 in Dello Schiavo and Lytvynov (2023) from the Mecke equation for Poisson processes).

Theorem 2.4 (Theorem 1.1 in Dello Schiavo and Lytvynov (2023))**.** *Let ρ be a finite diffuse measure on $(\mathbb{X}, \mathcal{X})$ and $\theta = \rho(\mathbb{X})$. A random measure ζ on $\mathbb{X}$ is a Dirichlet process with parameter measure ρ if and only if*

$$\mathbb{E}\left[\zeta(\mathbb{X}) f(\zeta)\right] = \mathbb{E}\left[\int_{\mathbb{X}} \int_0^1 f((1-u)\zeta + u\delta_x)(1-u)^{\theta-1}\, \mathrm{d}u\, \rho(\mathrm{d}x)\right]$$

for every measurable function $f \colon \mathbf{M}(\mathbb{X}) \to [0, \infty)$.

Moreover, if ζ is a Dirichlet process with parameter measure ρ, for each measurable function $f \colon \mathbf{M}_1(\mathbb{X}) \times \mathbb{X} \to [0, \infty)$, where $\mathbf{M}_1(\mathbb{X}) \subseteq \mathbf{M}(\mathbb{X})$ denotes the measurable subset of all probability measures, it holds

$$\mathbb{E}\left[\int_{\mathbb{X}} f(\zeta, x)\, \zeta(\mathrm{d}x)\right] = \mathbb{E}\left[\int_{\mathbb{X}} \int_0^1 f((1-u)\zeta + u\delta_x, x)(1-u)^{\theta-1}\, \mathrm{d}u\, \rho(\mathrm{d}x)\right]. \tag{2.4}$$

In the initial version of their work, published as the arXiv preprint arXiv:1706.07602 in 2017, Dello Schiavo and Lytvynov assumed the underlying space $\mathbb{X}$ to be a locally compact Polish space. Subsequently, Last (2020) achieved a more general formulation. In particular, Last (2020) considers an arbitrary measurable space $(\mathbb{X}, \mathcal{X})$ and does not assume the measure to be diffuse. Last (2020) refers to a probability measure ν on $\mathbb{X}$ as *good* if there exists a measurable set $B \subseteq \mathbb{X}$ such that $\nu(B) \in (0,1) \setminus \{\frac{1}{2}\}$. For instance, if $\nu = \frac{1}{2}(\delta_x + \delta_y)$ with distinct points $x, y \in \mathbb{X}$, then the measure ν is not good.

Theorem 2.5 (Theorem 1.1 and Theorem 2.1 in Last (2020))**.** *Let ν be a good probability measure on $\mathbb{X}$ and let G be a probability measure on $[0,1]$ with first moment $b_1 = \int_0^1 t\, G(\mathrm{d}t) \in (0,1)$. Suppose that a random measure ζ on $\mathbb{X}$ satisfies the identity*

$$\mathbb{E}\left[\int_{\mathbb{X}} f(\zeta, x)\, \zeta(\mathrm{d}x)\right] = \mathbb{E}\left[\int_{\mathbb{X}} \int_0^1 f((1-t)\zeta + t\delta_x, x)\, G(\mathrm{d}t)\, \nu(\mathrm{d}x)\right] \tag{2.5}$$

for all measurable $f\colon \mathbf{M}(\mathbb{X}) \times \mathbb{X} \to [0,\infty)$. Then ζ is a Dirichlet process with parameter measure $(1-b_1)b_1^{-1}\nu$. Moreover, in this case, G is the Beta*-distribution* $\mathrm{Beta}(1, (1-b_1)b_1^{-1})$*. Conversely, if ζ is a Dirichlet process with parameter measure ρ, then* (2.5) *holds for all measurable $f\colon \mathbf{M}(\mathbb{X}) \times \mathbb{X} \to [0,\infty)$ with $\nu = \rho(\mathbb{X})^{-1}\rho$ and $G = \mathrm{Beta}(1, \rho(\mathbb{X}))$.*

Comparing the Mecke-type identity for Dirichlet processes with the Mecke equation for Poisson processes reveals a key difference in the way new points on the right-hand side of the equation are incorporated. Whereas in the Poisson case, a new point is added to the process, in the Dirichlet case, the process is perturbed by forming a convex combination of the process and a new point. The weights are governed by a Beta-distribution, ensuring that the resulting measure remains a probability measure.

2.3. Survey of related literature

In the following section, we summarise relevant concepts from the literature. In particular, we recall moment formulas, the chaos decomposition from Peccati (2008), the integration by parts formula from Flint and Torrisi (2023) and the Fleming–Viot process with parent-independent mutation.

As an initial application, the Mecke-type equation from the previous section provides a means to compute the moments of integrals with respect to a Dirichlet process. While Corollary 3.5 in Dello Schiavo and Lytvynov (2023) states a recursive formula directly derived from the Mecke-type equation, we instead cite the non-recursive formulation given in Lemma 2.2 of Ethier (1990).

Proposition 2.6 (Lemma 2.2 in Ethier (1990), cf. Corollary 3.5 in Dello Schiavo and Lytvynov (2023))**.** *Let ζ be a Dirichlet process with parameter measure ρ and set $\theta = \rho(\mathbb{X})$. Suppose $n \in \mathbb{N}$ and let $f_1, \ldots, f_n\colon \mathbb{X} \to \mathbb{R}$ be measurable and bounded functions. It then holds*

$$\mathbb{E}\left[\prod_{i=1}^n \left(\int_{\mathbb{X}} f_i(x)\, \zeta(\mathrm{d}x)\right)\right] = \sum_{d=1}^n \sum_{\beta \in \pi(n,d)} c(\beta) \prod_{k=1}^d \left(\int_{\mathbb{X}} \prod_{i \in \beta_k} f_i(x)\, \rho(\mathrm{d}x)\right)$$

where $\pi(n,d)$ denotes the set of partitions of $\{1, \ldots, n\}$ into d non-empty subsets $\beta_1, \ldots, \beta_d$ and

$$c(\beta) = \frac{(|\beta_1| - 1)! \cdots (|\beta_d| - 1)!}{\theta(\theta+1)\cdots(\theta+n-1)}.$$

As this result already suggests, general expressions for such moments can become intricate and combinatorial in nature. Another formula for the moments of products of integrals, similar to the one considered in Proposition 2.6, can be found in Corollary 3.5 of Dello Schiavo and Quattrocchi (2023). Their expression is derived from moments of the Dirichlet distribution, which they deduced through a refined combinatorial analysis of exponent patterns. However, the explicit formulation involves further notation and combinatorial concepts such as multinomial coefficients for q-coloured set partitions, defined by Dello Schiavo and Quattrocchi (2023) in Definition 2.1, Definition 2.2 and Definition 2.3. For this reason, a detailed exposition is omitted here and the interested reader is referred to Dello Schiavo and Quattrocchi (2023) for a comprehensive treatment. Further formulas of this type involving cycle index polynomials of the symmetric group are used in Dello Schiavo (2019).

A chaos expansion for square-integrable random variables of a Dirichlet process is proven in Peccati (2008). We recall Theorem 1 from Peccati (2008), which states the decomposition.

Theorem 2.7 (Theorem 1 in Peccati (2008)). *Let ζ be a Dirichlet process with parameter measure ρ on the measurable space $(\mathbb{X}, \mathcal{X})$ with $\rho(\mathbb{X}) = \theta$. Each $F\colon \mathbf{M}(\mathbb{X}) \to \mathbb{R}$ that satisfies $\mathbb{E}[F(\zeta)^2] < \infty$ admits a unique representation of the type*

$$F(\zeta) = \mathbb{E}[F(\zeta)] + \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} h_n(x)\, \zeta^n(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.},$$

where the convergence is in $L^2(\mathbb{P})$ and, for every $n \in \mathbb{N}$, the function $h_n\colon \mathbb{X}^n \to \mathbb{R}$ is symmetric and satisfies

$$\mathbb{E}\left[h_n(X_1, \dots, X_n)^2\right] < \infty \qquad \text{and} \qquad \mathbb{E}\left(h_n(X_1, \dots, X_n) \,|\, X_1, \dots, X_{n-1}\right) = 0, \quad \mathbb{P}\text{-a.s.}$$

for a Pólya sequence $(X_n)_{n\in\mathbb{N}}$ with parameter ρ as defined in (2.2).

A characterisation of the kernel functions $h_n\colon \mathbb{X}^n \to \mathbb{R}$, $n \in \mathbb{N}$, appearing in the chaos expansion is provided in Theorem 2 of Peccati (2008).

Theorem 2.8 (Theorem 2 and Lemma 1 in Peccati (2008)). *If a function F admits a decomposition as specified in the previous Theorem, then, for each $n \in \mathbb{N}$, the function h_n fulfils*

$$h_n(x_1, \dots, x_n) = \sum_{k=1}^{n} \theta^{(n,k)} \sum_{1\le j_1 \le \dots \le j_k \le n} \mathbb{E}\left(F(\zeta) - \mathbb{E}[F(\zeta)] \,|\, X_1 = x_{j_1}, \dots, X_k = x_{j_k}\right)$$

for $(x_1, \dots, x_n)$ in a subset of $\mathbb{X}^n$ that has full measure under the probability measure induced by the vector $(X_1, \dots, X_n)$ from the Pólya sequence defined in (2.2). *The coefficients $\theta^{(n,k)}$ are for $n \in \mathbb{N}$ and $k \in \{1, \dots, n\}$ defined as*

$$\theta^{(n,k)} = \lim_{N\to\infty} \binom{N}{n} \binom{N-k}{n-k}^{-1} \theta_N^{(n,k)},$$

where the terms $\theta_N^{(n,k)}$ solve the recursion

$$\begin{cases} \theta_N^{(n,n)} = \Psi_N(n,n,n)^{-1}, \\ \sum_{i=q}^{n} \sum_{j=q}^{i} \theta_N^{(i,j)} \Psi_N(q,n,j) = 0, \quad q \in \{1, \dots, n-1\}, \end{cases}$$

and

$$\theta_N^{(N,k)} = -\sum_{s=k}^{N-1} \theta_N^{(s,k)}, \quad k \in \{1, \dots, N-1\}, \qquad \text{as well as} \qquad \theta_N^{(N,N)} = 1$$

with

$$\Psi_N(q,n,m) = \sum_{r=0}^{q} \binom{q}{r} \binom{N-n}{m-r} \mathbb{1}_{\{N-n\ge m-r\}} \frac{(m-r)!}{(m-q)!} \frac{\prod_{s=1}^{m-q} [\theta + q + s - 1]}{\prod_{s=1}^{m-r} [\theta + n + s - 1]}, \quad 1 \le q \le m \le n \le N.$$

An integration-by-parts formula involving what Dello Schiavo and Lytvynov (2023) refer to as "the discrete gradient" is established in Flint and Torrisi (2023). We recall their result. Let $\mathbf{P}(\mathbb{X})$ denote the set of all discrete probability measures on a locally compact Polish space $\mathbb{X}$ with Borel-σ-field $\mathcal{X}$. For a measurable function $F\colon \mathbf{P}(\mathbb{X}) \to \mathbb{R}$, Flint and Torrisi (2023) define for $x \in \mathbb{X}$ and $t \in [0,1]$ in equation (1) a gradient $D_{(x,t)}F$ of F by

$$D_{(x,t)}F(\mu) = F((1-t)\mu + t\delta_x) - F(\mu), \quad \mu \in \mathbf{P}(\mathbb{X}).$$

As noted by Flint and Torrisi on p. 704, their notion of a gradient corresponds to the difference operator in the Poisson setting (cf. chapter 18.1 in Last and Penrose (2017)), including an additional re-balancing mechanism to ensure that the perturbed measure remains a probability measure. Moreover, let ρ be a diffuse measure on $\mathbb{X}$ such that $\theta = \rho(\mathbb{X}) \in (0, \infty)$ and let the probability measure $\widehat{\rho}$ on $\mathbb{X} \times [0,1]$ be given by

$$\widehat{\rho}(A) = \int_0^1 \int_{\mathbb{X}} \mathbb{1}_A(x,t)\, \rho(\mathrm{d}x)\, (1-t)^{\theta-1}\, \mathrm{d}t$$

for all measurable $A \subseteq \mathbb{X} \times [0,1]$. Theorem 1 in Flint and Torrisi (2023) establishes the adjoint of D on $L^2(\mathbb{X} \times [0,1], \widehat{\rho})$.

Theorem 2.9 (Theorem 1 in Flint and Torrisi (2023))**.** *Let* $p, q \in [1, \infty]$ *with* $\frac{1}{p} + \frac{1}{q} = 1$ *and let* $F\colon \mathbf{P}(\mathbb{X}) \to \mathbb{R}$ *be a measurable function satisfying* $\mathbb{E}[F(\zeta)^p] < \infty$*. Furthermore, assume that* $h\colon \mathbf{P}(\mathbb{X}) \times \mathbb{X} \times [0,1] \to \mathbb{R}$ *is a measurable function satisfying*

$$\mathbb{E}\left[\int_{\mathbb{X}\times[0,1]} |h(\zeta, x, t)|^q \,\widehat{\rho}(\mathrm{d}(x,t))\right] < \infty.$$

Then there exists a random variable $\delta(h) \in L^1(\mathbb{P})$ *such that*

$$\mathbb{E}\left[\int_{\mathbb{X}\times[0,1]} h(\zeta, x, t) D_{(x,t)} F(\zeta)\,\widehat{\rho}(\mathrm{d}(x,t))\right] = \mathbb{E}\left[F(\zeta)\delta(h)\right].$$

In population genetics, the distribution of a Dirichlet process arises as the stationary distribution of the Fleming—Viot process with parent-independent mutation. We recall this process. Let $\mathbb{X}$ be a compact metric space with Borel σ-field $\mathcal{X}$ and denote by $\mathbf{M}_1(\mathbb{X})$ the set of all probability measures on $\mathbb{X}$. Furthermore, let $\theta > 0$ and $\nu_0 \in \mathbf{M}_1(\mathbb{X})$ with support $\mathbb{X}$. We set $\rho = \theta\nu_0$. A Fleming—Viot process with parent-independent mutation, a measure-valued stochastic process $(X_t)_{t\geq 0}$ with almost surely continuous paths, is the solution to the martingale problem $(\mathcal{F}C^\infty, L_\rho)$ with

$$\mathcal{F}C^\infty := \Big\{F\colon \mathbf{M}_1(\mathbb{X}) \to \mathbb{R} :\ F(\mu) = \varphi\left(\int_{\mathbb{X}} f_1(y)\,\mu(\mathrm{d}y), \dots, \int_{\mathbb{X}} f_d(y)\,\mu(\mathrm{d}y)\right), \\ \varphi \in C^\infty(\mathbb{R}^d),\ f_i \in C(\mathbb{X}), i \in \{1, \dots, d\}, d \in \mathbb{N}\Big\}$$

and

$$(L_\rho F)(\mu) := \frac{1}{2}\sum_{i,j=1}^d (\partial_i\partial_j\varphi)\left(\int_{\mathbb{X}} f_1(y)\,\mu(\mathrm{d}y), \dots, \int_{\mathbb{X}} f_d(y)\,\mu(\mathrm{d}y)\right)\mathrm{Cov}_\mu(f_i, f_j) \\ + \sum_{i=1}^d (\partial_i\varphi)\left(\int_{\mathbb{X}} f_1(y)\,\mu(\mathrm{d}y), \dots, \int_{\mathbb{X}} f_d(y)\,\mu(\mathrm{d}y)\right)\int_{\mathbb{X}} (Af_i)(x)\,\mu(\mathrm{d}x), \quad \mu \in \mathbf{M}_1(\mathbb{X}), \tag{2.6}$$

where the *mutation operator* A is for $f \in C(\mathbb{X})$ given by

$$Af(x) := \frac{1}{2}\int_{\mathbb{X}} f(y) - f(x)\,\rho(\mathrm{d}y), \quad x \in \mathbb{X}. \tag{2.7}$$

That means, for every $F \in \mathcal{F}C^\infty$, the process $(M_t^F)_{t\geq 0}$ defined by

$$M_t^F := F(X_t) - F(X_0) - \int_0^t (L_\rho F)(X_s)\,\mathrm{d}s, \quad t \geq 0,$$

is a martingale with respect to the canonical filtration $(\mathcal{F}_t)_{t\geq 0}$ with $\mathcal{F}_t = \sigma(X_s\colon s \leq t)$, $t \geq 0$ (cf. (3.12), (3.14) as well as (8.1) in Ethier and Kurtz (1993) for the description of the problem and Theorem 3.2 of Ethier and Kurtz (1993) for the well-posedness of the martingale problem and the uniqueness of the solution). According to Theorem 8.1 and Theorem 8.2 of Ethier and Kurtz (1993) (with proofs provided in the references cited therein), the distribution of a Dirichlet process with parameter measure ρ is the unique stationary reversible distribution of this Markov process.

MOTIVATION FOR THE FLEMING—VIOT PROCESS

This section provides a brief motivation for the Fleming–Viot process. It is not required for the present thesis and can be skipped. It is included for the benefit of readers who appreciate a more explicit conceptual bridge.

Population genetics seeks to understand how genetic variation is shaped and maintained over often long time scales, possibly ranging from decades to thousands of years. To achieve this, mathematical models that describe genetic diversity under specific assumptions can play an important role. These models not only enable the analysis of the mechanisms driving evolutionary change, but also provide a framework for investigating the impact of the underlying assumptions themselves. One of the simplest and best-known models is the Wright–Fisher model, named after Ronald A. Fisher and Sewall Wright (cf. Fisher (1930); Wright (1931)). A short description of the model can be found in the box below.

Neutral Wright–Fisher model (cf. Definition 2.1 in Etheridge (2011))

Consider a population of size $N \in \mathbb{N}$ evolving in discrete generations. For each $k \in \mathbb{N}_0$, generation $k+1$ is formed by randomly sampling N genes with replacement from generation k. That is, each gene in the next generation independently selects its parent at random from the individuals present in the previous generation.

Biologically, the Wright–Fisher model assumes non-overlapping generations (as, for example, in annual plants) of a fixed size and identical environmental conditions for all individuals. The population is considered neutral, meaning that all variants of a gene have equal reproductive success. Most species are either haploid, carrying only a single copy of each chromosome (e.g. bacteria), or diploid, carrying two copies of each chromosome (e.g. humans). The Wright–Fisher model assumes a haploid population, so each individual has exactly one parent. In diploid populations, although individuals have two parents, each gene can be traced back to a single parental copy. Therefore, it is common to model a diploid population of size N as a haploid population of size $2N$.

In its simplest form, the two-type Wright–Fisher model without mutation and without selection assumes the existence of two alleles (variants of a gene), denoted by A and a. The number of A-alleles in generation k forms a Markov chain with transition probabilities given by

$$p_{ij} = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(\frac{N-i}{N}\right)^{N-j}, \quad i,j \in \{0,\ldots,N\}.$$

Here, p_{ij} denotes the probability of transitioning from i copies of allele A in generation k to j copies in generation $k+1$, $i,j \in \{0,\ldots,N\}$.

For further properties and extensions, the interested reader is referred to the vast literature on this subject, e.g. Etheridge (2011); Ewens (2004); Hofrichter, Jost and Tran (2017).

The neutral K-type Wright–Fisher model with mutation considers $K \in \mathbb{N}$ distinct allelic types $A_1,\ldots,A_K$ within a population of constant size $N \in \mathbb{N}$. At each time step $n \in \mathbb{N}_0$, the state of the population is represented by a vector $Z^N(n) \in \{0,\ldots,N\}^K$, where the ith component $Z_i^N(n)$ denotes the number of genes of type A_i for $i \in \{1,\ldots,K\}$ and $Z_1^N(n) + \ldots + Z_K^N(n) = N$. Mutation is incorporated by assuming that allele A_i mutates to allele A_j at a rate $u_{ij} \in [0,1]$ for all $i,j \in \{1,\ldots,K\}$. The transition probabilities of the resulting Markov chain $(Z^N(n))_{n\in\mathbb{N}_0}$ are then given by a multinomial distribution. If the mutation probabilities satisfy

$$u_{ij} = \min\left\{\frac{\mu_{ij}}{N}, \frac{1}{K}\right\}, \quad i,j \in \{1,\ldots,K\},$$

for constants $\mu_{ij} \geq 0$ with $\mu_{ii} = 0$, $i,j \in \{1,\ldots,K\}$, it can be shown (cf. Theorem 1.1 in Chapter 10 of Ethier and Kurtz (2005)) that the rescaled process $(X^N(t))_{t\geq 0}$ with $X^N(t) := \frac{Z_1^N([Nt])}{N}$, $t \geq 0$, satisfies

$$\lim_{N\to\infty} \sup_{0\leq t\leq t_0} \sup_{x\in S} \left|\mathbb{E}\left(f\left(X_1^N(t),\ldots,X_{K-1}^N(t)\right)\middle|\left(X_1^N(0),\ldots,X_{K-1}^N(0)\right)=x\right)\right.$$
$$\left. -\mathbb{E}\left(f(X_1(t),\ldots,X_{K-1}(t))|(X_1(0),\ldots,X_{K-1}(0))=x\right)\right| = 0$$

for all $f \in C(\Delta_K)$ and $t_0 \geq 0$, where

$$S := \left\{\frac{1}{N}x \,:\, x \in \mathbb{N}_0^{K-1}, \sum_{i=1}^{K-1} x_i \leq N\right\} \quad \text{as well as} \quad \tilde{\Delta}_K := \left\{x \in [0,1]^{K-1} \,:\, \sum_{i=1}^{K-1} x_i \leq 1\right\}$$

and $(X(t))_{t\geq 0}$ is the K-type Wright–Fisher diffusion process. The Wright–Fisher diffusion process is a Markov process with sample paths in the space of continuous functions from $[0,\infty)$ into the simplex $\Delta_K = \{x \in [0,1]^K \,:\, x_1 + \ldots + x_K = 1\}$. It is characterised by its generator $L\colon C^2(\tilde{\Delta}_K) \to C(\tilde{\Delta}_K)$, given by

$$(L(f))(x) := \frac{1}{2}\sum_{i,j=1}^{K-1} x_i(\mathbb{1}_{\{i=j\}} - x_j)\left(\frac{\partial^2}{\partial_i\partial_j}f\right)(x) + \sum_{i=1}^{K-1}\left(\sum_{j=1}^{K-1}(\mu_{ji}x_j - \mu_{ij}x_i)\right)\left(\frac{\partial}{\partial_i}f\right)(x), \quad x \in \tilde{\Delta}_K.$$

For general positive mutation rates between distinct alleles, the Wright–Fisher diffusion process has a stationary distribution, but its closed form remains unknown and is an area of research (cf. e.g. Burden and Griffiths (2019), where the stationary distribution in the case of small mutation rates, which occur, for example, in Drosophila, is obtained). In parent-independent mutation, i.e. the probability of mutation depends solely on the target allele, irrespective of the ancestral allele, the stationary reversible distribution is a Dirichlet distribution.

From a biological perspective, mutations can give rise to novel alleles that have not previously occurred in the population. Modelling this process requires an extension to an infinite type space. Under suitable conditions on the mutation rates, the generator L can be extended when the type space $\mathbb{X}$ is countably infinite. A classical example is the stepwise mutation model from Ohta and Kimura (1973), where the set of possible alleles is the set of integers and allele $i \in \mathbb{Z}$ may mutate to $i-1$ or $i+1$. However, for uncountably infinite type spaces, a different approach is required. The approach introduced in Fleming and Viot (1979) is to consider a complete, separable metric space $\mathbb{X}$ as the set of possible types and to replace the simplex Δ_K used in the previous analysis with the space $\mathcal{P}(\mathbb{X})$ of all Borel probability measures on $\mathbb{X}$. The generator becomes

$$(L(\varphi))(\mu) = \frac{1}{2}\int_{\mathbb{X}}\int_{\mathbb{X}}\left(\frac{\partial^2}{\partial\mu(x)\partial\mu(y)}\varphi\right)(\mu)\,(\delta_x(\mathrm{d}y) - \mu(\mathrm{d}y))\,\mu(\mathrm{d}x) + \int_{\mathbb{X}}\left(A\left(\frac{\partial}{\partial\mu}\varphi\right)\right)(x)\,\mu(\mathrm{d}x), \quad \mu \in \mathcal{P}(\mathbb{X}),$$

where φ is an element of a suitable function space, $\frac{\partial}{\partial\mu(x)}\varphi(\mu) = \lim_{\varepsilon\to 0}\varepsilon^{-1}(\varphi(\mu+\varepsilon\delta_x) - \varphi(\mu))$ and A is an operator representing mutation. (It is also possible to include selection and recombination (cf. (3.12) in Ethier and Kurtz (1993).) Here, we followed the presentation in the article Ethier and Kurtz (1993). This article addresses general measure-valued diffusion processes arising in population genetics, named Fleming–Viot processes after Fleming and Viot. Ethier and Kurtz establish the existence of these processes under general assumptions and analyse their properties. Under suitable conditions, they show that Fleming–Viot processes approximate Wright–Fisher diffusions. We refer the interested reader to their comprehensive treatment or to Dawson (1993) for further details on measure-valued Markov processes.

2.4. General notation

In this section, we introduce notation used throughout this thesis.

We denote by $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{N}_0$ and $\mathbb{N}$ the real line, the set of integers, the set of nonnegative integers and the set of positive integers, respectively. Given $a, b \in \mathbb{R}$, we define $a \wedge b := \min\{a, b\}$. If $x \in \mathbb{R}$ and $n \in \mathbb{N}$, we denote by $x^{(n)}$ for $n \in \mathbb{N}$ and $x \in \mathbb{R}$ the rising factorial

$$x^{(n)} := x(x+1)\dots(x+n-1) = \prod_{l=1}^{n}(x+n-l).$$

Moreover, let $x^{(0)} := 1$. For an arbitrary set A and $r \in \mathbb{N}$, we define

$$[r]_0 := \{0,\dots,r\}, \quad [r] := [r]_0 \setminus \{0\} \quad \text{and} \quad A^{[r]} := \{(i_1,\dots,i_r) \in A^r\colon\ i_k \neq i_l \text{ if } k \neq l, k, l \in \{1,\dots,r\}\}$$

if A contains at least r elements. Otherwise, let $A^{[r]} := \emptyset$, the empty set. By $\mathbb{1}_{\{\cdot\}}$, we mean the indicator function.

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space. The Dirac measure at a point $x \in \mathbb{X}$ is denoted by δ_x. Given $n \in \mathbb{N}$ and $(x_1,\dots,x_n) \in \mathbb{X}^n$, we define $\mathbf{x_n} := (x_1,\dots,x_n) \in \mathbb{X}^n$ as well as $\delta_{\mathbf{x}_n} := \delta_{x_1} + \ldots + \delta_{x_n}$ and interpret $\mathbf{x}_0$ and $\delta_{\mathbf{x}_0}$ as void. Moreover, we sometimes write $\nu(f)$ instead of $\int_{\mathbb{X}} f(x)\,\nu(\mathrm{d}x)$ for a measure ν on $\mathbb{X}$ and an integrable function f. To enhance readability for those primarily interested in the statements, we refrain from using this notation within the statements themselves. Let $C_b(\mathbb{X})$ be the space of all bounded and continuous functions $f\colon \mathbb{X} \to \mathbb{R}$ and denote by $C^\infty(\mathbb{X})$ the space of all smooth functions $f\colon \mathbb{X} \to \mathbb{R}$. We write $\mathbf{M}(\mathbb{X})$ for the set of all finite measures on $\mathbb{X}$ equipped with the smallest σ-field that makes the mappings $\mathbf{M}(\mathbb{X}) \ni \mu \mapsto \mu(B)$, $B \in \mathcal{X}$, measurable and denote by $\mathbf{M}_1(\mathbb{X})$ the set of all probability measures on $\mathbb{X}$. Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, we denote by ζ a Dirichlet process, i.e. a random measure (a measurable mapping $\zeta \colon \Omega \to \mathbf{M}(\mathbb{X})$) satisfying (1.1). We refer to Kallenberg (2017) for a comprehensive treatment of random measures.

CHAPTER 3

A multivariate perspective on the Mecke-type equation

In this chapter, we begin by defining measures that will play a central role in the subsequent analysis. Properties of these measures, which are used in the development of a multivariate formulation of the Mecke-type equation, are established. Finally, a formula for evaluating certain integrals with respect to these measures, which will be used in later chapters, is obtained.

Throughout the chapter, let $(\mathbb{X}, \mathcal{X})$ be a measurable space.

3.1. A multivariate Mecke-type equation

In this section, we derive a multivariate version of the Mecke-type equation. We begin by introducing the measures involved.

Definition 3.1. Given a measure μ on $(\mathbb{X}, \mathcal{X})$ and $n \in \mathbb{N}$, let the measure $\mu^{[n]}$ on $\mathbb{X}^n$ be defined by

$$\mu^{[n]}(B) \coloneqq \int_{\mathbb{X}} \int_{\mathbb{X}} \dots \int_{\mathbb{X}} \mathbb{1}_B(x_1, \dots, x_n)\, (\mu + \delta_{x_1} + \dots + \delta_{x_{n-1}})(\mathrm{d}x_n) \dots (\mu + \delta_{x_1})(\mathrm{d}x_2)\, \mu(\mathrm{d}x_1), \quad B \in \mathcal{X}^{\otimes n}.$$

Furthermore, set $\mu^{[0]}(c) \coloneqq c$ for $c \in \mathbb{R}$.

We observe that in the special case $n = 1$, we have $\mu^{[1]} = \mu$. We illustrate this definition with an example.

Example 3.2. Let μ be a measure on $(\mathbb{X}, \mathcal{X})$ and $n \in \mathbb{N}$. Given pairwise disjoint $B_1, \dots, B_n \in \mathcal{X}$, we have

$$\mu^{[n]}(B_1 \times \dots \times B_n) = \prod_{i=1}^{n} \mu(B_i)$$

and

$$\mu^{[n]}(B_1^n) = \mu(B_1)(\mu(B_1) + 1) \dots (\mu(B_1) + n - 1). \qquad \circ$$

We recall the definitions $[n]_0 \coloneqq \{0, \dots, n\}$ and $[n] \coloneqq \{1, \dots, n\}$ for $n \in \mathbb{N}$. We begin with a lemma that characterises integrability with respect to $\mu^{[n]}$ for a σ-finite measure μ on $\mathbb{X}$ and $n \in \mathbb{N}$.

Lemma 3.3. *Let μ be a σ-finite measure on $\mathbb{X}$. Let $p \in [1, \infty)$ and $n \in \mathbb{N}$. A function $f \colon \mathbb{X}^n \to \mathbb{R}$ is an element of $L^p(\mu^{[n]})$ if and only if f is measurable and for all $m \in [n-1]_0$ and all permutations π of $[n]$ it holds*

$$\int_{\mathbb{X}^{n-m}} \int_{\mathbb{X}^m} |f(x_{\pi(1)}, \dots, x_{\pi(n)})|^p \prod_{k=1}^{m} \delta_{x_n}(\mathrm{d}x_k)\, \mu^{n-m}(\mathrm{d}(x_{m+1}, \dots, x_n)) < \infty.$$

Proof. It suffices to consider $p = 1$ since $f \in L^p$ if and only if $|f|^p \in L^1$. We proceed by induction over $n \in \mathbb{N}$. For $n = 1$, the claim follows immediately from the fact that $f \in L^1(\mu)$ if and only if f is measurable and $\int_{\mathbb{X}} |f(x)|\, \mu(\mathrm{d}x) < \infty$. Assume that the assertion holds for some $n \in \mathbb{N}$. A function $f \colon \mathbb{X}^{n+1} \to \mathbb{R}$ is an element of $L^1(\mu^{[n+1]})$ if and only if f is measurable and $\int_{\mathbb{X}^{n+1}} |f(x)|\, \mu^{[n+1]}(\mathrm{d}x) < \infty$. By definition of $\mu^{[n+1]}$ and $\mu^{[n]}$, we have

$$\int_{\mathbb{X}^{n+1}} |f(x)|\, \mu^{[n+1]}(\mathrm{d}x) = \int_{\mathbb{X}^n} \int_{\mathbb{X}} |f(\mathbf{x_n}, x_{n+1})|\, (\mu + \delta_{\mathbf{x_n}})(\mathrm{d}x_{n+1})\, \mu^{[n]}(\mathrm{d}\mathbf{x_n}).$$

Thus, $\int_{\mathbb{X}^{n+1}} |f(x)|\, \mu^{[n+1]}(\mathrm{d}x)$ is finite if and only if $g \in L^1(\mu^{[n]})$, where

$$g(x_1, \dots, x_n) \coloneqq \int_{\mathbb{X}} |f(x_1, \dots, x_n, y)|\, (\mu + \delta_{x_1} + \dots + \delta_{x_n})(\mathrm{d}y).$$

According to the induction hypothesis, this holds if and only if g is measurable and for all $\overline{m} \in [n-1]_0$ and all permutations $\overline{\pi}$ of $[n]$ it holds

$$\int_{\mathbb{X}^{n-\overline{m}}} \int_{\mathbb{X}^{\overline{m}}} |g(x_{\overline{\pi}(1)}, \dots, x_{\overline{\pi}(n)})| \prod_{k=1}^{\overline{m}} \delta_{x_n}(\mathrm{d}x_k)\, \mu^{n-\overline{m}}(\mathrm{d}(x_{\overline{m}+1}, \dots, x_n)) < \infty.$$

By the definition of g, the integral in the previous line is

$$\begin{aligned}
&\int_{\mathbb{X}^{n-\overline{m}}} \int_{\mathbb{X}^{\overline{m}}} \int_{\mathbb{X}} |f(x_{\overline{\pi}(1)}, \dots, x_{\overline{\pi}(n)}, y)| \, (\mu + \delta_{x_1} + \ldots + \delta_{x_n})(\mathrm{d}y) \prod_{k=1}^{\overline{m}} \delta_{x_n}(\mathrm{d}x_k)\, \mu^{n-\overline{m}}(\mathrm{d}(x_{\overline{m}+1}, \dots, x_n)) \\
&= \int_{\mathbb{X}^{n-\overline{m}}} \int_{\mathbb{X}^{\overline{m}}} \int_{\mathbb{X}} |f(x_{\overline{\pi}(1)}, \dots, x_{\overline{\pi}(n)}, y)| \, \mu(\mathrm{d}y) \prod_{k=1}^{\overline{m}} \delta_{x_n}(\mathrm{d}x_k)\, \mu^{n-\overline{m}}(\mathrm{d}(x_{\overline{m}+1}, \dots, x_n)) \\
&+ \sum_{j=1}^{n} \int_{\mathbb{X}^{n-\overline{m}}} \int_{\mathbb{X}^{\overline{m}}} |f(x_{\overline{\pi}(1)}, \dots, x_{\overline{\pi}(n)}, x_j)| \prod_{k=1}^{\overline{m}} \delta_{x_n}(\mathrm{d}x_k)\, \mu^{n-\overline{m}}(\mathrm{d}(x_{\overline{m}+1}, \dots, x_n)) =: I_{1,\overline{m}} + I_{2,\overline{m}}.
\end{aligned}$$

From this, we obtain the assertion (The cases $m = 0$ and $m = n$ of the assertion for f follow from $I_{1,0}$ and $I_{2,n-1}$, respectively. The case $m \in [n-1]$ follows from $I_{2,m-1}$ and $I_{1,m}$.). □

As the following example shows, the space $L^2(\mu^{[n]})$ may be smaller than $L^2(\mu^n)$.

Example 3.4. On $\mathbb{X} := [0, 1]$ we consider the restriction μ of the Lebesgue measure to $[0, 1]$ and $f : \mathbb{X}^2 \to \mathbb{R}$ defined by $f(x, y) := \mathbb{1}_{(0,1)^2}(x, y) \frac{1}{\sqrt[3]{xy}}$. Because of

$$\int_0^1 \int_0^1 f(x, y)^2 \,\mathrm{d}x \,\mathrm{d}y = \int_0^1 \int_0^1 \left(\frac{1}{\sqrt[3]{xy}} \right)^2 \mathrm{d}x \,\mathrm{d}y = \left(\int_0^1 x^{-\frac{2}{3}} \,\mathrm{d}x \right)^2 < \infty$$

we have $f \in L^2(\mu^2)$. On the other hand, the integral

$$\begin{aligned}
&\int_0^1 \int_0^1 f(x, y)^2 \mu^{[2]}(\mathrm{d}(x, y)) = \int_0^1 \int_0^1 f(x, y)^2 (\mu + \delta_x)(\mathrm{d}y) \mu(\mathrm{d}x) \\
&= \int_0^1 \int_0^1 f(x, y)^2 \,\mathrm{d}y \,\mathrm{d}x + \int_0^1 f(x, x)^2 \,\mathrm{d}x = \int_0^1 \int_0^1 f(x, y)^2 \,\mathrm{d}y \,\mathrm{d}x + \int_0^1 x^{-\frac{4}{3}} \,\mathrm{d}x
\end{aligned}$$

is not finite. ∘

We continue with a simple yet useful lemma that collects properties of the measures.

Lemma 3.5. *Let μ be a σ-finite measure on $\mathbb{X}$. Then the measure $\mu^{[n]}$ is symmetric for every $n \in \mathbb{N}$ and, for every $m, n \in \mathbb{N}$ and $B \in \mathcal{X}^{\otimes(n+m)}$, it holds*

$$\mu^{[n+m]}(B) = \int_{\mathbb{X}^m} \int_{\mathbb{X}^n} \mathbb{1}_B(x_1, \dots, x_{n+m}) \, (\mu + \delta_{x_1} + \ldots + \delta_{x_m})^{[n]}(\mathrm{d}(x_{m+1}, \dots, x_{m+n})) \; \mu^{[m]}(\mathrm{d}(x_1, \dots, x_m)). \tag{3.1}$$

Proof. We first establish the recursion formula (3.1), which will be a key ingredient in showing the symmetry of the measure. We use induction with respect to $n \in \mathbb{N}$. Let $A \in \mathcal{X}^{\otimes(1+m)}$ and $B \in \mathcal{X}^{\otimes(n+1+m)}$. By definition, we have

$$\begin{aligned}
\mu^{[1+m]}(A) &= \int_{\mathbb{X}} \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{1}_A(\mathbf{x_{m+1}}) \, (\mu + \delta_{\mathbf{x_m}})(\mathrm{d}x_{m+1}) \, (\mu + \delta_{\mathbf{x_{m-1}}})(\mathrm{d}x_m) \ldots (\mu + \delta_{x_1})(\mathrm{d}x_2) \mu(\mathrm{d}x_1) \\
&= \int_{\mathbb{X}^m} \int_{\mathbb{X}} \mathbb{1}_A(\mathbf{x_{m+1}}) \, (\mu + \delta_{\mathbf{x_m}})^{[1]}(\mathrm{d}x_{m+1}) \, \mu^{[m]}(\mathrm{d}\mathbf{x_{m+1}}),
\end{aligned}$$

the base case of the induction. Furthermore, it holds

$$\mu^{[n+1+m]}(B) = \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} \mathbb{1}_B(\mathbf{x_{m+n+1}}) \, (\mu + \delta_{\mathbf{x_{n+1+m}}})(\mathrm{d}x_{n+1+m}) \ldots (\mu + \delta_{x_1})(\mathrm{d}x_2) \, \mu(\mathrm{d}x_1)$$

$$
\begin{aligned}
&= \int_{\mathbb{X}^{n+m}} \int_{\mathbb{X}} \mathbb{1}_B(\mathbf{x_{m+n+1}})\,(\mu + \delta_{\mathbf{x_{n+m}}})(\mathrm{d}x_{n+1+m})\,\mu^{[n+m]}(\mathrm{d}\mathbf{x_{n+m}})) \\
&= \int_{\mathbb{X}^m} \int_{\mathbb{X}^n} \int_{\mathbb{X}} \mathbb{1}_B(\mathbf{x_{m+n+1}})\,(\mu + \delta_{\mathbf{x_{n+m}}})(\mathrm{d}x_{n+1+m})\,(\mu + \delta_{\mathbf{x_m}})^{[n]}(\mathrm{d}(x_{m+1}, \dots, x_{m+n}))\,\mu^{[m]}(\mathrm{d}\mathbf{x_m}) \\
&= \int_{\mathbb{X}^m} \int_{\mathbb{X}^{n+1}} \mathbb{1}_B(\mathbf{x_{m+n+1}})\,(\mu + \delta_{\mathbf{x_m}})^{[n+1]}(\mathrm{d}(x_{m+1}, \dots, x_{n+1+m}))\,\mu^{[m]}(\mathrm{d}\mathbf{x_m}),
\end{aligned}
$$

where the penultimate equality is a consequence of the induction hypothesis.

To establish the symmetry, we show the validity of

$$
\int_{\mathbb{X}^n} f(x_1, \dots, x_n)\,\mu^{[n]}(\mathrm{d}(x_1, \dots, x_n)) = \int_{\mathbb{X}^n} f(x_{\pi(1)}, \dots, x_{\pi(n)})\,\mu^{[n]}(\mathrm{d}(x_1, \dots, x_n))
$$

for all $n \in \mathbb{N}$, permutations π of $[n]$ and measurable $f\colon \mathbb{X}^n \to [0, \infty)$ by induction on n. If $n = 2$ and $f\colon \mathbb{X}^2 \to [0, \infty)$ is measurable, the definition of $\mu^{[2]}$ yields

$$
\begin{aligned}
&\int_{\mathbb{X}^2} f(x_1, x_2)\,\mu^{[2]}(\mathrm{d}(x_1, x_2)) = \int_{\mathbb{X}} \int_{\mathbb{X}} f(x_1, x_2)\,(\mu + \delta_{x_1})(\mathrm{d}x_2)\,\mu(\mathrm{d}x_1) \\
&= \int_{\mathbb{X}^2} f(x_1, x_2)\,\mu^2(\mathrm{d}(x_1, x_2)) + \int_{\mathbb{X}} f(x_1, x_1)\,\mu(\mathrm{d}x_1) = \int_{\mathbb{X}^2} f(x_2, x_1)\,\mu^{[2]}(\mathrm{d}(x_1, x_2)).
\end{aligned}
$$

For the induction step, let $n \in \mathbb{N}$ and assume the symmetry of $\mu^{[k]}$ for all natural numbers $k < n$. Let $f\colon \mathbb{X}^n \to [0, \infty)$ be measurable and π be a permutation of $[n]$. Define $l := \pi^{-1}(n)$ and assume $l > 1$ (if this is not the case, take $l = \pi^{-1}(n-1)$ and modify the following calculations as necessary). We decompose π as $\pi = \sigma \circ (l\ n)$ with $\sigma = \pi \circ (l\ n)$. Here, $(l\ n)$ denotes the transposition that exchanges the numbers l and n. It holds $\sigma(n) = n$ and thus σ permutes the numbers $1, \dots, n-1$. The recursion from (3.1) and the induction hypothesis yield

$$
\begin{aligned}
&\int_{\mathbb{X}^n} f(\mathbf{x_n})\,\mu^{[n]}(\mathrm{d}\mathbf{x_n}) \\
&= \int_{\mathbb{X}^{l-1}} \int_{\mathbb{X}^{n-l+1}} f(\mathbf{x_n})\,(\mu + \delta_{\mathbf{x_{l-1}}})^{[n-l+1]}(\mathrm{d}(x_l, \dots, x_n))\,\mu^{[l-1]}(\mathrm{d}\mathbf{x_{l-1}}) \\
&= \int_{\mathbb{X}^{l-1}} \int_{\mathbb{X}^{n-l+1}} f(\mathbf{x_{l-1}}, x_n, x_{l+1}, \dots, x_{n-1}, x_l)\,(\mu + \delta_{\mathbf{x_{l-1}}})^{[n-l+1]}(\mathrm{d}(x_l, \dots, x_n))\,\mu^{[l-1]}(\mathrm{d}\mathbf{x_{l-1}}) \\
&= \int_{\mathbb{X}^n} f(\mathbf{x_{l-1}}, x_n, x_{l+1}, \dots, x_{n-1}, x_l)\,\mu^{[n]}(\mathrm{d}\mathbf{x_n}) \\
&= \int_{\mathbb{X}^{n-1}} \int_{\mathbb{X}} f(\mathbf{x_{l-1}}, x_n, x_{l+1}, \dots, x_{n-1}, x_l)\,(\mu + \delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\,\mu^{[n-1]}(\mathrm{d}\mathbf{x_{n-1}}). \\
&= \int_{\mathbb{X}^{n-1}} F(\mathbf{x_{n-1}})\,\mu^{[n-1]}(\mathrm{d}\mathbf{x_{n-1}}),
\end{aligned}
$$

where $F : \mathbb{X}^{n-1} \to [0, \infty)$ is defined by

$$
F(x_1, \dots, x_{n-1}) := \int_{\mathbb{X}} f(x_1, \dots, x_{l-1}, x_n, x_{l+1}, \dots, x_{n-1}, x_l)\,(\mu + \delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n).
$$

Let $\hat{\sigma}$ be the permutation of $[n-1]$ that satisfies $\hat{\sigma}(i) = \sigma(i)$ for every $i \in [n-1]$. The induction hypothesis gives

$$
\begin{aligned}
&\int_{\mathbb{X}^{n-1}} F(\mathbf{x_{n-1}})\,\mu^{[n-1]}(\mathrm{d}\mathbf{x_{n-1}}) = \int_{\mathbb{X}^{n-1}} F(x_{\hat{\sigma}(1)}, \dots, x_{\hat{\sigma}(n-1)})\,\mu^{[n-1]}(\mathrm{d}\mathbf{x_{n-1}}) \\
&= \int_{\mathbb{X}^{n-1}} \int_{\mathbb{X}} f(x_{\sigma(1)}, \dots, x_{\sigma(l-1)}, x_n, x_{\sigma(l+1)}, \dots, x_{\sigma(n-1)}, x_{\sigma(l)})\,(\mu + \delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\,\mu^{[n-1]}(\mathrm{d}\mathbf{x_{n-1}}) \\
&= \int_{\mathbb{X}^n} f(x_{\pi(1)}, \dots, x_{\pi(n)})\,\mu^{[n]}(\mathrm{d}\mathbf{x_n}).
\end{aligned}
$$

□

The following superposition lemma (cf. Lemma 2.1 in Ethier and Griffiths (1993) for the weighted sum of two independent Dirichlet processes) is a key ingredient in the multivariate version of the Mecke-type equation.

Lemma 3.6. *Let $n \in \mathbb{N}$ and ζ_j be independent Dirichlet processes with parameter measures ρ_j, $j \in [n]$. Further, let Z be independent of $(\zeta_1, \dots, \zeta_n)$ and follow a $\mathrm{Dir}(\rho_1(\mathbb{X}), \dots, \rho_n(\mathbb{X}))$ distribution. Then*

$$\sum_{j=1}^{n} Z_j \zeta_j \tag{3.2}$$

is a Dirichlet process with parameter measure $\rho = \sum_{j=1}^{n} \rho_j$.

Proof. We make use of the construction of Dirichlet processes with Poisson processes (cf. (2.1) in Section 2.1) and the superposition theorem for independent Poisson processes (cf. Theorem 3.3 in Last and Penrose (2017)).

Let $\rho := \sum_{j=1}^{n} \rho_j$. Furthermore, let η_j, $j \in [n]$, be independent Poisson processes on $\mathbb{X} \times (0, \infty)$ with intensity measures $\mathbb{E}\eta_j(\mathrm{d}(x, r)) = r^{-1} e^{-r}\, \mathrm{d}r\, \rho_j(\mathrm{d}x)$. Let $j \in [n]$. We define a random measure ξ_j on $\mathbb{X}$ by

$$\xi_j(B) := \int_{\mathbb{X} \times (0,\infty)} \mathbb{1}_B(x)\, \eta_j(\mathrm{d}(x, r)), \quad B \in \mathcal{X},$$

and set

$$\zeta_j^* := \frac{\xi_j}{\xi_j(\mathbb{X})}.$$

The random variable $\xi_j(\mathbb{X})$ follows a Gamma distribution with shape parameter $\rho_j(\mathbb{X})$ and scale parameter 1. Moreover, ζ_j^* is a Dirichlet process independent of $\xi_j(\mathbb{X})$ with parameter measure ρ_j and is thus equal in distribution to ζ_j. By the independence of the underlying Poisson processes, $(\xi_1(\mathbb{X}), \zeta_1^*), \dots, (\xi_n(\mathbb{X}), \zeta_n^*)$ are independent. Since $\xi_j(\mathbb{X})$ and ζ_j are independent (cf. e.g. § 4 in Ferguson (1973) or Lemma 1 in Tsilevich and Vershik (1999)), $j \in [n]$, we obtain that $\xi_1(\mathbb{X}), \dots, \xi_n(\mathbb{X}), \zeta_1^*, \dots, \zeta_n^*$ are independent. Let

$$\zeta := \frac{\sum_{j=1}^{n} \xi_j}{\sum_{j=1}^{n} \xi_j(\mathbb{X})}.$$

From the superposition theorem for independent Poisson processes (cf. Theorem 3.3 in Last and Penrose (2017)) it follows that ζ is a Dirichlet process with parameter measure ρ. Further, it holds

$$\zeta = \sum_{j=1}^{n} \frac{\xi_j}{\xi_j(\mathbb{X})} \frac{\xi_j(\mathbb{X})}{\sum_{j=1}^{n} \xi_j(\mathbb{X})} = \sum_{j=1}^{n} Z_j^* \zeta_j^*,$$

where

$$Z_j^* := \frac{\xi_j(\mathbb{X})}{\sum_{j=1}^{n} \xi_j(\mathbb{X})}, \quad j \in [n].$$

The random vector $Z^* := (Z_1^*, \dots, Z_n^*)$ has a Dirichlet distribution with parameter vector $(\rho_1(\mathbb{X}), \dots, \rho_n(\mathbb{X}))$. All in all, $Z^*, \zeta_1^*, \dots, \zeta_n^*$ are independent and $(Z^*, \zeta_1^*, \dots, \zeta_n^*)$ is in distribution equal to $(Z, \zeta_1, \dots, \zeta_n)$. We conclude

$$\zeta = \sum_{j=1}^{n} Z_j^* \zeta_j^* \overset{\mathrm{d}}{=} \sum_{j=1}^{n} Z_j \zeta_j. \qquad \square$$

The preceding lemma (or Sethuraman's fixed point equation (2.3)) shows that for a Dirichlet process ζ with parameter measure ρ, an independent $\mathrm{Beta}(1, \theta)$-distributed random variable Z and $x \in \mathbb{X}$, the random measure $(1 - Z)\zeta + Z\delta_x$ is a Dirichlet process with parameter measure $\rho + \delta_x$. This gives rise to the following different formulation of the Mecke-type equation from Dello Schiavo and Lytvynov (2023) and Last (2020) recalled in (2.4) and (2.5). Let ζ_ρ denote a Dirichlet process with parameter measure ρ. For a measurable function $f \colon \mathbf{M}(\mathbb{X}) \times \mathbb{X} \to [0, \infty)$, where $\mathbf{M}(\mathbb{X})$ denotes the set of finite measures on $\mathbb{X}$, equation (2.5) can be written as

$$\mathbb{E}\left[\int_{\mathbb{X}} f(\zeta_\rho, x)\, \zeta_\rho(\mathrm{d}x)\right] = \frac{1}{\rho(\mathbb{X})} \mathbb{E}\left[\int_{\mathbb{X}} f(\zeta_{\rho+\delta_x}, x)\, \rho(\mathrm{d}x)\right].$$

Using this point of view, it is possible to state a multivariate version of the Meck-type equation. From Section 2.4, we recall the notation $x^{(n)}$ for the rising factorial

$$x^{(n)} = x(x+1)\dots(x+n-1) = \prod_{l=1}^{n} (x + n - l), \quad n \in \mathbb{N}, x \in \mathbb{R}.$$

Theorem 3.7. *Let ζ be a Dirichlet process with parameter measure ρ, let $n \in \mathbb{N}$ and $f\colon \mathbf{M}(\mathbb{X}) \times \mathbb{X}^n \to [0,\infty)$ be a measurable function. Then*

$$\mathbb{E}\left[\int_{\mathbb{X}^n} f(\zeta_\rho, x_1, \dots, x_n)\, \zeta_\rho^n(\mathrm{d}(x_1,\dots,x_n))\right] = \frac{1}{\theta^{(n)}}\mathbb{E}\left[\int_{\mathbb{X}^n} f(\zeta_{\rho+\delta_{x_1}+\ldots+\delta_{x_n}}, x_1,\dots,x_n)\, \rho^{[n]}(\mathrm{d}(x_1,\dots,x_n))\right]. \quad (3.3)$$

Proof. We proceed by induction using the Mecke-type equation (2.4) from Dello Schiavo and Lytvynov (2023) as the base case. For the induction step, let $n \in \mathbb{N}$ and $f\colon \mathbf{M}(\mathbb{X}) \times \mathbb{X}^n \to [0,\infty)$ be measurable. We define $H\colon \mathbf{M}(\mathbb{X}) \times \mathbb{X} \to [0,\infty)$ by

$$H(\mu, x_1) := \int_{\mathbb{X}^n} f(\mu, x_1, x_2, \dots, x_{n+1})\, \mu^n(\mathrm{d}(x_2,\dots,x_{n+1})).$$

The Mecke-type equation from Dello Schiavo and Lytvynov (2023) (or the one from Last (2020)) yields

$$\begin{aligned}\mathbb{E}&\left[\int_{\mathbb{X}^{n+1}} f(\zeta_\rho, x_1,\dots,x_{n+1})\, \zeta_\rho^{n+1}(\mathrm{d}(x_1,\dots,x_{n+1}))\right] = \mathbb{E}\left[\int_{\mathbb{X}} H(\zeta_\rho, x_1)\, \zeta_\rho(\mathrm{d}x_1)\right] \\ &= \frac{1}{\theta}\mathbb{E}\left[\int_{\mathbb{X}} H(\zeta_{\rho+\delta_{x_1}}, x_1)\, \rho(\mathrm{d}x_1)\right] = \frac{1}{\theta}\mathbb{E}\left[\int_{\mathbb{X}}\int_{\mathbb{X}^n} f(\zeta_{\rho+\delta_{x_1}}, x_1,\dots,x_{n+1})\, \zeta^n_{\rho+\delta_{x_1}}(\mathrm{d}(x_2,\dots,x_{n+1}))\, \rho(\mathrm{d}x_1)\right].\end{aligned}$$

By the induction hypothesis and the recursion form Lemma 3.5, this is equal to

$$\begin{aligned}&\frac{1}{\theta}\mathbb{E}\left[\int_{\mathbb{X}} \frac{1}{(\theta+1)^{(n)}}\int_{\mathbb{X}^n} f(\zeta_{\rho+\delta_{x_1}+\ldots+\delta_{x_{n+1}}}, x_1,\dots,x_{n+1})\, (\rho+\delta_{x_1})^{[n]}(\mathrm{d}(x_2,\dots,x_{n+1}))\, \rho(\mathrm{d}x_1)\right] \\ &= \frac{1}{\theta^{(n+1)}}\mathbb{E}\left[\int_{\mathbb{X}^{n+1}} f(\zeta_{\rho+\delta_{x_1}+\ldots+\delta_{x_{n+1}}}, x_1,\dots,x_{n+1})\, \rho^{[n+1]}(\mathrm{d}(x_1,\dots,x_{n+1}))\right].\end{aligned}$$ □

A direct consequence of this multivariate Mecke-type equation is a formula for the moment measures of a Dirichlet process.

Corollary 3.8. *Let $n \in \mathbb{N}$. A Dirichlet process ζ with parameter measure ρ satisfies*

$$\mathbb{E}\left[\zeta^n(B)\right] = \frac{1}{\theta^{(n)}}\rho^{[n]}(B), \quad B \in \mathcal{X}^{\otimes n}. \quad (3.4)$$

3.2. Integral formula for subsequent applications

As we will later encounter integrals with respect to the sum of a measure and Dirac measures, we provide a formula for their evaluation.

We fix some notation. Let $m, r \in \mathbb{N}$ with $m \geq r$ and $(i_1,\dots,i_r) \in [m]^{[r]}$. Suppose $1 \leq j_1 < j_2 < \ldots < j_{m-r} \leq m$ are such that $\{j_1,\dots,j_{m-r}\} = [m] \setminus \{i_1,\dots,i_r\}$. Define the mapping $f_{i_1,\dots,i_r}\colon \mathbb{X}^m \to \mathbb{R}$ for a function $f\colon \mathbb{X}^m \to \mathbb{R}$ by

$$f_{i_1,\dots,i_r}(x_1,\dots,x_m) := f(\widetilde{x}_1,\dots,\widetilde{x}_m), \quad (3.5)$$

where

$$\widetilde{x}_{i_1} = x_1,\ \dots,\ \widetilde{x}_{i_r} = x_r, \quad \widetilde{x}_{j_1} = x_{r+1},\ \dots,\ \widetilde{x}_{j_{m-r}} = x_m.$$

That is, in order to compute $f_{i_1,\dots,i_r}(x_1,\dots,x_m)$ for $(x_1,\dots,x_m) \in \mathbb{X}^m$, the function f is evaluated at the point whose i_k-th coordinate is x_k for $k \in [r]$ and whose remaining coordinates are filled with $x_{r+1},\dots,x_m$. Finally, for $k \in \mathbb{N}$ and $x_1,\dots,x_k, z_1,\dots,z_{m-r} \in \mathbb{X}$ let

$$f^k_{i_1,\dots,i_r}(x_1,\dots,x_k,z_1,\dots,z_{m-r}) := \sum_{1\leq j_1\leq\ldots\leq j_r\leq k} f_{i_1,\dots,i_r}(x_{j_1},\dots,x_{j_r},z_1,\dots,z_{m-r}) \quad (3.6)$$

if $m > r$. If $m = r$, define

$$f^k_{i_1,\dots,i_r}(x_1,\dots,x_k) := \sum_{1\leq j_1\leq\ldots\leq j_r\leq k} f_{i_1,\dots,i_r}(x_{j_1},\dots,x_{j_r}). \quad (3.7)$$

We illustrate these concepts with an example.

Example 3.9. Let $f\colon \mathbb{X}^5 \to \mathbb{R}$. For $x_1, x_2, z_1, z_2 \in \mathbb{X}$, it then holds

$$\begin{aligned} f^2_{5,1,4}(x_1,x_2,z_1,z_2) &= \sum_{1\le j_1\le j_2\le j_3\le 2} f_{5,1,4}(x_{j_1},x_{j_2},x_{j_3},z_1,z_2) \\ &= f_{5,1,4}(x_1,x_1,x_1,z_1,z_2) + f_{5,1,4}(x_1,x_1,x_2,z_1,z_2) + f_{5,1,4}(x_1,x_2,x_2,z_1,z_2) + f_{5,1,4}(x_2,x_2,x_2,z_1,z_2) \\ &= f(x_1,z_1,z_2,x_1,x_1) + f(x_1,z_1,z_2,x_2,x_1) + f(x_2,z_1,z_2,x_2,x_1) + f(x_2,z_1,z_2,x_2,x_2). \end{aligned}$$

$\circ$

A special case arises when the considered function is given as a product of functions, each depending on a different variable. Specifically, let $m, r \in \mathbb{N}$ with $m \ge r$ and $f_1,\dots,f_m\colon \mathbb{X}\to\mathbb{R}$. The tensor product $(\otimes_{j=1}^m f_j)\colon \mathbb{X}^m \to \mathbb{R}$ is defined by

$$(\bigotimes_{j=1}^m f_j)(x_1,\dots,x_r) := \prod_{j=1}^m f_j(x_j). \tag{3.8}$$

For $(i_1,\dots,i_r)\in[m]^{[r]}$ define

$$f_{\otimes i_1,\dots,i_r} := \bigotimes_{j=1}^r f_{i_j} \qquad \text{and} \qquad f^{\otimes i_1,\dots,i_r} := \bigotimes_{j\notin\{i_1,\dots,i_r\}} f_j, \tag{3.9}$$

where $f^{\otimes i_1,\dots,i_m} := 1$. Given $k\in\mathbb{N}$ and $x_1,\dots,x_k\in\mathbb{X}$, let

$$f^k_{\otimes i_1,\dots,i_r}(x_1,\dots,x_k) := \sum_{1\le j_1\le\dots\le j_r\le k} f_{\otimes i_1,\dots,i_r}(x_{j_1},\dots,x_{j_r}). \tag{3.10}$$

We begin with a recursive formula for integrals with respect to the sum of a measure and a Dirac measure.

Lemma 3.10. *Let μ be a σ-finite measure on $\mathbb{X}$ and $x\in\mathbb{X}$. Then for each $m\in\mathbb{N}$ and each $g\colon\mathbb{X}^m\to\mathbb{R}$ that is integrable with respect to $(\mu+\delta_x)^{[m]}$, using the notation from* (3.5)*, it holds*

$$\begin{aligned} &\int_{\mathbb{X}^m} g(y_1,\dots,y_m)\,(\mu+\delta_x)^{[m]}(\mathrm{d}(y_1,\dots,y_m)) \\ &\qquad = \int_{\mathbb{X}^m} g(y)\,\mu^{[m]}(\mathrm{d}y) + \sum_{i=1}^m \int_{\mathbb{X}^{m-1}} g_i(x,y_1,\dots,y_{m-1})\,(\mu+\delta_x)^{[m-1]}(\mathrm{d}(y_1,\dots,y_{m-1})). \end{aligned}$$

Proof. By assumption, the integrals on the right-hand side are finite. We proceed by induction over $m\in\mathbb{N}$. The base case is

$$\int_{\mathbb{X}} g(y)\,(\mu+\delta_x)(\mathrm{d}y) = \int_{\mathbb{X}} g(y)\,\mu(\mathrm{d}y) + g(x)$$

for integrable functions $g\colon\mathbb{X}\to\mathbb{R}$. We now assume the validity of the assertion for all functions that are integrable with respect to $(\mu+\delta_x)^{[m]}$. Let $g\colon\mathbb{X}^{m+1}\to\mathbb{R}$ be integrable with respect to $(\mu+\delta_x)^{[m+1]}$. From the recursion (3.1) in Lemma 3.5 we obtain

$$\begin{aligned} \int_{\mathbb{X}^{m+1}} g(y)\,(\mu+\delta_x)^{[m+1]}(\mathrm{d}y) &= \int_{\mathbb{X}^m}\int_{\mathbb{X}} g(\mathbf{y_{m+1}})\,(\mu+\delta_x+\delta_{\mathbf{y_m}})(\mathrm{d}y_{m+1})\,(\mu+\delta_x)^{[m]}(\mathrm{d}\mathbf{y_m}) \\ &= \int_{\mathbb{X}^m} F(\mathbf{y_m})\,(\mu+\delta_x)^{[m]}(\mathrm{d}\mathbf{y_m}) \end{aligned}$$

with $F\colon\mathbb{X}^m\to\mathbb{R}$ defined by

$$F(\mathbf{y_m}) := \int_{\mathbb{X}} g(\mathbf{y_{m+1}})\,(\mu+\delta_x+\delta_{\mathbf{y_m}})(\mathrm{d}y_{m+1}) = \int_{\mathbb{X}} g(\mathbf{y_{m+1}})\,(\mu+\delta_{\mathbf{y_m}})(\mathrm{d}y_{m+1}) + g(\mathbf{y_m},x).$$

The induction hypothesis and the recursion (3.1) yield

$$\begin{aligned} &\int_{\mathbb{X}^{m+1}} g(y)\,(\mu+\delta_x)^{[m+1]}(\mathrm{d}y) = \int_{\mathbb{X}^m} F(\mathbf{y_m})\,\mu^{[m]}(\mathrm{d}\mathbf{y_m}) + \sum_{i=1}^m \int_{\mathbb{X}^{m-1}} F_i(x,\mathbf{y_{m-1}})\,(\mu+\delta_x)^{[m-1]}(\mathrm{d}\mathbf{y_{m-1}}) \\ &= \int_{\mathbb{X}^{m+1}} g(\mathbf{y_{m+1}})\,\mu^{[m+1]}(\mathrm{d}\mathbf{y_{m+1}}) + \int_{\mathbb{X}^m} g(\mathbf{y_m},x)\,\mu^{[m]}(\mathrm{d}\mathbf{y_m}) \end{aligned}$$

$$+\sum_{i=1}^{m}\int_{\mathbb{X}^m}\int_{\mathbb{X}} g_i(x,\mathbf{y_{m-1}},y_{m+1})\,(\mu+\delta_{\mathbf{y_{m-1}}}+\delta_x)(\mathrm{d}y_{m+1})\,(\mu+\delta_x)^{[m-1]}(\mathrm{d}\mathbf{y_{m-1}})$$
$$+\sum_{i=1}^{m}\int_{\mathbb{X}^{m-1}} g_i(x,\mathbf{y_{m-1}},x)\,(\mu+\delta_x)^{[m-1]}(\mathrm{d}\mathbf{y_{m-1}}),$$

which is equal to

$$\int_{\mathbb{X}^{m+1}} g(\mathbf{y_{m+1}})\,\mu^{[m+1]}(\mathrm{d}\mathbf{y_{m+1}})+\int_{\mathbb{X}^m} g_{m+1}(x,\mathbf{y_m})\,\mu^{[m]}(\mathrm{d}\mathbf{y_m})$$
$$+\sum_{i=1}^{m}\int_{\mathbb{X}^m} g_i(x,\mathbf{y_m})\,(\mu+\delta_x)^{[m]}(\mathrm{d}\mathbf{y_m})+\sum_{i=1}^{m}\int_{\mathbb{X}^{m-1}} g_{m+1,i}(x,x,\mathbf{y_{m-1}})\,(\mu+\delta_x)^{[m-1]}(\mathrm{d}\mathbf{y_{m-1}}).$$

An application of the induction hypothesis to the function $\mathbf{y_m}\mapsto g_{m+1}(x,\mathbf{y_m})$ shows

$$\int_{\mathbb{X}^m} g_{m+1}(x,\mathbf{y_m})\,(\mu+\delta_x)^{[m]}(\mathrm{d}\mathbf{y_m})=\int_{\mathbb{X}^m} g_{m+1}(x,\mathbf{y_m})\,\mu^{[m]}(\mathrm{d}\mathbf{y_m})$$
$$+\sum_{i=1}^{m}\int_{\mathbb{X}^{m-1}} g_{m+1,i}(x,x,\mathbf{y_{m-1}})\,(\mu+\delta_x)^{[m-1]}(\mathrm{d}\mathbf{y_{m-1}}).$$

Hence, we conclude

$$\int_{\mathbb{X}^{m+1}} g(y)\,(\mu+\delta_x)^{[m+1]}(\mathrm{d}y)=\int_{\mathbb{X}^{m+1}} g(\mathbf{y_{m+1}})\,\mu^{[m+1]}(\mathrm{d}\mathbf{y_{m+1}})+\sum_{i=1}^{m}\int_{\mathbb{X}^m} g_i(x,\mathbf{y_m})\,(\mu+\delta_x)^{[m]}(\mathrm{d}\mathbf{y_m})$$
$$+\int_{\mathbb{X}^m} g_{m+1}(x,\mathbf{y_m})\,(\mu+\delta_x)^{[m]}(\mathrm{d}\mathbf{y_m}).\ \square$$

The next proposition considers integrals with respect to the sum of a measure and several Dirac measures. We recall the notation introduced in (3.6) and (3.7).

Proposition 3.11. *Let μ be a σ-finite measure. Then for all $m,k\in\mathbb{N}$, $x_1,\dots,x_k\in\mathbb{X}$ and functions $f\colon\mathbb{X}^m\to\mathbb{R}$ that are integrable with respect to $(\mu+\delta_{x_1}+\ldots+\delta_{x_k})^{[m]}$ it holds*

$$\int_{\mathbb{X}^m} f(z)\,(\mu+\delta_{x_1}+\ldots+\delta_{x_k})^{[m]}(\mathrm{d}z) \tag{3.11}$$
$$=\int_{\mathbb{X}^m} f(z)\,\mu^{[m]}(\mathrm{d}z)+\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\int_{\mathbb{X}^{m-r}} f^k_{i_1,\dots,i_r}(x_1,\dots,x_k,z_1,\dots,z_{m-r})\,\mu^{[m-r]}(\mathrm{d}(z_1,\dots,z_{m-r})).$$

In (3.11), consistent with our notation $\mu^{[0]}(c)=c$, $c\in\mathbb{R}$, we interpret

$$\sum_{(i_1,\dots,i_m)\in[m]^{[m]}}\int_{\mathbb{X}^0} f^k_{i_1,\dots,i_m}(\mathbf{x_k},\mathbf{z_0})\,\mu^{[0]}(\mathrm{d}\mathbf{z_0})=\sum_{(i_1,\dots,i_m)\in[m]^{[m]}} f^k_{i_1,\dots,i_m}(\mathbf{x_k}).$$

Proof. By assumption, the right-hand side is finite. We prove the assertion by induction on m. If $m=1$, we have

$$(\mu+\delta_{\mathbf{x_k}})\,(f)=\mu\,(f)+\sum_{i=1}^{k} f(x_i)=\mu\,(f)+f^k_1(x_1,\dots,x_k)$$

by definition of f^k_1. For the induction step, we assume that the assertion holds for some $m\in\mathbb{N}$, all $k\in\mathbb{N}$ and all $x_1,\dots,x_k\in\mathbb{X}$ as well as all integrable $f\colon\mathbb{X}^m\to\mathbb{R}$. To establish the formula for $m+1$, we use induction on $k\in\mathbb{N}$. By an application of Lemma 3.10, for integrable $f\colon\mathbb{X}^{m+1}\to\mathbb{R}$, we find

$$(\mu+\delta_{x_1})^{[m+1]}(f)=\mu^{[m+1]}(f)+\sum_{i=1}^{m+1}\int_{\mathbb{X}^m} f_i(x_1,\mathbf{y_m})\,(\mu+\delta_{x_1})^{[m]}(\mathrm{d}\mathbf{y_m}).$$

Let $i\in[m+1]$. The induction hypothesis in the induction on m applied to the function $\mathbf{y_m}\mapsto f^{(i)}(\mathbf{y_m}):=f_i(x_1,\mathbf{y_m})$ shows that the second term on the right-hand side is

$$\sum_{i=1}^{m+1}\int_{\mathbb{X}^m} f^{(i)}(\mathbf{y_m})\,\mu^{[m]}(\mathrm{d}\mathbf{y_m})+\sum_{i=1}^{m+1}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\int_{\mathbb{X}^{m-r}} (f^{(i)})^1_{i_1,\dots,i_r}(x_1,\mathbf{z_{m-r}})\,\mu^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})$$

$$= \sum_{i=1}^{m+1} \int_{\mathbb{X}^m} f_i(x_1, \mathbf{y_m})\, \mu^{[m]}(\mathrm{d}\mathbf{y_m}) + \sum_{r=1}^{m} \sum_{(i_1,\dots,i_{r+1})\in[m+1]^{[r+1]}} \int_{\mathbb{X}^{m-r}} f^1_{i_1,\dots,i_{r+1}}(x_1, \mathbf{z_{m-r}})\, \mu^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})$$
$$= \sum_{r=1}^{m+1} \sum_{(i_1,\dots,i_r)\in[m+1]^{[r]}} \int_{\mathbb{X}^{m-r}} f^1_{i_1,\dots,i_r}(x_1, \mathbf{z_{m-r}})\, \mu^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}}),$$

which finishes the base case of the induction on k. Next, we assume that the assertion holds for some $k \in \mathbb{N}$. Lemma 3.10 for an integrable function $f\colon \mathbb{X}^{m+1} \to \mathbb{R}$ yields

$$(\mu + \delta_{\mathbf{x_{k+1}}})^{[m+1]}(f) = (\mu + \delta_{\mathbf{x_k}})^{[m+1]}(f) + \sum_{i=1}^{m+1} \int_{\mathbb{X}^m} f_i(x_{k+1}, \mathbf{y_m})\, (\mu + \delta_{\mathbf{x_{k+1}}})^{[m]}(\mathrm{d}\mathbf{y_m})$$
$$= (\mu + \delta_{\mathbf{x_k}})^{[m+1]}(f) + \sum_{i=1}^{m+1} \int_{\mathbb{X}^m} f^{(i)}(\mathbf{y_m})\, (\mu + \delta_{\mathbf{x_{k+1}}})^{[m]}(\mathrm{d}\mathbf{y_m}),$$

where $f^{(i)}(\mathbf{y_m}) := f_i(x_{k+1}, \mathbf{y_m})$ for $\mathbf{y_m} \in \mathbb{X}^m$ and $i \in [m+1]$. Using the induction hypothesis in the induction on k, we obtain

$$(\mu + \delta_{\mathbf{x_k}})^{[m+1]}(f) = \mu^{[m+1]}(f) + \sum_{r=1}^{m+1} \sum_{(i_1,\dots,i_r)\in[m+1]^{[r]}} \int_{\mathbb{X}^{m+1-r}} f^k_{i_1,\dots,i_r}(\mathbf{x_k}, \mathbf{z_{m+1-r}})\, \mu^{[m+1-r]}(\mathrm{d}\mathbf{z_{m+1-r}}).$$

Let $i \in [m+1]$. The induction hypothesis in the induction on m yields

$$\int_{\mathbb{X}^m} f^{(i)}(\mathbf{y_m})\, (\mu + \delta_{\mathbf{x_{k+1}}})^{[m]}(\mathrm{d}\mathbf{y_m})$$
$$= \mu^{[m]}(f^{(i)}) + \sum_{r=1}^{m} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} \int_{\mathbb{X}^{m-r}} (f^{(i)})^{k+1}_{i_1,\dots,i_r}(\mathbf{x_{k+1}}, \mathbf{z_{m-r}})\, \mu^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}}).$$

Combining these findings, we find

$$(\mu + \delta_{\mathbf{x_{k+1}}})^{[m+1]}(f)$$
$$= \mu^{[m+1]}(f) + \sum_{r=1}^{m+1} \sum_{(i_1,\dots,i_r)\in[m+1]^{[r]}} \int_{\mathbb{X}^{m+1-r}} f^k_{i_1,\dots,i_r}(\mathbf{x_k}, \mathbf{z_{m+1-r}})\, \mu^{[m+1-r]}(\mathrm{d}\mathbf{z_{m+1-r}}) + \sum_{i=1}^{m+1} \mu^{[m]}(f^{(i)})$$
$$+ \sum_{i=1}^{m+1} \sum_{r=1}^{m} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} \int_{\mathbb{X}^{m-r}} (f^{(i)})^{k+1}_{i_1,\dots,i_r}(\mathbf{x_{k+1}}, \mathbf{z_{m-r}})\, \mu^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}}).$$

By definition, for $r \in [m+1]$, $(i_1, \dots, i_r) \in [m+1]^{[r]}$ and $\mathbf{z_{m+1-r}} \in \mathbb{X}^{m+1-r}$, we have

$$f^{k+1}_{i_1,\dots,i_r}(\mathbf{x_{k+1}}, \mathbf{z_{m+1-r}}) = \sum_{1\le j_1\le\dots\le j_r\le k+1} f_{i_1,\dots,i_r}(x_{j_1}, \dots, x_{j_r}, \mathbf{z_{m+1-r}})$$
$$= \sum_{1\le j_1\le\dots\le j_r\le k} f_{i_1,\dots,i_r}(x_{j_1}, \dots, x_{j_r}, \mathbf{z_{m+1-r}}) + \sum_{1\le j_1\le\dots\le j_{r-1}\le k+1} f_{i_1,\dots,i_r}(x_{j_1}, \dots, x_{j_{r-1}}, x_{k+1}, \mathbf{z_{m+1-r}})$$
$$= f^k_{i_1,\dots,i_r}(\mathbf{x_k}, \mathbf{z_{m+1-r}}) + \sum_{1\le j_1\le\dots\le j_{r-1}\le k+1} f_{i_1,\dots,i_r}(x_{j_1}, \dots, x_{j_{r-1}}, x_{k+1}, \mathbf{z_{m+1-r}})$$

if $r \ge 2$. In the case $r = 1$, it holds

$$f^{k+1}_{i_1}(\mathbf{x_{k+1}}, \mathbf{z_m}) = \sum_{1\le j_1\le k+1} f_{i_1}(x_{j_1}, \mathbf{z_m}) = \sum_{1\le j_1\le k} f_{i_1}(x_{j_1}, \mathbf{z_m}) + f_{i_1}(x_{k+1}, \mathbf{z_m}) = f^k_{i_1}(\mathbf{x_k}, \mathbf{z_m}) + f^{(i_1)}(\mathbf{z_m}).$$

Thus, the claim is now a consequence of

$$\sum_{r=1}^{m+1} \sum_{(i_1,\dots,i_r)\in[m+1]^{[r]}} \int_{\mathbb{X}^{m+1-r}} f^{k+1}_{i_1,\dots,i_r}(\mathbf{x_{k+1}}, \mathbf{z_{m+1-r}})\, \mu^{[m+1-r]}(\mathrm{d}\mathbf{z_{m+1-r}})$$

$$= \sum_{r=1}^{m+1} \sum_{(i_1,\dots,i_r)\in[m+1]^{[r]}} \int_{\mathbb{X}^{m+1-r}} f^k_{i_1,\dots,i_r}(\mathbf{x_k}, \mathbf{z_{m+1-r}})\, \mu^{[m+1-r]}(\mathrm{d}\mathbf{z_{m+1-r}}) + \sum_{i_1=1}^{m+1} \int_{\mathbb{X}^m} f^{(i_1)}(\mathbf{z_m})\, \mu^{[m]}(\mathrm{d}\mathbf{z_m})$$
$$+ \sum_{r=2}^{m+1} \sum_{(i_1,\dots,i_r)\in[m+1]^{[r]}} \int_{\mathbb{X}^{m+1-r}} \sum_{1\le j_1\le\ldots\le j_{r-1}\le k+1} f_{i_1,\dots,i_r}(x_{j_1},\dots,x_{j_{r-1}}, x_{k+1}, \mathbf{z_{m+1-r}})\, \mu^{[m+1-r]}(\mathrm{d}\mathbf{z_{m+1-r}}).$$

and the observation that

$$\sum_{r=2}^{m+1} \sum_{(i_1,\dots,i_r)\in[m+1]^{[r]}} \int_{\mathbb{X}^{m+1-r}} \sum_{1\le j_1\le\ldots\le j_{r-1}\le k+1} f_{i_1,\dots,i_r}(x_{j_1},\dots,x_{j_{r-1}}, x_{k+1}, \mathbf{z_{m+1-r}})\, \mu^{[m+1-r]}(\mathrm{d}\mathbf{z_{m+1-r}})$$
$$= \sum_{r=1}^{m} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} \sum_{i=1}^{m+1} \int_{\mathbb{X}^{m-r}} \sum_{1\le j_1\le\ldots\le j_r\le k+1} (f^{(i)})_{i_1,\dots,i_r}(x_{j_1},\dots,x_{j_r}, \mathbf{z_{m-r}})\, \mu^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}}). \qquad \square$$

In the special case, when the considered function is given as a product, the proposition simplifies. We recall the definitions (3.9) and (3.10).

Corollary 3.12. *Let μ be a σ-finite measure on $\mathbb{X}$. Then*

$$\int_{\mathbb{X}^m} (\bigotimes_{i=1}^m f_i)(z)\, (\mu + \delta_{x_1} + \ldots + \delta_{x_k})^{[m]}(\mathrm{d}z)$$
$$= \int_{\mathbb{X}^m} (\bigotimes_{i=1}^m f_i)(z)\, \mu^{[m]}(\mathrm{d}z) + \sum_{r=1}^{m} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} f^k_{\otimes i_1,\dots,i_r}(x_1,\dots,x_k) \int_{\mathbb{X}^{m-r}} f^{\otimes i_1,\dots,i_r}(z)\, \mu^{[m-r]}(\mathrm{d}z)$$

holds true if $k, m \in \mathbb{N}$, $x_1,\dots,x_k \in \mathbb{X}$ and $f_1,\dots,f_m\colon \mathbb{X}\to\mathbb{R}$ are measurable functions such that $\otimes_{i=1}^m f_i$ is integrable with respect to $(\mu + \delta_{x_1} + \ldots + \delta_{x_k})^{[m]}$.

CHAPTER 4

Chaos expansion

Orthogonal decompositions constitute a fundamental tool with widespread applications across mathematics. In the context of stochastic analysis, the orthogonal decomposition of a square-integrable random variable into an infinite orthogonal sum is commonly referred to as the chaos expansion. This concept was initially introduced in the setting of the Wiener process by Wiener (1938). Subsequently, Itô (1951) proved that this decomposition can be represented in terms of iterated stochastic integrals. Over time, this methodology has found widespread applications and has been extended beyond the Gaussian setting (cf. e.g. Last and Penrose (2011) for Poisson processes).

As recalled in Chapter 2.3, Peccati (2008) establishes a chaos expansion for random variables that are square-integrable with respect to the distribution of a Dirichlet process. In this chapter, we give an alternative, constructive proof of the chaos expansion obtained by Peccati (2008), including explicit formulas for the projections. To this end, the first two sections are devoted to the definition of relevant function spaces and an analysis of their structural properties. The chaos expansion is then the subject of the final section. We consider a measurable space $(\mathbb{X}, \mathcal{X})$ with a finite measure $\rho \neq 0$ and set $\theta \coloneqq \rho(\mathbb{X})$. Furthermore, let ζ be a Dirichlet process with parameter measure ρ and define

$$L^2(\zeta) \coloneqq \{F\colon \mathbf{M}(\mathbb{X}) \to \mathbb{R} \mid F \text{ is measurable and } \mathbb{E}[F(\zeta)^2] < \infty\}.$$

4.1. The spaces $\mathbb{H}_n$

This section is dedicated to the study of the spaces $\mathbb{H}_n$, $n \in \mathbb{N}$, which consist of functions that will later serve as integrands in the projection onto the nth chaos. At first, we define these spaces and then study their properties.

Definition 4.1. Let $n \in \mathbb{N}$. A function $g \in L^2(\rho^{[n]})$ is said to be *symmetric* if

$$g(x_1, \dots, x_n) = g(x_{\pi(1)}, \dots, x_{\pi(n)}) \quad \rho^{[n]}\text{-a.e. } (x_1, \dots, x_n) \in \mathbb{X}^n$$

for every permutation π of $[n]$. Let $\mathbb{H}_n$ denote the set of all symmetric functions $g \in L^2(\rho^{[n]})$ satisfying

$$\int_{\mathbb{X}} g(x_1, \dots, x_{n-1}, x)\,(\rho + \delta_{x_1} + \ldots + \delta_{x_{n-1}})(\mathrm{d}x) = 0 \tag{4.1}$$

for $\rho^{[n-1]}$-almost all $(x_1, \dots, x_{n-1}) \in \mathbb{X}^{n-1}$. Further, let $\mathbb{H}_0 \coloneqq \mathbb{R}$.

We note that in the special case $n = 1$, the space $\mathbb{H}_1$ consists of all functions $g \in L^2(\rho)$ that satisfy the condition $\int_{\mathbb{X}} g(x)\,\rho(\mathrm{d}x) = 0$. Moreover, we recall the recursion for the measures $\rho^{[n]}$, $n \in \mathbb{N}$, from (3.1) in Lemma 3.5.

Since for each $n \in \mathbb{N}$ the measure $\frac{1}{\theta^{(n)}}\rho^{[n]}$ is the joint distribution of the first n elements of the Pólya sequence $(X_n)_{n\in\mathbb{N}}$ from (2.2), we have for $g \in \mathbb{H}_n$ that

$$\mathbb{E}(g(X_1, \dots, X_n) | X_1, \dots, X_{n-1}) = 0, \quad \mathbb{P}\text{-a.s.},$$

or, equivalently,

$$\mathbb{E}[g(X_1, \dots, X_n)h(X_1, \dots, X_{n-1})] = 0$$

for every measurable and bounded function $h\colon \mathbb{X}^{n-1} \to [0, \infty)$. The last equation gives rise to an idea for a characterisation of the spaces $\mathbb{H}_n$, $n \in \mathbb{N}$, if $n > 1$.

Lemma 4.2. *Let $n \in \mathbb{N}$, $n > 1$. A symmetric function $g \in L^2(\rho^{[n]})$ is an element of $\mathbb{H}_n$ if and only if for all measurable and bounded functions $h\colon \mathbb{X}^{n-1} \to [0, \infty)$ it holds*

$$\int_{\mathbb{X}^n} h(x_1, \dots, x_{n-1})g(x_1, \dots, x_n)\,\rho^{[n]}(\mathrm{d}(x_1, \dots, x_n)) = 0.$$

Proof. If $g \in \mathbb{H}_n$ and $h\colon \mathbb{X}^{n-1} \to [0,\infty)$ is measurable and bounded, using the recursion (3.1) from Lemma 3.5, we have

$$\int_{\mathbb{X}^n} h(\mathbf{x_{n-1}})g(\mathbf{x_n})\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}) = \int_{\mathbb{X}^{n-1}} h(\mathbf{x_{n-1}})\left(\int_{\mathbb{X}} g(\mathbf{x_{n-1}},x_n)\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\right)\rho^{[n-1]}(\mathrm{d}\mathbf{x_{n-1}}) = 0.$$

On the other hand, let $g \in L^2(\rho^{[n]})$ be a symmetric function satisfying the equation from the statement for all measurable functions h. We assume $g \notin \mathbb{H}_n$, i.e. we assume that there exists a measurable set $B \subseteq \mathbb{X}^{n-1}$ with $\rho^{[n-1]}(B) > 0$ and

$$\int_{\mathbb{X}} g(\mathbf{x_{n-1}},x)\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x) \neq 0, \quad \mathbf{x_{n-1}} \in B.$$

For $h = \mathbb{1}_B$, using once again the recursion (3.1) from Lemma 3.5, it then holds

$$\int_{\mathbb{X}^n} h(\mathbf{x_{n-1}})g(\mathbf{x_n})\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}) = \int_B\int_{\mathbb{X}} g(\mathbf{x_n})\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\,\rho^{[n-1]}(\mathrm{d}\mathbf{x_{n-1}}) \neq 0.$$

This contradicts the assumption. □

The preceding lemma allows us to establish that the spaces $\mathbb{H}_n$, $n \in \mathbb{N}$, are closed.

Lemma 4.3. *Given $n \in \mathbb{N}$, the set $\mathbb{H}_n$ is a closed subset of $L^2(\rho^{[n]})$.*

Proof. The claim holds for $n = 0$. Let $n \geq 1$ and $(g_m)_{m\in\mathbb{N}}$ be a sequence in $\mathbb{H}_n$ with $g_m \to g$, $m \to \infty$, for some function $g \in L^2(\rho^{[n]})$. To begin with, we observe that g is symmetric since the subspace of symmetric functions is closed. (For every permutation π of $[n]$, the mapping $S_\pi\colon L^2(\rho^{[n]}) \to L^2(\rho^{[n]})$ defined by $(S_\pi f)(x_1,\dots,x_n) = f(x_{\pi(1)},\dots,x_{\pi(n)})$, $(x_1,\dots,x_n) \in \mathbb{X}^n$, is an isometry. Hence, the set of all symmetric functions, $\bigcap_{\pi \text{ permutation of } [n]} \ker(\mathrm{id} - S_\pi)$, where ker denotes the kernel of a linear mapping, is closed.) Let $h\colon \mathbb{X}^{n-1} \to [0,\infty)$ be measurable and bounded. Since $g_m \in \mathbb{H}_n$ for each $m \in \mathbb{N}$, Lemma 4.2 implies

$$\begin{aligned}\int_{\mathbb{X}^n} h(\mathbf{x_{n-1}})g(\mathbf{x_n})\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}) &= \int_{\mathbb{X}^n} h(\mathbf{x_{n-1}})(g(\mathbf{x_n}) - g_m(\mathbf{x_n}))\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}) + \int_{\mathbb{X}^n} h(\mathbf{x_{n-1}})g_m(\mathbf{x_n})\,\rho^{[n]}(\mathrm{d}\mathbf{x_n})\\ &= \int_{\mathbb{X}^n} h(\mathbf{x_{n-1}})(g(\mathbf{x_n}) - g_m(\mathbf{x_n}))\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}).\end{aligned}$$

By the Cauchy–Schwarz inequality, the absolute value of this integral is bounded by

$$\left(\int_{\mathbb{X}^n} h(\mathbf{x_{n-1}})^2\,\rho^{[n]}(\mathrm{d}\mathbf{x_n})\right)^{\frac{1}{2}}\left(\int_{\mathbb{X}^n} (g(\mathbf{x_n}) - g_m(\mathbf{x_n}))^2\,\rho^{[n]}(\mathrm{d}\mathbf{x_n})\right)^{\frac{1}{2}},$$

which converges to zero as $m \to \infty$. Thus, the claim is a consequence of Lemma 4.2. □

We proceed with a general lemma that facilitates the evaluation of integrals. As immediate corollaries, we obtain formulas for integrating products of functions from $\mathbb{H}_m$ and $\mathbb{H}_n$, $m,n \in \mathbb{N}$, required in the subsequent analysis.

Lemma 4.4. *Let $m,n,r \in \mathbb{N}_0$ so that $r < n$. Further, let $f\colon \mathbb{X}^{m+r+n} \to \mathbb{R}$ be such that $f \in L^1(\rho^{[m+r+n]})$ and $f(x_1,\dots,x_{m+r},\cdot) \in \mathbb{H}_n$ for $\rho^{[m+r]}$-almost all $(x_1,\dots,x_{m+r}) \in \mathbb{X}^{m+r}$. Then*

$$\begin{aligned}&\int_{\mathbb{X}^{m+n}} f(x_1,\dots,x_m,y_1,\dots,y_r,y_1,\dots,y_n)\,\rho^{[m+n]}(\mathrm{d}(x_1,\dots,x_m,y_1,\dots,y_n))\\ &= \sum_{(i_1,\dots,i_{n-r})\in[m]^{[n-r]}} \int f(x_1,\dots,x_m,y_1,\dots,y_r,y_1,\dots,y_r,x_{i_1},\dots,x_{i_{n-r}})\,\rho^{[m+r]}(\mathrm{d}(x_1,\dots,x_m,y_1,\dots,y_r))\end{aligned}$$

is valid whenever $m \geq n-r$. If $m < n-r$, the integral on the left-hand side vanishes.

Proof. We first consider the case $m \geq n-r$. The recursion (3.1) from Lemma 3.5 leads to

$$\begin{aligned}&\int_{\mathbb{X}^{m+n}} f(\mathbf{x_m},\mathbf{y_r},\mathbf{y_n})\,\rho^{[m+n]}(\mathrm{d}(\mathbf{x_m},\mathbf{y_n}))\\ &= \int_{\mathbb{X}^{n-1}}\int_{\mathbb{X}^m}\int_{\mathbb{X}} f(\mathbf{x_m},\mathbf{y_r},\mathbf{y_n})\,(\rho+\delta_{\mathbf{x_m}}+\delta_{\mathbf{y_{n-1}}})(\mathrm{d}y_n)\,(\rho+\delta_{\mathbf{y_{n-1}}})^{[m]}(\mathrm{d}\mathbf{x_m})\,\rho^{[n-1]}(\mathrm{d}\mathbf{y_{n-1}})\end{aligned}$$

$$= \int_{\mathbb{X}^{n-1}} \int_{\mathbb{X}^m} \sum_{i_1=1}^{m} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_{n-1}}, x_{i_1}) \, (\rho + \delta_{\mathbf{y_{n-1}}})^{[m]}(\mathrm{d}\mathbf{x_m}) \, \rho^{[n-1]}(\mathrm{d}\mathbf{y_{n-1}}).$$

If $n \geq 2$, a repetition of this procedure shows that the integral equals

$$\int_{\mathbb{X}^{n-2}} \int_{\mathbb{X}^m} \sum_{i_1=1}^{m} \sum_{i_2=1, i_2 \neq i_1}^{m} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_{n-2}}, x_{i_2}, x_{i_1}) \, (\rho + \delta_{\mathbf{y_{n-2}}})^{[m]}(\mathrm{d}\mathbf{x_m}) \, \rho^{[n-2]}(\mathrm{d}\mathbf{y_{n-2}})$$
$$= \int_{\mathbb{X}^{m+n-2}} \sum_{(i_1,i_2) \in [m]^{[2]}} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_{n-2}}, x_{i_2}, x_{i_1}) \, \rho^{[m+n-2]}(\mathrm{d}(\mathbf{x_m}, \mathbf{y_{n-2}})).$$

Inductively, we obtain

$$\int_{\mathbb{X}^{m+n}} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_n}) \, \rho^{[m+n]}(\mathrm{d}(\mathbf{x_m}, \mathbf{y_n}))$$
$$= \int_{\mathbb{X}^{m+r}} \sum_{(i_1,\dots,i_{n-r}) \in [m]^{[n-r]}} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_r}, x_{i_1}, \dots, x_{i_{n-r}}) \, \rho^{[m+r]}(\mathrm{d}(\mathbf{x_m}, \mathbf{y_r})).$$

In particular, if $m = n - r$, this equals

$$m! \int_{\mathbb{X}^{m+r}} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_r}, \mathbf{x_m}) \, \rho^{[m+r]}(\mathrm{d}(\mathbf{x_m}, \mathbf{y_r}))$$

by the symmetry of functions in $\mathbb{H}_n$. If $m < n - r$, the defining properties of $\mathbb{H}_n$ provide

$$\int_{\mathbb{X}^{m+n}} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_n}) \, \rho^{[m+n]}(\mathrm{d}(\mathbf{x_m}, \mathbf{y_n}))$$
$$= \int_{\mathbb{X}^n} \sum_{(i_1,\dots,i_m) \in [m]^{[m]}} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_{n-m}}, x_{i_1}, \dots, x_{i_m}) \, \rho^{[n]}(\mathrm{d}(\mathbf{x_m}, \mathbf{y_{n-m}}))$$
$$= m! \int_{\mathbb{X}^{n-1}} \int_{\mathbb{X}} f(\mathbf{x_m}, \mathbf{y_r}, \mathbf{y_{n-m-1}}, y_{n-m}, \mathbf{x_m}) \, (\rho + \delta_{\mathbf{x_m}} + \delta_{\mathbf{y_{n-m-1}}})(\mathrm{d}y_{n-m}) \, \rho^{[n-1]}(\mathrm{d}(\mathbf{x_m}, \mathbf{y_{n-m-1}})) = 0. \square$$

The following corollary gives the first of the previously announced formulas.

Corollary 4.5. *Let $n, l, k \in \mathbb{N}_0$ with $n \geq l \geq k$. Then functions $g \in \mathbb{H}_n$ and $h \in L^2(\rho^{[l]})$ satisfy*

$$\int_{\mathbb{X}^{l+n-k}} g(y_1, \dots, y_n) h(x_1, \dots, x_{l-k}, y_1, \dots, y_k) \, \rho^{[l+n-k]}(\mathrm{d}(x_1, \dots, x_{l-k}, y_1, \dots, y_n))$$
$$= \mathbb{1}_{\{l=n\}} (n-k)! \int_{\mathbb{X}^n} g(z) h(z) \, \rho^{[n]}(\mathrm{d}z).$$

Proof. If $k = n$, it holds $n = l = k$ and there is nothing to show. Now, assume $k < n$. Then we are in the setting of Lemma 4.4 with $m = l - k$, n, $r = k$ and the function $f \colon \mathbb{X}^{l+n} \to \mathbb{R}$ given by

$$f(\mathbf{x_l}, \mathbf{y_n}) := h(\mathbf{x_l}) g(\mathbf{y_n}).$$

Since g and h are square-integrable and ρ is a finite measure, g and h are in $L^1(\rho^{[n]})$ and $L^1(\rho^{[l]})$, respectively. Thus, the Cauchy–Schwarz inequality shows that f is an element of $L^1(\rho^{[l+n]})$. Applying Lemma 4.4 to

$$\int_{\mathbb{X}^{l-k+n}} h(\mathbf{x_{l-k}}, \mathbf{y_k}) g(\mathbf{y_n}) \, \rho^{[l-k+n]}(\mathrm{d}(\mathbf{x_{l-k}}, \mathbf{y_n})) = \int_{\mathbb{X}^{l-k+n}} f(\mathbf{x_{l-k}}, \mathbf{y_k}, \mathbf{y_n}) \, \rho^{[l-k+n]}(\mathrm{d}(\mathbf{x_{l-k}}, \mathbf{y_n})),$$

we obtain

$$\sum_{(i_1,\dots,i_{n-k}) \in [l-k]^{[n-k]}} \int_{\mathbb{X}^l} f(\mathbf{x_{l-k}}, \mathbf{y_k}, \mathbf{y_k}, x_{i_1}, \dots, x_{i_{n-k}}) \, \rho^{[l]}(\mathrm{d}(\mathbf{x_{l-k}}, \mathbf{y_k}))$$
$$= \sum_{(i_1,\dots,i_{n-k}) \in [l-k]^{[n-k]}} \int_{\mathbb{X}^l} h(\mathbf{x_{l-k}}, \mathbf{y_k}) g(\mathbf{y_k}, x_{i_1}, \dots, x_{i_{n-k}}) \, \rho^{[l]}(\mathrm{d}(\mathbf{x_{l-k}}, \mathbf{y_k})) \quad (4.2)$$

if $l - k \geq n - k$. Otherwise, the integral vanishes. By the assumption $n \geq l \geq k$ and the symmetry of g, the assertion follows. $\square$

Another formula used later is stated in the subsequent corollary.

Corollary 4.6. *Let $k, l \in \mathbb{N}_0$ with $k \geq l$ and $g \in L^2(\rho^{[k+1]})$, $h \in L^2(\rho^{[l+1]})$ such that $g(x, \cdot) \in \mathbb{H}_k$ and $h(x, \cdot) \in \mathbb{H}_l$ for all $x \in \mathbb{X}$. Then*

$$\begin{aligned}
&\int_{\mathbb{X}^{k+l+1}} g(x, x_1, \ldots, x_k) h(x, y_1, \ldots, y_l)\, \rho^{[k+l+1]}(\mathrm{d}(x, x_1, \ldots, x_k, y_1, \ldots, y_l)) \\
&= \mathbb{1}_{\{k=l+1\}} k! \int_{\mathbb{X}^k} g(x_1, x_1, \ldots, x_k) h(x_1, \ldots, x_k)\, \rho^{[k]}(\mathrm{d}(x_1, \ldots, x_k)) \\
&\quad + \mathbb{1}_{\{k=l\}} k! \Bigg(\int_{\mathbb{X}^{k+1}} g(z) h(z)\, \rho^{[k+1]}(\mathrm{d}z) \\
&\qquad\qquad + k \int_{\mathbb{X}^k} g(x_1, x_1, \ldots, x_k) h(x_1, x_1, \ldots, x_k)\, \rho^{[k]}(\mathrm{d}(x_1, \ldots, x_k)) \Bigg).
\end{aligned}$$

Proof. If $k = 0$, then $l = 0$ as well and there is nothing to prove. Therefore, we assume $k \geq 1$.

First, assume $l \geq 1$. An application of Lemma 4.4 (with $m = k+1$, $n = l$, $r = 0$) to the function $f\colon \mathbb{X}^{k+l+1} \to \mathbb{R}$ defined by

$$f(\mathbf{x_{k+1}}, \mathbf{y_l}) := g(\mathbf{x_{k+1}}) h(x_1, \mathbf{y_l})$$

yields

$$\begin{aligned}
&\int_{\mathbb{X}^{k+l+1}} g(\mathbf{x_{k+1}}) h(x_1, \mathbf{y_l})\, \rho^{[k+l+1]}(\mathrm{d}(\mathbf{x_{k+1}}, \mathbf{y_l})) \\
&\qquad = \sum_{(i_1, \ldots, i_l) \in [k+1]^{[l]}} \int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}}) h(x_1, x_{i_1}, \ldots, x_{i_l})\, \rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}}). \qquad (4.3)
\end{aligned}$$

Let $(i_1, \ldots, i_l) \in [k+1]^{[l]}$. If there exists $j \in \{2, \ldots, k+1\} \setminus \{i_1, \ldots, i_l\}$, the defining properties of $\mathbb{H}_k$ give

$$\begin{aligned}
&\int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}}) h(x_1, x_{i_1}, \ldots, x_{i_l})\, \rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}}) \\
&= \int_{\mathbb{X}^k} \int_{\mathbb{X}} g(\mathbf{x_{k+1}}) \left(\rho + \delta_{\mathbf{x_{j-1}}} + \delta_{x_{j+1}} + \ldots + \delta_{x_{k+1}}\right)(\mathrm{d}x_j) h(x_1, x_{i_1}, \ldots, x_{i_l})\, \rho^{[k]}(\mathrm{d}(\mathbf{x_{j-1}}, x_{j+1}, \ldots, x_{k+1})) \\
&= \int_{\mathbb{X}^k} g(x_1, x_1, \mathbf{x_{j-1}}, x_{j+1}, \ldots, x_{k+1}) h(x_1, x_{i_1}, \ldots, x_{i_l})\, \rho^{[k]}(\mathrm{d}(\mathbf{x_{j-1}}, x_{j+1}, \ldots, x_{k+1})). \qquad (4.4)
\end{aligned}$$

If, in addition, there exists $i \in \{2, \ldots, k+1\} \setminus \{j, i_1, \ldots, i_l\}$, we obtain

$$\begin{aligned}
&\int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}}) h(x_1, x_{i_1}, \ldots, x_{i_l})\, \rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}}) \\
&= \int_{\mathbb{X}^{k-1}} \int_{\mathbb{X}} g(x_1, x_1, \mathbf{x_{j-1}}, x_{j+1}, \ldots, x_{k+1}) \left(\rho + \delta_{\mathbf{x_{k+1\setminus i,j}}}\right)(\mathrm{d}x_i) h(x_1, x_{i_1}, \ldots, x_{i_l})\, \rho^{[k-1]}(\mathrm{d}\mathbf{x_{k+1\setminus i,j}}) \\
&= 0,
\end{aligned}$$

where $\mathbf{x_{k+1\setminus i,j}}$ denotes the vector $\mathbf{x_{k+1}}$ with the ith and jth entry omitted. Thus, if $k \geq l+2$,

$$\int_{\mathbb{X}^{k+l+1}} g(\mathbf{x_{k+1}}) h(x_1, \mathbf{y_l})\, \rho^{[k+l+1]}(\mathrm{d}(\mathbf{x_{k+1}}, \mathbf{y_l})) = 0.$$

If $k = l+1$, the only summands remaining in (4.3) are the ones corresponding to tuples $(i_1, \ldots, i_l) \in [k+1]^{[l]}$ not containing 1 (because, in this case, $\{2, \ldots, k+1\} \setminus \{i_1, \ldots, i_l\}$ contains exactly one element). Since there are $k!$ tuples of this kind, using the symmetry of h in its last arguments, we obtain

$$\int_{\mathbb{X}^{k+l+1}} g(\mathbf{x_{k+1}}) h(x_1, \mathbf{y_l})\, \rho^{[k+l+1]}(\mathrm{d}(\mathbf{x_{k+1}}, \mathbf{y_l})) = k! \int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}}) h(x_1, x_2, \ldots, x_k)\, \rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}}).$$

If $k = l$, the right-hand side of (4.3) becomes

$$\sum_{(i_1, \ldots, i_k) \in [k+1]^{[k]}} \int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}}) h(x_1, x_{i_1}, \ldots, x_{i_k})\, \rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}}).$$

The set $[k+1]^{[k]}$ contains $k \cdot k!$ tuples containing 1 and $k!$ tuples not containing 1. An application of (4.4) and exploiting the symmetry of h yields

$$\sum_{(i_1,\dots,i_k)\in[k+1]^{[k]}} \int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}})h(x_1,x_{i_1},\dots,x_{i_k})\,\rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}})$$
$$= kk!\int_{\mathbb{X}^k} g(x_1,x_1,\dots,x_k)h(x_1,x_1,\dots,x_k)\,\rho^{[k]}(\mathrm{d}(x_1,\dots,x_k)) + k!\int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}})h(\mathbf{x_{k+1}})\,\rho^{[k+1]}(\mathrm{d}z).$$

If $l=0$, by Lemma 4.4 (with $m=1$, $n=k$, $r=0$ and $f(\mathbf{x_{k+1}}) \coloneqq g(\mathbf{x_{k+1}})h(x_1)$, $\mathbf{x_{k+1}} \in \mathbb{X}^{k+1}$) we have

$$\int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}})h(x_1)\,\rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}}) = 0$$

if $k \ge 2$ and, if $k=1$,

$$\int_{\mathbb{X}^{k+1}} g(\mathbf{x_{k+1}})h(x_1)\,\rho^{[k+1]}(\mathrm{d}\mathbf{x_{k+1}}) = \int_{\mathbb{X}} g(x_1,x_1)h(x_1)\,\rho(\mathrm{d}x_1). \qquad \square$$

Several other formulas needed in the subsequent chapter are collected in the next corollary.

Corollary 4.7. *Let $m \in \mathbb{N}_0$, $n \in \mathbb{N}$, $f \in \mathbb{H}_n$ and $h \in L^2(\rho^{[m+1]})$ such that $h(x,\cdot) \in \mathbb{H}_m$ for all $x \in \mathbb{X}$. Then*

$$\begin{aligned}
&\int_{\mathbb{X}^{m+n+1}} h(y)f(z)\,\rho^{[m+n+1]}(\mathrm{d}(y,z))\\
&\quad= \mathbb{1}_{\{m=n+1\}} m! \int_{\mathbb{X}^m} h(x_1,x_1,\dots,x_m)f(x_2,\dots,x_m)\,\rho^{[m]}(\mathrm{d}(x_1,\dots,x_m))\\
&\qquad+ \mathbb{1}_{\{m=n\}}\Big(m!\int_{\mathbb{X}^{m+1}} h(x_1,\dots,x_{m+1})f(x_2,\dots,x_{m+1})\,\rho^{[m+1]}(\mathrm{d}(x_1,\dots,x_{m+1}))\\
&\qquad\qquad+mm!\int_{\mathbb{X}^m} h(x_1,x_1,\dots,x_m)f(x_1,\dots,x_m)\,\rho^{[m]}(\mathrm{d}(x_1,\dots,x_m))\Big)\\
&\qquad+ \mathbb{1}_{\{m+1=n\}}(m+1)!\int_{\mathbb{X}^{m+1}} h(x_1,\dots,x_{m+1})f(x_1,\dots,x_{m+1})\,\rho^{[m+1]}(\mathrm{d}(x_1,\dots,x_{m+1}))
\end{aligned}$$

as well as

$$\begin{aligned}
&\int_{\mathbb{X}^{m+n}} h(x,x_1,\dots,x_m)f(x,y_1,\dots,y_{n-1})\,\rho^{[m+n]}(\mathrm{d}(x,x_1,\dots,x_m,y_1,\dots,y_{n-1}))\\
&\quad= \mathbb{1}_{\{m=n\}} m! \int_{\mathbb{X}^m} h(x_1,x_1,\dots,x_m)f(x_1,\dots,x_m)\,\rho^{[m]}(\mathrm{d}(x_1,\dots,x_m))\\
&\qquad+ \mathbb{1}_{\{m=n-1\}} m! \int_{\mathbb{X}^{m+1}} h(x_1,\dots,x_{m+1})f(x_1,\dots,x_{m+1})\,\rho^{[m+1]}(\mathrm{d}(x_1,\dots,x_{m+1})).
\end{aligned}$$

and, if m is positive,

$$\begin{aligned}
&\int_{\mathbb{X}^{m+n}}\int_{\mathbb{X}} h(x,x_1,\dots,x_m)\,(\rho+\delta_{x_1}+\ldots+\delta_{x_m})(\mathrm{d}x)f(y_1,\dots,y_n)\,\rho^{[m+n]}(\mathrm{d}(x_1,\dots,x_m,y_1,\dots,y_n))\\
&\quad= \mathbb{1}_{\{m=n\}} m! \int_{\mathbb{X}^{m+1}} h(x_{m+1},x_1,\dots,x_m)f(x_1,\dots,x_m)\,\rho^{[m+1]}(\mathrm{d}(x_1,\dots,x_{m+1}))\\
&\qquad+ \mathbb{1}_{\{m=n+1\}} m! \int_{\mathbb{X}^m} h(x_m,x_1,\dots,x_m)f(x_1,\dots,x_{m-1})\,\rho^{[m]}(\mathrm{d}(x_1,\dots,x_m)).
\end{aligned}$$

Proof. Corollary 4.6 applied to the functions h and $g\colon \mathbb{X}^{n+1} \to \mathbb{R}$ defined by $g(\mathbf{x_{n+1}}) \coloneqq f(x_2,\dots,x_{n+1})$ yields the first assertion.

For the second integral, the case $n \ge m+1$ follows from Corollary 4.5 (with $k=1$). If $n \le m$, arguing as in the proof of Corollary 4.5 (cf. (4.2)) yields

$$\int_{\mathbb{X}^{m+n}} h(y_1,\mathbf{x_m})f(\mathbf{y_n})\,\rho^{[m+n]}(\mathrm{d}(\mathbf{x_m},\mathbf{y_n}))$$

$$= \sum_{(i_1,\dots,i_{n-1})\in[m]^{[n-1]}} \int_{\mathbb{X}^{m+1}} h(y_1,\mathbf{x_m}) f(y_1, x_{i_1},\dots,x_{i_{n-1}})\, \rho^{[m+1]}(\mathrm{d}(\mathbf{x_m},y_1)).$$

By the properties of $h(x,\cdot)\in\mathbb{H}_m$, $x\in\mathbb{X}$, disintegrating as in the proof of Corollary 4.6 (cf. (4.4)) yields that the sum vanishes unless $m=n$ and which case the integral takes the asserted value.

To conclude, we consider the third integral. An application of Lemma 4.4 to $\widetilde{f}\colon \mathbb{X}^{m+n}\to\mathbb{R}$, given by

$$\widetilde{f}(\mathbf{x_m},\mathbf{y_n}) := \int_{\mathbb{X}} h(t,\mathbf{x_m})\,(\rho+\delta_{\mathbf{x_m}})(\mathrm{d}t) f(\mathbf{y_n})$$

gives

$$\begin{aligned} &\int_{\mathbb{X}^{m+n}}\int_{\mathbb{X}} h(x,\mathbf{x_m})\,(\rho+\delta_{\mathbf{x_m}})(\mathrm{d}x) f(\mathbf{y_n})\,\rho^{[m+n]}(\mathrm{d}(\mathbf{x_m},\mathbf{y_n}))\\ &\quad= \sum_{(i_1,\dots,i_n)\in[m]^{[n]}} \int_{\mathbb{X}^m}\int_{\mathbb{X}} h(x,\mathbf{x_m})\,(\rho+\delta_{\mathbf{x_m}})(\mathrm{d}x) f(x_{i_1},\dots,x_{i_n})\,\rho^{[m]}(\mathrm{d}\mathbf{x_m}) \end{aligned} \tag{4.5}$$

unless $m<n$ in which case the integral is zero. (Note that $\widetilde{f}\in L^1(\rho^{[m+n]})$ since its L^1-norm is bounded by

$$\begin{aligned} &\int_{\mathbb{X}^{m+n}}\int_{\mathbb{X}} |h(t,\mathbf{x_m})f(\mathbf{y_n})|\,(\rho+\delta_{\mathbf{x_m}})(\mathrm{d}t)\,\rho^{[m+n]}(\mathrm{d}(\mathbf{x_m},\mathbf{y_n}))\\ &\qquad\qquad \le \int_{\mathbb{X}^{m+n+1}} |h(\mathbf{x_{m+1}})f(\mathbf{y_n})|\,\rho^{[m+n+1]}(\mathrm{d}(\mathbf{x_{m+1}},\mathbf{y_n})) \end{aligned}$$

which by the Cauchy–Schwarz inequality is bounded by the product of the L^2-norms of f and h.) By the recursion (3.1) from Lemma 3.5, the integral (4.5) equals

$$\sum_{(i_1,\dots,i_n)\in[m]^{[n]}} \int_{\mathbb{X}^{m+1}} h(x,\mathbf{x_m}) f(x_{i_1},\dots,x_{i_n})\,\rho^{[m+1]}(\mathrm{d}(x,\mathbf{x_m})).$$

Using the properties of h and the symmetry of f yields the assertion. □

A natural question arises whether functions of the type considered in Lemma 4.4, that is, measurable functions which, upon fixing certain arguments, belong to the space $\mathbb{H}_n$ for some $n\in\mathbb{N}$, also belong to a space $\mathbb{H}_m$ for some $m>n$ when regarded as functions of all variables. However, the following example shows that this is not necessarily the case.

Example 4.8. Let $B\in\mathcal{X}$ and $h\colon\mathbb{X}^2\to\mathbb{R}$ be defined by

$$h(x,y) := \frac{1}{\rho(\mathbb{X})+1}(\rho(B)+\mathbb{1}_B(y))\mathbb{1}_B(x) - \frac{1}{\rho(\mathbb{X})}\rho(B)\mathbb{1}_B(x).$$

Because of

$$\int_{\mathbb{X}} h(x,y)\,\rho(\mathrm{d}y) = \frac{\rho(\mathbb{X})}{\rho(\mathbb{X})+1}\rho(B)\mathbb{1}_B(x) + \frac{1}{\rho(\mathbb{X})+1}\rho(B)\mathbb{1}_B(x) - \rho(B)\mathbb{1}_B(x) = 0, \quad x\in\mathbb{X},$$

we have $h(x,\cdot)\in\mathbb{H}_1$ for all $x\in\mathbb{X}$. However, the function h is not symmetric and thus cannot be an element of $\mathbb{H}_2$. Denoting its symmetrisation by $\widetilde{h}$, we obtain

$$\begin{aligned} &\int_{\mathbb{X}} \widetilde{h}(x,y)\,(\rho+\delta_x)(\mathrm{d}y) = \int_{\mathbb{X}} \frac{1}{2}\,(h(x,y)+h(y,x))\,\,(\rho+\delta_x)(\mathrm{d}y)\\ &= \frac{1}{2}\int_{\mathbb{X}} \frac{1}{\rho(\mathbb{X})+1}(\rho(B)\mathbb{1}_B(x)+\rho(B)\mathbb{1}_B(y)+2\cdot\mathbb{1}_B(y)\mathbb{1}_B(x)) - \frac{1}{\rho(\mathbb{X})}\rho(B)\,(\mathbb{1}_B(x)+\mathbb{1}_B(y))\,\,(\rho+\delta_x)(\mathrm{d}y)\\ &= \frac{1}{2}\left(\rho(B)\mathbb{1}_B(x) + \frac{\rho(B)(\rho(B)+\mathbb{1}_B(x))}{\rho(\mathbb{X})+1} + \frac{2(\rho(B)+\mathbb{1}_B(x))\mathbb{1}_B(x)}{\rho(\mathbb{X})+1} - \frac{\rho(B)((\rho(\mathbb{X})+2)\mathbb{1}_B(x)+\rho(B))}{\rho(\mathbb{X})}\right)\\ &= \frac{1}{2}\left(\frac{\rho(\mathbb{X})-2}{\rho(\mathbb{X})^{(2)}}\rho(B)\mathbb{1}_B(x) - \frac{1}{\rho(\mathbb{X})^{(2)}}\rho(B)^2 + \frac{2}{\rho(\mathbb{X})+1}\mathbb{1}_B(x)\right), \quad x\in\mathbb{X}, \end{aligned}$$

and conclude that the symmetrisation does not need to be an element of $\mathbb{H}_2$ either. Moreover, also the function $f\colon \mathbb{X} \to \mathbb{R}$ given by

$$y \mapsto \int_{\mathbb{X}} h(x,y)\,(\rho + \delta_y)(\mathrm{d}x)$$

may not be an element of $\mathbb{H}_1$: Indeed, we have

$$\begin{aligned} f(y) &= \int_{\mathbb{X}} \frac{1}{\rho(\mathbb{X})+1}(\rho(B) + \mathbb{1}_B(y))\mathbb{1}_B(x) - \frac{1}{\rho(\mathbb{X})}\rho(B)\mathbb{1}_B(x)\,(\rho+\delta_y)(\mathrm{d}x) \\ &= \frac{1}{\rho(\mathbb{X})+1}(\rho(B)+\mathbb{1}_B(y))^2 - \frac{1}{\rho(\mathbb{X})}\rho(B)(\rho(B)+\mathbb{1}_B(y)), \quad y \in \mathbb{X}, \end{aligned}$$

and

$$\begin{aligned} \int_{\mathbb{X}} f(y)\,\rho(\mathrm{d}y) &= \int_{\mathbb{X}} \frac{1}{\rho(\mathbb{X})+1}(\rho(B)^2 + 2\rho(B)\mathbb{1}_B(y) + \mathbb{1}_B(y)) - \frac{1}{\rho(\mathbb{X})}\rho(B)(\rho(B)+\mathbb{1}_B(y))\,\rho(\mathrm{d}y) \\ &= \frac{\rho(B)^2\rho(\mathbb{X})}{\rho(\mathbb{X})+1} + \frac{2\rho(B)^2}{\rho(\mathbb{X})+1} + \frac{\rho(B)}{\rho(\mathbb{X})+1} - \rho(B)^2 - \frac{\rho(B)^2}{\rho(\mathbb{X})} = \frac{\rho(B)}{\rho(\mathbb{X})+1} - \frac{\rho(B)^2}{\rho(\mathbb{X})^{(2)}}. \qquad \circ \end{aligned}$$

Finally, we consider the situation in which a function, upon fixing one of its arguments, lies in $\mathbb{H}_m$, $m \in \mathbb{N}$, while the full function in all variables is an element of $\mathbb{H}_{m+1}$. The following lemma reveals a characteristic property of the function in this case.

Lemma 4.9. *Let $m \in \mathbb{N}$ and $f\colon \mathbb{X}^{m+1} \to \mathbb{R}$ be such that $f \in L^2(\rho^{[m+1]})$ and $f(x,\cdot) \in \mathbb{H}_m$ for ρ-almost all $x \in \mathbb{X}$. Then f is an element of $\mathbb{H}_{m+1}$ if and only if f is symmetric and $f(y_m, y_1, \dots, y_m) = 0$ for $\rho^{[m]}$-almost all $(y_1,\dots,y_m) \in \mathbb{X}^m$.*

Proof. By assumption, there exists a null set $N_1 \in \mathcal{X}$ such that $f(x,\cdot) \in \mathbb{H}_m$ for $x \in \mathbb{X}\setminus N_1$, i.e. $f(x,\cdot)$ is symmetric and there is a $\rho^{[m-1]}$-null set $N_{m-1} \subseteq \mathbb{X}^{m-1}$ such that

$$0 = \int_{\mathbb{X}} f(x, \mathbf{y_{m-1}}, y_m)\,(\rho + \delta_{\mathbf{y_{m-1}}})(\mathrm{d}y_m) = \int_{\mathbb{X}} f(x,\mathbf{y_{m-1}},y_m)\,\rho(\mathrm{d}y_m) + \sum_{r=1}^{m-1} f(x,\mathbf{y_{m-1}},y_r) \tag{4.6}$$

for all $x \in \mathbb{X}\setminus N_1$ and $\mathbf{y_{m-1}} \in \mathbb{X}^{m-1}\setminus N_{m-1}$.

First, let $f \in \mathbb{H}_{m+1}$. Then the function f is symmetric and a $\rho^{[m]}$-null set $N_m \subseteq \mathbb{X}^m$ with

$$0 = \int_{\mathbb{X}} f(\mathbf{y_m}, y_{m+1})\,(\rho+\delta_{\mathbf{y_m}})(\mathrm{d}y_{m+1}) = \int_{\mathbb{X}} f(\mathbf{y_m},y_{m+1})\,\rho(\mathrm{d}y_{m+1}) + \sum_{r=1}^{m} f(\mathbf{y_m}, y_r)$$

for all $\mathbf{y_m} \in \mathbb{X}^m \setminus N_m$ can be found. Together with (4.6) this shows

$$f(\mathbf{y_m}, y_m) = 0$$

for $\mathbf{y_m} \in (\mathbb{X}^m\setminus N_m) \cap ((\mathbb{X}\setminus N_1)\times(\mathbb{X}^{m-1}\setminus N_{m-1}))$. As the complement of this set is a $\rho^{[m]}$-null set, the necessity is established.

Second, assume that f is symmetric and satisfies $f(y_m,y_1,\dots,y_m) = 0$ for $(y_1,\dots,y_m) \in \mathbb{X}^m \setminus \widetilde{N}_m$ where $\widetilde{N}_m$ is a $\rho^{[m]}$-null set. For $(y_1,\dots,y_m) \in (\mathbb{X}^m\setminus\widetilde{N}_m)\cap((\mathbb{X}\setminus N_1)\times(\mathbb{X}^{m-1}\setminus N_{m-1}))$, which is again a set with full measure, equation (4.6) becomes

$$0 = \int_{\mathbb{X}} f(\mathbf{y_m}, y_{m+1})\,(\rho + \delta_{y_2} + \ldots + \delta_{y_m})(\mathrm{d}y_{m+1}) = \int_{\mathbb{X}} f(\mathbf{y_m},y_{m+1})\,\rho(\mathrm{d}y_{m+1}) + \sum_{r=2}^{m} f(\mathbf{y_m},y_r)$$

and it thus holds

$$\int_{\mathbb{X}} f(\mathbf{y_m}, y_{m+1})\,(\rho+\delta_{\mathbf{y_m}})(\mathrm{d}y_{m+1}) = \int_{\mathbb{X}} f(\mathbf{y_m},y_{m+1})\,\rho(\mathrm{d}y_{m+1}) + \sum_{r=1}^{m} f(\mathbf{y_m},y_r) = 0. \qquad \square$$

4.2. THE SPACES $\mathbb{F}_n$

We now consider spaces $\mathbb{F}_n$, $n \in \mathbb{N}$, which are formed by integrating functions from $\mathbb{H}_n$ with respect to a Dirichlet process with parameter measure ρ. In this section, we define the spaces and then discuss several of their properties.

Definition 4.10. Let $n \in \mathbb{N}$. Define the space $\mathbb{F}_n$ by

$$\mathbb{F}_n := \left\{ \int_{\mathbb{X}^n} g(x)\, \zeta^n(\mathrm{d}x) \colon\ g \in \mathbb{H}_n \right\} \subseteq L^2(\mathbb{P}).$$

Moreover, let $\mathbb{F}_0 := \mathbb{R}$.

As a first step, we note that Corollary 3.8, together with Corollary 4.5, yields an isometry relation for the elements of these spaces.

Corollary 4.11. *Let $m, n \in \mathbb{N}$. For functions $g \in \mathbb{H}_m$ and $h \in \mathbb{H}_n$ it holds true that*

$$\mathbb{E}\left[\int_{\mathbb{X}^m} g(x)\, \zeta^m(\mathrm{d}x) \int_{\mathbb{X}^n} h(y)\, \zeta^n(\mathrm{d}y)\right] = \mathbb{1}_{\{m=n\}} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} g(x) h(x)\, \rho^{[n]}(\mathrm{d}x).$$

Proof. By Corollary 3.8, we have

$$\begin{aligned}\mathbb{E}\left[\int_{\mathbb{X}^m} g(x)\, \zeta^m(\mathrm{d}x) \int_{\mathbb{X}^n} h(y)\, \zeta^n(\mathrm{d}y)\right] &= \mathbb{E}\left[\int_{\mathbb{X}^{m+n}} g(x)h(y)\, \zeta^{m+n}(\mathrm{d}(x,y))\right] \\ &= \frac{1}{\theta^{(m+n)}} \int_{\mathbb{X}^{m+n}} g(x)h(y)\, \rho^{[m+n]}(\mathrm{d}(x,y)).\end{aligned}$$

According to Corollary 4.5 (with $k = 0$), this is equal to

$$\mathbb{1}_{\{m=n\}} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^m} g(x) h(x)\, \rho^{[n]}(\mathrm{d}x). \qquad \square$$

Next, we show that the coincidence of two random variables in $\mathbb{F}_n$, $n \in \mathbb{N}$, implies the $\rho^{[n]}$-almost everywhere coincidence of the functions appearing in their respective representations. This result will later be used to show the almost everywhere uniqueness of the projections onto these spaces.

Corollary 4.12. *Let $n \in \mathbb{N}$ and $h, \widetilde{h} \in \mathbb{H}_n$ be such that*

$$\int_{\mathbb{X}^n} h(x)\, \zeta^n(\mathrm{d}x) = \int_{\mathbb{X}^n} \widetilde{h}(x)\, \zeta^n(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.}$$

Then $h = \widetilde{h}$ holds $\rho^{[n]}$-almost everywhere.

Proof. Let $g \in L^2(\rho^{[n]})$. The assumption and Corollary 4.11 yield

$$0 = \mathbb{E}\left[\int_{\mathbb{X}^n} g(x)\, \zeta^n(\mathrm{d}x) \int_{\mathbb{X}^n} (h(y) - \widetilde{h}(y))\, \zeta^n(\mathrm{d}y)\right] = \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} g(x)(h(x) - \widetilde{h}(x))\, \rho^{[n]}(\mathrm{d}x). \qquad \square$$

The next lemma provides yet another property of the spaces $\mathbb{F}_n$, $n \in \mathbb{N}$. The method employed in the proof is a standard procedure in functional analysis. However, it is included here to help illustrate the concepts.

Lemma 4.13. *Let $m, n \in \mathbb{N}$. The space $\mathbb{F}_m$ is a closed subset of $L^2(\mathbb{P})$ and is orthogonal to $\mathbb{F}_n$, $n \neq m$, in $L^2(\mathbb{P})$.*

Proof. The orthogonality is a consequence of Corollary 4.11. Let $m \in \mathbb{N}$ and let $(F_n)_{n\in\mathbb{N}}$ with

$$F_n = \int_{\mathbb{X}^m} g_n(x)\, \zeta^m(\mathrm{d}x), \quad n \in \mathbb{N},$$

be a sequence in $\mathbb{F}_m$ satisfying $F_n \to F$ in $L^2(\mathbb{P})$ as $n \to \infty$ for a random variable $F \in L^2(\mathbb{P})$. Corollary 4.11 yields

$$\mathbb{E}\left[(F_k - F_l)^2\right] = \mathbb{E}\left[\left(\int_{\mathbb{X}^m} g_k(x) - g_l(x)\, \zeta^m(\mathrm{d}x)\right)^2\right] = \frac{m!}{\theta^{(2m)}} \int_{\mathbb{X}^m} (g_k(x) - g_l(x))^2\, \rho^{[m]}(\mathrm{d}x)$$

for $k, l \in \mathbb{N}$. Because $(F_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbb{P})$, this shows that the functions g_n, $n \in \mathbb{N}$, form a Cauchy sequence in $\mathbb{H}_m$. Since this space is closed by Lemma 4.3, there exists a limit $g \in \mathbb{H}_m$. Using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ for $a, b \in \mathbb{R}$, we obtain

$$\mathbb{E}\left[\left(F - \int_{\mathbb{X}^m} g(x)\,\zeta^m(\mathrm{d}x)\right)^2\right] \leq 2\mathbb{E}\left[(F - F_n)^2\right] + 2\mathbb{E}\left[\left(F_n - \int_{\mathbb{X}^m} g(x)\,\zeta^m(\mathrm{d}x)\right)^2\right].$$

The convergence $F_n \to F$, $n \to \infty$, in $L^2(\mathbb{P})$ implies that the first expectation on the right-hand side tends to zero. By Corollary 4.11 the second expectation is equal to

$$\mathbb{E}\left[\left(\int_{\mathbb{X}^m} g_n(x)\,\zeta^m(\mathrm{d}x) - \int_{\mathbb{X}^m} g(x)\,\zeta^m(\mathrm{d}x)\right)^2\right] = \frac{m!}{\theta^{(2m)}} \int_{\mathbb{X}^m} (g_n(x) - g(x))^2\,\rho^{[m]}(\mathrm{d}x) \to 0, \quad n \to \infty.$$

Thus, it follows that

$$F = \int_{\mathbb{X}^m} g(x)\,\zeta^m(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.}$$

□

4.3. Projections onto the spaces $\mathbb{F}_n$

In this section, we derive formulas for the projections of a square integrable function of a Dirichlet process onto the spaces $\mathbb{F}_n$, $n \in \mathbb{N}$. The general formula is given in Proposition 4.15. However, we begin with a discussion of the underlying ideas and special cases. Readers primarily interested in the results may wish to skip the preliminary discussion and proceed directly to Proposition 4.15.

Let $F \in L^2(\zeta)$, i.e. $F\colon \mathbf{M}(\mathbb{X}) \to \mathbb{R}$ is measurable and satisfies

$$\mathbb{E}[F(\zeta)^2] < \infty.$$

Let $n \in \mathbb{N}_0$. To determine the orthogonal projection of F onto $\mathbb{F}_n$, we seek $h_n \in \mathbb{H}_n$ such that

$$\mathbb{E}\left[F(\zeta)\int_{\mathbb{X}^n} g(x)\,\zeta^n(\mathrm{d}x)\right] = \mathbb{E}\left[\int_{\mathbb{X}^n} h_n(y)\,\zeta^n(\mathrm{d}y)\int_{\mathbb{X}^n} g(x)\,\zeta^n(\mathrm{d}x)\right] \tag{4.7}$$

for all $g \in \mathbb{H}_n$. By Theorem 3.7, the left-hand side of (4.7) becomes

$$\mathbb{E}\left[F(\zeta)\int_{\mathbb{X}^n} g(x)\,\zeta^n(\mathrm{d}x)\right] = \frac{1}{\theta^{(n)}}\int_{\mathbb{X}^n} T_{F,n}(x_1,\dots,x_n)g(x_1,\dots,x_n)\,\rho^{[n]}(\mathrm{d}(x_1,\dots,x_n)),$$

where $T_{F,0} := \mathbb{E}\left[F(\zeta)\right]$ and

$$T_{F,k}(x_1,\dots,x_k) := \mathbb{E}\left[F(\zeta_{\rho+\delta_{x_1}+\dots+\delta_{x_k}})\right], \quad (x_1,\dots,x_k) \in \mathbb{X}^k, \quad k \in \mathbb{N}. \tag{4.8}$$

Using Corollary 4.11, the right-hand side of (4.7) is

$$\mathbb{E}\left[\int_{\mathbb{X}^n} h_n(y)\,\zeta^n(\mathrm{d}y)\int_{\mathbb{X}^n} g(y)\,\zeta^n(\mathrm{d}y)\right] = \frac{n!}{\theta^{(2n)}}\int_{\mathbb{X}^n} g(x)h_n(x)\,\rho^{[n]}(\mathrm{d}x).$$

Thus, equation (4.7) is equivalent to

$$\frac{(\theta+n)^{(n)}}{n!}\int_{\mathbb{X}^n} T_{F,n}(x)g(x)\,\rho^{[n]}(\mathrm{d}x) = \int_{\mathbb{X}^n} g(x)h_n(x)\,\rho^{[n]}(\mathrm{d}x). \tag{4.9}$$

In the case $n = 0$, the function $h_0 = \mathbb{E}[F(\zeta)]$ satisfies this equation. For $n \in \mathbb{N}$, we proceed with the following approach inspired by the Gram–Schmidt process

$$h_n(x_1,\dots,x_n) := \frac{(\theta+n)^{(n)}}{n!}T_{F,n}(x_1,\dots,x_n) + \sum_{j=0}^{n-1}\sum_{i_1<\dots<i_j} w_j(x_{i_1},\dots,x_{i_j}), \tag{4.10}$$

where $w_j \in L^2(\rho^{[j]})$ for $j \in [n-1]$ are symmetric functions and $w_0 \in \mathbb{R}$ is a constant. For each $g \in \mathbb{H}_n$, we then have

$$\int_{\mathbb{X}^n}\left(h_n(x) - \frac{(\theta+n)^{(n)}}{n!}T_{F,n}(x)\right)g(x)\,\rho^{[n]}(\mathrm{d}x) = \sum_{j=0}^{n-1}\sum_{i_1<\dots<i_j}\int_{\mathbb{X}^n} w_j(x_{i_1},\dots,x_{i_j})g(x)\,\rho^{[n]}(\mathrm{d}x) = 0$$

by Lemma 4.2. Hence, h_n fulfils condition (4.9) and it remains to determine the functions w_j, $j \in [n-1]$, such that $h_n \in \mathbb{H}_n$, i.e. such that

$$\int_{\mathbb{X}} h_n(x_1, \ldots, x_{n-1}, x)\,(\rho + \delta_{x_1} + \ldots + \delta_{x_{n-1}})(\mathrm{d}x) = 0, \quad \rho^{[n-1]}\text{-a.e. } (x_1, \ldots, x_{n-1}) \in \mathbb{X}^{n-1}. \tag{4.11}$$

Before addressing this task, we record a useful recursive property of the functions $T_{F,n}$.

Lemma 4.14. *It holds $\int_{\mathbb{X}} T_{F,1}(x)\,\rho(\mathrm{d}x) = \theta T_{F,0}$. For each $k \in \mathbb{N}$, the recursion*

$$\int_{\mathbb{X}} T_{F,k}(x_1, \ldots, x_{k-1}, x)\,(\rho + \delta_{x_1} + \ldots + \delta_{x_{k-1}})(\mathrm{d}x) = (\theta + k - 1) T_{F,k-1}(x_1, \ldots, x_{k-1})$$

is valid for $\rho^{[k-1]}$-almost all $(x_1, \ldots, x_{k-1}) \in \mathbb{X}^{k-1}$.

Proof. If $k = 1$, the definition of $T_{F,1}$, the Mecke-type equation from Theorem 3.7 and the definition of $T_{F,0}$ yield

$$\int_{\mathbb{X}} T_{F,1}(x)\,\rho(\mathrm{d}x) = \int_{\mathbb{X}} \mathbb{E}\left[F(\zeta_{\rho+\delta_x})\right]\,\rho(\mathrm{d}x) = \theta \mathbb{E}\left[F(\zeta_\rho)\right] = \theta T_{F,0}.$$

Let $k \geq 2$ and $g \in L^2(\rho^{[k-1]})$. By the definitions of $T_{F,k}$ and $\rho^{[k]}$, we have

$$\int_{\mathbb{X}^{k-1}} \int_{\mathbb{X}} T_{F,k}(\mathbf{x_{k-1}}, x_k)\,(\rho + \delta_{\mathbf{x_{k-1}}})(\mathrm{d}x_k) g(\mathbf{x_{k-1}})\,\rho^{[k-1]}(\mathrm{d}\mathbf{x_{k-1}}) = \int_{\mathbb{X}^k} \mathbb{E}\left[F(\zeta_{\rho+\delta_{\mathbf{x_k}}})\right] g(\mathbf{x_{k-1}})\,\rho^{[k]}(\mathrm{d}\mathbf{x_k}).$$

According to the Mecke-type equation from Theorem 3.7 and the fact that ζ_ρ is a probability measure, this is equal to

$$\theta^{(k)} \mathbb{E}\left[\int_{\mathbb{X}^k} F(\zeta_\rho) g(\mathbf{x_{k-1}})\,\zeta_\rho^k(\mathrm{d}\mathbf{x_k})\right] = \theta^{(k)} \mathbb{E}\left[\int_{\mathbb{X}^{k-1}} F(\zeta_\rho) g(\mathbf{x_{k-1}})\,\zeta_\rho^{k-1}(\mathrm{d}\mathbf{x_{k-1}})\right].$$

Using the Mecke-type equation from Theorem 3.7 and the definition of $T_{F,k-1}$, this equals

$$\frac{\theta^{(k)}}{\theta^{(k-1)}} \int_{\mathbb{X}^{k-1}} \mathbb{E}\left[F(\zeta_{\rho+\delta_{\mathbf{x_{k-1}}}})\right] g(\mathbf{x_{k-1}})\,\rho^{[k-1]}(\mathrm{d}\mathbf{x_{k-1}}) = (\theta + k - 1) \int_{\mathbb{X}^{k-1}} T_{F,k-1}(\mathbf{x_{k-1}}) g(\mathbf{x_{k-1}})\,\rho^{[k-1]}(\mathrm{d}\mathbf{x_{k-1}}).$$

□

Before deriving a general formula for the projection onto $\mathbb{F}_n$, $n \in \mathbb{N}$, we will first consider the cases $n \in \{1, 2, 3\}$ to illustrate the general approach.

1) $n = 1$

In this case, we aim to determine

$$h_1(x) = (\theta + 1)^{(1)} T_{F,1}(x) + w_0 = (\theta + 1) T_{F,1}(x) + w_0, \quad x \in \mathbb{X}.$$

From condition (4.11), we obtain

$$0 = \int_{\mathbb{X}} h_1(x)\,\rho(\mathrm{d}x) = \int_{\mathbb{X}} (\theta + 1) T_{F,1}(x)\,\rho(\mathrm{d}x) + \theta w_0.$$

Moreover, Lemma 4.14 yields

$$\int_{\mathbb{X}} T_{F,1}(x)\,\rho(\mathrm{d}x) = \theta \mathbb{E}[F(\zeta_\rho)].$$

Hence, choosing $w_0 := -(\theta + 1)\mathbb{E}[F(\zeta_\rho)]$ results in

$$h_1(x) = (\theta + 1)(T_{F,1}(x) - \mathbb{E}[F(\zeta_\rho)]), \quad x \in \mathbb{X}.$$

2) $n = 2$

The approach from (4.10) in this case reads

$$h_2(x_1, x_2) = \frac{(\theta+2)(\theta+3)}{2} T_{F,2}(x_1, x_2) + w_1(x_1) + w_1(x_2) + w_0, \quad (x_1, x_2) \in \mathbb{X}^2.$$

Hence, we seek $w_0 \in \mathbb{R}$ and $w_1 \in L^2(\rho)$ such that

$$0 = \int_{\mathbb{X}} h_2(x_1, x_2)\,(\rho + \delta_{x_1})(\mathrm{d}x_2), \quad \rho\text{-almost all } x_1 \in \mathbb{X}.$$

Given $x_1 \in \mathbb{X}$, this integral is equal to

$$\int_{\mathbb{X}} \frac{(\theta+2)(\theta+3)}{2} T_{F,2}(x_1, x_2)\,(\rho + \delta_{x_1})(\mathrm{d}x_2) + (\theta+2)w_1(x_1) + \int_{\mathbb{X}} w_1(x_2)\,\rho(\mathrm{d}x_2) + (\theta+1)w_0$$

and vanishes for the choice

$$w_1(x_1) := -\frac{\theta+3}{2} \int_{\mathbb{X}} T_{F,2}(x_1, x_2)\,(\rho + \delta_{x_1})(\mathrm{d}x_2), \quad x_1 \in \mathbb{X}, \quad \text{and} \quad w_0 := -\frac{1}{\theta+1} \int_{\mathbb{X}} w_1(x)\,\rho(\mathrm{d}x).$$

Using Lemma 4.14, these terms simplify to

$$w_1(x_1) = -\frac{(\theta+3)(\theta+1)}{2} T_{F,1}(x_1), \quad x_1 \in \mathbb{X},$$

and

$$w_0 = \frac{1}{\theta+1} \frac{(\theta+3)(\theta+1)}{2} \int_{\mathbb{X}} T_{F,1}(x)\,\rho(\mathrm{d}x) = \frac{\theta+3}{2} \theta T_{F,0} = \frac{(\theta+3)\theta}{2} \theta \mathbb{E}[F(\zeta_\rho)].$$

Thus, we obtain

$$h_2(x_1, x_2) = \frac{\theta+3}{2} \left((\theta+2) T_{F,2}(x_1, x_2) - (\theta+1) T_{F,1}(x_1) - (\theta+1) T_{F,1}(x_2) + \theta T_{F,0} \right), \quad (x_1, x_2) \in \mathbb{X}^2.$$

3) $n = 3$

We again use (4.10), i.e. we consider

$$\begin{aligned} h_3(x_1, x_2, x_3) = &\frac{(\theta+3)(\theta+4)(\theta+5)}{3!} T_{F,3}(x_1, x_2, x_3) + w_2(x_1, x_2) + w_2(x_1, x_3) + w_2(x_2, x_3) \\ &+ w_1(x_1) + w_1(x_2) + w_1(x_3) + w_0, \quad (x_1, x_2, x_3) \in \mathbb{X}^3, \end{aligned}$$

and aim at determining $w_0 \in \mathbb{R}$, $w_1 \in L^2(\rho)$ and a symmetric function $w_2 \in L^2(\rho^{[2]})$ satisfying (4.11). In this setting, (4.11) requires

$$\begin{aligned} 0 &= \int_{\mathbb{X}} h_3(x_1, x_2, x_3)\,(\rho + \delta_{x_1} + \delta_{x_2})(\mathrm{d}x_3) \\ &= \int_{\mathbb{X}} \frac{(\theta+3)(\theta+4)(\theta+5)}{3!} T_{F,3}(x_1, x_2, x_3)\,(\rho + \delta_{x_1} + \delta_{x_2})(\mathrm{d}x_3) + (\theta+4) w_2(x_1, x_2) \\ &\quad + \int_{\mathbb{X}} w_2(x_1, x_3)\,(\rho + \delta_{x_1})(\mathrm{d}x_3) + \int_{\mathbb{X}} w_2(x_2, x_3)\,(\rho + \delta_{x_2})(\mathrm{d}x_3) + (\theta+3) w_1(x_1) + (\theta+3) w_1(x_2) \\ &\quad + \int_{\mathbb{X}} w_1(x_3)\,\rho(\mathrm{d}x_3) + (\theta+2) w_0, \quad \rho^{[2]}\text{-almost all } (x_1, x_2) \in \mathbb{X}^2. \end{aligned}$$

Let

$$\begin{aligned} w_2(x_1, x_2) &:= -\frac{(\theta+3)^{(3)}}{3!(\theta+4)} \int_{\mathbb{X}} T_{F,3}(x_1, x_2, y)\,(\rho + \delta_{x_1} + \delta_{x_2})(\mathrm{d}y), \quad (x_1, x_2) \in \mathbb{X}^2, \\ w_1(x) &:= -\frac{1}{\theta+3} \int_{\mathbb{X}} w_2(x, y)\,(\rho + \delta_x)(\mathrm{d}y), \quad x \in \mathbb{X}, \end{aligned}$$

$$w_0 := -\frac{1}{\theta+2}\int_{\mathbb{X}} w_1(y)\,\rho(\mathrm{d}y).$$

By Lemma 4.14, we have

$$\begin{aligned} w_2(x_1,x_2) &= -\frac{(\theta+5)(\theta+3)(\theta+2)}{3!}T_{F,2}(x_1,x_2), \quad (x_1,x_2)\in\mathbb{X}^2,\\ w_1(x) &= \frac{(\theta+5)(\theta+2)}{3!}\int_{\mathbb{X}} T_{F,2}(x,y)\,(\rho+\delta_x)(\mathrm{d}y) = \frac{(\theta+5)(\theta+2)(\theta+1)}{3!}T_{F,1}(x), \quad x\in\mathbb{X},\\ w_0 &= -\frac{(\theta+5)(\theta+1)}{3!}\int_{\mathbb{X}} T_{F,1}(y)\,\rho(\mathrm{d}y) = -\frac{(\theta+5)(\theta+1)\theta}{3!}T_{F,0}, \end{aligned}$$

from which it follows that

$$\begin{aligned} h_3(x_1,x_2,x_3) = \frac{\theta+5}{3!}&((\theta+3)^{(2)}T_{F,3}(x_1,x_2,x_3) - (\theta+2)^{(2)}(T_{F,2}(x_1,x_2)+T_{F,2}(x_1,x_3)+T_{F,2}(x_2,x_3))\\ &+(\theta+1)^{(2)}(T_{F,1}(x_1)+T_{F,1}(x_2)+T_{F,1}(x_3)) - (\theta+1)\theta T_{F,0}), \quad (x_1,x_2,x_3)\in\mathbb{X}^3. \end{aligned}$$

The formula for general $n\in\mathbb{N}$ is the subject of the following proposition.

Proposition 4.15. *Let $F\in L^2(\zeta)$. The projection of F onto $\mathbb{F}_0$ is $\mathbb{E}[F(\zeta)]$. The projection onto $\mathbb{F}_n$, $n\in\mathbb{N}$, is $\int_{\mathbb{X}^n} f_n(x)\,\zeta^n(\mathrm{d}x)$, where $f_n\in\mathbb{H}_n$ is for $\rho^{[n]}$-almost all $(x_1,\dots,x_n)\in\mathbb{X}^n$ given by*

$$\begin{aligned} f_n(x_1,\dots,x_n) := \frac{\theta+2n-1}{n!}\Bigg(&(-1)^n\theta^{(n-1)}\mathbb{E}\left[F(\zeta_\rho)\right]\\ &+\sum_{j=1}^n(-1)^{n-j}(\theta+j)^{(n-1)}\sum_{1\le i_1<\ldots<i_j\le n}\mathbb{E}\left[F(\zeta_{\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_j}}})\right]\Bigg). \end{aligned} \tag{4.12}$$

Proof. The first claim follows from $\mathbb{E}[F(\zeta)c]=c\mathbb{E}[F(\zeta)]$ for all $c\in\mathbb{F}_0=\mathbb{R}$. We note that the inequality

$$(c_1+\ldots+c_k)^2\le k(c_1^2+\ldots+c_k^2)$$

holds for all $c_1,\dots,c_k\in\mathbb{R}$ and $k\in\mathbb{N}$. Let $n\in\mathbb{N}$. By the aforementioned inequality, the integral $\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x)$ is bounded by

$$\begin{aligned} &\frac{2(\theta+2n-1)^2}{n!^2}\int_{\mathbb{X}^n}\left(\theta^{(n-1)}\mathbb{E}\left[F(\zeta_\rho)\right]\right)^2\rho^{[n]}(\mathrm{d}\mathbf{x_n})\\ &+\frac{2(\theta+2n-1)^2}{n!^2}\int_{\mathbb{X}^n}\left(\sum_{j=1}^n(-1)^{n-j}(\theta+j)^{(n-1)}\sum_{1\le i_1<\ldots<i_j\le n}\mathbb{E}\left[F(\zeta_{\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_j}}})\right]\right)^2\rho^{[n]}(\mathrm{d}\mathbf{x_n}). \end{aligned}$$

The first integral in this expression is finite since $\rho^{[n]}(\mathbb{X})$ is finite and the integrand is finite by Jensen's inequality. An upper bound for the second integral is

$$\int_{\mathbb{X}^n} n\sum_{j=1}^n\left((\theta+j)^{(n-1)}\right)^2\binom{n}{j}\sum_{1\le i_1<\ldots<i_j\le n}\mathbb{E}\left[F(\zeta_{\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_j}}})^2\right]\rho^{[n]}(\mathrm{d}\mathbf{x_n}).$$

According to the Mecke-type equation in Theorem 3.7, this integral is equal to the finite value

$$n\sum_{j=1}^n\left((\theta+j)^{(n-1)}\right)^2\binom{n}{j}^2\theta^{(n)}\mathbb{E}\left[F(\zeta_\rho)^2\right].$$

Hence, we conclude that f_n is an element of $L^2(\rho^{[n]})$. Moreover, f_n is symmetric. In order to show $f_n\in\mathbb{H}_n$, we thus further need to establish

$$\int_{\mathbb{X}} f_n(\mathbf{x_{n-1}},x_n)\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)=0 \qquad \text{for } \rho^{[n-1]}\text{-a.e. } (x_1,\dots,x_{n-1})\in\mathbb{X}^{n-1}. \tag{4.13}$$

Let $(x_1,\dots,x_{n-1}) \in \mathbb{X}^{n-1}$. By definition of f_n and with the notation from (4.8), the integral in (4.13) equals

$$\begin{aligned}
&\frac{\theta+2n-1}{n!}\int_{\mathbb{X}}(-1)^n\theta^{(n-1)}T_{F,0}+\sum_{j=1}^{n}(-1)^{n-j}(\theta+j)^{(n-1)}\sum_{i_1<\ldots<i_j}T_{F,j}(x_{i_1},\dots,x_{i_j})\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\\
&=\frac{\theta+2n-1}{n!}\int_{\mathbb{X}}(-1)^n\theta^{(n-1)}T_{F,0}+(\theta+n)^{(n-1)}T_{F,n}(\mathbf{x_n})\\
&\qquad\qquad\qquad+\sum_{j=1}^{n-1}\sum_{i_1<\ldots<i_j}(-1)^{n-j}(\theta+j)^{(n-1)}T_{F,j}(x_{i_1},\dots,x_{i_j})\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n).
\end{aligned}$$

The integral in the first line is, according to Lemma 4.14, equal to

$$\begin{aligned}
&\int_{\mathbb{X}}(-1)^n\theta^{(n-1)}T_{F,0}+(\theta+n)^{(n-1)}T_{F,n}(\mathbf{x_n})\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\\
&=(-1)^n\theta^{(n-1)}(\theta+n-1)T_{F,0}+(\theta+n-1)(\theta+n)^{(n-1)}T_{F,n-1}(\mathbf{x_{n-1}})\\
&=(-1)^n\theta^{(n)}T_{F,0}+(\theta+n-1)^{(n)}T_{F,n-1}(\mathbf{x_{n-1}}),
\end{aligned}$$

while a closer examination of the last line reveals that the sum appearing there can be further decomposed, leading to

$$\begin{aligned}
&\int_{\mathbb{X}}\sum_{j=1}^{n-1}\sum_{1\le i_1<\ldots<i_j\le n}(-1)^{n-j}(\theta+j)^{(n-1)}T_{F,j}(x_{i_1},\dots,x_{i_j})\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\\
&=\sum_{j=1}^{n-1}\sum_{1\le i_1<\ldots<i_j\le n-1}(-1)^{n-j}(\theta+j)^{(n-1)}T_{F,j}(x_{i_1},\dots,x_{i_j})(\theta+n-1)\\
&\quad+\int_{\mathbb{X}}(-1)^{n-1}(\theta+1)^{(n-1)}T_{F,1}(x_n)\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n)\\
&\quad+\sum_{j=2}^{n-1}\sum_{1\le i_1<\ldots<i_{j-1}\le n-1}\int_{\mathbb{X}}(-1)^{n-j}(\theta+j)^{(n-1)}T_{F,j}(x_{i_1},\dots,x_{i_{j-1}},x_n)\,(\rho+\delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n).
\end{aligned}$$

We proceed with the computation of the two remaining integrals in the last two lines of the above expression. These are equal to

$$\begin{aligned}
&\int_{\mathbb{X}}(-1)^{n-1}(\theta+1)^{(n-1)}T_{F,1}(x_n)\,\rho(\mathrm{d}x_n)+\sum_{l=1}^{n}(-1)^{n-1}(\theta+1)^{(n-1)}T_{F,1}(x_l)\\
&+\sum_{j=2}^{n-1}\sum_{1\le i_1<\ldots<i_{j-1}\le n-1}\sum_{\substack{l=1,\\ l\notin\{i_1,\dots,i_{j-1}\}}}^{n-1}(-1)^{n-j}(\theta+j)^{(n-1)}T_{F,j}(x_{i_1},\dots,x_{i_{j-1}},x_l)\\
&+\sum_{j=2}^{n-1}\sum_{1\le i_1<\ldots<i_{j-1}\le n-1}\int_{\mathbb{X}}(-1)^{n-j}(\theta+j)^{(n-1)}T_{F,j}(x_{i_1},\dots,x_{i_{j-1}},x_n)\,(\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_{j-1}}})(\mathrm{d}x_n).
\end{aligned} \tag{4.14}$$

By the recursion from Lemma 4.14, we have

$$\begin{aligned}
&\int_{\mathbb{X}}(-1)^{n-1}(\theta+1)^{(n-1)}T_{F,1}(x_n)\,\rho(\mathrm{d}x_n)\\
&\quad+\sum_{j=2}^{n-1}\sum_{1\le i_1<\ldots<i_{j-1}\le n-1}\int_{\mathbb{X}}(-1)^{n-j}(\theta+j)^{(n-1)}T_{F,j}(x_{i_1},\dots,x_{i_{j-1}},x_n)\,(\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_{j-1}}})(\mathrm{d}x_n)\\
&=(-1)^{n-1}\theta^{(n)}T_{F,0}+\sum_{j=2}^{n-1}\sum_{1\le i_1<\ldots<i_{j-1}\le n-1}(-1)^{n-j}(\theta+j-1)^{(n)}T_{F,j-1}(x_{i_1},\dots,x_{i_{j-1}}).
\end{aligned}$$

For $j\in\{2,\dots,n-1\}$, the symmetry of $T_{F,j}$ and combinatorial counting (The sum on the left-hand side in the following expression involves selecting a $(j-1)$-tuple of distinct elements of the set $[n-1]$ and then inserting an additional integer $l \in [n-1]$ into it, where l is not in the set $\{i_1, \dots, i_{j-1}\}$. Because of the symmetry of $T_{F,j}$, the value of the function remains unchanged regardless of where l is inserted. This leads to the sum on the right-hand side, where all possible j-tuples of integers are considered. The factor j accounts for the j ways to insert l into the $(j-1)$-tuple.) yield

$$\sum_{1 \le i_1 < \ldots < i_{j-1} \le n-1} \sum_{\substack{l=1,\\ l \notin \{i_1,\ldots,i_{j-1}\}}}^{n-1} (-1)^{n-j} (\theta + j)^{(n-1)} T_{F,j}(x_{i_1}, \ldots, x_{i_{j-1}}, x_l) = j \sum_{1 \le i_1 < \ldots < i_j \le n-1} (-1)^{n-j} (\theta + j)^{(n-1)} T_{F,j}(x_{i_1}, \ldots, x_{i_j}).$$

Hence, (4.14) becomes

$$\begin{aligned}
&\sum_{l=1}^{n} (-1)^{n-1} ((\theta + 1)^{(n-1)} + (\theta + 1)^{(n)}) T_{F,1}(x_l) + (-1)^{n-1} \theta^{(n)} T_{F,0} \\
&+ \sum_{j=2}^{n-1} j \sum_{1 \le i_1 < \ldots < i_j \le n-1} (-1)^{n-j} (\theta + j)^{(n-1)} T_{F,j}(x_{i_1}, \ldots, x_{i_j}) \\
&+ \sum_{l=2}^{n-2} \sum_{1 \le i_1 < \ldots < i_l \le n-1} (-1)^{n-l-1} (\theta + l)^{(n)} T_{F,l}(x_{i_1}, \ldots, x_{i_l}),
\end{aligned}$$

which can be simplified to

$$\begin{aligned}
&\sum_{l=1}^{n} (-1)^{n} (\theta + 1)^{(n-1)} (\theta + n - 1) T_{F,1}(x_l) - (n-1)(\theta + n - 1)^{(n-1)} T_{F,n-1}(\mathbf{x_{n-1}}) \\
&- (-1)^n \theta^{(n)} T_{F,0} - \sum_{j=2}^{n-2} \sum_{1 \le i_1 < \ldots < i_j \le n-1} (-1)^{n-j} (\theta + j)^{(n-1)} (\theta + n - 1) T_{F,j}(x_{i_1}, \ldots, x_{i_j}).
\end{aligned}$$

Combining all our findings, we arrive at

$$\begin{aligned}
&\frac{n!}{\theta + 2n - 1} \int_{\mathbb{X}} f_n(\mathbf{x_{n-1}}, x_n) \, (\rho + \delta_{\mathbf{x_{n-1}}})(\mathrm{d}x_n) \\
=&(\theta + n - 1)^{(n)} T_{F,n-1}(\mathbf{x_{n-1}}) + \sum_{j=1}^{n-1} \sum_{1 \le i_1 < \ldots < i_j \le n-1} (-1)^{n-j} (\theta + j)^{(n-1)} T_{F,j}(x_{i_1}, \ldots, x_{i_j}) (\theta + n - 1) \\
&+ (-1)^n \theta^{(n)} T_{F,0} + \sum_{l=1}^{n} (-1)^n (\theta + 1)^{(n-1)} (\theta + n - 1) T_{F,1}(x_l) - (n-1)(\theta + n - 1)^{(n-1)} T_{F,n-1}(\mathbf{x_{n-1}}) \\
&- (-1)^n \theta^{(n)} T_{F,0} - \sum_{j=2}^{n-2} \sum_{1 \le i_1 < \ldots < i_j \le n-1} (-1)^{n-j} (\theta + j)^{(n-1)} (\theta + n - 1) T_{F,j}(x_{i_1}, \ldots, x_{i_j}).
\end{aligned}$$

By collecting like terms, this is equal to

$$\begin{aligned}
&T_{F,n-1}(\mathbf{x_{n-1}}) \Big((\theta + n - 1)^{(n)} - (\theta + n - 1)^{(n-1)} (\theta + n - 1) - (n-1)(\theta + n - 1)^{(n-1)} \Big) \\
&- \sum_{1 \le i_1 \le n-1} (-1)^n (\theta + 1)^{(n-1)} T_{F,1}(x_{i_1}) (\theta + n - 1) + \sum_{l=1}^{n} (-1)^{n-1} (\theta + 1)^{(n-1)} (\theta + n - 1) T_{F,1}(x_l) \\
&= 0.
\end{aligned}$$

Hence, f_n is an element of $\mathbb{H}_n$. As a final step, we show that $\int_{\mathbb{X}^n} f_n(x) \, \zeta^n(x)$ is indeed the projection of F onto $\mathbb{F}_n$ in $L^2(\mathbb{P})$. The almost sure uniqueness of the projection follows from Corollary 4.12. Let $g \in \mathbb{H}_n$. On the one hand, the Mecke-type equation from Theorem 3.7 yields

$$\mathbb{E}\left[F(\zeta_\rho) \int_{\mathbb{X}^n} g(x) \, \zeta_\rho^n(\mathrm{d}x) \right] = \frac{1}{\theta^{(n)}} \int_{\mathbb{X}^n} \mathbb{E}\left[F(\zeta_{\rho + \delta_{\mathbf{x_n}}}) \right] g(\mathbf{x_n}) \, \rho^{[n]}(\mathrm{d}\mathbf{x_n}). \tag{4.15}$$

On the other hand, Corollary 4.11 implies

$$\mathbb{E}\left[\int_{\mathbb{X}^n} f_n(y)\,\zeta_\rho^n(\mathrm{d}y)\int_{\mathbb{X}^n} g(x)\,\zeta_\rho^n(\mathrm{d}x)\right] = \frac{n!}{\theta^{(2n)}}\int_{\mathbb{X}^n} f_n(x)g(x)\,\rho^{[n]}(\mathrm{d}x).$$

By the definition of f_n, the right-hand side equals

$$\frac{n!}{\theta^{(2n)}}\int_{\mathbb{X}^n}\frac{\theta+2n-1}{n!}\left((-1)^n\theta^{(n-1)}T_{F,0}+\sum_{j=1}^{n}(-1)^{n-j}(\theta+j)^{(n-1)}\sum_{i_1<\ldots<i_j}T_{F,j}(x_{i_1},\ldots,x_{i_j})\right) g(\mathbf{x_n})\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}).$$

Since g is an element of $\mathbb{H}_n$ and thus Corollary 4.5 (for $k=0$) is applicable, this reduces to

$$\frac{n!}{\theta^{(2n)}}\int_{\mathbb{X}^n}\left(\frac{\theta+2n-1}{n!}(-1)^{n-n}(\theta+n)^{(n-1)}T_{F,n}(\mathbf{x_n})\right) g(\mathbf{x_n})\,\rho^{[n]}(\mathrm{d}\mathbf{x_n})$$

which is, because of $T_{F,n}(\mathbf{x_n}) = \mathbb{E}[F(\zeta_{\rho+\delta_{\mathbf{x_n}}})]$, in turn equal to (4.15). Therefore, $\int_{\mathbb{X}^n} f_n(x)\,\zeta^n(\mathrm{d}x)$ is indeed the orthogonal projection of F onto $\mathbb{F}_n$. □

A structural aspect of the projection formula is discussed in the subsequent remark.

Remark 4.16. The formula for the function f_n, $n\in\mathbb{N}$, from (4.12) in the previous statement can, provided the summand for $j=0$ is interpreted appropriately, i.e. the second sum is for $j=0$ interpreted as $\mathbb{E}[F(\zeta_\rho)]$, be written as follows

$$f_n(x_1,\ldots,x_n) = \frac{\theta+2n-1}{n!}\sum_{j=0}^{n}(-1)^{n-j}(\theta+j)^{(n-1)}\sum_{1\le i_1<\ldots<i_j\le n}\mathbb{E}\left[F(\zeta_{\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_j}}})\right] \tag{4.16}$$

for $\rho^{[n]}$-almost all $(x_1,\ldots,x_n)\in\mathbb{X}^n$. We thus see that the computation of f_n involves the integration of $F\in L^2(\zeta)$ with respect to the Palm measures (in the sense of the definition on p. 212 in Kallenberg (2017)) of ζ. A comparison with the Poisson case (c.f. e.g. Chapter 18 in Last and Penrose (2017)) reveals a structural similarity: In both settings, alternating sums of Palm expectations are considered. ◇

We finish this section with an example that gives the orthogonal projections of particular random variables.

Example 4.17. Let $m\in\mathbb{N}$ and $f\in L^2(\rho^{[m]})$. Then $\int_{\mathbb{X}^m} f(x)\,\zeta^m(\mathrm{d}x)$ is contained in the spaces $\mathbb{F}_i$, $i\in[m]_0$, that is

$$\int_{\mathbb{X}^m} f(x)\,\zeta^m(\mathrm{d}x)\in\bigoplus_{i=0}^{m}\mathbb{F}_i. \tag{4.17}$$

Proof. Let $F\colon\mathbf{M}(\mathbb{X})\to\mathbb{R}$ be given by $F(\mu) := \mu^m(f)$, where we used our notation $\mu^m(f)=\int_{\mathbb{X}^m} f(x)\,\mu^m(\mathrm{d}x)$ (cf. Section 2.4). By the assumption on f and Corollary 3.8 it holds $F\in L^2(\zeta)$. To begin with, we note that for each $n\in\mathbb{N}$ such that $n>m$ and for each $g\in\mathbb{H}_n$, we have

$$\mathbb{E}\left[\int_{\mathbb{X}^m} f(x)\,\zeta^m(\mathrm{d}x)\int_{\mathbb{X}^n} g(y)\,\zeta^n(\mathrm{d}y)\right] = \frac{1}{\theta^{(m+n)}}\int_{\mathbb{X}^{m+n}} f(x)g(y)\,\rho^{[m+n]}(\mathrm{d}(x,y)) = 0$$

due to Corollary 3.8 and Corollary 4.5 (with $k=0$). Hence, F is an element of the orthogonal complement of $\mathbb{F}_j$ $j>m$. Next, we calculate the projections of F onto $\mathbb{F}_k$, $k\in[m]_0$. Let $k\in[m]$. According to (4.16), the projection of F onto $\mathbb{F}_k$ is $\int_{\mathbb{X}^k} f_k(x)\,\zeta^k(\mathrm{d}x)$ where f_k is for $\rho^{[k]}$-almost all $(x_1,\ldots,x_k)\in\mathbb{X}^k$ given by

$$f_k(x_1,\ldots,x_k) = \frac{\theta+2k-1}{k!}\sum_{j=0}^{k}(-1)^{k-j}(\theta+j)^{(k-1)}\sum_{1\le i_1<\ldots<i_j\le k}\mathbb{E}\left[F(\zeta_{\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_j}}})\right].$$

By Corollary 3.8, this is equal to

$$\frac{\theta+2k-1}{k!}\left((-1)^k\frac{\theta^{(k-1)}}{\theta^{(m)}}\rho^{[m]}(f)+\sum_{j=1}^{k}(-1)^{k-j}\frac{(\theta+j)^{(k-1)}}{(\theta+j)^{(m)}}\sum_{1\le i_1<\ldots<i_j\le k}(\rho+\delta_{x_{i_1}}+\ldots+\delta_{x_{i_j}})^{[m]}(f)\right).$$

The projection of F onto $\mathbb{F}_0$ is

$$\mathbb{E}[F(\zeta)] = \mathbb{E}\left[\int_{\mathbb{X}^m} f(x)\,\zeta^m(\mathrm{d}x)\right] = \frac{1}{\theta^{(m)}}\rho^{[m]}(f).$$

We now show

$$F(\zeta) = \mathbb{E}[F(\zeta)] + \sum_{k=1}^{m} \int_{\mathbb{X}^k} f_k(x)\,\zeta^k(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.} \tag{4.18}$$

The right-hand side of (4.18) is

$$\frac{1}{\theta^{(m)}}\rho^{[m]}(f) + \sum_{k=1}^{m} \frac{\theta+2k-1}{k!}(-1)^k \frac{\theta^{(k-1)}}{\theta^{(m)}}\rho^{[m]}(f)$$
$$+ \sum_{k=1}^{m} \frac{\theta+2k-1}{k!} \sum_{j=1}^{k} (-1)^{k-j} \frac{(\theta+j)^{(k-1)}}{(\theta+j)^{(m)}} \binom{k}{j} \int_{\mathbb{X}^k} (\rho+\delta_{\mathbf{x_j}})^{[m]}(f)\,\zeta^k(\mathrm{d}\mathbf{x_k}).$$

By Proposition 3.11, for $k \in [m]$, $j \in [k]$ and $\mathbf{x_j} \in \mathbb{X}^j$, we have

$$\int_{\mathbb{X}^k} (\rho+\delta_{\mathbf{x_j}})^{[m]}(f)\,\zeta^k(\mathrm{d}\mathbf{x_k})$$
$$= \rho^{[m]}(f) + \sum_{r=1}^{m} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} \int_{\mathbb{X}^j}\int_{\mathbb{X}^{m-r}} f^j_{i_1,\dots,i_r}(\mathbf{x_j},\mathbf{z_{m-r}})\,\rho^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})\,\zeta^j(\mathrm{d}\mathbf{x_j}),$$

using the notation introduced in (3.6) and (3.7). Hence, the right-hand side of (4.18) becomes

$$\sum_{j=1}^{m}\sum_{k=j}^{m} \frac{(\theta+2k-1)}{(k-j)!j!}(-1)^{k-j}\frac{(\theta+j)^{(k-1)}}{(\theta+j)^{(m)}} \sum_{r=1}^{m} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} \int_{\mathbb{X}^j}\int_{\mathbb{X}^{m-r}} f^j_{i_1,\dots,i_r}(\mathbf{x_j},\mathbf{z_{m-r}})\rho^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})\zeta^j(\mathrm{d}\mathbf{x_j})$$
$$+ \left(\frac{1}{\theta^{(m)}} + \sum_{k=1}^{m}\frac{\theta+2k-1}{k!}(-1)^k\frac{\theta^{(k-1)}}{\theta^{(m)}} + \sum_{j=1}^{m}\sum_{k=j}^{m}\frac{\theta+2k-1}{(k-j)!j!}(-1)^{k-j}\frac{(\theta+j)^{(k-1)}}{(\theta+j)^{(m)}}\right)\rho^{[m]}(f). \tag{4.19}$$

From Lemma A.2, we obtain the following summation formula

$$\sum_{k=j}^{m} (-1)^{k-j}\frac{\theta+2k-1}{(k-j)!}(\theta+j)^{(k-1)} = (-1)^{m-j}\frac{(\theta+j)^{(m)}}{(m-j)!}$$

for each $m \in \mathbb{N}$ and $j \in [m-1]$. We now use this formula to simplify the second sum in the term inside the parentheses in the second line of (4.19). This yields the following simplified form of the entire term inside the parentheses

$$\frac{1}{\theta^{(m)}} + \sum_{k=1}^{m}\frac{\theta+2k-1}{k!}(-1)^k\frac{\theta^{(k-1)}}{\theta^{(m)}} + \sum_{j=1}^{m}\frac{1}{(m-j)!j!}(-1)^{m-j}.$$

Applying the binomial theorem to the last sum, we can further simplify this term to

$$\frac{1}{\theta^{(m)}} + \sum_{k=1}^{m}\frac{\theta+2k-1}{k!}(-1)^k\frac{\theta^{(k-1)}}{\theta^{(m)}} - \frac{(-1)^m}{m!}.$$

Lemma A.1 shows that this expression (and thus the entire second line in (4.19)) vanishes. Using the above stated summation formula from Lemma A.2 also for the sum in the first line of (4.19), the entire expression in (4.19) reduces to

$$\sum_{r=1}^{m} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} \sum_{j=1}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \int_{\mathbb{X}^j}\int_{\mathbb{X}^{m-r}} f^j_{i_1,\dots,i_r}(\mathbf{x_j},\mathbf{z_{m-r}})\,\rho^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})\,\zeta^j(\mathrm{d}\mathbf{x_j}).$$

Let $r \in [m]$. By the definition of $f^j_{i_1,\dots,i_r}$, $j \in [m]$, it holds

$$\sum_{j=1}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \int_{\mathbb{X}^j} \int_{\mathbb{X}^{m-r}} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} f^j_{i_1,\dots,i_r}(\mathbf{x_j}, \mathbf{z_{m-r}})\, \rho^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})\, \zeta^j(\mathrm{d}\mathbf{x_j})$$
$$= \sum_{j=1}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \sum_{1\le l_1\le\dots\le l_r\le j} \int_{\mathbb{X}^j} \int_{\mathbb{X}^{m-r}} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} f_{i_1,\dots,i_r}(x_{l_1},\dots,x_{l_r}, \mathbf{z_{m-r}})\, \rho^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})\, \zeta^j(\mathrm{d}\mathbf{x_j}). \tag{4.20}$$

We now differentiate based on the number of identical arguments in the integrand, that is, we decompose the set of all sequences $(l_1,\dots,l_r) \in \mathbb{N}^r$ with $1 \le l_1 \le \dots \le l_r \le j$ for $j \in [m]$ according to the number of repeated entries. To this end, let $j \in [m]$, $k \in [j \wedge r]$ and $\lambda_1,\dots,\lambda_k \in \mathbb{N}$ with

$$\sum_{i=1}^{k} \lambda_i = r.$$

We define λ as the multiset (a multiset is a generalisation of a set that allows multiple occurrences of the same element; formally, a multiset over a set S can be defined as a function $m\colon S \to \mathbb{N}_0$, where $m(\nu)$ gives the number of times $\nu \in S$ appears in the multiset) containing $\lambda_1,\dots,\lambda_k$. A sequence $(l_1,\dots,l_r) \in \mathbb{N}^r$ is said to have multiplicity structure λ if there exist exactly k distinct values among the entries and each distinct value appears $\lambda_1,\dots,\lambda_k$ times, respectively. Formally, this means that the multiset consisting of $l_1,\dots,l_r$ can be partitioned into k equivalence classes of equal entries, with cardinalities $\lambda_1,\dots,\lambda_k$. Moreover, we set

$$B_\lambda := \{(l_1,\dots,l_r) \in \mathbb{N}^r\colon\ 1 \le l_1 \le \dots \le l_r \le j \text{ and the multiset with elements } l_1,\dots,l_r \text{ has multiplicity structure } \lambda\}.$$

We observe that the number of elements of B_λ is

$$\binom{j}{k} \frac{k!}{\prod_{i=1}^{M(\lambda)} m_\lambda(\nu_i)!},$$

where $M(\lambda)$ is the number of distinct values in λ, which are denoted by $\nu_1,\dots,\nu_{M(\lambda)}$, each occurring $m_\lambda(\nu_i)$ times in λ, $i \in [M(\lambda)]$. We thus have the disjoint decomposition

$$\{(l_1,\dots,l_r) \in \mathbb{N}^r\colon\ 1 \le l_1 \le \dots \le l_r \le j\} = \bigcup_{k=1}^{j\wedge r} \bigcup_{\substack{\lambda \text{ multiset of } k \text{ natural numbers} \\ \text{that sum up to } r}} B_\lambda$$

and (4.20) becomes

$$\sum_{j=1}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \sum_{k=1}^{j\wedge r} \sum_{\lambda} \sum_{(l_1,\dots,l_r)\in B_\lambda} \iint \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} f_{i_1,\dots,i_r}(x_{l_1},\dots,x_{l_r}, \mathbf{z_{m-r}})\, \rho^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})\, \zeta^j(\mathrm{d}\mathbf{x_j}), \tag{4.21}$$

where $\sum_\lambda$ is used to express the sum over all multisets λ consisting of k natural numbers that sum to r. Note that, for $\mathbf{z_{m-r}} \in \mathbb{X}^{m-r}$, the function

$$\mathbb{X}^r \ni (x_1,\dots,x_r) \mapsto \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} f_{i_1,\dots,i_r}(x_1,\dots,x_r, \mathbf{z_{m-r}})$$

is symmetric. Hence, for fixed $j \in [m]$, $k \in [j \wedge r]$, a multiset λ with elements $\lambda_1,\dots,\lambda_k$ with sum r and $(l_1,\dots,l_r) \in B_\lambda$, the integral in the corresponding summand in (4.21) simplifies to

$$\int_{\mathbb{X}^k} \int_{\mathbb{X}^{m-r}} \sum_{(i_1,\dots,i_r)\in[m]^{[r]}} f_{i_1,\dots,i_r}(\underbrace{x_1,\dots,x_1}_{\lambda_1 \text{ times}},\dots,\underbrace{x_k,\dots,x_k}_{\lambda_k \text{ times}}, \mathbf{z_{m-r}})\, \rho^{[m-r]}(\mathrm{d}\mathbf{z_{m-r}})\, \zeta^k(\mathrm{d}\mathbf{x_k}) =: I_\lambda.$$

We note that I_λ now depends only on λ and, consequently, on k and r. Inserting these expressions into (4.20) yields

$$\sum_{j=1}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \sum_{k=1}^{j\wedge r} \sum_{\lambda} \sum_{(l_1,\dots,l_r)\in B_\lambda} I_\lambda = \sum_{k=1}^{r} \sum_{\lambda} I_\lambda \sum_{j=k}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \binom{j}{k} \frac{k!}{\prod_{i=1}^{M(\lambda)} m_\lambda(\nu_i)!}.$$

We now evaluate

$$\sum_{j=k}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \binom{j}{k} k!$$

for fixed $k \in [r]$. The binomial theorem yields

$$\sum_{j=k}^{m} \frac{(-1)^{m-j}}{(m-j)!j!} \binom{j}{k} k! = \sum_{j=k}^{m} \frac{(-1)^{m-j}}{(m-j)!(j-k)!} = \sum_{i=0}^{m-k} \frac{(-1)^{m-k-i}}{(m-k-i)!i!} = \frac{1}{(m-k)!}(1-1)^{m-k}.$$

This expression is zero unless $k = m$ is satisfied, which can only occur when $r = m$ and $j = m$. If this is the case, the corresponding multiset λ consists of m times the value 1, i.e. $M(\lambda) = 1$, $m_\lambda(1) = m$ and $B_\lambda = \{(1, 2, \dots, m-1, m)\}$. Putting everything together and using the symmetry of ζ^m for the final step, we conclude, $\mathbb{P}$-a.s., that

$$\mathbb{E}[F(\zeta)] + \sum_{k=1}^{m} \int_{\mathbb{X}^k} f_k(x)\, \zeta^k(\mathrm{d}x) = \frac{1}{m!} \int_{\mathbb{X}^m} \sum_{(i_1,\dots,i_m)\in[m]^{[m]}} f_{i_1,\dots,i_m}(\mathbf{x_m})\, \zeta^m(\mathrm{d}\mathbf{x_m}) = \int_{\mathbb{X}^m} f(z)\, \zeta^m(\mathrm{d}z). \quad \square$$

4.4. The chaos expansion

Having completed all necessary preparations, we can prove the chaos expansion in this section.

Theorem 4.18. *Every $F \in L^2(\zeta)$ admits a unique representation*

$$F(\zeta) = \mathbb{E}[F(\zeta)] + \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} f_n(x)\, \zeta^n(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.}, \tag{4.22}$$

where the convergence is in $L^2(\mathbb{P})$ and f_n, $n \in \mathbb{N}$, is $\rho^{[n]}$-a.e. given by (4.16).

Proof. Uniqueness of the projections is a consequence of Lemma 4.12. By Proposition 4.15, the projection of F onto $\mathbb{F}_n$, $n \in \mathbb{N}$, is $\int_{\mathbb{X}^n} f_n(x)\, \zeta^n(\mathrm{d}x)$ and the projection of F onto $\mathbb{F}_0$ is $\mathbb{E}[F(\zeta)]$. It remains to show

$$\{F(\zeta)\colon\ F \in L^2(\zeta)\} = \bigoplus_{n=0}^{\infty} \mathbb{F}_n.$$

As a direct sum of closed subspaces, $\bigoplus_{n=0}^{\infty} \mathbb{F}_n$ is closed in $L^2(\mathbb{P})$. It contains, by definition, random variables of the form

$$\int_{\mathbb{X}^n} h(x)\, \zeta^n(\mathrm{d}x),$$

with $h \in \mathbb{H}_n$ and $n \in \mathbb{N}$. Example 4.17 further shows that

$$\int_{\mathbb{X}^n} g(x)\, \zeta^n(\mathrm{d}x) \in \bigoplus_{i=0}^{n} \mathbb{F}_i$$

for each $g \in L^2(\rho^{[n]})$, $n \in \mathbb{N}$. The linear hull of such random variables is dense in $\{F(\zeta)\colon\ F \in L^2(\zeta)\}$ according to Lemma 2 from Peccati (2008). $\square$

The functions f_n, $n \in \mathbb{N}$, in the above expansion (4.22) are also called *kernel functions*. Moreover, given $F, G \in L^2(\zeta)$ with kernel functions f_n and g_n, $n \in \mathbb{N}$, respectively, we obtain the isometry formula

$$\mathbb{E}\left[F(\zeta)G(\zeta)\right] = \mathbb{E}\left[F(\zeta)\right] \mathbb{E}[G(\zeta)] + \sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x) g_n(x)\, \rho^{[n]}(\mathrm{d}x). \tag{4.23}$$

Building on the chaos decomposition, the following remark addresses the associated Fock space.

Remark 4.19. Let $\mathbb{H}$ be the vector space of all sequences $g = (g_k)_{k\in\mathbb{N}_0}$ such that $g_0 \in \mathbb{R}$, $g_k \in L^2(\rho^{[k]})$ for all $k \in \mathbb{N}$ and

$$\sum_{k=0}^{\infty} \frac{k!}{\theta^{(2k)}} \int_{\mathbb{X}^k} g_k(x)^2 \, \rho^{[k]}(\mathrm{d}x) < \infty.$$

Equipped with the norm

$$\|g\|_{\mathbb{H}} = \left(\sum_{k=0}^{\infty} \frac{k!}{\theta^{(2k)}} \int_{\mathbb{X}^k} g_k(x)^2 \, \rho^{[k]}(\mathrm{d}x) \right)^{\frac{1}{2}}, \qquad g = (g_k)_{k\in\mathbb{N}_0} \in \mathbb{H},$$

the space $\mathbb{H}$ becomes a Hilbert space as a countable direct sum of Hilbert spaces. Moreover, the space $\mathbb{H}$ is isometrically isomorphic to $\bigoplus_{k=0}^{\infty} \mathbb{F}_k$ via the mapping

$$(f_k)_{k\in\mathbb{N}_0} \mapsto \sum_{k=0}^{\infty} \int_{\mathbb{X}^k} f_k(x) \, \zeta^k(\mathrm{d}x).$$

◇

CHAPTER 5

Malliavin operators

We now turn to a powerful set of analytical tools known as Malliavin calculus or the stochastic calculus of variations. Originally developed by Paul Malliavin in Malliavin (1976) as an infinite-dimensional integration by parts technique in the Brownian motion case, this theory has since evolved into an important component of modern stochastic analysis for different processes and has found widespread applications (cf. e.g. Privault and Schoutens (2002) for Rademacher sequences, Nualart (2006) for the Wiener space and fractional Brownian motion, Di Nunno, Øksendal and Proske (2009) for Lévy processes and Last (2016) for Poisson processes).

In this chapter, the chaos expansion is employed to construct the fundamental operators of Malliavin calculus: a gradient ∇, a divergence δ and a generator L. Each of these operators is discussed in a dedicated section. Throughout the chapter, let $(\mathbb{X}, \mathcal{X}, \rho)$ be a finite measure space with $\theta := \rho(\mathbb{X}) > 0$ and denote by ζ a Dirichlet process on $\mathbb{X}$ with parameter measure ρ.

5.1. The gradient

In this section, we construct a gradient operator on a subset of $L^2(\zeta)$.

Let C_ζ be the Campbell measure of ζ, i.e. the probability measure on $\Omega \times \mathbb{X}$ defined by

$$C_\zeta(A) := \int_\Omega \int_\mathbb{X} \mathbb{1}_A(\omega, x)\, \zeta(\omega, \mathrm{d}x)\, \mathbb{P}(\mathrm{d}\omega), \quad A \in \mathcal{A} \otimes \mathcal{X}.$$

We begin by specifying the set of random variables for which the gradient will be defined.

Definition 5.1. Let $\mathrm{dom}(\nabla)$ denote the set of all $F \in L^2(\zeta)$ with chaos expansion (4.22) such that the kernel functions additionally satisfy

$$\sum_{n=1}^{\infty} \frac{(\theta + n - 1)n n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x) < \infty. \tag{5.1}$$

We note that, since random variables of the form $\int_{\mathbb{X}^n} f(x)\, \zeta^n(\mathrm{d}x)$ with $n \in \mathbb{N}$ and $f \in L^2(\rho^{[n]})$ are elements of $\mathrm{dom}(\nabla)$, this set is dense in $L^2(\zeta)$ (cf. Lemma 2 from Peccati (2008)).

The subsequent lemma lays the foundation for the definition of the gradient.

Lemma 5.2. *Let $F \in \mathrm{dom}(\nabla)$ with chaos expansion* (4.22). *Then*

$$\Omega \times \mathbb{X} \ni (\omega, x) \mapsto \sum_{n=1}^{\infty} n \left(\int_{\mathbb{X}^{n-1}} f_n(x, y_1, \dots, y_{n-1})\, \zeta^{n-1}(\omega, \mathrm{d}(y_1, \dots, y_{n-1})) - \int_{\mathbb{X}^n} f_n(y)\, \zeta^n(\omega, \mathrm{d}y) \right)$$

converges in $L^2(C_\zeta)$.

Proof. Given $n \in \mathbb{N}$, let $H_n \colon \Omega \times \mathbb{X} \to \mathbb{R}$ be defined by

$$\begin{aligned} H_n(\omega, x) &:= \int_{\mathbb{X}^{n-1}} f_n(x, \mathbf{y_{n-1}})\, \zeta^{n-1}(\omega, \mathrm{d}\mathbf{y_{n-1}}) - \int_{\mathbb{X}^n} f_n(\mathbf{y})\, \zeta^n(\omega, \mathrm{d}\mathbf{y}), \quad n > 1, \\ H_1(\omega, x) &:= f_1(x) - \int_\mathbb{X} f_1(y)\, \zeta(\omega, \mathrm{d}y). \end{aligned}$$

Let $m, n \in \mathbb{N}$. It holds

$$\int_{\Omega \times \mathbb{X}} H_m(\omega, x) H_n(\omega, x)\, C_\zeta(\mathrm{d}(\omega, x)) = \mathbb{E}\left[\int_\mathbb{X} H_m(x) H_n(x)\, \zeta(\mathrm{d}x) \right]$$

$$= \mathbb{E}\left[\int_{\mathbb{X}} \left(\int_{\mathbb{X}^{m-1}} f_m(x, \mathbf{y_{m-1}})\, \zeta^{m-1}(\mathrm{d}\mathbf{y_{m-1}}) - \int_{\mathbb{X}^m} f_m(\mathbf{y})\, \zeta^m(\mathrm{d}\mathbf{y})\right) \left(\int_{\mathbb{X}^{n-1}} f_n(x, \mathbf{z_{n-1}})\, \zeta^{n-1}(\mathrm{d}\mathbf{z_{n-1}}) - \int_{\mathbb{X}^n} f_n(\mathbf{z})\, \zeta^n(\mathrm{d}\mathbf{z})\right) \zeta(\mathrm{d}x)\right].$$

By Corollary 3.8, this expectation is equal to

$$\frac{1}{\theta^{(m+n-1)}} \int_{\mathbb{X}^{m+n-1}} f_m(x, \mathbf{y_{m-1}}) f_n(x, \mathbf{z_{n-1}})\, \rho^{[m+n-1]}(\mathrm{d}(x, \mathbf{y_{m-1}}, \mathbf{z_{n-1}})) - \frac{1}{\theta^{(m+n)}} \int_{\mathbb{X}^{m+n}} f_m(\mathbf{y_m}) f_n(\mathbf{z_n})\, \rho^{[m+n]}(\mathrm{d}(\mathbf{y_m}, \mathbf{z_n})).$$

According to Corollary 4.5, this reduces to

$$\mathbb{1}_{\{m=n\}} \left(\frac{(n-1)!}{\theta^{(2n-1)}} \int_{\mathbb{X}^n} f_n(\mathbf{x})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x}) - \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(\mathbf{x})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x})\right) = \mathbb{1}_{\{m=n\}} \frac{(n-1)!}{\theta^{(2n)}} (\theta + n - 1) \int_{\mathbb{X}^n} f_n(\mathbf{x})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x}).$$

Hence, for $n_0 \in \mathbb{N}$, we obtain

$$\begin{aligned} \int_{\Omega\times\mathbb{X}} \left(\sum_{n=1}^{n_0} nH_n(\omega, x)\right)^2 C_\zeta(\mathrm{d}(\omega, x)) &= \sum_{n=1}^{n_0} \sum_{m=1}^{n_0} \mathbb{E}\left[\int_{\mathbb{X}} nmH_n(x)H_m(x)\, \zeta(\mathrm{d}x)\right] \\ &= \sum_{n=1}^{n_0} \frac{n^2(n-1)!}{\theta^{(2n)}} (\theta + n - 1) \int_{\mathbb{X}^n} f_n(\mathbf{x})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x}). \end{aligned}$$

Since F is an element of $\mathrm{dom}(\nabla)$, the series

$$\sum_{n=1}^{\infty} \frac{nn!}{\theta^{(2n)}} (\theta + n - 1) \int_{\mathbb{X}^n} f_n(\mathbf{x})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x})$$

converges. We conclude that

$$\left(\sum_{n=1}^{n_0} nH_n\right)_{n_0\in\mathbb{N}}$$

is a Cauchy sequence in $L^2(C_\zeta)$ and therefore convergent. Moreover, the following identity holds

$$\mathbb{E}\left[\int_{\mathbb{X}} \left(\sum_{n=1}^{\infty} nH_n(x)\right)^2 \zeta(\mathrm{d}x)\right] = \sum_{n=1}^{\infty} \frac{nn!}{\theta^{(2n)}} (\theta + n - 1) \int_{\mathbb{X}^n} f_n(\mathbf{x})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x}). \tag{5.2}$$

□

We now introduce the *gradient*.

Definition 5.3. Let $\nabla\colon \mathrm{dom}(\nabla) \to L^2(C_\zeta)$ be defined by

$$(\nabla F)(\omega, x) \coloneqq \sum_{n=1}^{\infty} n \left(\int_{\mathbb{X}^{n-1}} f_n(x, y_1, \dots, y_{n-1})\, \zeta^{n-1}(\omega, \mathrm{d}(y_1, \dots, y_{n-1})) - \int_{\mathbb{X}^n} f_n(y)\, \zeta^n(\omega, \mathrm{d}y)\right),$$

where the chaos expansion of $F \in \mathrm{dom}(\nabla)$ is given by (4.22).

We work with a measurable version of the gradient (This follows from the fact that convergence in $L^2(C_\zeta)$ implies the existence of a subsequence that converges C_ζ-a.e. Since each partial sum is measurable, the pointwise limit of any such subsequence is measurable as well. Therefore, the limit function, which is equal to the L^2-limit, is measurable.) and occasionally suppress the dependence on Ω, writing simply $\nabla_x F$, $x \in \mathbb{X}$, for the random variable $\omega \mapsto (\nabla F)(\omega, x)$.

The following corollary gathers two immediate properties of the gradient.

Corollary 5.4. *Let $F, G \in \mathrm{dom}(\nabla)$ and denote the kernel functions in their chaos expansions from* (4.22) *by f_n and g_n, $n \in \mathbb{N}$, respectively. Then*

$$\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)(\nabla_x G)\, \zeta(\mathrm{d}x)\right] = \sum_{n=1}^{\infty} \frac{nn!}{\theta^{(2n)}}(\theta + n - 1) \int_{\mathbb{X}^n} f_n(x) g_n(x)\, \rho^{[n]}(\mathrm{d}x). \tag{5.3}$$

Moreover, the gradient is centred in the sense that

$$\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x) = 0, \quad \mathbb{P}\text{-a.s.} \tag{5.4}$$

Proof. The identity (5.2) implies (5.3) (note that the summands in the defining series of the gradient are pairwise orthogonal in $L^2(C_\zeta)$). To see the second claim, we may define for $m \in \mathbb{N}$ and $x \in \mathbb{X}$ the random variable H_x^m by truncating the infinite series defining $\nabla_x F$ at m. By the definition and the fact that ζ is a probability measure, we then have $\int_{\mathbb{X}} H_x^m\, \zeta(\mathrm{d}x) = 0$. Jensen's inequality (for the probability measure ζ) yields

$$\mathbb{E}\left[\left(\int_{\mathbb{X}} (H_x^m - \nabla_x F)\, \zeta(\mathrm{d}x)\right)^2\right] \le \mathbb{E}\left[\int_{\mathbb{X}} (H_x^m - \nabla_x F)^2\, \zeta(\mathrm{d}x)\right]$$

The assertion is now a consequence of the convergence of $(H_m)_{m\in\mathbb{N}}$ to ∇F in $L^2(C_\zeta)$ established in Lemma 5.2. □

The next result shows that the gradient operator is closed. This property will play a key role in Chapter 6.

Lemma 5.5. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{dom}(\nabla)$ and assume that $(\nabla F_n)_{n\in\mathbb{N}}$ forms a Cauchy sequence in $L^2(C_\zeta)$, i.e.*

$$\lim_{m,n\to\infty} \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F_m - \nabla_x F_n)^2\, \zeta(\mathrm{d}x)\right] = 0.$$

Then there exists $F \in \mathrm{dom}(\nabla)$ with

$$\lim_{n\to\infty} \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F - \nabla_x F_n)^2\, \zeta(\mathrm{d}x)\right] = 0. \tag{5.5}$$

If additionally $F_n \to \widetilde{F}$ as $n \to \infty$ in $L^2(\zeta)$ for some $\widetilde{F} \in L^2(\zeta)$, the limit function F in (5.5) *can be chosen in such a way that $F = \widetilde{F}$ holds $\mathbb{P}$-almost surely.*

Proof. Let for each $n \in \mathbb{N}$ the chaos expansion (4.22) of $F_n \in L^2(\zeta)$ be given by

$$F_n(\zeta) = \mathbb{E}[F_n(\zeta)] + \sum_{k=1}^{\infty} \int_{\mathbb{X}^k} f_{n,k}(x)\, \zeta^k(\mathrm{d}x).$$

The assumed convergence and (5.3) yield

$$\begin{aligned} 0 &= \lim_{m,n\to\infty} \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F_m - \nabla_x F_n)^2\, \zeta(\mathrm{d}x)\right] \\ &= \lim_{m,n\to\infty} \sum_{k=1}^{\infty} \frac{kk!}{\theta^{(2k)}}(\theta + k - 1) \int_{\mathbb{X}^k} (f_{m,k}(x) - f_{n,k}(x))^2\, \rho^{[k]}(\mathrm{d}x). \end{aligned}$$

Similar to the Fock space in Remark 4.19, let $\widetilde{\mathbb{H}}$ be the vector space of all sequences $g = (g_k)_{k\in\mathbb{N}}$ such that $g_k \in L^2(\rho^{[k]})$ for all $k \in \mathbb{N}$ and

$$\sum_{k=1}^{\infty} \frac{kk!(\theta + k - 1)}{\theta^{(2k)}} \int_{\mathbb{X}^k} g_k(x)^2\, \rho^{[k]}(\mathrm{d}x) < \infty.$$

In contrast to Remark 4.19, we now include an additional weighting factor in each summand. Equipped with the norm

$$\|g\|_{\widetilde{\mathbb{H}}} = \left(\sum_{k=1}^{\infty} \frac{kk!(\theta + k - 1)}{\theta^{(2k)}} \int_{\mathbb{X}^k} g_k(x)^2\, \rho^{[k]}(\mathrm{d}x)\right)^{\frac{1}{2}}, \qquad g = (g_k)_{k\in\mathbb{N}} \in \widetilde{\mathbb{H}},$$

the space $\widetilde{\mathbb{H}}$ becomes a Hilbert space as a countable direct sum of Hilbert spaces. The sequence $(f_n)_{n\in\mathbb{N}} = ((f_{n,k})_{k\in\mathbb{N}})_{n\in\mathbb{N}}$ from above is a Cauchy sequence in $\widetilde{\mathbb{H}}$. Hence, there exists a limit $f = (f_k)_{k\in\mathbb{N}}$ in $\widetilde{\mathbb{H}}$, i.e.

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{kk!(\theta+k-1)}{\theta^{(2k)}} \int_{\mathbb{X}^k} (f_{n,k}(x) - f_k(x))^2 \, \rho^{[k]}(\mathrm{d}x) = 0.$$

Since for each $k \in \mathbb{N}$, we have $f_{n,k} \to f_k$ as $n \to \infty$ and $f_{n,k} \in \mathbb{H}_k$ for all $n \in \mathbb{N}$ the fact that $\mathbb{H}_k$ is closed by Lemma 4.3 implies $f_k \in \mathbb{H}_k$. Thus, as $f \in \widetilde{\mathbb{H}}$, for each constant $c \in \mathbb{R}$, the function F_c with chaos expansion

$$F_c(\zeta) = c + \sum_{k=1}^{\infty} \int_{\mathbb{X}^k} f_k(x)\, \zeta^k(\mathrm{d}x),$$

belongs to $\mathrm{dom}(\nabla)$ and satisfies (5.5).

If $(F_n)_{n\in\mathbb{N}}$ additionally converges to some $\widetilde{F}$ in $L^2(\zeta)$ and the kernel functions of $\widetilde{F}$ are denoted by $\widetilde{f}_n$, $n \in \mathbb{N}$, we have by (4.23) that

$$\mathbb{E}[(F_n(\zeta) - \widetilde{F}(\zeta))G(\zeta)] = \sum_{k=0}^{\infty} \frac{k!}{\theta^{(2k)}} \int_{\mathbb{X}^k} (f_{n,k}(x) - \widetilde{f}_k(x)) g_k(x)\, \rho^{[k]}(\mathrm{d}x)$$

for each $n \in \mathbb{N}$ and every $G \in L^2(\zeta)$ with kernel functions g_k, $k \in \mathbb{N}$. By the convergence $F_n \to \widetilde{F}$, $n \to \infty$, and the Cauchy–Schwarz inequality, we thus obtain

$$\int_{\mathbb{X}^k} (f_{n,k}(x) - \widetilde{f}_k(x)) g(x)\, \rho^{[k]}(\mathrm{d}x) \to 0, \quad n \to \infty,$$

for each $k \in \mathbb{N}$ and $g \in \mathbb{H}_k$, implying $f_k = \widetilde{f}_k$, $\rho^{[k]}$-a.e. for each $k \in \mathbb{N}$. Choosing $c = \mathbb{E}[\widetilde{F}(\zeta)]$ yields that $F_{\mathbb{E}[\widetilde{F}(\zeta)]} = \widetilde{F}$ holds $\mathbb{P}$-a.s. □

5.2. The divergence

In this section, we introduce the adjoint in $L^2(C_\zeta)$ of the gradient from the previous section.

Definition 5.6. Let $\mathrm{dom}(\delta)$ denote the set of all measurable functions $H \colon \mathbf{M}(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}$ such that $\mathbb{E}[\int_{\mathbb{X}} H(\zeta, x)^2 \, \zeta(\mathrm{d}x)] < \infty$ and for which there exists a constant $c \geq 0$ such that

$$\left| \mathbb{E}\left[\int_{\mathbb{X}} H(x) \nabla_x F \, \zeta(\mathrm{d}x) \right] \right| \leq c\, \mathbb{E}[F(\zeta)^2]^{\frac{1}{2}} \tag{5.6}$$

for all $F \in \mathrm{dom}(\nabla)$.

The next lemma provides the necessary groundwork for defining the divergence.

Lemma 5.7. *Let $H \in \mathrm{dom}(\delta)$. Then there exists an $\mathbb{P}$-a.s. uniquely determined $\sigma(\zeta)$-measurable $\delta(H) \in L^2(\mathbb{P})$ fulfilling*

$$\mathbb{E}\left[\delta(H)F\right] = \mathbb{E}\left[\int_{\mathbb{X}} H(x) \nabla_x F \, \zeta(\mathrm{d}x) \right] \tag{5.7}$$

for all $F \in \mathrm{dom}(\nabla)$.

Proof. By condition (5.6), the linear mapping $\mathrm{dom}(\nabla) \ni F \mapsto \mathbb{E}\left[\int_{\mathbb{X}} H(x)\nabla_x F\, \zeta(\mathrm{d}x)\right]$ is continuous and can thus be extended to a linear mapping from $L^2(\zeta)$ to $\mathbb{R}$. The Riesz representation theorem yields a unique element $\delta(H) \in L^2(\mathbb{P})$ satisfying (5.7). □

Definition 5.8. The operator δ is called *divergence operator* and equation (5.7) is referred to as *integration by parts* or *partial integration.*

We note that the operator δ is linear. Choosing $F \equiv 1$ in (5.7) shows $\mathbb{E}[\delta(H)] = 0$ since the kernel functions from the chaos decomposition (4.22) of F are given by $f_n = 0$, $n \in \mathbb{N}$, and thus, by definition, $\nabla F = 0$.

For certain functions in $L^2(C_\zeta)$, it is possible to define an operator δ' based on the chaos expansion. This alternative construction will be pursued in the subsequent analysis. In particular, we will show that δ and δ' coincide on the set $\mathrm{dom}(\delta')$. We begin with a lemma that provides a criterion for the convergence of a series in $L^2(C_\zeta)$. The functions that can be constructed in this manner will form the domain of the operator δ'.

Lemma 5.9. *Let $h_n\colon \mathbb{X}^{n+1} \to \mathbb{R}$, $n \in \mathbb{N}_0$, be measurable functions such that $h_n(x,\cdot) \in \mathbb{H}_n$ for each $x \in \mathbb{X}$ and*

$$\sum_{n=1}^{\infty} \frac{(\theta+n)^2(n+1)!}{\theta^{(2n+2)}} \int_{\mathbb{X}^n} h_n(z)^2\, \rho^{[n+1]}(\mathrm{d}z) < \infty. \tag{5.8}$$

Then the series

$$(\omega, x) \mapsto h_0(x) + \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} h_n(x,y)\, \zeta^n(\omega, \mathrm{d}y) \tag{5.9}$$

converges in $L^2(C_\zeta)$.

Proof. Let $m, n \in \mathbb{N}$ with $m \geq n$. Define $H_n\colon \Omega \times \mathbb{X} \to \mathbb{R}$ by

$$H_n(\omega, x) := \int_{\mathbb{X}^n} h_n(x, \mathbf{y_n})\, \zeta^n(\omega, \mathrm{d}\mathbf{y_n}).$$

By Corollary 3.8, we obtain

$$\int_{\Omega\times\mathbb{X}} H_m(\omega,x) H_n(\omega,x)\, C_\zeta(\mathrm{d}(\omega,x)) = \frac{1}{\theta^{(m+n+1)}} \int_{\mathbb{X}^{m+n+1}} h_m(x,\mathbf{y_m}) h_n(x,\mathbf{z_n})\, \rho^{[m+n+1]}(\mathrm{d}(x,\mathbf{y_m},\mathbf{z_n})).$$

According to Corollary 4.6, we have

$$\begin{aligned}
&\int_{\mathbb{X}^{m+n+1}} h_m(x,\mathbf{y_m}) h_n(x,\mathbf{z_n})\, \rho^{[m+n+1]}(\mathrm{d}(x,\mathbf{y_m},\mathbf{z_n})) \\
&= \mathbb{1}_{\{m=n\}} m! \left(\int_{\mathbb{X}^{m+1}} h_m(\mathbf{x_{m+1}})^2\, \rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}}) + \sum_{r=1}^{m} \int_{\mathbb{X}^m} h_m(x_r, \mathbf{x_m})^2\, \rho^{[m]}(\mathrm{d}\mathbf{x_m}) \right) \\
&\quad + \mathbb{1}_{\{m=n+1\}} (n+1)! \int_{\mathbb{X}^{n+1}} h_{n+1}(x_1, \mathbf{x_{n+1}}) h_n(\mathbf{x_{n+1}})\, \rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}).
\end{aligned}$$

Hence, for $n_0, n_1 \in \mathbb{N}$ with $n_0 \leq n_1$, it follows

$$\begin{aligned}
&\mathbb{E}\left[\int_{\mathbb{X}} \left(\sum_{n=n_0}^{n_1} \int_{\mathbb{X}^n} h_n(x,y)\, \zeta^n(\mathrm{d}y) \right)^2 \zeta(\mathrm{d}x) \right] = \sum_{m,n=n_0}^{n_1} \mathbb{E}\left[\int_{\mathbb{X}} H_m(x) H_n(x)\, \zeta(\mathrm{d}x) \right] \\
&= \sum_{n=n_0}^{n_1} \frac{n!}{\theta^{(2n+1)}} \left(\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\, \rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}) + \sum_{r=1}^{n} \int_{\mathbb{X}^n} h_n(x_r, \mathbf{x_n})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \right) \\
&\quad + \sum_{n=n_0}^{n_1-1} \frac{(n+1)!}{\theta^{(2n+2)}} \int_{\mathbb{X}^{n+1}} h_{n+1}(x_1, \mathbf{x_{n+1}}) h_n(\mathbf{x_{n+1}})\, \rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}) =: S_1 + S_2.
\end{aligned}$$

We treat the sums S_1 and S_2 individually and establish upper bounds for both. On the one hand, because of

$$\begin{aligned}
\sum_{r=1}^{n} \int_{\mathbb{X}^n} h_n(x_r, \mathbf{x_n})^2\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) &= \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(y, \mathbf{x_n})^2\, \delta_{\mathbf{x_n}}(\mathrm{d}y)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \\
&\leq \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(y,\mathbf{x_n})^2\, (\rho + \delta_{\mathbf{x_n}})(\mathrm{d}y)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) = \int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\, \rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}), \quad n \in \mathbb{N},
\end{aligned}$$

where we used the symmetry of h_n in its last n arguments, $n \in \mathbb{N}$, we have

$$S_1 \leq 2 \sum_{n=n_0}^{n_1} \frac{n!}{\theta^{(2n+1)}} \int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\, \rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}).$$

On the other hand, the Cauchy–Schwarz inequality in $L^2(\rho^{[n+1]})$ yields

$$\int_{\mathbb{X}^{n+1}} h_{n+1}(x_1, \mathbf{x_{n+1}}) h_n(\mathbf{x_{n+1}})\, \rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})$$

$$\leq \left(\int_{\mathbb{X}^{n+1}} h_{n+1}(x_1,\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\right)^{\frac{1}{2}} \left(\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\right)^{\frac{1}{2}}, \quad n\in\mathbb{N}.$$

Inserting this in S_2 yields the upper bound

$$\sum_{n=n_0}^{n_1-1}\left(\frac{(n+1)!}{\theta^{(2n+2)}}\int_{\mathbb{X}^{n+1}} h_{n+1}(x_1,\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\right)^{\frac{1}{2}} \left(\frac{(n+1)!}{\theta^{(2n+2)}}\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\right)^{\frac{1}{2}}.$$

By the Cauchy–Schwarz inequality, this sum is bounded by

$$\left(\sum_{n=n_0}^{n_1-1}\frac{(n+1)!}{\theta^{(2n+2)}}\int_{\mathbb{X}^{n+1}} h_{n+1}(x_1,\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\right)^{\frac{1}{2}} \left(\sum_{n=n_0}^{n_1-1}\frac{(n+1)!}{\theta^{(2n+2)}}\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\right)^{\frac{1}{2}}.$$

From the symmetry of h_{n+1}, $n\in\mathbb{N}$, in the last $n+1$ arguments, we obtain

$$\begin{aligned}(n+1)&\int_{\mathbb{X}^{n+1}} h_{n+1}(x_1,\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}) = \int_{\mathbb{X}^{n+1}}\int_{\mathbb{X}} h_{n+1}(y,\mathbf{x_{n+1}})^2\,\delta_{\mathbf{x_{n+1}}}(\mathrm{d}y)\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\\ &\leq \int_{\mathbb{X}^{n+1}}\int_{\mathbb{X}} h_{n+1}(y,\mathbf{x_{n+1}})^2\,(\rho+\delta_{\mathbf{x_{n+1}}})(\mathrm{d}y)\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}) = \int_{\mathbb{X}^{n+2}} h_{n+1}(\mathbf{x_{n+2}})^2\,\rho^{[n+2]}(\mathrm{d}\mathbf{x_{n+2}}), \quad n\in\mathbb{N}.\end{aligned}$$

Thus, an upper bound for S_2 is

$$\sum_{n=n_0}^{n_1}\frac{(n+1)!}{\theta^{(2n+2)}}\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}).$$

Combining our findings, we arrive at

$$\begin{aligned}&\mathbb{E}\left[\int_{\mathbb{X}}\left(\sum_{n=n_0}^{n_1}\int_{\mathbb{X}^n} h_n(x,y)\,\zeta^n(\mathrm{d}y)\right)^2\zeta(\mathrm{d}x)\right]\\ &\leq 2\sum_{n=n_0}^{n_1}\frac{n!}{\theta^{(2n+1)}}\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}) + \sum_{n=n_0}^{n_1}\frac{(n+1)!}{\theta^{(2n+2)}}\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}})\\ &= \sum_{n=n_0}^{n_1}\frac{n!(2\theta+5n+3)}{\theta^{(2n+2)}}\int_{\mathbb{X}^{n+1}} h_n(\mathbf{x_{n+1}})^2\,\rho^{[n+1]}(\mathrm{d}\mathbf{x_{n+1}}).\end{aligned}$$

By (5.8), the statement is thus a consequence of the inequality

$$\frac{n!(2\theta+5n+3)}{\theta^{(2n+2)}} \leq \frac{(\theta+n)^2(n+1)!}{\theta^{(2n+2)}}, \quad n\in\mathbb{N}, n\geq 3,$$

yielding that the series under consideration is a Cauchy sequence in $L^2(C_\zeta)$. The validity of this inequality can be established as follows. The inequality is equivalent to $0 \leq n^3+(1+2\theta)n^2+(\theta^2+2\theta-5)n+\theta^2-2\theta-3$. The right-hand side of this expression is monotonically increasing in $n\in\mathbb{N}$, and it is positive for all $\theta>0$ when $n\geq 3$. □

We can now define the domain of δ'.

Definition 5.10. Let $h_n\colon \mathbb{X}^{n+1}\to\mathbb{R}$, $n\in\mathbb{N}_0$, be measurable functions such that $h_n(x,\cdot)\in\mathbb{H}_n$ holds for each $x\in\mathbb{X}$ and (5.8) is satisfied. Let $H\colon \mathbf{M}(\mathbb{X})\times\mathbb{X}\to\mathbb{R}$ be a measurable function such that the series of functions from (5.9) converges in $L^2(C_\zeta)$ to $(\omega,x)\mapsto H(\zeta(\omega),x)$. The set of all such functions H is denoted by $\mathrm{dom}(\delta')$.

In the following lemma, we establish the convergence of a series in $L^2(\mathbb{P})$, which will later be used to define the operator δ'.

Lemma 5.11. *Let $h_n\colon \mathbb{X}^{n+1}\to\mathbb{R}$, $n\in\mathbb{N}_0$, be measurable functions such that $h_n(x,\cdot)\in\mathbb{H}_n$ for each $x\in\mathbb{X}$ and* (5.8) *is satisfied. Then the series*

$$\sum_{n=1}^{\infty}\left((\theta+n)\int_{\mathbb{X}^{n+1}} h_n(z)\,\zeta^{n+1}(\mathrm{d}z) - \int_{\mathbb{X}^n}\int_{\mathbb{X}} h_n(x,y_1,\dots,y_n)\,(\rho+\delta_{y_1}+\ldots+\delta_{y_n})(\mathrm{d}x)\,\zeta^n(\mathrm{d}(y_1,\dots,y_n))\right)$$

converges in $L^2(\mathbb{P})$.

Proof. Given $n \in \mathbb{N}$, we define

$$X_n := \int_{\mathbb{X}^{n+1}} h_n(z)\,\zeta^{n+1}(\mathrm{d}z) \qquad \text{and} \qquad Y_n := \frac{1}{\theta+n}\int_{\mathbb{X}^n}\int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,(\rho+\delta_{\mathbf{y_n}})(\mathrm{d}x)\,\zeta^n(\mathrm{d}\mathbf{y_n}),$$

as well as $Z_n := (\theta+n)(X_n - Y_n)$. Then the series to be considered is

$$\sum_{n=1}^{\infty}(\theta+n)(X_n-Y_n) = \sum_{n=1}^{\infty} Z_n.$$

Let $m, n \in \mathbb{N}$. Applying Jensen's inequality (for the probability measure ζ) as well as Corollary 3.8, we obtain

$$\mathbb{E}[X_n^2] = \mathbb{E}\left[\left(\int_{\mathbb{X}^{n+1}} h_n(z)\,\zeta^{n+1}(\mathrm{d}z)\right)^2\right] \le \mathbb{E}\left[\int_{\mathbb{X}^{n+1}} h_n(z)^2\,\zeta^{n+1}(\mathrm{d}z)\right] = \frac{1}{\theta^{(n+1)}}\int_{\mathbb{X}^{n+1}} h_n(z)^2\,\rho^{[n+1]}(\mathrm{d}z).$$

An application of Jensen's inequality to the probability measure $(\theta+n)^{-1}(\rho+\delta_{\mathbf{y_n}})\,\zeta^n(\mathrm{d}\mathbf{y_n})$, together with Corollary 3.8, leads to

$$\begin{aligned}\mathbb{E}[Y_n^2] &= \mathbb{E}\left[\left(\frac{1}{\theta+n}\int_{\mathbb{X}^n}\int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,(\rho+\delta_{\mathbf{y_n}})(\mathrm{d}x)\,\zeta^n(\mathrm{d}\mathbf{y_n})\right)^2\right]\\ &\le \mathbb{E}\left[\int_{\mathbb{X}^n}\frac{1}{\theta+n}\int_{\mathbb{X}} h_n(x,\mathbf{y_n})^2\,(\rho+\delta_{\mathbf{y_n}})(\mathrm{d}x)\,\zeta^n(\mathrm{d}\mathbf{y_n})\right] = \frac{1}{\theta^{(n+1)}}\int_{\mathbb{X}^{n+1}} h_n(z)^2\,\rho^{[n+1]}(\mathrm{d}z).\end{aligned}$$

Hence, X_n and Y_n are elements of $L^2(\zeta)$. Moreover, let $g \in \mathbb{H}_m$ for $m \in \mathbb{N}$. Corollary 3.8 yields

$$\mathbb{E}\left[X_n\int_{\mathbb{X}^m} g(x)\,\zeta^m(\mathrm{d}x)\right] = \frac{1}{\theta^{(m+n+1)}}\int_{\mathbb{X}^{m+n+1}} h_n(z)g(x)\,\rho^{[m+n+1]}(\mathrm{d}(x,z)) \tag{5.10}$$

and

$$\mathbb{E}\left[Y_n\int_{\mathbb{X}^m} g(x)\,\zeta^m(\mathrm{d}x)\right] = \frac{1}{\theta^{(m+n)}}\int_{\mathbb{X}^{m+n}} g(\mathbf{x_m})\frac{1}{\theta+n}\int_{\mathbb{X}} h_n(t,\mathbf{z_n})\,(\rho+\delta_{\mathbf{z_n}})(\mathrm{d}t)\,\rho^{[m+n]}(\mathrm{d}(\mathbf{x_m},\mathbf{z_n})). \tag{5.11}$$

By Corollary 4.7, the expressions in (5.10) and (5.11) vanish unless $m \in \{n+1, n, n-1\}$ and $m \in \{n, n-1\}$, respectively. Thus, we conclude $X_n \in \mathbb{F}_{n-1} \oplus \mathbb{F}_n \oplus \mathbb{F}_{n+1}$ and $Y_n \in \mathbb{F}_{n-1} \oplus \mathbb{F}_n$, which implies $Z_n \in \mathbb{F}_{n-1} \oplus \mathbb{F}_n \oplus \mathbb{F}_{n+1}$.

We show that the series from the statement converges in $L^2(\mathbb{P})$. To this end, we show that the sequence of its partial sums forms a Cauchy sequence. Let $n_0, n_1 \in \mathbb{N}$ with $n_0 \le n_1$. By the established orthogonality, it holds

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{n=n_0}^{n_1} Z_n\right)^2\right] &= \sum_{m,n=n_0}^{n_1} \mathbb{1}_{\{|m-n|<3\}}\mathbb{E}\left[Z_m Z_n\right]\\ &= \sum_{n=n_0}^{n_1}\mathbb{E}\left[Z_n^2\right] + 2\sum_{n=n_0}^{n_1-1}\mathbb{E}\left[Z_n Z_{n+1}\right] + 2\sum_{n=n_0}^{n_1-2}\mathbb{E}\left[Z_n Z_{n+2}\right].\end{aligned}$$

From the inequality $2\mathbb{E}\left[Z_n Z_m\right] \le \mathbb{E}[Z_n^2] + \mathbb{E}[Z_m^2]$, $m, n \in \mathbb{N}$, it follows that it suffices to consider

$$\sum_{n=n_0}^{n_1}\mathbb{E}\left[Z_n^2\right].$$

Let $n \in \mathbb{N}$. Corollary 3.8 yields

$$\begin{aligned}\mathbb{E}\left[Z_n^2\right] &= (\theta+n)^2\mathbb{E}\left[\left(\int_{\mathbb{X}^{n+1}} h_n(z)\,\zeta^{n+1}(\mathrm{d}z) - \frac{1}{\theta+n}\int_{\mathbb{X}^n}\int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,(\rho+\delta_{\mathbf{y_n}})(\mathrm{d}x)\,\zeta^n(\mathrm{d}\mathbf{y_n})\right)^2\right]\\ &= \frac{(\theta+n)^2}{\theta^{(2n+2)}}\int_{\mathbb{X}^{2n+2}}(h_n\otimes h_n)(z)\,\rho^{[2n+2]}(\mathrm{d}z)\end{aligned}$$

$$
\begin{aligned}
&- \frac{2(\theta+n)^2}{\theta^{(2n+1)}(\theta+n)} \int_{\mathbb{X}^{2n+1}} h_n(\mathbf{z_{n+1}}) \left(\int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,(\rho+\delta_{\mathbf{y_n}})(\mathrm{d}x) \right) \rho^{[2n+1]}(\mathrm{d}(\mathbf{y_n},\mathbf{z_{n+1}})) \qquad (5.12)\\
&+ \frac{(\theta+n)^2}{\theta^{(2n)}(\theta+n)^2} \int_{\mathbb{X}^{2n}} \left(\int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,(\rho+\delta_{\mathbf{y_n}})(\mathrm{d}x) \right) \left(\int_{\mathbb{X}} h_n(x,\mathbf{z_n})\,(\rho+\delta_{\mathbf{z_n}})(\mathrm{d}x) \right) \rho^{[2n]}(\mathrm{d}(\mathbf{y_n},\mathbf{z_n})).
\end{aligned}
$$

By the recursive definition of $\rho^{[2n+2]}$, it holds

$$
\begin{aligned}
&\int_{\mathbb{X}^{2n+2}} (h_n \otimes h_n)(z)\,\rho^{[2n+2]}(\mathrm{d}z)\\
&= \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}} \int_{\mathbb{X}} h_n(x,\mathbf{x_n}) h_n(y,\mathbf{y_n})\,(\rho+\delta_y+\delta_{\mathbf{x_n}}+\delta_{\mathbf{y_n}})(\mathrm{d}x)\,(\rho+\delta_{\mathbf{x_n}}+\delta_{\mathbf{y_n}})(\mathrm{d}y)\,\rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n}))\\
&= \int_{\mathbb{X}^{2n}} \Bigg(\int_{\mathbb{X}^2} h_n(x,\mathbf{x_n}) h_n(y,\mathbf{y_n})\,\rho^2(\mathrm{d}(x,y)) + 2\int_{\mathbb{X}} h_n(x,\mathbf{x_n})\,\rho(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{y_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}y)\\
&\qquad + 2\int_{\mathbb{X}} h_n(x,\mathbf{x_n})\,\rho(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{y_n})\,\delta_{\mathbf{y_n}}(\mathrm{d}y) + \int_{\mathbb{X}} h_n(y,\mathbf{x_n}) h_n(y,\mathbf{y_n})\,\rho(\mathrm{d}y)\\
&\qquad + 2\int_{\mathbb{X}} h_n(y,\mathbf{x_n}) h_n(y,\mathbf{y_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}y) + 2\int_{\mathbb{X}} h_n(x,\mathbf{x_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{y_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}y)\\
&\quad + \int_{\mathbb{X}} h_n(x,\mathbf{x_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{y_n})\,\delta_{\mathbf{y_n}}(\mathrm{d}y) + \int_{\mathbb{X}} h_n(x,\mathbf{x_n})\,\delta_{\mathbf{y_n}}(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{y_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}y) \Bigg) \rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n}))\\
&=: I_1 + 2I_2 + 2I_3 + I_4 + 2I_5 + 2I_6 + I_7 + I_8,
\end{aligned}
$$

where, e.g.

$$
I_1 = \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}^2} h_n(x,\mathbf{x_n}) h_n(y,\mathbf{y_n})\,\rho^2(\mathrm{d}(x,y))\,\rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n})).
$$

We can apply the recursion (3.1) from Lemma 3.5 to simplify the integral in the second to last line of (5.12) to

$$
\begin{aligned}
&\int_{\mathbb{X}^{2n+1}} h_n(\mathbf{z_{n+1}}) \left(\int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,\rho(\mathrm{d}x) \right) \rho^{[2n+1]}(\mathrm{d}(\mathbf{z_{n+1}},\mathbf{y_n}))\\
&\qquad\qquad + \sum_{r=1}^{n} \int_{\mathbb{X}^{2n+1}} h_n(\mathbf{z_{n+1}}) h_n(y_r,\mathbf{y_n})\,\rho^{[2n+1]}(\mathrm{d}(\mathbf{z_{n+1}},\mathbf{y_n}))\\
&= \int_{\mathbb{X}^{2n}} \Bigg(\int_{\mathbb{X}^2} h_n(z,\mathbf{z_n}) h_n(x,\mathbf{y_n})\,\rho^2(\mathrm{d}(x,z)) + \sum_{r=1}^{n} \int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,\rho(\mathrm{d}x)\,(2h_n(z_r,\mathbf{z_n}) + h_n(y_r,\mathbf{z_n}))\\
&\qquad\qquad + \sum_{i,j=1}^{n} (h_n(z_i,\mathbf{z_n}) h_n(y_j,\mathbf{y_n}) + h_n(y_i,\mathbf{z_n}) h_n(y_j,\mathbf{y_n})) \Bigg) \rho^{[2n]}(\mathrm{d}(\mathbf{y_n},\mathbf{z_n})).\\
&= I_1 + 2I_3 + I_2 + I_7 + I_6.
\end{aligned}
$$

Likewise, the integral in the last line of (5.12) equals

$$
\int_{\mathbb{X}^{2n}} \left(\int_{\mathbb{X}} h_n(x,\mathbf{y_n})\,(\rho+\delta_{\mathbf{y_n}})(\mathrm{d}x) \right) \left(\int_{\mathbb{X}} h_n(x,\mathbf{z_n})\,(\rho+\delta_{\mathbf{z_n}})(\mathrm{d}x) \right) \rho^{[2n]}(\mathrm{d}(\mathbf{y_n},\mathbf{z_n})) = I_1 + 2I_3 + I_7.
$$

Consequently, (5.12) becomes

$$
\frac{n^2-\theta}{\theta^{(2n+2)}}(I_1+I_7) - \frac{2(n+1)(\theta+n)}{\theta^{(2n+2)}}(I_2+I_6) + \frac{2(n^2-\theta)}{\theta^{(2n+2)}} I_3 + \frac{(\theta+n)^2}{\theta^{(2n+2)}} I_4 + \frac{2(\theta+n)^2}{\theta^{(2n+2)}} I_5 + \frac{(\theta+n)^2}{\theta^{(2n+2)}} I_8.
$$

We now show that the terms $|I_1|, \dots, |I_8|$ are each bounded by a multiple of

$$
\frac{(n+1)!}{\theta^{(2n+2)}} \int_{\mathbb{X}^{n+1}} h_n(z)^2\,\rho^{[n+1]}(\mathrm{d}z)
$$

and it thus holds

$$
\mathbb{E}\left[Z_n^2\right] \le C \frac{(n+1)!(\theta+n)^2}{\theta^{(2n+2)}} \int_{\mathbb{X}^{n+1}} h_n(z)^2\,\rho^{[n+1]}(\mathrm{d}z)
$$

for a suitable constant $C > 0$, independent of $n \in \mathbb{N}$. The integral I_1 can be simplified with the help of Corollary 4.5 since $h_n(x,\cdot)$ is an element of $\mathbb{H}_n$ for each $x \in \mathbb{X}$. Thus, we obtain

$$I_1 = n! \int_{\mathbb{X}^n} \int_{\mathbb{X}^2} h_n(x,\mathbf{x_n}) h_n(y,\mathbf{x_n})\, \rho^2(\mathrm{d}(x,y))\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) = n! \int_{\mathbb{X}^n} \left(\int_{\mathbb{X}} h_n(x,\mathbf{x_n})\, \rho(\mathrm{d}x) \right)^2 \rho^{[n]}(\mathrm{d}\mathbf{x_n}).$$

Applying Jensen's inequality for the probability measure $\theta^{-1}\rho$ yields

$$|I_1| = n!\theta^2 \int_{\mathbb{X}^n} \left(\int_{\mathbb{X}} h_n(x,\mathbf{x_n})\, \frac{1}{\theta}\rho(\mathrm{d}x) \right)^2 \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \le n!\theta \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x,\mathbf{x_n})^2\, \rho(\mathrm{d}x)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}).$$

An upper bound for this integral is

$$n!\theta \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x,\mathbf{x_n})^2\, (\rho + \delta_{\mathbf{x_n}})(\mathrm{d}x)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) = n!\theta \int_{\mathbb{X}^{n+1}} h_n(z)^2\, \rho^{[n+1]}(\mathrm{d}z).$$

Moreover, $h_n(x,\cdot) \in \mathbb{H}_n$ for each $x \in \mathbb{X}$ also implies

$$\begin{aligned} I_2 &= \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}} h_n(x,\mathbf{x_n})\, \rho(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{y_n})\, \delta_{\mathbf{x_n}}(\mathrm{d}y)\, \rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n})) \\ &= n! \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x,\mathbf{x_n})\, \rho(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{x_n})\, \delta_{\mathbf{x_n}}(\mathrm{d}y)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}). \end{aligned}$$

From the Cauchy–Schwarz inequality, it follows

$$|I_2| \le n! \left(\int_{\mathbb{X}^n} \left(\int_{\mathbb{X}} |h_n(x,\mathbf{x_n})|\, \rho(\mathrm{d}x) \right)^2 \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \right)^{\frac{1}{2}} \left(\int_{\mathbb{X}^n} \left(\int_{\mathbb{X}} |h_n(y,\mathbf{x_n})|\, \delta_{\mathbf{x_n}}(\mathrm{d}y) \right)^2 \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \right)^{\frac{1}{2}}.$$

By Jensen's inequality for the probability measures $\theta^{-1}\rho$ and $n^{-1}\delta_{\mathbf{x_n}}$ for $\mathbf{x_n} \in \mathbb{X}^n$, an upper bound for this term is

$$n!\sqrt{n\theta} \left(\int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x,\mathbf{x_n})^2\, \rho(\mathrm{d}x)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \right)^{\frac{1}{2}} \left(\int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(y,\mathbf{x_n})^2\, \delta_{\mathbf{x_n}}(\mathrm{d}y)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \right)^{\frac{1}{2}}.$$

Using the recursive definition of $\rho^{[n+1]}$, this in turn is smaller than or equal to

$$\begin{aligned} n!\sqrt{n\theta} \left(\int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x,\mathbf{x_n})^2\, (\rho + \delta_{\mathbf{x_n}})(\mathrm{d}x)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \right)^{\frac{1}{2}} \left(\int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(y,\mathbf{x_n})^2\, (\rho + \delta_{\mathbf{x_n}})(\mathrm{d}y)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \right)^{\frac{1}{2}} \\ = n!\sqrt{n\theta} \int_{\mathbb{X}^{n+1}} h_n(z)^2\, \rho^{[n+1]}(\mathrm{d}z). \end{aligned}$$

Regarding I_3, Corollary 4.5 yields

$$I_3 = n! \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x,\mathbf{y_n})\, \rho(\mathrm{d}x) \int_{\mathbb{X}} h_n(y,\mathbf{y_n})\, \delta_{\mathbf{y_n}}(\mathrm{d}y)\, \rho^{[n]}(\mathrm{d}\mathbf{y_n}) = I_2.$$

Thus, an upper bound for the absolute value of this expression is

$$n!\sqrt{n\theta} \int_{\mathbb{X}^{n+1}} h_n(z)^2\, \rho^{[n+1]}(\mathrm{d}z).$$

In the case of the fourth integral I_4, due to $h_n(x,\cdot) \in \mathbb{H}_n$, $x \in \mathbb{X}$, Corollary 4.5 provides

$$|I_4| = \left| \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}} h_n(y,\mathbf{x_n}) h_n(y,\mathbf{y_n})\, \rho(\mathrm{d}y)\, \rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n})) \right| = n! \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(y,\mathbf{x_n})^2\, \rho(\mathrm{d}y)\, \rho^{[n]}(\mathrm{d}\mathbf{x_n}),$$

which is bounded by

$$n! \int_{\mathbb{X}^{n+1}} h_n(z)^2\, \rho^{[n+1]}(\mathrm{d}z).$$

Moreover, using $h_n(y,\cdot) \in \mathbb{H}_n$ for each $y \in \mathbb{X}$ results in

$$
\begin{aligned}
|I_5| &= \left| \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}} h_n(y, \mathbf{x_n}) h_n(y, \mathbf{y_n}) \, \delta_{\mathbf{x_n}}(\mathrm{d}y) \, \rho^{[2n]}(\mathrm{d}(\mathbf{x_n}, \mathbf{y_n})) \right| = n! \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(y, \mathbf{x_n})^2 \, \delta_{\mathbf{x_n}}(\mathrm{d}y) \, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \\
&\leq n! \int_{\mathbb{X}^{n+1}} h_n(z)^2 \, \rho^{[n+1]}(\mathrm{d}z)
\end{aligned}
$$

and

$$
\begin{aligned}
I_6 &= \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}} h_n(x, \mathbf{x_n}) \, \delta_{\mathbf{x_n}}(\mathrm{d}x) \int_{\mathbb{X}} h_n(y, \mathbf{y_n}) \, \delta_{\mathbf{x_n}}(\mathrm{d}y) \, \rho^{[2n]}(\mathrm{d}(\mathbf{x_n}, \mathbf{y_n})) \\
&= n! \int_{\mathbb{X}^n} \left(\int_{\mathbb{X}} h_n(x, \mathbf{x_n}) \, \delta_{\mathbf{x_n}}(\mathrm{d}x) \right)^2 \rho^{[n]}(\mathrm{d}\mathbf{x_n}) = n! n^2 \int_{\mathbb{X}^n} \left(\int_{\mathbb{X}} h_n(x, \mathbf{x_n}) \, \frac{1}{n} \delta_{\mathbf{x_n}}(\mathrm{d}x) \right)^2 \rho^{[n]}(\mathrm{d}\mathbf{x_n}).
\end{aligned}
$$

Jensen's inequality and the recursive definition of $\rho^{[n+1]}$ yield the upper bound

$$
|I_6| \leq n! n \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x, \mathbf{x_n})^2 \, \delta_{\mathbf{x_n}}(\mathrm{d}x) \, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) \leq n! n \int_{\mathbb{X}^{n+1}} h_n(z)^2 \, \rho^{[n+1]}(\mathrm{d}z)
$$

on the absolute value of this integral. Using $h_n(y,\cdot) \in \mathbb{H}_n$ for each $y \in \mathbb{X}$ once more, the penultimate integral is

$$
\begin{aligned}
I_7 &= \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}} h_n(x, \mathbf{x_n}) \, \delta_{\mathbf{x_n}}(\mathrm{d}x) \int_{\mathbb{X}} h_n(y, \mathbf{y_n}) \, \delta_{\mathbf{y_n}}(\mathrm{d}y) \, \rho^{[2n]}(\mathrm{d}(\mathbf{x_n}, \mathbf{y_n})) \\
&= \sum_{i=1}^{n} \int_{\mathbb{X}^{2n}} \int_{\mathbb{X}} h_n(x, \mathbf{x_n}) \, \delta_{\mathbf{x_n}}(\mathrm{d}x) h_n(y_i, \mathbf{y_n}) \, \rho^{[2n]}(\mathrm{d}(\mathbf{x_n}, \mathbf{y_n})) \\
&= \sum_{i=1}^{n} \sum_{(i_1,\ldots,i_{n-1}) \in [n]^{[n-1]}} \int_{\mathbb{X}^{n+1}} \int_{\mathbb{X}} h_n(x, \mathbf{x_n}) \, \delta_{\mathbf{x_n}}(\mathrm{d}x) h_n(y_i, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) \, \rho^{[n+1]}(\mathrm{d}(\mathbf{x_n}, y_i)) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{(i_1,\ldots,i_{n-1}) \in [n]^{[n-1]}} \int_{\mathbb{X}^{n+1}} h_n(x_j, \mathbf{x_n}) h_n(y_i, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) \, \rho^{[n+1]}(\mathrm{d}(\mathbf{x_n}, y_i)).
\end{aligned}
\tag{5.13}
$$

In order to further simplify this expression, we consider first a fixed summand. To this end, let $i, j \in [n]$ and $(i_1, \ldots, i_{n-1}) \in [n]^{[n-1]}$. If $j \in \{i_1, \ldots, i_{n-1}\}$, let $k \in [n] \setminus \{i_1, \ldots, i_{n-1}\}$. (In fact, in this case we have $[n] = \{k\} \cup \{i_1, \ldots, i_{n-1}\}$.) It then holds

$$
\begin{aligned}
&\int_{\mathbb{X}^{n+1}} h_n(x_j, \mathbf{x_n}) h_n(y_i, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) \, \rho^{[n+1]}(\mathrm{d}(\mathbf{x_n}, y_i)) \\
&= \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x_j, \mathbf{x_n}) \, (\rho + \delta_{y_i} + \delta_{x_1} + \ldots + \delta_{x_{k-1}} + \delta_{x_{k+1}} + \ldots + \delta_{x_n})(\mathrm{d}x_k) \\
&\qquad\qquad h_n(y_i, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) \, \rho^{[n]}(\mathrm{d}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n, y_i)) \\
&= \int_{\mathbb{X}^n} h_n(x_j, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) h_n(y_i, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) \, \rho^{[n]}(\mathrm{d}(x_{i_1}, \ldots, x_{i_{n-1}}, y_i)).
\end{aligned}
$$

By the Cauchy–Schwarz inequality, the absolute value of this integral is bounded by

$$
\left(\int_{\mathbb{X}^n} h_n(x_j, y_i, x_{i_1}, \ldots, x_{i_{n-1}})^2 \, \rho^{[n]}(\mathrm{d}(x_{i_1}, \ldots, x_{i_{n-1}}, y_i)) \right)^{\frac{1}{2}}
\left(\int_{\mathbb{X}^n} h_n(y_i, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) \, \rho^{[n]}(\mathrm{d}(x_{i_1}, \ldots, x_{i_{n-1}}, y_i)) \right)^{\frac{1}{2}}.
$$

In (5.13), we have $nn(n-1)(n-1)!$ summands of this type. If $j \notin \{i_1, \ldots, i_{n-1}\}$, a direct application of the Cauchy–Schwarz inequality gives

$$
\left| \int_{\mathbb{X}^{n+1}} h_n(x_j, \mathbf{x_n}) h_n(y_i, y_i, x_{i_1}, \ldots, x_{i_{n-1}}) \, \rho^{[n+1]}(\mathrm{d}(\mathbf{x_n}, y_i)) \right|
$$

$$\le \left(\int_{\mathbb{X}^{n+1}} h_n(x_j,\mathbf{x_n})^2\,\rho^{[n+1]}(\mathrm{d}(\mathbf{x_n},y_i))\right)^{\frac12}\left(\int_{\mathbb{X}^{n+1}} h_n(y_i,y_i,x_{i_1},\dots,x_{i_{n-1}})^2\,\rho^{[n+1]}(\mathrm{d}(\mathbf{x_n},y_i))\right)^{\frac12}$$
$$= (\theta+n)\left(\int_{\mathbb{X}^{n}} h_n(x_j,\mathbf{x_n})^2\,\rho^{[n]}(\mathrm{d}(\mathbf{x_n}))\right)^{\frac12}\left(\int_{\mathbb{X}^{n}} h_n(y_i,y_i,\mathbf{x_{n-1}})^2\,\rho^{[n]}(\mathrm{d}(\mathbf{x_{n-1}},y_i))\right)^{\frac12}.$$

There are $nn(n-1)!$ summands of this type in (5.13). We note that this case distinction also remains valid in the case $n=1$ (in this case we have $[n]^{[n-1]}=\emptyset$). Plugging these two cases in (5.13), employing the triangle inequality and using the symmetry of h_n in its last n arguments, leads to

$$\left|\int_{\mathbb{X}^{2n}}\int_{\mathbb{X}} h_n(x,\mathbf{x_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}x)\int_{\mathbb{X}} h_n(y,\mathbf{y_n})\,\delta_{\mathbf{y_n}}(\mathrm{d}y)\,\rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n}))\right|$$
$$\le \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{(i_1,\dots,i_{n-1})\in[n]^{[n-1]}}\left|\int_{\mathbb{X}^{n+1}} h_n(x_j,\mathbf{x_n})h_n(y_i,y_i,x_{i_1},\dots,x_{i_{n-1}})\,\rho^{[n+1]}(\mathrm{d}(\mathbf{x_n},y_i))\right|$$
$$\le nn(n-1)(n-1)!\left(\int_{\mathbb{X}^{n}} h_n(x_1,\mathbf{x_n})^2\,\rho^{[n]}(\mathrm{d}\mathbf{x_n})\right)^{\frac12}\left(\int_{\mathbb{X}^{n}} h_n(x_1,\mathbf{x_n})^2\,\rho^{[n]}(\mathrm{d}\mathbf{x_n})\right)^{\frac12}$$
$$+\,nn(n-1)!(\theta+n)\left(\int_{\mathbb{X}^{n}} h_n(x_1,\mathbf{x_n})^2\,\rho^{[n]}(\mathrm{d}(\mathbf{x_n}))\right)^{\frac12}\left(\int_{\mathbb{X}^{n}} h_n(x_1,\mathbf{x_n})^2\,\rho^{[n]}(\mathrm{d}\mathbf{x_n})\right)^{\frac12}.$$

This expression is equal to

$$nn!(\theta+2n-1)\int_{\mathbb{X}^{n}} h_n(x_1,\mathbf{x_n})^2\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}) = n!(\theta+2n-1)\int_{\mathbb{X}^{n}}\sum_{r=1}^{n} h_n(x_r,\mathbf{x_n})^2\,\rho^{[n]}(\mathrm{d}\mathbf{x_n}),$$

which in turn is smaller than or equal to

$$n!(\theta+2n-1)\int_{\mathbb{X}^{n+1}} h_n(z)^2\,\rho^{[n+1]}(\mathrm{d}z).$$

The last integral I_8 can be treated analogously. Since $h_n(y,\cdot)\in\mathbb{H}_n$ for each $y\in\mathbb{X}$, it holds

$$I_8 = \int_{\mathbb{X}^{2n}}\int_{\mathbb{X}} h_n(x,\mathbf{x_n})\,\delta_{\mathbf{y_n}}(\mathrm{d}x)\int_{\mathbb{X}} h_n(y,\mathbf{y_n})\,\delta_{\mathbf{x_n}}(\mathrm{d}y)\,\rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n}))$$
$$= \sum_{i,j=1}^{n}\int_{\mathbb{X}^{2n}} h_n(y_i,\mathbf{x_n})h_n(x_j,\mathbf{y_n})\,\rho^{[2n]}(\mathrm{d}(\mathbf{x_n},\mathbf{y_n}))$$
$$= \sum_{i,j=1}^{n}\sum_{(i_1,\dots,i_{n-1})\in[n]^{[n-1]}}\int_{\mathbb{X}^{n+1}} h_n(y_i,\mathbf{x_n})h_n(x_j,y_i,x_{i_1},\dots,x_{i_{n-1}})\,\rho^{[n+1]}(\mathrm{d}(\mathbf{x_n},y_i)). \tag{5.14}$$

Again, we start by examining one summand. Therefore, we fix $i,j\in[n]$ and $(i_1,\dots,i_{n-1})\in[n]^{[n-1]}$. If $j\in\{i_1,\dots,i_{n-1}\}$, let $k\in[n]\setminus\{i_1,\dots,i_{n-1}\}$. It follows that

$$\int_{\mathbb{X}^{n+1}} h_n(y_i,\mathbf{x_n})h_n(x_j,y_i,x_{i_1},\dots,x_{i_{n-1}})\,\rho^{[n+1]}(\mathrm{d}(\mathbf{x_n},y_i))$$
$$= \int_{\mathbb{X}^{n}}\int_{\mathbb{X}} h_n(y_i,\mathbf{x_n})(\rho+\delta_{y_i}+\delta_{x_1}+\ldots+\delta_{x_{k-1}}+\delta_{x_{k+1}}+\ldots+\delta_{x_n})(\mathrm{d}x_k)$$
$$h_n(x_j,y_i,x_{i_1},\dots,x_{i_{n-1}})\,\rho^{[n]}(\mathrm{d}(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n,y_i))$$
$$= \int_{\mathbb{X}^{n}} h_n(x_j,y_i,x_{i_1},\dots,x_{i_{n-1}})h_n(y_i,y_i,x_{i_1},\dots,x_{i_{n-1}})\,\rho^{[n]}(\mathrm{d}(x_{i_1},\dots,x_{i_{n-1}},y_i)).$$

This term is the same as the one we encountered in this case for the penultimate summand. If $j\notin\{i_1,\dots,i_{n-1}\}$, we have $[n]=\{j\}\cup\{i_1,\dots,i_{n-1}\}$ and can again apply the Cauchy–Schwarz inequality directly. It holds

$$\left|\int_{\mathbb{X}^{n+1}} h_n(y_i,\mathbf{x_n})h_n(x_j,y_i,x_{i_1},\dots,x_{i_{n-1}})\,\rho^{[n+1]}(\mathrm{d}(\mathbf{x_n},y_i))\right|$$

$$\leq \left(\int_{\mathbb{X}^{n+1}} h_n(y_i, \mathbf{x_n})^2 \, \rho^{[n+1]}(\mathrm{d}(\mathbf{x_n}, y_i)) \right)^{\frac{1}{2}} \left(\int_{\mathbb{X}^{n+1}} h_n(x_j, y_i, x_{i_1}, \ldots, x_{i_{n-1}})^2 \, \rho^{[n+1]}(\mathrm{d}(\mathbf{x_n}, y_i)) \right)^{\frac{1}{2}}$$
$$= \int_{\mathbb{X}^{n+1}} h_n(z)^2 \, \rho^{[n+1]}(\mathrm{d}z).$$

Inserting these findings in (5.14) yields, with again the help of the triangle inequality, the following upper bound

$$\begin{aligned}|I_8| &\leq nn(n-1)(n-1)! \int_{\mathbb{X}^n} h_n(x_1, \mathbf{x_n})^2 \, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) + nn(n-1)! \int_{\mathbb{X}^{n+1}} h_n(z)^2 \, \rho^{[n+1]}(\mathrm{d}z) \\ &= n!(n-1) \int_{\mathbb{X}^n} \sum_{r=1}^{n} h_n(x_r, \mathbf{x_n})^2 \, \rho^{[n]}(\mathrm{d}\mathbf{x_n}) + nn! \int_{\mathbb{X}^{n+1}} h_n(z)^2 \, \rho^{[n+1]}(\mathrm{d}z) \\ &\leq n!(2n-1) \int_{\mathbb{X}^{n+1}} h_n(z)^2 \, \rho^{[n+1]}(\mathrm{d}z).\end{aligned}$$

All in all, we have shown that $\mathbb{E}[Z_n^2]$ is smaller than or equal to

$$\left(\frac{(n^2+\theta)(2\theta+2n-1)}{\theta^{(2n+2)}} + \frac{2(n+1)(\sqrt{n\theta}+n)(\theta+n)}{\theta^{(2n+2)}} + \frac{2(n^2+\theta)\sqrt{n\theta}}{\theta^{(2n+2)}} + \frac{2(n+1)(\theta+n)^2}{\theta^{(2n+2)}} \right) n! \rho^{[n+1]}(h_n^2)$$

and hence conclude

$$\mathbb{E}\left[\left(\sum_{n=n_0}^{n_1} Z_n \right)^2 \right] \leq 5 \sum_{n=n_0}^{n_1} \mathbb{E}\left[Z_n^2\right] \leq C \sum_{n=n_0}^{n_1} \frac{(\theta+n)^2(n+1)!}{\theta^{(2n+2)}} \int_{\mathbb{X}^{n+1}} h_n(z)^2 \, \rho^{[n+1]}(\mathrm{d}z)$$

for a suitable constant $C > 0$. □

Definition 5.12. For $H \in \mathrm{dom}(\delta')$, i.e. H is the limit in (5.9) where measurable functions $h_n \colon \mathbb{X}^{n+1} \to \mathbb{R}$, $n \in \mathbb{N}_0$, which satisfy $h_n(x, \cdot) \in \mathbb{H}_n$ for each $x \in \mathbb{X}$ and (5.8), are considered, let

$$\delta'(H) := \sum_{n=0}^{\infty} \Bigg((\theta+n) \int_{\mathbb{X}^{n+1}} h_n(z) \, \zeta^{n+1}(\mathrm{d}z) - \int_{\mathbb{X}^n} \int_{\mathbb{X}} h_n(x, y_1, \ldots, y_n) \, (\rho + \delta_{y_1} + \ldots + \delta_{y_n})(\mathrm{d}x) \, \zeta^n(\mathrm{d}(y_1, \ldots, y_n)) \Bigg).$$

By the preceding lemma, it holds $\delta'(H) \in L^2(\mathbb{P})$, i.e. $\mathbb{E}[\delta'(H)^2] < \infty$.

Remark 5.13. One might be tempted to set

$$\overline{h}_n(x, y_1, \ldots, y_n) := h_n(x, y_1, \ldots, y_n) - \frac{1}{\theta+n} \int_{\mathbb{X}} h_n(t, y_1, \ldots, y_n) \, (\rho + \delta_{y_1} + \ldots + \delta_{y_n})(\mathrm{d}t)$$

for $(x, y_1, \ldots, y_n) \in \mathbb{X}^{n+1}$ and $n \in \mathbb{N}$ in the above expression defining $\delta'(H)$. Using that ζ is a probability measure, $\delta'(H)$ can then be written as

$$\delta'(H) = \sum_{n=0}^{\infty} (\theta+n) \int_{\mathbb{X}^{n+1}} \overline{h}_n(z) \, \zeta^{n+1}(\mathrm{d}z).$$

However, one has to exercise caution since $\overline{h}_n$ may not be an element of $\mathbb{H}_{n+1}$, $n \in \mathbb{N}$, as can be seen in the following example. ◇

Example 5.14. We continue Example 4.8, i.e. let $B \in \mathcal{X}$ and

$$h(x, y) := \frac{1}{\theta+1}(\rho(B) + \mathbb{1}_B(y))\mathbb{1}_B(x) - \frac{1}{\theta}\rho(B)\mathbb{1}_B(x), \quad (x, y) \in \mathbb{X}^2.$$

Example 4.8 showed $h(x, \cdot) \in \mathbb{H}_1$ for $x \in \mathbb{X}$. Furthermore, given $(x, y) \in \mathbb{X}^2$, it holds

$$h(x, y) - \frac{1}{\theta+1} \int_{\mathbb{X}} h(x, y) \, (\rho + \delta_y)(\mathrm{d}x)$$

$$= h(x,y) - \frac{1}{\theta+1}\int_{\mathbb{X}} \frac{1}{\theta+1}(\rho(B)+\mathbb{1}_B(y))\mathbb{1}_B(x) - \frac{1}{\theta}\rho(B)\mathbb{1}_B(x)\,(\rho+\delta_y)(\mathrm{d}x)$$
$$= -\frac{1}{\theta^{(2)}}\rho(B)\mathbb{1}_B(x) + \frac{1}{\theta+1}\mathbb{1}_B(x)\mathbb{1}_B(y) - \left(\frac{1}{\theta+1}(\rho(B)+\mathbb{1}_B(y))\right)^2 + \frac{1}{\theta^{(2)}}\rho(B)(\rho(B)+\mathbb{1}_B(y))$$
$$= -\frac{1}{\theta^{(2)}}\rho(B)\mathbb{1}_B(x) + \frac{1}{\theta+1}\mathbb{1}_B(x)\mathbb{1}_B(y) + \frac{\rho(B)^2}{\theta^{(2)}(\theta+1)} - \frac{\theta-1}{\theta^{(2)}(\theta+1)}\rho(B)\mathbb{1}_B(y) - \frac{1}{(\theta+1)^2}\mathbb{1}_B(y).$$

This function is not symmetric and hence for this mapping h, the function $\overline{h}$ introduced in the previous remark does not belong to $\mathbb{H}_2$. For the symmetrization $\widetilde{\overline{h}}$, we have

$$\int_{\mathbb{X}} \widetilde{\overline{h}}(x,y)(\rho+\delta_x)(\mathrm{d}y)$$
$$= \frac{1}{2}\int_{\mathbb{X}} -\frac{1}{\theta^{(2)}}\rho(B)\mathbb{1}_B(x) + \frac{2}{\theta+1}\mathbb{1}_B(x)\mathbb{1}_B(y) + \frac{2\rho(B)^2}{\theta^{(2)}(\theta+1)} - \frac{\theta-1}{\theta^{(2)}(\theta+1)}\rho(B)\mathbb{1}_B(y) - \frac{1}{(\theta+1)^2}\mathbb{1}_B(y)$$
$$\qquad - \frac{1}{\theta^{(2)}}\rho(B)\mathbb{1}_B(y) - \frac{\theta-1}{\theta^{(2)}(\theta+1)}\rho(B)\mathbb{1}_B(x) - \frac{1}{(\theta+1)^2}\mathbb{1}_B(x)\,(\rho+\delta_x)(\mathrm{d}y)$$
$$= \frac{1}{2}\Bigg(-\frac{1}{\theta}\rho(B)\mathbb{1}_B(x) + \frac{2(\rho(B)+1)}{\theta+1}\mathbb{1}_B(x) + \frac{2\rho(B)^2}{\theta^{(2)}} - \frac{\theta-1}{\theta^{(2)}(\theta+1)}\rho(B)(\rho(B)+\mathbb{1}_B(x))$$
$$\qquad - \frac{1}{(\theta+1)^2}(\rho(B)+\mathbb{1}_B(x)) - \frac{1}{\theta^{(2)}}\rho(B)(\rho(B)+\mathbb{1}_B(x)) - \frac{\theta-1}{\theta^{(2)}}\rho(B)\mathbb{1}_B(x) - \frac{1}{(\theta+1)}\mathbb{1}_B(x)\Bigg)$$
$$= \frac{-1}{(\theta+1)^2}\rho(B)\mathbb{1}_B(x) + \frac{\theta}{2(\theta+1)^2}\mathbb{1}_B(x) + \rho(B)\frac{\theta+2}{2\theta^{(2)}(\theta+1)} + \frac{1}{\theta^{(2)}(\theta+1)}\rho(B)^2, \quad x\in\mathbb{X},$$

indicating that it may likewise not belong to $\mathbb{H}_2$. ◦

The next theorem shows that δ' satisfies an integration by parts formula.

Theorem 5.15. *Let $F \in \mathrm{dom}(\nabla)$ and $H \in \mathrm{dom}(\delta')$. Then*

$$\mathbb{E}\left[\delta'(H)F\right] = \mathbb{E}\left[\int_{\mathbb{X}} H(x)(\nabla_x F)\zeta(\mathrm{d}x)\right]. \tag{5.15}$$

Proof. Before proving the formula, we establish the finiteness of both expectations. On the one hand, since $F \in \mathrm{dom}(\nabla)$, we have in particular $F \in L^2(\zeta)$. Hence, it holds $\mathbb{E}[F(\zeta)^2] < \infty$. Thus, from the Cauchy–Schwarz inequality and Lemma 5.11 we obtain

$$\mathbb{E}\left[|\delta'(H)F|\right] \le \left(\mathbb{E}\left[\delta'(H)^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[F(\zeta)^2\right]\right)^{\frac{1}{2}} < \infty.$$

On the other hand, by Lemma 5.2 and definition, both ∇F and H are elements of $L^2(C_\zeta)$. The finiteness of the expectation on the right-hand side is then again a consequence of the Cauchy–Schwarz inequality.

We consider both sides of the equation separately. Let $F \in \mathrm{dom}(\nabla)$ with chaos decomposition (4.22) and let $h_n\colon \mathbb{X}^{n+1} \to \mathbb{R}$, $n \in \mathbb{N}_0$, be measurable functions such that $h_n(x,\cdot) \in \mathbb{H}_n$ for each $x \in \mathbb{X}$ and (5.8) is satisfied. Given $m_0, n_0 \in \mathbb{N}$, we have

$$\mathbb{E}\Bigg[\left(\sum_{m=0}^{m_0}\left((\theta+m)\int_{\mathbb{X}^{m+1}} h_m(y)\,\zeta^{m+1}(\mathrm{d}y) - \int_{\mathbb{X}^m}\int_{\mathbb{X}} h_m(x,\mathbf{y_m})\,(\rho+\delta_{\mathbf{y_m}})(\mathrm{d}x)\,\zeta^m(\mathrm{d}\mathbf{y_m})\right)\right)$$
$$\qquad\left(\mathbb{E}[F(\zeta)] + \sum_{n=1}^{n_0}\int_{\mathbb{X}^n} f_n(z)\,\zeta^n(\mathrm{d}z)\right)\Bigg]$$
$$= \mathbb{E}[F(\zeta)]\mathbb{E}\left[\left(\sum_{m=0}^{m_0}\left((\theta+m)\int_{\mathbb{X}^{m+1}} h_m(y)\,\zeta^{m+1}(\mathrm{d}y) - \int_{\mathbb{X}^m}\int_{\mathbb{X}} h_m(x,\mathbf{y_m})\,(\rho+\delta_{\mathbf{y_m}})(\mathrm{d}x)\,\zeta^m(\mathrm{d}\mathbf{y_m})\right)\right)\right]$$
$$+ \sum_{m=0}^{m_0}\sum_{n=1}^{n_0}(\theta+m)\mathbb{E}\left[\left(\int h_m(y)\,\zeta^{m+1}(\mathrm{d}y) - \iint h_m(x,\mathbf{y_m})\,(\rho+\delta_{\mathbf{y_m}})(\mathrm{d}x)\,\zeta^m(\mathrm{d}\mathbf{y_m})\right)\int f_n(z)\,\zeta^n(\mathrm{d}z)\right].$$

By Corollary 3.8, this is equal to

$$\sum_{m=0}^{m_0}\sum_{n=1}^{n_0}\frac{\theta+m}{\theta^{(m+n+1)}}\int_{\mathbb{X}^{m+n+1}}h_m(\mathbf{y_{m+1}})f_n(\mathbf{z_n})\,\rho^{[m+n+1]}(\mathrm{d}(\mathbf{y_{m+1}},\mathbf{z_n}))$$
$$-\sum_{m=0}^{m_0}\sum_{n=1}^{n_0}\frac{1}{\theta^{(m+n)}}\int_{\mathbb{X}^{m+n}}\int_{\mathbb{X}}h_m(x,\mathbf{y_m})\,(\rho+\delta_{\mathbf{y_m}})(\mathrm{d}x)f_n(\mathbf{z_n})\,\rho^{[m+n]}(\mathrm{d}(\mathbf{y_m},\mathbf{z_n})). \quad (5.16)$$

According to Corollary 4.7, the first line of (5.16) is

$$\sum_{m=2}^{m_0\wedge(n_0+1)}\frac{(\theta+m)m!}{\theta^{(2m)}}\int_{\mathbb{X}^m}h_m(x_1,\mathbf{x_m})f_{m-1}(x_2,\dots,x_m)\,\rho^{[m]}(\mathrm{d}\mathbf{x_m})$$
$$+\sum_{m=1}^{m_0\wedge n_0}\frac{(\theta+m)m!}{\theta^{(2m+1)}}\Bigg(\int_{\mathbb{X}^{m+1}}h_m(\mathbf{x_{m+1}})f_m(x_2,\dots,x_{m+1})\,\rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}})$$
$$+m\int_{\mathbb{X}^m}h_m(x_1,\mathbf{x_m})f_m(\mathbf{x_m})\,\rho^{[m]}(\mathrm{d}\mathbf{x_m})\Bigg)$$
$$+\sum_{m=0}^{m_0\wedge(n_0-1)}\frac{(\theta+m)(m+1)!}{\theta^{(2m+2)}}\int_{\mathbb{X}^{m+1}}h_m(\mathbf{x_{m+1}})f_{m+1}(\mathbf{x_{m+1}})\,\rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}}),$$

while the second equals

$$-\sum_{m=2}^{m_0\wedge(n_0+1)}\frac{m!}{\theta^{(2m-1)}}\int_{\mathbb{X}^m}h_m(x_m,\mathbf{x_m})f_{m-1}(\mathbf{x_{m-1}})\,\rho^{[m]}(\mathrm{d}\mathbf{x_m})$$
$$-\sum_{m=1}^{m_0\wedge n_0}\frac{m!}{\theta^{(2m)}}\int_{\mathbb{X}^{m+1}}h_m(x_{m+1},\mathbf{x_m})f_m(\mathbf{x_m})\,\rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}}).$$

We summarise these terms and obtain that (5.16) is equal to

$$\sum_{m=2}^{m_0\wedge(n_0+1)}\frac{(1-m)m!}{\theta^{(2m)}}\int_{\mathbb{X}^m}h_m(x_m,\mathbf{x_m})f_{m-1}(\mathbf{x_{m-1}})\,\rho^{[m]}(\mathrm{d}\mathbf{x_m})$$
$$-\sum_{m=1}^{m_0\wedge n_0}\Bigg(\frac{mm!}{\theta^{(2m+1)}}\int_{\mathbb{X}^{m+1}}h_m(\mathbf{x_{m+1}})f_m(x_2,\dots,x_{m+1})\,\rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}})$$
$$-\frac{m(\theta+m)m!}{\theta^{(2m+1)}}\int_{\mathbb{X}^m}h_m(x_1,\mathbf{x_m})f_m(\mathbf{x_m})\,\rho^{[m]}(\mathrm{d}\mathbf{x_m})\Bigg) \quad (5.17)$$
$$+\sum_{m=0}^{m_0\wedge(n_0-1)}\frac{(\theta+m)(m+1)!}{\theta^{(2m+2)}}\int_{\mathbb{X}^{m+1}}h_m(\mathbf{x_{m+1}})f_{m+1}(\mathbf{x_{m+1}})\,\rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}}),$$

ending the study of the left-hand side of the integration by parts formula.

Let $m_0, n_0 \in \mathbb{N}$. We now consider

$$\mathbb{E}\left[\int_{\mathbb{X}}\left(\sum_{m=0}^{m_0}\int_{\mathbb{X}^m}h_m(x,\mathbf{y_m})\,\zeta^m(\mathrm{d}\mathbf{y_m})\right)\left(\sum_{n=1}^{n_0}n\left(\int_{\mathbb{X}^{n-1}}f_n(x,\mathbf{y_{n-1}})\,\zeta^{n-1}(\mathrm{d}\mathbf{y_{n-1}})-\zeta^n(f_n)\right)\right)\zeta(\mathrm{d}x)\right]$$
$$=\sum_{m=0}^{m_0}\sum_{n=1}^{n_0}n\mathbb{E}\Bigg[\int_{\mathbb{X}^{m+n}}h_m(x,\mathbf{x_m})f_n(x,\mathbf{y_{n-1}})\,\zeta^{m+n}(\mathrm{d}(x,\mathbf{x_m},\mathbf{y_{n-1}}))$$
$$-\int_{\mathbb{X}^{m+n+1}}h_m(\mathbf{x_{m+1}})f_n(\mathbf{y_n})\,\zeta^{m+n+1}(\mathrm{d}(\mathbf{x_{m+1}},\mathbf{y_n}))\Bigg],$$

where, in the first line, $\zeta^n(f_n)=\int_{\mathbb{X}^n}f_n(x)\,\zeta^n(\mathrm{d}x)$, $n\in[n_0]$, cf. Section 2.4. By Corollary 3.8, this is equal to

$$\sum_{m=0}^{m_0}\sum_{n=1}^{n_0}\frac{n}{\theta^{(m+n)}}\int_{\mathbb{X}^{m+n}}h_m(x,\mathbf{x_m})f_n(x,\mathbf{y_{n-1}})\,\rho^{[m+n]}(\mathrm{d}(x,\mathbf{x_m},\mathbf{y_{n-1}}))$$

$$- \sum_{m=0}^{m_0} \sum_{n=1}^{n_0} \frac{n}{\theta^{(m+n+1)}} \int_{\mathbb{X}^{m+n+1}} h_m(\mathbf{x_{m+1}}) f_n(\mathbf{y_n})\, \rho^{[m+n+1]}(\mathrm{d}(\mathbf{x_{m+1}}, \mathbf{y_n})). \quad (5.18)$$

Using Corollary 4.7, we obtain that the first sum equals

$$\sum_{m=1}^{m_0 \wedge n_0} \frac{m m!}{\theta^{(2m)}} \int_{\mathbb{X}^m} h_m(x_1, \mathbf{x_m}) f_m(\mathbf{x_m})\, \rho^{[m]}(\mathrm{d}\mathbf{x_m}) + \sum_{m=0}^{m_0 \wedge (n_0 - 1)} \frac{(m+1)m!}{\theta^{(2m+1)}} \int_{\mathbb{X}^{m+1}} h_m(\mathbf{x_{m+1}}) f_{m+1}(\mathbf{x_{m+1}})\, \rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}}),$$

while the second is

$$\begin{aligned} &- \sum_{m=1}^{m_0 \wedge (n_0+1)} \frac{(m-1)m!}{\theta^{(2m)}} \int_{\mathbb{X}^m} h_m(x_1, \mathbf{x_m}) f_{m-1}(x_2, \ldots, x_m)\, \rho^{[m]}(\mathrm{d}\mathbf{x_m}) \\ &- \sum_{m=1}^{m_0 \wedge n_0} \frac{m m!}{\theta^{(2m+1)}} \Big(\int_{\mathbb{X}^{m+1}} h_m(\mathbf{x_{m+1}}) f_m(x_2, \ldots, x_{m+1})\, \rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}}) \\ &\qquad + m \int_{\mathbb{X}^m} h_m(x_1, \mathbf{x_m}) f_m(\mathbf{x_m})\, \rho^{[m]}(\mathrm{d}\mathbf{x_m}) \Big) \\ &- \sum_{m=0}^{m_0 \wedge (n_0-1)} \frac{(m+1)(m+1)!}{\theta^{(2m+2)}} \int_{\mathbb{X}^{m+1}} h_m(\mathbf{x_{m+1}}) f_{m+1}(\mathbf{x_{m+1}})\, \rho^{[m+1]}(\mathrm{d}\mathbf{x_{m+1}}). \end{aligned}$$

After summarising the terms, (5.18) becomes (5.17).

Let $m_0, n_0 \in \mathbb{N}$ and, given $x \in \mathbb{X}$, let

$$F_{n_0}(\zeta) := \mathbb{E}[F(\zeta)] + \sum_{n=1}^{n_0} \int_{\mathbb{X}^n} f_n(z)\, \zeta^n(\mathrm{d}z) \qquad \text{and} \qquad H_{m_0}(x) := \sum_{m=0}^{m_0} \int_{\mathbb{X}^m} h_m(x, \mathbf{y_m})\, \zeta^m(\mathrm{d}\mathbf{y_m}).$$

We then have

$$\nabla_x F_{n_0} = \sum_{n=1}^{n_0} n \left(\int_{\mathbb{X}^{n-1}} f_n(x, \mathbf{y_{n-1}})\, \zeta^{n-1}(\mathrm{d}\mathbf{y_{n-1}}) - \int_{\mathbb{X}^n} f_n(y)\, \zeta^n(\mathrm{d}y) \right)$$

and

$$\delta'(H_{m_0}) = \sum_{m=0}^{m_0} \left((\theta + m) \int_{\mathbb{X}^{m+1}} h_m(y)\, \zeta^{m+1}(\mathrm{d}y) - \int_{\mathbb{X}^m} \int_{\mathbb{X}} h_m(x, \mathbf{y_m})\, (\rho + \delta_{\mathbf{y_m}})(\mathrm{d}x)\, \zeta^m(\mathrm{d}\mathbf{y_m}) \right).$$

By Theorem 3.7 and Lemma 5.11, it holds $F_{n_0}(\zeta) \to F(\zeta)$ and $\delta'(H_{n_0}) \to \delta(H)$ in $L^2(\mathbb{P})$ for $n_0 \to \infty$, respectively. Hence, with

$$\begin{aligned} |\mathbb{E}\,[\delta'(H_{m_0}) F_{n_0}] - \mathbb{E}[\delta'(H) F]| &= |\mathbb{E}\,[(\delta'(H_{m_0}) - \delta'(H))\, F_{n_0}] + \mathbb{E}\,[\delta'(H)\,(F_{n_0} - F)]| \\ &\le \mathbb{E}\left[(\delta'(H_{m_0}) - \delta'(H))^2\right]^{\frac{1}{2}} \mathbb{E}\left[|F_{n_0}|^2\right]^{\frac{1}{2}} + \mathbb{E}\left[\delta'(H)^2\right]^{\frac{1}{2}} \mathbb{E}\left[(F_{n_0} - F)^2\right]^{\frac{1}{2}}, \end{aligned}$$

where we used the Cauchy–Schwarz inequality, we conclude

$$\lim_{m_0, n_0 \to \infty} \mathbb{E}\,[\delta'(H_{m_0}) F_{n_0}] = \mathbb{E}[\delta'(H) F].$$

On the other hand, by Lemma 5.2 and Lemma 5.9, the convergences $\nabla F_{n_0} \to \nabla F$ and $H_{n_0} \to H$ in $L^2(C_\zeta)$ for $n_0 \to \infty$ hold true. Therefore, from the Cauchy–Schwarz inequality we obtain

$$\begin{aligned} &\left| \mathbb{E}\left[\int_{\mathbb{X}} H_{m_0}(x) (\nabla_x F_{n_0})\, \zeta(\mathrm{d}x) \right] - \mathbb{E}\left[\int_{\mathbb{X}} H(x) (\nabla_x F)\, \zeta(\mathrm{d}x) \right] \right| \\ &= \left| \mathbb{E}\left[\int_{\mathbb{X}} (H_{m_0}(x) - H(x))\, (\nabla_x F_{n_0})\, \zeta(\mathrm{d}x) \right] + \mathbb{E}\left[\int_{\mathbb{X}} H(x)\, (\nabla_x F_{n_0} - \nabla_x F)\, \zeta(\mathrm{d}x) \right] \right| \end{aligned}$$

$$\leq \mathbb{E}\left[\int_{\mathbb{X}} (H_{m_0}(x) - H(x))^2 \, \zeta(\mathrm{d}x)\right]^{\frac{1}{2}} \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F_{n_0})^2 \, \zeta(\mathrm{d}x)\right]^{\frac{1}{2}} + \mathbb{E}\left[\int_{\mathbb{X}} H(x)^2 \, \zeta(\mathrm{d}x)\right]^{\frac{1}{2}} \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F_{n_0} - \nabla_x F)^2 \, \zeta(\mathrm{d}x)\right]^{\frac{1}{2}},$$

and consequently

$$\lim_{m_0, n_0 \to \infty} \mathbb{E}\left[\int_{\mathbb{X}} H_{m_0}(x)(\nabla_x F_{n_0}) \, \zeta(\mathrm{d}x)\right] = \mathbb{E}\left[\int_{\mathbb{X}} H(x)(\nabla_x F) \, \zeta(\mathrm{d}x)\right]. \qquad \square$$

The preceding theorem shows that δ' coincides with δ on $\operatorname{dom}(\delta')$.

Corollary 5.16. *It holds* $\operatorname{dom}(\delta') \subseteq \operatorname{dom}(\delta)$ *and for* $H \in \operatorname{dom}(\delta')$*, the random variables* $\delta(H)$ *and* $\delta'(H)$ *coincide* $\mathbb{P}$*-a.s.*

Proof. Since $\delta'(H)$ satisfies partial integration, the claim follows from the definition of $\delta(H)$. $\square$

5.3. The generator

In this section, we consider the third operator in Malliavin calculus, the generator L. The designation will become clear in Chapter 6, where we show that the operator generates a semigroup that arises naturally in the context of Fleming—Viot processes. Again, we begin by specifying the class of functions on which this operator is defined.

Definition 5.17. Let $\operatorname{dom}(L)$ stand for the set of all $F \in L^2(\zeta)$ such that the functions from its chaos expansion (4.22) satisfy

$$\sum_{n=1}^{\infty} \frac{(\theta + n - 1)^2 n^2 n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x) < \infty. \tag{5.19}$$

A comparison with (5.1) in the definition of $\operatorname{dom}(\nabla)$, where the convergence of the series

$$\sum_{n=1}^{\infty} \frac{(\theta + n - 1)n n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x),$$

is required, reveals that $\operatorname{dom}(L) \subseteq \operatorname{dom}(\nabla)$.

The next lemma shows the convergence of a series used later in the definition of the operator L.

Lemma 5.18. *Let* $F \in \operatorname{dom}(L)$ *with chaos expansion* (4.22)*. Then the series*

$$\sum_{n=1}^{\infty} n(\theta + n - 1) \int_{\mathbb{X}^n} f_n(x) \, \zeta^n(\mathrm{d}x)$$

converges in $L^2(\mathbb{P})$.

Proof. Let $m_0, n_0 \in \mathbb{N}$ with $m_0 \leq n_0$. Corollary 4.11 yields

$$\mathbb{E}\left[\left(\sum_{n=m_0}^{n_0} n(\theta + n - 1) \int_{\mathbb{X}^n} f_n(x) \zeta^n(\mathrm{d}x)\right)^2\right] = \sum_{n=m_0}^{n_0} \frac{n^2(\theta + n - 1)^2 n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x).$$

Consequently, the partial sums of the series under consideration form a Cauchy sequence in $L^2(\mathbb{P})$ if and only if $F \in \operatorname{dom}(L)$. $\square$

We can now define the operator L.

Definition 5.19. For $F \in \operatorname{dom}(L)$ with chaos expansion (4.22) let

$$L(F) := -\sum_{n=1}^{\infty} n(\theta + n - 1) \int_{\mathbb{X}^n} f_n(x) \, \zeta^n(\mathrm{d}x).$$

For reasons that will become clear in the next chapter, we refer to the linear mapping $L\colon \operatorname{dom}(L) \to L^2(\mathbb{P})$ as *Fleming–Viot operator* or *generator*. It plays a similar role as the Ornstein–Uhlenbeck operator in a Gaussian or Poisson context. This fact is supported by our next theorem.

Theorem 5.20. *Let $F \in \operatorname{dom}(\nabla)$. Then F belongs to $\operatorname{dom}(L)$ if and only if $\nabla F \in \operatorname{dom}(\delta)$. Moreover, in this case, it holds*

$$\delta(\nabla F) = -L(F), \quad \mathbb{P}\text{-}a.s. \tag{5.20}$$

Proof. Let $F \in \operatorname{dom}(\nabla)$ with chaos expansion (4.22). On the one hand, if $F \in \operatorname{dom}(L)$, we infer from (5.3) that

$$\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)(\nabla_x G)\,\zeta(\mathrm{d}x)\right] = \sum_{n=1}^{\infty} \frac{nn!(\theta+n-1)}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x) g_n(x)\,\rho^{[n]}(\mathrm{d}x)$$

holds for every $G \in \operatorname{dom}(\nabla)$ with chaos expansion $G(\zeta) = \mathbb{E}[G(\zeta)] + \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} g_n(x)\,\zeta^n(\mathrm{d}x)$. The absolute value of this expression is, according to the Cauchy–Schwarz inequality, bounded by

$$\sum_{n=1}^{\infty} \left(\frac{n\sqrt{n!}(\theta+n-1)}{\sqrt{\theta^{(2n)}}} \int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x)\right)^{\frac12} \left(\frac{\sqrt{n!}}{\sqrt{\theta^{(2n)}}} \int_{\mathbb{X}^n} g_n(x)^2\,\rho^{[n]}(\mathrm{d}x)\right)^{\frac12}.$$

Furthermore, by the Cauchy–Schwarz inequality, an upper bound for this series is

$$\left(\sum_{n=1}^{\infty} \frac{n^2 n!(\theta+n-1)}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x)\right)^{\frac12} \left(\sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} g_n(x)^2\,\rho^{[n]}(\mathrm{d}x)\right)^{\frac12}.$$

The assumptions $F \in \operatorname{dom}(L)$ and $G \in \operatorname{dom}(\nabla)$ imply that this equals $c\mathbb{E}[G(\zeta)]^{\frac12}$ with $c = \mathbb{E}[(LF)^2]^{\frac12} < \infty$. Thus, ∇F is an element of $\operatorname{dom}(\delta)$.

On the other hand, if $\nabla F \in \operatorname{dom}(\delta)$, we obtain from the integration by parts formula the equality

$$\mathbb{E}\left[\delta(\nabla F)G\right] = \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)(\nabla_x G)\,\zeta(\mathrm{d}x)\right]$$

for all $G \in \operatorname{dom}(\nabla)$. We set $H \coloneqq \delta(\nabla F)$. Let the chaos expansions of H and $G \in \operatorname{dom}(\nabla)$ be given by

$$H(\zeta) = \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} h_n(x)\,\zeta^n(\mathrm{d}x) \qquad \text{and} \qquad G(\zeta) = \mathbb{E}[G(\zeta)] + \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} g_n(x)\,\zeta^n(\mathrm{d}x),$$

respectively. The isometry properties from (4.23) and (5.3) yield

$$\mathbb{E}\left[\delta(\nabla F)G\right] = \sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} h_n(x) g_n(x)\,\rho^{[n]}(\mathrm{d}x) \tag{5.21}$$

as well as

$$\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)(\nabla_x G)\,\zeta(\mathrm{d}x)\right] = \sum_{n=1}^{\infty} \frac{nn!(\theta+n-1)}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x) g_n(x)\,\rho^{[n]}(\mathrm{d}x). \tag{5.22}$$

Since these identities hold for all $G \in \operatorname{dom}(\nabla)$, by choosing, for example, $g_n = 0$ for all $n \in \mathbb{N}$ except for one, we obtain

$$n(\theta+n-1) \int_{\mathbb{X}^n} f_n(x) g(x)\,\rho^{[n]}(\mathrm{d}x) = \int_{\mathbb{X}^n} h_n(x) g(x)\,\rho^{[n]}(\mathrm{d}x)$$

for each $n \in \mathbb{N}$ and each $g \in \mathbb{H}_n$. Equivalently,

$$\int_{\mathbb{X}^n} (n(\theta+n-1) f_n(x) - h_n(x)) g(x)\,\rho^{[n]}(\mathrm{d}x) = 0$$

for each $n \in \mathbb{N}$ and each $g \in \mathbb{H}_n$. Thus,

$$h_n = n(\theta+n-1) f_n, \quad \rho^{[n]}\text{-a.e.}, n \in \mathbb{N}.$$

Since $H \in L^2(\zeta)$, the convergence of

$$\sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} h_n(x)^2 \, \rho^{[n]}(\mathrm{d}x) = \sum_{n=1}^{\infty} \frac{n! n^2 (\theta+n-1)^2}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x)$$

yields $F \in \mathrm{dom}(L)$. Finally, it follows

$$\mathbb{E}\left[\delta(\nabla F) G\right] = \sum_{n=1}^{\infty} \frac{n! n(\theta+n-1)}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x) g_n(x) \, \rho^{[n]}(\mathrm{d}x) = \mathbb{E}[(-L(F))G]. \qquad \square$$

The following remark gathers immediate properties of L.

Remark 5.21. Given $F, G \in \mathrm{dom}(L)$, by Theorem 5.20 and the partial integration formula (5.7), we obtain

$$\mathbb{E}\left[(-L(F))G\right] = \mathbb{E}\left[\delta(\nabla F) G\right] = \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)(\nabla_x G) \, \zeta(\mathrm{d}x)\right] = \mathbb{E}\left[F \delta(\nabla G)\right] = \mathbb{E}\left[(-L(G))F\right], \tag{5.23}$$

showing that L is symmetric and negative semi-definite.

Since $L(F)$ is the L^2-limit of centred random variables, we further infer $\mathbb{E}[L(F)] = 0$ for each $F \in \mathrm{dom}(L)$. ◇

5.4. The uncentred gradient

An examination of the definition of the gradient ∇ prompts consideration of the second term in each summand in the defining series, which ensures that each summand is centred with respect to ζ. It is also possible to define a gradient operator that omits this centring term. This alternative approach is explored in the following section. However, we will see that this approach not only complicates calculations but also lacks the link to the Fleming–Viot processes with parent-independent mutation studied in population genetics.

Definition 5.22. Let $\mathrm{dom}(\nabla^0)$ be the set of all $F \in L^2(\zeta)$ with chaos expansion (4.22) such that the kernel functions satisfy

$$\sum_{n=1}^{\infty} \frac{n n!}{\theta^{(2n-1)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x) < \infty. \tag{5.24}$$

The subsequent remark shows that the sets $\mathrm{dom}(\nabla)$ and $\mathrm{dom}(\nabla^0)$ coincide.

Remark 5.23. The series in (5.1) converges if and only if the series in (5.24) converges. To see this, note first that if $F \in \mathrm{dom}(\nabla^0)$ with chaos expansion (4.22), the inequality

$$\frac{(\theta+n-1) n n!}{\theta^{(2n)}} \leq \frac{n n!}{\theta^{(2n-1)}}, \quad n \in \mathbb{N},$$

implies the convergence of

$$\sum_{n=1}^{\infty} \frac{(\theta+n-1) n n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x)$$

and thus shows $F \in \mathrm{dom}(\nabla)$. Conversely, since the sequence $(\frac{\theta+n-1}{\theta+2n-1})_{n \in \mathbb{N}}$ converges to $\frac{1}{2}$, there exists some $n_0 \in \mathbb{N}$ such that

$$\frac{1}{4} \leq \frac{\theta+n-1}{\theta+2n-1}, \quad n \geq n_0,$$

is valid, and we conclude

$$\frac{1}{4} \frac{n n!}{\theta^{(2n-1)}} \leq \frac{(\theta+n-1) n n!}{\theta^{(2n)}}, \quad n \geq n_0. \qquad ◇$$

As for the gradient ∇, the convergence of the series in the definition of the domain guarantees the convergence of a series in $L^2(C_\zeta)$.

Lemma 5.24. *Let $F \in \mathrm{dom}(\nabla^0)$ with chaos expansion* (4.22)*. Then*

$$\Omega \times \mathbb{X} \ni (\omega, x) \mapsto \sum_{n=1}^{\infty} n \int_{\mathbb{X}^{n-1}} f_n(x, y_1, \dots, y_{n-1})\, \zeta^{n-1}(\omega, \mathrm{d}(y_1, \dots, y_{n-1}))$$

converges in $L^2(C_\zeta)$.

Proof. Let $n_0 \in \mathbb{N}$. Similar to the proof of Lemma 5.2, it holds

$$\mathbb{E}\left[\int_{\mathbb{X}} \left(\sum_{n=1}^{n_0} n \int_{\mathbb{X}^{n-1}} f_n(x, \mathbf{y_{n-1}})\, \zeta^{n-1}(\mathrm{d}\mathbf{y_{n-1}})\right)^2 \zeta(\mathrm{d}x)\right] = \sum_{n=1}^{n_0} \frac{n^2(n-1)!}{\theta^{(2n-1)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x). \tag{5.25}$$

□

The definition of the *uncentred gradient* reads as follows.

Definition 5.25. Let $\nabla^0 \colon \mathrm{dom}(\nabla^0) \to L^2(C_\zeta)$, be defined by

$$(\nabla^0 F)(\omega, x) \coloneqq \sum_{n=1}^{\infty} n \int_{\mathbb{X}^{n-1}} f_n(x, y_1, \dots, y_{n-1})\, \zeta^{n-1}(\omega, \mathrm{d}(y_1, \dots, y_{n-1})),$$

where the kernel functions from the chaos expansion (4.22) of $F \in \mathrm{dom}(\nabla^0)$ are denoted by f_n, $n \in \mathbb{N}$.

Although the definition of the uncentred gradient may seem simpler at first glance, computations in terms of the chaos expansion involving the uncentred gradient can become more complicated. For example, formulas for partial integration become more intricate when the centring is omitted.

Example 5.26. Let $h \colon \mathbb{X}^3 \to \mathbb{R}$ be square-integrable with respect to $\rho^{[3]}$ and assume $h(x, \cdot) \in \mathbb{H}_2$ for each $x \in \mathbb{X}$. Moreover, let $F \in \mathrm{dom}(\nabla^0)$ with chaos expansion (4.22). Let $n_0 \in \mathbb{N}$, $n_0 \geq 3$. By Corollary 3.8 and Corollary 4.7, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_{\mathbb{X}} \left(\int_{\mathbb{X}^2} h(x, \mathbf{z_2})\, \zeta^2(\mathrm{d}\mathbf{z_2})\right) \left(\sum_{n=1}^{n_0} n \int_{\mathbb{X}^{n-1}} f_n(x, \mathbf{y_{n-1}})\, \zeta^{n-1}(\mathrm{d}\mathbf{y_{n-1}})\right) \zeta(\mathrm{d}x)\right] \\
&= \sum_{n=1}^{n_0} \frac{n}{\theta^{(n+2)}} \int_{\mathbb{X}^{n+2}} h(x, \mathbf{z_2}) f_n(x, \mathbf{y_{n-1}})\, \rho^{[n+2]}(\mathrm{d}(x, \mathbf{z_2}, \mathbf{y_{n-1}})) \\
&= \frac{2 \cdot 2!}{\theta^{(4)}} \int_{\mathbb{X}^2} h(x, x, y) f_2(x, y)\, \rho^{[2]}(\mathrm{d}(x, y)) + \frac{3 \cdot 2!}{\theta^{(5)}} \int_{\mathbb{X}^3} h(x, y, z) f_3(x, y, z)\, \rho^{[3]}(\mathrm{d}(x, y, z)).
\end{aligned}$$

Taking the limit $n_0 \to \infty$, which is feasible since $\nabla^0 F \in L^2(C_\zeta)$, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\int_{\mathbb{X}} \left(\int_{\mathbb{X}^2} h(x, \mathbf{z_2})\, \zeta^2(\mathrm{d}\mathbf{z_2})\right) \left(\nabla^0_x F\right) \zeta(\mathrm{d}x)\right] \\
&\qquad = \frac{2 \cdot 2!}{\theta^{(4)}} \int_{\mathbb{X}^2} h(x, x, y) f_2(x, y)\, \rho^{[2]}(\mathrm{d}(x, y)) + \frac{3 \cdot 2!}{\theta^{(5)}} \int_{\mathbb{X}^3} h(x, y, z) f_3(x, y, z)\, \rho^{[3]}(\mathrm{d}(x, y, z)).
\end{aligned}$$

Because of

$$\begin{aligned}
&\mathbb{E}\left[\int_{\mathbb{X}^3} h(z)\, \zeta^3(\mathrm{d}z) \left(\mathbb{E}[F(\zeta)] + \sum_{n=1}^{n_0} \int_{\mathbb{X}^n} f_n(y)\, \zeta^n(\mathrm{d}y)\right)\right] = \sum_{n=1}^{n_0} \frac{1}{\theta^{(n+3)}} \int_{\mathbb{X}^{n+3}} h(\mathbf{z_3}) f_n(\mathbf{y_n})\, \rho^{[n+3]}(\mathrm{d}(\mathbf{y_n}, \mathbf{z_3})) \\
&= \frac{2!}{\theta^{(4)}} \int_{\mathbb{X}^2} h(x, x, y) f_1(y)\, \rho^{[2]}(\mathrm{d}(x, y)) + \frac{2!}{\theta^{(5)}} \int_{\mathbb{X}^3} h(x, y, z) f_2(y, z)\, \rho^{[3]}(\mathrm{d}(x, y, z)) \\
&\qquad + \frac{2 \cdot 2!}{\theta^{(5)}} \int_{\mathbb{X}^2} h(x, x, y) f_2(x, y)\, \rho^{[2]}(\mathrm{d}(x, y)) + \frac{3!}{\theta^{(6)}} \int_{\mathbb{X}^3} h(x, y, z) f_3(x, y, z)\, \rho^{[3]}(\mathrm{d}(x, y, z))
\end{aligned}$$

for each $n_0 \in \mathbb{N}$, $n_0 \geq 3$, by again Corollary 4.7, we conclude that in this case an adjoint or divergence $\delta^0(h)$, which satisfies

$$\mathbb{E}\left[\int_{\mathbb{X}} \left(\int_{\mathbb{X}^2} h(x, \mathbf{z_2})\, \zeta^2(\mathrm{d}\mathbf{z_2})\right) \left(\nabla^0_x F\right) \zeta(\mathrm{d}x)\right] = \mathbb{E}\left[F \delta^0(h)\right],$$

has the form

$$\delta^0(h) = (\theta+5)\int_{\mathbb{X}^3} h(z)\,\zeta^3(\mathrm{d}z) - \frac{\theta+5}{\theta+4}\int_{\mathbb{X}^2}\int_{\mathbb{X}} h(x,\mathbf{y_2})\,\rho(\mathrm{d}x)\,\zeta^2(\mathrm{d}\mathbf{y_2})$$
$$- \frac{2(\theta+6)}{\theta+4}\int_{\mathbb{X}^2} h(x,x,y)\zeta^2(\mathrm{d}(x,y)) - \frac{4}{(\theta+2)^{(3)}}\int_{\mathbb{X}}\int_{\mathbb{X}} h(x,x,y)\,(\rho+\delta_y)(\mathrm{d}x)\,\zeta(\mathrm{d}y), \quad \mathbb{P}\text{-a.s.}$$

(Note that, according to Corollary 5.4, the adjoint to the centred gradient is in this case

$$(\theta+2)\int_{\mathbb{X}^3} h(x)\,\zeta^3(\mathrm{d}x) - \int_{\mathbb{X}^2}\int_{\mathbb{X}} h(t,y,z)\,(\rho+\delta_y+\delta_z)(\mathrm{d}t)\,\zeta^2(\mathrm{d}(y,z)), \quad \mathbb{P}\text{-a.s.})$$

In general, if a random variable $\delta^0(h_m)$, which satisfies

$$\mathbb{E}\left[\int_{\mathbb{X}}\left(\int_{\mathbb{X}^m} h_m(x,\mathbf{z_m})\,\zeta^m(\mathrm{d}\mathbf{z_m})\right)\left(\nabla_x^0 F\right)\,\zeta(\mathrm{d}x)\right] = \mathbb{E}\left[F\delta^0(h_m)\right],$$

for some $m \in \mathbb{N}_0$ and a mapping $h_m \in L^2(\rho^{[m+1]})$ with $h_m(x,\cdot) \in \mathbb{H}_m$ for all $x \in \mathbb{X}$, is to be established, a similar calculation leads to the formula

$$\delta^0(h_m) = (\theta+2m-1)\int_{\mathbb{X}^{m+1}} h_m(z)\,\zeta^{m+1}(\mathrm{d}z) - \frac{\theta+2m-1}{\theta+2m}\int_{\mathbb{X}^m}\int_{\mathbb{X}} h_m(x,\mathbf{y_m})\,\rho(\mathrm{d}x)\,\zeta^m(\mathrm{d}\mathbf{y_m})$$
$$- \frac{m(\theta+2m+2)}{\theta+2m}\int_{\mathbb{X}^m} h_m(x,x,\mathbf{y_{m-1}})\zeta^m(\mathrm{d}(x,\mathbf{y_{m-1}}))$$
$$- \frac{2m}{(\theta+2m-2)^{(3)}}\int_{\mathbb{X}^{m-1}}\int_{\mathbb{X}} h_m(x,x,\mathbf{y_{m-1}})\,(\rho+\delta_{\mathbf{y_{m-1}}})(\mathrm{d}x)\,\zeta^{m-1}(\mathrm{d}\mathbf{y_{m-1}}), \quad \mathbb{P}\text{-a.s.}. \qquad \circ$$

This example not only illustrates that the adjoint in chaos expansion of the uncentred gradient is more involved but also raises the question of an associated generator.

Definition 5.27. Let $\mathrm{dom}(L^0)$ be the set of all $F \in L^2(\zeta)$ with chaos expansion (4.22) such that

$$\sum_{n=1}^{\infty} \frac{(\theta+2n-1)^2 n^2 n!}{\theta^{(2n)}}\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x) < \infty.$$

For such F, let

$$L^0(F) := -\sum_{n=1}^{\infty} n(\theta+2n-1)\int_{\mathbb{X}^n} f_n(x)\,\zeta^n(\mathrm{d}x).$$

We note that the condition imposed on the kernel functions in $\mathrm{dom}(L^0)$ ensures the convergence of $L^0(F)$ for $F \in \mathrm{dom}(L^0)$. As in the case of the gradient, compared to L, the only modification is a factor of 2 in each summand. While L and $\mathrm{dom}(L)$ involve terms of the form $-n(\theta+n-1)$ and $(\theta^{(2n)})^{-1}(\theta+n-1)^2 nn!$, the expressions for L^0 take up $-n(\theta+2n-1)$ and $(\theta^{(2n)})^{-1}(\theta+2n-1)^2 nn!$ for each $n \in \mathbb{N}$, respectively. Thus, the domains of L and L^0 also coincide.

The next result is analogous to (5.23) for the centred case.

Lemma 5.28. *The set* $\mathrm{dom}(L^0)$ *is a subset of* $\mathrm{dom}(\nabla^0)$*. Furthermore, given* $F \in \mathrm{dom}(L^0)$ *and* $G \in \mathrm{dom}(\nabla^0)$*, it holds*

$$\mathbb{E}[(-L^0(F))G] = \mathbb{E}\left[\int_{\mathbb{X}}(\nabla_x^0 F)(\nabla_x^0 G)\,\zeta(\mathrm{d}x)\right].$$

Proof. The inequality $1 \le n(\theta+2n-1)$ for $n \in \mathbb{N}$ implies the inclusion $\mathrm{dom}(L^0) \subseteq \mathrm{dom}(\nabla^0)$.

We denote the kernel functions of $F \in \mathrm{dom}(L^0)$ and $G \in \mathrm{dom}(\nabla^0)$ by f_n and g_n, $n \in \mathbb{N}$, respectively. From equation (5.25) and polarisation, we obtain

$$\mathbb{E}\left[\int_{\mathbb{X}}(\nabla_x^0 F)(\nabla_x^0 G)\,\zeta(\mathrm{d}x)\right] = \sum_{n=1}^{\infty}\frac{nn!}{\theta^{(2n-1)}}\int_{\mathbb{X}^n} f_n(x)g_n(x)\,\rho^{[n]}(\mathrm{d}x).$$

Let $n_0, m_0 \in \mathbb{N}$. Corollary 4.11 yields

$$\mathbb{E}\left[\left(\sum_{n=1}^{n_0} n(\theta+2n-1)\int_{\mathbb{X}^n} f_n(x)\,\zeta^n(\mathrm{d}x)\right)\left(\sum_{n=0}^{m_0}\int_{\mathbb{X}^n} g_n(x)\,\zeta^n(\mathrm{d}x)\right)\right]$$
$$= \sum_{n=1}^{m_0\wedge n_0} \frac{(\theta+2n-1)nn!}{\theta^{(2n)}}\int_{\mathbb{X}^n} f_n(x)g_n(x)\,\rho^{[n]}(\mathrm{d}x).$$

By the Cauchy–Schwarz inequality, an upper bound for this expression is

$$\left(\sum_{n=1}^{m_0\wedge n_0} \frac{(\theta+2n-1)nn!}{\theta^{(2n)}}\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x)\right)^{\frac{1}{2}}\left(\sum_{n=1}^{m_0\wedge n_0} \frac{(\theta+2n-1)nn!}{\theta^{(2n)}}\int_{\mathbb{X}^n} g_n(x)^2\,\rho^{[n]}(\mathrm{d}x)\right)^{\frac{1}{2}}.$$

Since $F, G \in \mathrm{dom}(\nabla^0)$, it follows

$$\mathbb{E}\left[(-L^0(F))G\right] = \sum_{n=1}^{\infty} \frac{(\theta+2n-1)nn!}{\theta^{(2n)}}\int_{\mathbb{X}^n} f_n(x)g_n(x)\,\rho^{[n]}(\mathrm{d}x). \qquad \square$$

Chapter 6

Connection to Markov processes from population genetics

In this chapter, the gradient and the generator from the previous chapter are connected to the operators considered in the study of Markov processes in population genetics. To this end, first some concepts used in the study of Fleming–Viot processes are introduced. In dedicated sections, the gradient ∇, a bilinear form arising from the scalar product in $L^2(C_\zeta)$ and the generator L are related to the operators considered in the context of Fleming–Viot processes with parent-independent mutation.

Throughout this chapter, $\mathbb{X}$ is assumed to be a locally compact Polish space equipped with the Borel σ-field $\mathcal{X}$. By $\mathbf{M}_1(\mathbb{X})$, the set of all probability measures on $\mathbb{X}$ is denoted. Furthermore, let $\theta > 0$ and $\nu_0 \in \mathbf{M}_1(\mathbb{X})$ with support $\mathbb{X}$. Let $\rho \coloneqq \theta\nu_0$ and consider a Dirichlet process ζ on $\mathbb{X}$ with parameter measure ρ.

6.1. Relevant concepts

In this section, concepts used in the studies of Fleming–Viot processes are reviewed.

Let $C_b(\mathbb{X})$ be the space of all bounded and continuous functions $f\colon \mathbb{X} \to \mathbb{R}$ and let $\mathbb{S}$ denote the space of all functions $F\colon \mathbf{M}_1(\mathbb{X}) \to \mathbb{R}$ of the form

$$F(\mu) = \varphi\left(\int_{\mathbb{X}} g_1(y)\,\mu(\mathrm{d}y), \dots, \int_{\mathbb{X}} g_d(y)\,\mu(\mathrm{d}y)\right),$$

where $d \in \mathbb{N}$, $g_1, \dots, g_d \in C_b(\mathbb{X})$ and $\varphi \in C^\infty(\mathbb{R}^d)$. We first revisit a notion of gradient commonly used in the study of Fleming–Viot processes (cf. e.g. Overbeck, Röckner and Schmuland (1995) or Shao (2011) for the definition given here). For $F \in \mathbb{S}$ define $\nabla^*\colon \Omega \times \mathbb{X} \to \mathbb{R}$ by

$$(\nabla^* F)(\omega, x) \coloneqq \sum_{i=1}^{d} (\partial_i\varphi)\left(\int_{\mathbb{X}} g_1(y)\,\zeta(\omega,\mathrm{d}y), \dots, \int_{\mathbb{X}} g_d(y)\,\zeta(\omega,\mathrm{d}y)\right)\left(g_i(x) - \int_{\mathbb{X}} g_i(y)\,\zeta(\omega,\mathrm{d}y)\right). \quad (6.1)$$

Similarly to before, we denote the random variable $\omega \mapsto \nabla^* F(\omega, x)$ by $\nabla^*_x F$. We further introduce a bilinear form $\mathcal{E}^*\colon \mathbb{S} \times \mathbb{S} \to \mathbb{R}$ by

$$\mathcal{E}^*(F, G) \coloneqq \mathbb{E}[\mathrm{Cov}_\zeta(\nabla^* F, \nabla^* G)], \quad F, G \in \mathbb{S}. \quad (6.2)$$

Here, we use the notation

$$\mathrm{Cov}_\zeta(H, \widetilde{H}) \coloneqq \int_{\mathbb{X}} H_x \widetilde{H}_x\,\zeta(\mathrm{d}x) - \int_{\mathbb{X}} H_x\,\zeta(\mathrm{d}x)\int_{\mathbb{X}} \widetilde{H}_x\,\zeta(\mathrm{d}x)$$

for measurable functions $(\omega, x) \mapsto H_x(\omega)$ and $(\omega, x) \mapsto \widetilde{H}_x(\omega)$ in $L^2(C_\zeta)$ and set $\mathrm{Var}_\zeta(H) \coloneqq \mathrm{Cov}_\zeta(H, H)$.

6.2. The gradient

In this section, a connection between the gradient ∇ from Chapter 5 and the gradient ∇^* is drawn.

We recall the tensor product from (3.8).

Lemma 6.1. *Let $m \in \mathbb{N}$ and let $h\colon \mathbb{X}^m \to \mathbb{R}$ be measurable and of the form $h = h_1 \otimes \ldots \otimes h_m$, where each $h_i\colon \mathbb{X} \to \mathbb{R}$, $i \in [m]$, is bounded. Let $F\colon \mathbf{M}_1(\mathbb{X}) \to \mathbb{R}$ be defined by $F(\mu) = \int_{\mathbb{X}^m} h(y)\,\mu^m(\mathrm{d}y)$. It then holds $F(\zeta) \in \mathrm{dom}(\nabla)$ and*

$$\nabla F = \nabla^* F, \quad C_\zeta\text{-a.e.}$$

Proof. At first, we note that $F \in L^2(\zeta)$ since h is bounded. The proof is divided into several steps. At first, $\nabla^* F$ is calculated. In the next part, we consider ∇F, where we distinguish the cases $m = 1$ and $m \geq 2$. While the claim follows easily in the case $m = 1$, for $m \geq 2$, in a first step we consider the terms incorporating x. The terms independent of x are then collected in a final step.

(a) Since F can be represented as

$$F(\mu) = \int_{\mathbb{X}} h_1(y)\,\mu(\mathrm{d}y) \cdot \ldots \cdot \int_{\mathbb{X}} h_m(y)\,\mu(\mathrm{d}y) = \varphi\left(\int_{\mathbb{X}} g_1(y)\,\mu(\mathrm{d}y), \ldots, \int_{\mathbb{X}} g_m(y)\,\mu(\mathrm{d}y)\right), \quad \mu \in \mathbf{M}_1(\mathbb{X}),$$

with $g_i = h_i$, $i \in [m]$, and $\varphi\colon \mathbb{R}^m \to \mathbb{R}$, $\varphi(x_1, \ldots, x_m) = \prod_{i=1}^m x_i$, it is an element of $\mathbb{S}$ and the explicit formula yields

$$\begin{aligned}\nabla_x^* F &= \sum_{i=1}^m (\partial_i \varphi)\left(\int_{\mathbb{X}} h_1(y)\,\zeta(\mathrm{d}y), \ldots, \int_{\mathbb{X}} h_m(y)\,\zeta(\mathrm{d}y)\right)\left(h_i(x) - \int_{\mathbb{X}} h_i(y)\,\zeta(\mathrm{d}y)\right)\\ &= \sum_{i=1}^m h_i(x) \prod_{\substack{j=1\\ j\neq i}}^m \left(\int_{\mathbb{X}} h_j(y)\,\zeta(\mathrm{d}y)\right) - m\int_{\mathbb{X}^m} h(y)\,\zeta^m(\mathrm{d}y), \quad x \in \mathbb{X}.\end{aligned}$$

(b) As a next step, we consider ∇F.

First, assume $m = 1$, i.e. $F(\mu) = \int_{\mathbb{X}} h(y))\,\mu(\mathrm{d}y)$. The chaos expansion of F is given by

$$F(\zeta) = \mathbb{E}[F(\zeta)] + \int_{\mathbb{X}} f_1(x)\,\zeta(\mathrm{d}x), \quad \mathbb{P}\text{-a.s.}$$

with

$$f_1(x) = (\theta+1)\left(\frac{1}{\theta+1}\left(\int_{\mathbb{X}} h(y)\,\rho(\mathrm{d}y) + h(x)\right) - \frac{1}{\theta}\int_{\mathbb{X}} h(y)\,\rho(\mathrm{d}y)\right), \quad \rho\text{-almost all } x \in \mathbb{X}.$$

Because of

$$\nabla_x F = f_1(x) - \int_{\mathbb{X}} f_1(x)\,\zeta(\mathrm{d}x) = h(x) - \int_{\mathbb{X}} h(y)\,\zeta(\mathrm{d}y), \quad x \in \mathbb{X},$$

the claim follows in this case.

Let $m \geq 2$ and $x \in \mathbb{X}$. By Example 4.17, the chaos expansion of F is finite, i.e. $\int_{\mathbb{X}^m} h(y)\,\zeta^m(\mathrm{d}y) \in \bigoplus_{i=0}^m \mathbb{F}_i$. Consequently, F belongs to $\mathrm{dom}(\nabla)$. By definition, we have

$$\nabla_x F = \sum_{n=1}^m n\int_{\mathbb{X}^{n-1}} f_n(x, \mathbf{y_{n-1}})\,\zeta^{n-1}(\mathrm{d}\mathbf{y_{n-1}}) - \sum_{n=1}^m n\int_{\mathbb{X}^n} f_n(\mathbf{y_n})\,\zeta^n(\mathrm{d}\mathbf{y_n}),$$

where the chaos expansion of F is given by (4.22). According to (4.16), the function f_n is for $\rho^{[n]}$-almost all $(x_1, \ldots, x_n) \in \mathbb{X}^n$ given by

$$f_n(x_1, \ldots, x_n) = \frac{\theta + 2n - 1}{n!}\sum_{j=0}^n (-1)^{n-j}(\theta+j)^{(n-1)} \sum_{1 \leq i_1 < \ldots < i_j \leq n} \mathbb{E}\left[F(\zeta_{\rho + \delta_{x_{i_1}} + \ldots + \delta_{x_{i_j}}})\right].$$

We recall the notation introduced in Section 2.4. We decompose ∇F into two parts: one that contains all terms involving x (these terms will be considered in step 1)) and another that consists of the remaining terms independent of x (these terms will be considered in step 2)). As the second sum in the definition of ∇F is independent of x, we consider it in step 2) and now decompose the first part. This gives

$$\begin{aligned}&\sum_{n=1}^m n\int_{\mathbb{X}^{n-1}} f_n(x, \mathbf{y_{n-1}})\,\zeta^{n-1}(\mathrm{d}\mathbf{y_{n-1}})\\ &= \sum_{n=1}^m \frac{\theta+2n-1}{(n-1)!}\Bigg[(-1)^n \frac{\theta^{(n-1)}}{\theta^{(m)}}\rho^{[m]}(h) + (-1)^{n-1}\frac{(\theta+1)^{(n-1)}}{(\theta+1)^{(m)}}(\rho + \delta_x)^{[m]}(h)\\ &\qquad + \sum_{j=1}^{n-1}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} \sum_{1 < i_1 < \ldots < i_j \leq n} \int (\rho + \delta_{y_{i_1}} + \ldots + \delta_{y_{i_j}})^{[m]}(h)\,\zeta^n(\mathrm{d}\mathbf{y_n})\\ &\qquad + \sum_{j=2}^{n}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} \sum_{1 < i_2 < \ldots < i_j \leq n} \int (\rho + \delta_x + \delta_{y_{i_2}} + \ldots + \delta_{y_{i_j}})^{[m]}(h)\,\zeta^n(\mathrm{d}\mathbf{y_n})\Bigg],\end{aligned}$$

where the sums in the last two lines are empty sums in the case $n = 1$. By Corollary 3.12, using the notation introduced in (3.9), for $n \in \{2, \dots, m\}$, the second to last sum in the square brackets is equal to

$$\sum_{j=1}^{n-1}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\binom{n-1}{j}\left(\rho^{[m]}(h)+\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^j(h^j_{\otimes i_1,\dots,i_r})\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r})\right)$$

while the last sum is

$$\sum_{j=2}^{n}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\binom{n-1}{j-1}\Bigg(\rho^{[m]}(h)$$
$$+\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\int_{\mathbb{X}^{j-1}}h^j_{\otimes i_1,\dots,i_r}(\mathbf{y_{j-1}},x)\,\zeta^{j-1}(\mathrm{d}\mathbf{y_{j-1}})\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r})\Bigg).$$

1) In this step, the term containing x is examined, i.e.

$$\sum_{n=2}^{m}\frac{\theta+2n-1}{(n-1)!}\sum_{j=2}^{n}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\binom{n-1}{j-1}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^{j-1}(h^j_{\otimes i_1,\dots,i_r}(\cdot,x))\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r})$$
$$+\sum_{n=1}^{m}\frac{\theta+2n-1}{(n-1)!}(-1)^{n-1}\frac{(\theta+1)^{(n-1)}}{(\theta+1)^{(m)}}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}h^1_{\otimes i_1,\dots,i_r}(x)\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r}).$$

Interchanging the order of summation in the first sum gives

$$\sum_{j=2}^{m}\sum_{n=j}^{m}\frac{\theta+2n-1}{(j-1)!(n-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^{j-1}(h^j_{\otimes i_1,\dots,i_r}(\cdot,x))\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r})$$
$$+\sum_{n=1}^{m}\frac{\theta+2n-1}{(n-1)!}(-1)^{n-1}\frac{(\theta+1)^{(n-1)}}{(\theta+1)^{(m)}}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}h^1_{\otimes i_1,\dots,i_r}(x)\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r}).$$

In both lines, applying Lemma A.2 to the sum over n and inserting the definition of $h^j_{\otimes i_1,\dots,i_r}$ for $j, r \in [m]$ and $(i_1, \dots, i_r) \in [m]^{[r]}$ yields that the above equals

$$\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\Bigg(\sum_{j=2}^{m}\frac{(-1)^{m-j}}{(j-1)!(m-j)!}\sum_{1\le l_1\le\dots\le l_r\le j}\int h_{\otimes i_1,\dots,i_r}(y_{l_1},\dots,y_{l_r})\,\zeta^{j-1}(\mathrm{d}\mathbf{y_{j-1}})$$
$$+\frac{(-1)^{m-1}}{(m-1)!}h_{\otimes i_1,\dots,i_r}(x,\dots,x)\Bigg)\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r}),\quad(6.3)$$

where we set $y_j \coloneqq x$ in every summand in the first line, $j \in \{2, \dots, m\}$. Let $r \in \mathbb{N}$ and $(i_1, \dots, i_r) \in [m]^{[r]}$. We now focus on the expression within the brackets. As in Example 4.17, we distinguish cases based on the multiplicity of identical arguments in the integrand. However, here, we also account for the number $i \in [r]_0$ of times $x = y_j$, $j \in [m]$, occurs. With this aim, for $j \in \{2, \dots, m\}$, we decompose

$$\{(l_1,\dots,l_r)\in\mathbb{N}^r\colon\ 1\le l_1\le\dots\le l_r\le j\}=B_r\cup\bigcup_{i=0}^{r-1}\bigcup_{k=1}^{(j-1)\wedge(r-i)}\bigcup_{\substack{\lambda\text{ multiset of }k\text{ natural numbers}\\\text{that sum up to }r-i}}B_{\lambda,i},$$

where

$$B_r \coloneqq \{(l_1,\dots,l_r)\in\mathbb{N}^r\colon\ l_1=\dots=l_r=j\}\,,$$
$$B_{\lambda,0} \coloneqq \{(l_1,\dots,l_r)\in\mathbb{N}^r\colon\ 1\le l_1\le\dots\le l_r\le j-1,$$
$$\text{and the multiset with elements } l_1,\dots,l_r \text{ has multiplicity structure } \lambda\}$$

for a multiset λ containing $\lambda_1, \ldots, \lambda_k \in \mathbb{N}$ with sum r and

$$B_{\lambda,i} := \{(l_1, \ldots, l_r) \in \mathbb{N}^r : 1 \le l_1 \le \ldots \le l_{r-i} \le j-1,\ l_{r-i+1} = \ldots = l_r = j, \text{ and the multiset with elements } l_1, \ldots, l_{r-i} \text{ has multiplicity structure } \lambda\}$$

for $i \in [r-1]_0$ and a multiset λ containing $\lambda_1, \ldots, \lambda_k \in \mathbb{N}$ with sum $r-i$. For such a multiset λ with $M(\lambda)$ distinct values, denoted by $\nu_1, \ldots, \nu_{M(\lambda)}$, each occurring $m_\lambda(\nu_h)$ times in λ, $h \in [M(\lambda)]$, the set $B_{\lambda,i}$ contains

$$\binom{j-1}{k} \frac{k!}{\prod_{h=1}^{M(\lambda)} m_\lambda(\nu_h)!}$$

elements, $i \in [r-1]$. With this decomposition, for fixed $r \in [m]$ and $(i_1, \ldots, i_r) \in [m]^{[r]}$, the term in the brackets in (6.3) becomes

$$\sum_{j=2}^{m} \frac{(-1)^{m-j}}{(j-1)!(m-j)!} \sum_{i=0}^{r-1} \sum_{k=1}^{(j-1)\wedge(r-i)} \sum_{\lambda} \sum_{(l_1,\ldots,l_r)\in B_{\lambda,i}} \int h_{\otimes i_1,\ldots,i_r}(y_{l_1}, \ldots, y_{l_r})\, \zeta^{j-1}(\mathrm{d}\mathbf{y_{j-1}}) + \sum_{j=2}^{m} \frac{(-1)^{m-j}}{(j-1)!(m-j)!} h_{\otimes i_1,\ldots,i_r}(x, \ldots, x) + \frac{(-1)^{m-1}}{(m-1)!} h_{\otimes i_1,\ldots,i_r}(x, \ldots, x),$$

where again $\sum_\lambda$ in the first line is used to express the sum over all multisets λ as specified above. Since

$$\mathbb{X}^r \ni (t_1, \ldots, t_r) \mapsto \sum_{(i_1,\ldots,i_r)\in[m]^{[r]}} h_{\otimes i_1,\ldots,i_r}(t_1, \ldots, t_r) \rho^{[m-r]}(h^{\otimes i_1,\ldots,i_r})$$

is a symmetric function for $r \in [m]$, we have

$$\int_{\mathbb{X}^{j-1}} \sum_{(i_1,\ldots,i_r)\in[m]^{[r]}} h_{\otimes i_1,\ldots,i_r}(y_{l_1}, \ldots, y_{l_r}) \rho^{[m-r]}(h^{\otimes i_1,\ldots,i_r})\, \zeta^{j-1}(\mathrm{d}\mathbf{y_{j-1}}) = \int_{\mathbb{X}^k} \sum_{(i_1,\ldots,i_r)\in[m]^{[r]}} h_{\otimes i_1,\ldots,i_r}(\underbrace{y_1, \ldots, y_1}_{\lambda_1 \text{ times}}, \ldots, \underbrace{y_k, \ldots, y_k}_{\lambda_k \text{ times}}, \underbrace{x, \ldots, x}_{i \text{ times}}) \rho^{[m-r]}(h^{\otimes i_1,\ldots,i_r})\, \zeta^k(\mathrm{d}\mathbf{y_k}) =: I_{\lambda,i,r}$$

for fixed $j \in \{2, \ldots, m\}$, $i \in [r-1]_0$, $k \in [(j-1)\wedge(r-i)]$, a multiset λ with elements $\lambda_1, \ldots, \lambda_k$ whose sum is $r-i$ and $(l_1, \ldots, l_r) \in B_{\lambda,i}$. The integral now depends solely on λ, which in turn is determined by k, i, and r. Substituting this expression into (6.3) yields

$$\begin{aligned} &\sum_{r=1}^{m}\sum_{j=2}^{m} \frac{(-1)^{m-j}}{(j-1)!(m-j)!} \sum_{i=0}^{r-1} \sum_{k=1}^{(j-1)\wedge(r-i)} \sum_{\lambda} \sum_{(l_1,\ldots,l_r)\in B_{\lambda,i}} I_{\lambda,i,r} \\ &\qquad + \sum_{j=1}^{m} \frac{(-1)^{m-j}}{(j-1)!(m-j)!} \sum_{r=1}^{m} \sum_{(i_1,\ldots,i_r)\in[m]^{[r]}} h_{\otimes i_1,\ldots,i_r}(x, \ldots, x)\rho^{[m-r]}(h^{\otimes i_1,\ldots,i_r}) \\ &= \sum_{r=1}^{m}\sum_{i=0}^{r-1}\sum_{k=1}^{r-i}\sum_{\lambda} I_{\lambda,i,r} \sum_{j=k+1}^{m} \binom{j-1}{k} \frac{k!}{\prod_{h=1}^{M(\lambda)} m_\lambda(\nu_h)!} \frac{(-1)^{m-j}}{(j-1)!(m-j)!} \\ &\qquad + \sum_{j=1}^{m} \frac{(-1)^{m-j}}{(j-1)!(m-j)!} \sum_{r=1}^{m} \sum_{(i_1,\ldots,i_r)\in[m]^{[r]}} h_{\otimes i_1,\ldots,i_r}(x, \ldots, x)\rho^{[m-r]}(h^{\otimes i_1,\ldots,i_r}). \end{aligned}$$

Let $i \in [r-1]_0$ and $k \in [(m-1)\wedge(r-i)]_0$. We proceed by analysing the sum

$$\sum_{j=k+1}^{m} \frac{(-1)^{m-j}}{(j-1-k)!(m-j)!}.$$

According to an index shift and the binomial theorem, it is equal to

$$\sum_{j=0}^{m-k-1} \frac{(-1)^{m-k-1-j}}{j!(m-k-1-j)!} = \frac{1}{(m-k-1)!}(1-1)^{m-k-1}$$

which evaluates to zero unless $k = m-1$, in which case the sum equals 1. If $k = m-1$, we have $(r,i) = (m-1,0)$, $(r,i) = (m,0)$ or $(r,i) = (m,1)$. In the first two cases, we obtain a multiple of $I_{\lambda,0}$ for suitable multisets λ. Being independent of x, these integrals will be considered in step 2). In the latter case, $r = m$ and $i = 1$, the remaining term is

$$\sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers} \\ \text{that sum up to } m-1}} I_{\lambda,1,m} \frac{1}{\prod_{h=1}^{M(\lambda)} m_\lambda(\nu_h)!} = \frac{1}{(m-1)!} I_{\kappa,1,m},$$

where the multiset κ consists of $m-1$ times the element 1. This expression is equal to

$$\begin{aligned} &\frac{1}{(m-1)!} \int_{\mathbb{X}^{m-1}} \sum_{(i_1,\ldots,i_m)\in[m]^{[m]}} h_{\otimes i_1,\ldots,i_m}(y_1,\ldots,y_{m-1},x)\,\zeta^{m-1}(\mathrm{d}\mathbf{y_{m-1}}) \\ &= \frac{1}{(m-1)!}(m-1)! \sum_{i_1=1}^{m} h_{i_1}(x) \int_{\mathbb{X}^{m-1}} h^{\otimes i_1}(y_1,\ldots,y_{m-1})\,\zeta^{m-1}(\mathrm{d}\mathbf{y_{m-1}}) = \nabla_x^* F + m \int_{\mathbb{X}^m} h(y)\,\zeta^m(\mathrm{d}y), \end{aligned}$$

which concludes step 1).

2) In this step, the terms independent of x in the above calculation of ∇F are studied. These terms are

$$-\sum_{n=1}^{m} n \int_{\mathbb{X}^n} f_n(\mathbf{y_n})\,\zeta^n(\mathrm{d}\mathbf{y_n}),$$

the second sum in the definition of ∇F, and

$$\begin{aligned} &\sum_{n=1}^{m} \frac{\theta+2n-1}{(n-1)!} \Bigg[\Bigg(\frac{\theta^{(n-1)}}{\theta^{(m)}} - \frac{(\theta+1)^{(n-1)}}{(\theta+1)^{(m)}} + \mathbb{1}_{\{n\geq 2\}} \sum_{j=1}^{n-1} (-1)^j \frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} \binom{n-1}{j} \Bigg) (-1)^n \rho^{[m]}(h) \\ &+ \mathbb{1}_{\{n\geq 2\}} \sum_{j=1}^{n-1} (-1)^{n-j} \frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} \binom{n-1}{j} \sum_{r=1}^{m} \sum_{(i_1,\ldots,i_r)\in[m]^{[r]}} \zeta^j(h^j_{\otimes i_1,\ldots,i_r}) \rho^{[m-r]}(h^{\otimes i_1,\ldots,i_r}) \\ &\qquad + \mathbb{1}_{\{n\geq 2\}} \sum_{j=2}^{n} (-1)^{n-j} \frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} \binom{n-1}{j-1} \rho^{[m]}(h) \Bigg] \\ &+ \sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers} \\ \text{that sum up to } m-1}} I_{\lambda,0,m-1} \frac{1}{\prod_{h=1}^{M(\lambda)} m_\lambda(\nu_h)!} \\ &+ \sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers} \\ \text{that sum up to } m}} I_{\lambda,0,m} \frac{1}{\prod_{h=1}^{M(\lambda)} m_\lambda(\nu_h)!}, \end{aligned} \tag{6.4}$$

from the above calculations. We consider them separately, starting with (6.4). The only multiset containing $m-1$ natural numbers with sum $m-1$ is the set consisting of $m-1$ times the value 1. Hence, the second to last term in (6.4) is equal to

$$\begin{aligned} &\frac{1}{(m-1)!} \int_{\mathbb{X}^{m-1}} \sum_{(i_1,\ldots,i_{m-1})\in[m]^{[m-1]}} h_{\otimes i_1,\ldots,i_{m-1}}(y_1,\ldots,y_{m-1})\,\zeta^{m-1}(\mathrm{d}\mathbf{y_{m-1}}) \rho^{[1]}(h^{\otimes i_1,\ldots,i_{m-1}}) \\ &\qquad = \sum_{i_1=1}^{m} \rho(h_{i_1}) \zeta^{m-1}(h^{\otimes i_1}). \end{aligned}$$

A multiset of $m-1$ natural numbers whose sum is m contains $m-2$ times the value 1 and once the value 2. Thus, we obtain

$$\sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers} \\ \text{that sum up to } m}} I_{\lambda,0,m} \frac{1}{\prod_{h=1}^{M(\lambda)} m_\lambda(\nu_h)!} = \frac{1}{(m-2)!} \sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers} \\ \text{that sum up to } m}} I_{\lambda,0,m}.$$

Furthermore, because of

$$\sum_{n=2}^{m} \frac{\theta+2n-1}{(n-1)!}(-1)^n \left[\frac{\theta^{(n-1)}}{\theta^{(m)}} - \frac{(\theta+1)^{(n-1)}}{(\theta+1)^{(m)}} + \sum_{j=1}^{n-1}(-1)^j \frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\binom{n-1}{j} \right.$$
$$\left. + \sum_{j=2}^{n}(-1)^j \frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\binom{n-1}{j-1}\right] - (\theta+1)\left(\frac{1}{\theta^{(m)}} - \frac{1}{(\theta+1)^{(m)}}\right)$$
$$= -\sum_{n=1}^{m} \frac{\theta+2n-1}{(n-1)!}(-1)^{n-1}\frac{\theta^{(n-1)}}{\theta^{(m)}} + \sum_{j=1}^{m-1}\frac{1}{j!}\sum_{n=j+1}^{m}\frac{\theta+2n-1}{(n-1-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}$$
$$+ \sum_{j=1}^{m}\frac{1}{(j-1)!}\sum_{n=j}^{m}\frac{\theta+2n-1}{(n-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}},$$

which, when combining the first two sums and using Lemma A.2 as well as the binomial theorem in order to evaluate the last, reduces to

$$\sum_{j=0}^{m-1}\frac{1}{j!}\sum_{n=j+1}^{m}\frac{\theta+2n-1}{(n-1-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} + \sum_{j=1}^{m}\frac{1}{(j-1)!(m-j)!}(-1)^{m-j}$$
$$= \sum_{j=0}^{m-1}\frac{1}{j!}\sum_{n=j+1}^{m}\frac{\theta+2n-1}{(n-1-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} + 0,$$

expression (6.4) becomes

$$\sum_{j=0}^{m-1}\frac{1}{j!}\sum_{n=j+1}^{m}\frac{\theta+2n-1}{(n-1-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\rho^{[m]}(h) + \sum_{i_1=1}^{m}\rho(h_{i_1})\zeta^{m-1}(h^{\otimes i_1})$$
$$+ \sum_{j=1}^{m-1}\frac{1}{j!}\sum_{n=j+1}^{m}\frac{\theta+2n-1}{(n-1-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^j(h^j_{\otimes i_1,\dots,i_r})\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r})$$
$$+ \frac{1}{(m-2)!}\sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers}\\ \text{that sum up to } m}} I_{\lambda,0,m}. \tag{6.5}$$

On the other hand, by Corollary 3.12, we have

$$-\sum_{n=1}^{m} n\int_{\mathbb{X}^n} f_n(\mathbf{y_n})\,\zeta^n(\mathrm{d}\mathbf{y_n}) = -\sum_{n=1}^{m}\frac{\theta+2n-1}{(n-1)!}\left(\sum_{j=0}^{n}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\binom{n}{j}\rho^{[m]}(h)\right.$$
$$\left. + \sum_{j=1}^{n}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\binom{n}{j}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^j(h^j_{\otimes i_1,\dots,i_r})\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r})\right).$$

This is equal to

$$-\sum_{j=0}^{m}\frac{1}{j!}\sum_{n=j}^{m}\frac{(\theta+2n-1)n}{(n-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\rho^{[m]}(h)$$
$$-\sum_{j=1}^{m}\frac{1}{j!}\sum_{n=j}^{m}\frac{(\theta+2n-1)n}{(n-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^j(h^j_{\otimes i_1,\dots,i_r})\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r}). \tag{6.6}$$

Combining (6.5) and (6.6) leads to

$$\sum_{j=0}^{m}\frac{1}{j!}C_j\rho^{[m]}(h) + \sum_{j=1}^{m}\frac{1}{j!}C_j\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^j(h^j_{\otimes i_1,\dots,i_r})\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r}) + \sum_{i_1=1}^{m}\rho(h_{i_1})\zeta^{m-1}(h^{\otimes i_1})$$

$$+\frac{1}{(m-2)!}\sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers}\\ \text{that sum up to } m}} I_{\lambda,0,m}, \tag{6.7}$$

with

$$C_j := \sum_{n=j+1}^{m} \frac{\theta+2n-1}{(n-1-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}}\left(1-\frac{n}{n-j}\right)-(\theta+2j-1)j\frac{(\theta+j)^{(j-1)}}{(\theta+j)^{(m)}},\ \ j\in[m-1]_0,$$

and

$$C_m = -(\theta+2m-1)m\frac{(\theta+m)^{(m-1)}}{(\theta+m)^{(m)}} = -m(\theta+2m-1)\frac{1}{\theta+2m-1} = -m.$$

For $j \in [m-1]_0$, Lemma A.2 implies

$$C_j = -j\sum_{n=j}^{m}\frac{(\theta+2n-1)}{(n-j)!}(-1)^{n-j}\frac{(\theta+j)^{(n-1)}}{(\theta+j)^{(m)}} = (-1)^{m-j-1}\frac{j}{(m-j)!}.$$

Because of

$$\sum_{j=0}^{m}\frac{1}{j!}C_j = \sum_{j=1}^{m-1}\frac{(-1)^{m-j-1}}{(j-1)!(m-j)!}-\frac{1}{(m-1)!} = \sum_{j=1}^{m}\frac{(-1)^{m-j-1}}{(j-1)!(m-j)!} = -(1-1)^{m-1} = 0$$

by the binomial theorem, we obtain that (6.7) equals

$$\sum_{j=1}^{m}\frac{(-1)^{m-j-1}}{(m-j)!(j-1)!}\sum_{r=1}^{m}\sum_{(i_1,\dots,i_r)\in[m]^{[r]}}\zeta^j(h^j_{\otimes i_1,\dots,i_r})\rho^{[m-r]}(h^{\otimes i_1,\dots,i_r})+\sum_{i_1=1}^{m}\rho(h_{i_1})\zeta^{m-1}(h^{\otimes i_1})$$
$$+\frac{1}{(m-2)!}\sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers}\\ \text{that sum up to } m}} I_{\lambda,0,m}.$$

The first term in this expression is

$$\sum_{j=1}^{m}\frac{(-1)^{m-j-1}}{(m-j)!(j-1)!}\sum_{r=1}^{m}\sum_{k=1}^{r\wedge j}\sum_{\lambda}\sum_{(l_1,\dots,l_r)\in B_\lambda} I_{\lambda,0,r}, \tag{6.8}$$

where $\sum_\lambda$ is the sum over all multisets λ consisting of k natural numbers that sum to r and

$$B_\lambda := \{(l_1,\dots,l_r)\in\mathbb{N}^r:\ 1\le l_1\le\ldots\le l_r\le j$$
$$\text{and the multiset with elements } l_1,\dots,l_r \text{ has multiplicity structure } \lambda\}.$$

By symmetry, (6.8) equals

$$\sum_{r=1}^{m}\sum_{k=1}^{r}\sum_{\lambda} I_{\lambda,r}\sum_{j=k}^{m}\frac{(-1)^{m-j-1}j}{(m-j)!(j-k)!}\frac{1}{\prod_{i=1}^{M(\lambda)} m_\lambda(\nu_i)!}.$$

Let $r\in[m]$ and $k\in[r]$. We now proceed to evaluate

$$\sum_{j=k}^{m}\frac{(-1)^{m-j-1}j}{(m-j)!(j-k)!}.$$

Through a series of index shifts and applications of the binomial theorem, this sum simplifies to

$$\sum_{j=0}^{m-k}\frac{(-1)^{m-k-j-1}j}{(m-k-j)!j!}-k\sum_{j=0}^{m-k}\frac{(-1)^{m-k-j}}{(m-k-j)!j!} = \sum_{j=1}^{m-k}\frac{(-1)^{m-k-j-1}}{(m-k-j)!(j-1)!}-k\frac{(1-1)^{m-k}}{(m-k)!}$$
$$= \sum_{j=0}^{m-k-1}\frac{(-1)^{m-k-j}}{(m-k-1-j)!j!}-k\frac{(1-1)^{m-k}}{(m-k)!} = -\mathbb{1}_{\{m-k-1\ge 0\}}\frac{(1-1)^{m-k-1}}{(m-k-1)!}-k\frac{(1-1)^{m-k}}{(m-k)!}.$$

Thus, the only terms remaining in (6.8) are those corresponding to $k = m$ and $k = m-1$, i.e. the cases when $(r,k) \in \{(m,m),(m-1,m-1),(m,m-1)\}$. In the case $r = k = m$, the expression in (6.8) is

$$-m \sum_{\substack{\lambda \text{ multiset of } m \text{ natural numbers}\\ \text{that sum up to } m}} I_{\lambda,0,m} \frac{1}{\prod_{i=1}^{M(\lambda)} m_\lambda(\nu_i)!} = -\frac{m}{m!} I_{\kappa',0,m}$$

$$= -\frac{m}{m!}\int_{\mathbb{X}^m} \sum_{(i_1,\dots,i_m)\in[m]^{[m]}} h_{\otimes i_1,\dots,i_m}(y_1,\dots,y_m)\,\zeta^m(\mathrm{d}\mathbf{y_m}) = -m\int_{\mathbb{X}^m} h(y)\,\zeta^m(\mathrm{d}y),$$

with κ' denoting the multiset consisting of m times the value 1 – the only multiset of m natural numbers summing to m. If $r = m-1$ and $k = m-1$, we obtain

$$\sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers}\\ \text{that sum up to } m-1}} I_{\lambda,0,m-1} \frac{(-1)}{\prod_{i=1}^{M(\lambda)} m_\lambda(\nu_i)!} = -\frac{1}{(m-1)!} I_{\kappa,0,m-1} = -\sum_{i_1=1}^{m} \rho(h_{i_1})\zeta^{m-1}(h^{\otimes i_1}).$$

Finally, if $r = m$ and $k = m-1$,

$$\sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers}\\ \text{that sum up to } m}} I_{\lambda,0,m} \frac{(-1)}{\prod_{h=1}^{M(\lambda)} m_\lambda(\nu_h)!} = -\frac{1}{(m-2)!} \sum_{\substack{\lambda \text{ multiset of } m-1 \text{ natural numbers}\\ \text{that sum up to } m}} I_{\lambda,0,m}. \qquad \square$$

Building on the previous result, we establish a connection between $\nabla^* F$ and ∇F for arbitrary $F \in \mathbb{S}$.

Lemma 6.2. *Let $F \in \mathbb{S}$. Then F is an element of* $\mathrm{dom}(\nabla)$ *and it holds*

$$\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x^* F - \nabla_x F)^2\,\zeta(\mathrm{d}x)\right] = 0. \tag{6.9}$$

Proof. Let $F \in \mathbb{S}$ with representation $F(\mu) = \varphi(\mu(g_1),\dots,\mu(g_d))$, $\mu \in \mathbf{M}_1(\mathbb{X})$, for some $d \in \mathbb{N}$, $\varphi \in C^\infty(\mathbb{R}^d)$ and measurable as well as bounded $g_i\colon \mathbb{X} \to \mathbb{R}$, $i \in [d]$. The boundedness of $g_1,\dots,g_k$ implies the existence of a constant $c > 0$ such that $|g_j(x)| \le c$ holds for all $x \in \mathbb{X}$ and $j \in [k]$. By the Weierstrass approximation theorem (cf. Theorem 1.6.2 in Narasimhan (1985)), for each $n \in \mathbb{N}$, there exists a polynomial $\varphi_n\colon \mathbb{R}^k \to \mathbb{R}$ such that

$$\lim_{n\to\infty} \varphi_n = \varphi \qquad \text{and} \qquad \lim_{n\to\infty} \partial_j\varphi_n = \partial_j\varphi, \quad \text{for all } j \in [k],$$

uniformly on $[-c,c]^k$. Given $n \in \mathbb{N}$, let $F_n\colon \mathbf{M}_1(\mathbb{X}) \to \mathbb{R}$, be defined by

$$F_n(\mu) := \varphi_n(\mu(g_1),\dots,\mu(g_d)).$$

Because of

$$\mathbb{E}[(F_n(\zeta) - F(\zeta))^2] = \mathbb{E}[((\varphi_n - \varphi)(\zeta((g_1),\dots,\zeta(g_n))^2] \to 0, \quad n \to \infty,$$

by dominated convergence, we have $F_n \to F$ as $n \to \infty$ in $L^2(\zeta)$. Furthermore, dominated convergence also yields

$$\mathbb{E}\left[\int_{\mathbb{X}} (\nabla^* F_n - \nabla^* F)^2\,\zeta(\mathrm{d}x)\right] = \mathbb{E}\left[\int_{\mathbb{X}} \left(\sum_{i=1}^{k} (\partial_i(\varphi_n - \varphi))(\zeta(g_1),\dots,\zeta(g_n))(g_i - \zeta(g_i))\right)^2 \zeta(\mathrm{d}x)\right]$$
$$\to 0, \quad n \to \infty.$$

Since each φ_n is a polynomial, the corresponding function F_n, $n \in \mathbb{N}$, can be written as a finite sum of functions of the form considered in Lemma 6.1. Hence, by Lemma 6.1, it follows that for each $n \in \mathbb{N}$ we have $F_n \in \mathrm{dom}(\nabla)$ and

$$\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x^* F_m - \nabla_x^* F_n)^2\,\zeta(\mathrm{d}x)\right] = \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F_m - \nabla_x F_n)^2\,\zeta(\mathrm{d}x)\right], \quad m \in \mathbb{N}.$$

As the left-hand side tends to zero for $m,n \to \infty$, Lemma 5.5 yields $F \in \mathrm{dom}(\nabla)$ and

$$\lim_{n\to\infty} \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F - \nabla_x F_n)^2\,\zeta(\mathrm{d}x)\right] = 0.$$

The assertion follows from the triangle inequality in $L^2(C_\zeta)$. $\square$

The above lemma also states a connection to the "the discrete gradient". (This name is coined in Dello Schiavo and Lytvynov (2023).) The precise connection as well as the connection between the integration-by-parts formula from Flint and Torrisi (2023) in $L^2(\widehat{\rho})$ and the partial integration in $L^2(C_\zeta)$ introduced in the previous chapter is subject to further research.

6.3. The Dirichlet form

In this section, we introduce a bilinear form $\mathcal{E}$ defined on $\mathrm{dom}(\nabla) \times \mathrm{dom}(\nabla)$, which we will identify with the Dirichlet form associated with the generator L_ρ from (2.6). For an introduction to the theory of Dirichlet forms, we refer to Fukushima, Oshima and Takeda (1994) or Ma and Röckner (1992).

Definition 6.3. Let $\mathcal{E}\colon \mathrm{dom}(\nabla) \times \mathrm{dom}(\nabla) \to \mathbb{R}$ be defined by

$$\mathcal{E}(F,G) := \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)(\nabla_x G)\, \zeta(\mathrm{d}x)\right], \quad F, G \in \mathrm{dom}(\nabla), \tag{6.10}$$

and set $\mathcal{E}(F) := \mathcal{E}(F,F)$, $F \in \mathrm{dom}(\nabla)$.

By Corollary 5.4, the gradient is centred with respect to ζ and we can thus write $\mathcal{E}(F,G)$ for $F, G \in \mathrm{dom}(\nabla)$ as

$$\mathcal{E}(F,G) = \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)(\nabla_x G)\, \zeta(\mathrm{d}x)\right] - \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x) \int_{\mathbb{X}} (\nabla_x G)\, \zeta(\mathrm{d}x)\right] = \mathbb{E}\left[\mathrm{Cov}_\zeta(\nabla F, \nabla G)\right] \tag{6.11}$$

or, in terms of the chaos expansion,

$$\mathcal{E}(F,G) = \sum_{n=1}^{\infty} \frac{n n!}{\theta^{(2n)}} (\theta + n - 1) \int_{\mathbb{X}^n} f_n(x) g_n(x)\, \rho^{[n]}(\mathrm{d}x) \tag{6.12}$$

for $F, G \in \mathrm{dom}(\nabla)$ (cf. equation (5.3)), where the kernel functions of F and G are denoted by f_n and g_n, $n \in \mathbb{N}$, respectively. From (5.23), it follows further that

$$\mathcal{E}(F,G) = \mathbb{E}[(-LF)G], \quad F \in \mathrm{dom}(L), G \in \mathrm{dom}(\nabla). \tag{6.13}$$

Moreover, we note that, due to the properties of the covariance, $\mathcal{E}$ is bilinear, symmetric and positive semi–definit.

We now establish that $\mathcal{E}$ is a closed form (cf. e.g. property ($\mathcal{E}$.3) on p. 4 of Fukushima, Oshima and Takeda (1994)), i.e. the space $\mathrm{dom}(\nabla)$ with the metric $\mathrm{dom}(\nabla) \times \mathrm{dom}(\nabla) \ni (F,G) \mapsto \mathcal{E}_1(F,G) := \mathcal{E}(F,G) + \mathbb{E}[F(\zeta)G(\zeta)]$ is complete.

Lemma 6.4. *Let $(F_n)_{n\in\mathbb{N}}$ be a sequence in $\mathrm{dom}(\nabla)$ and assume that $\lim_{m,n\to\infty} \mathcal{E}(F_m - F_n) = 0$, i.e. $(\nabla F_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(C_\zeta)$. Then there exists $F \in \mathrm{dom}(\nabla)$ with*

$$\lim_{n\to\infty} \mathcal{E}(F - F_n) = 0.$$

Proof. The assertion is a consequence of (5.4) and Lemma 5.5. □

The above lemma is a key ingredient in the proof of the next result. We recall the definition of $\mathcal{E}^*$ from (6.2), namely

$$\mathcal{E}^*(F,G) = \mathbb{E}[\mathrm{Cov}_\zeta(\nabla^* F, \nabla^* G)], \quad F, G \in \mathbb{S}.$$

Theorem 6.5. *The operator $(\mathrm{dom}(\nabla) \times \mathrm{dom}(\nabla), \mathcal{E})$ is the closure of $(\mathbb{S} \times \mathbb{S}, \mathcal{E}^*)$.*

Proof. By the polarisation identity,

$$\mathcal{E}(F,G) = \frac{1}{4}\left(\mathcal{E}(F+G) - \mathcal{E}(F-G)\right), \quad F, G \in \mathrm{dom}(\nabla),$$

it suffices to consider the form $\mathcal{E}$ as acting on a single argument, i.e. in the following analysis we consider the mapping $F \mapsto \mathcal{E}(F)$.

Let $F \in \mathbb{S}$. According to Lemma 6.2, it holds $\mathbb{E}\left[\operatorname{Var}_\zeta(\nabla^* F - \nabla F)\right] = 0$. Thus, by the symmetry of $\mathbb{E}\left[\operatorname{Cov}_\zeta(\cdot,\cdot)\right]$, it follows

$$\begin{aligned}\mathbb{E}\left[\operatorname{Var}_\zeta(\nabla F)\right] - \mathbb{E}\left[\operatorname{Var}_\zeta(\nabla^* F)\right] &= \mathbb{E}\left[\operatorname{Cov}_\zeta(\nabla F - \nabla^* F, \nabla F)\right] - \mathbb{E}\left[\operatorname{Cov}_\zeta(\nabla^* F - \nabla F, \nabla^* F)\right] \\ &= \mathbb{E}\left[\operatorname{Cov}_\zeta(\nabla F - \nabla^* F, \nabla F + \nabla^* F)\right].\end{aligned}$$

Applying the Cauchy Schwarz inequality leads to

$$\left|\mathbb{E}\left[\operatorname{Var}_\zeta(\nabla F)\right] - \mathbb{E}\left[\operatorname{Var}_\zeta(\nabla^* F)\right]\right| \le \mathbb{E}\left[\operatorname{Var}_\zeta(\nabla F - \nabla^* F)\right]^{\frac{1}{2}} \mathbb{E}\left[\operatorname{Var}_\zeta(\nabla F + \nabla^* F)\right]^{\frac{1}{2}} = 0.$$

Hence, $\mathcal{E}(F) = \mathcal{E}^*(F)$ for $F \in \mathbb{S}$. Furthermore, $\mathcal{E}$ is closed by Lemma 6.4.

We now show that the closure of the graph of $\mathcal{E}^*$ is the graph of $\mathcal{E}$. On the one hand, based on the reasoning above and the inclusion $\mathbb{S} \subseteq \operatorname{dom}(\nabla)$ from Lemma 6.2, the graph of $\mathcal{E}^*$ is contained within the graph of $\mathcal{E}$. Since the graph of $\mathcal{E}$ is closed, this inclusion also extends to the closure of the graph of $\mathcal{E}^*$. On the other hand, we show that $\mathbb{S}$ is dense in the Hilbert space $(\operatorname{dom}(\nabla), \mathcal{E}_1)$, where

$$\mathcal{E}_1(F,G) := \mathbb{E}[F(\zeta)G(\zeta)] + \mathcal{E}(F,G), \quad F, G \in \operatorname{dom}(\nabla).$$

(The completeness of the space is a consequence of the fact that $\mathcal{E}$ is closed (cf. p. 4 in Fukushima, Oshima and Takeda (1994)).) We already know from Lemma 5.5 and Lemma 6.2 that $\operatorname{dom}(\nabla)$ is closed and that $\mathbb{S} \subseteq \operatorname{dom}(\nabla)$. It remains to show that each $F \in \operatorname{dom}(\nabla)$ can be approximated by a sequence from $\mathbb{S}$. Let $F \in \operatorname{dom}(\nabla)$. By (4.23) and (6.12) we have

$$\mathcal{E}_1(F,F) = \mathbb{E}[F(\zeta)^2] + \sum_{n=1}^{\infty} \frac{n! + (\theta + n - 1)nn!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x).$$

Let $\varepsilon > 0$ and choose $k \in \mathbb{N}$ such that

$$\sum_{n=k+1}^{\infty} \frac{n! + (\theta + n - 1)nn!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x) \le \varepsilon.$$

Let

$$F_0 := \mathbb{E}[F(\zeta)^2] + \sum_{n=1}^{k} \int_{\mathbb{X}^n} f_n(x) \, \zeta^n(\mathrm{d}x)$$

and $\mathbb{S}_0$ be the subspace of $\mathbb{S}$ spanned by the functions

$$\mu \mapsto \int_{\mathbb{X}} h_1(x)\, \mu(\mathrm{d}x) \cdot \ldots \cdot \int_{\mathbb{X}} h_m(x)\, \mu(\mathrm{d}x),$$

with $m \in \mathbb{N}$ and $h_1, \ldots, h_m \in C_b(\mathbb{X})$. Since $\{G(\zeta) : G \in \mathbb{S}_0\}$ is dense in $L^2(\zeta)$ (cf. Lemma 2 in Peccati (2008)), there exists $\widetilde{G} \in \mathbb{S}_0$ such that

$$\mathbb{E}\left[(F_0 - \widetilde{G}(\zeta))^2\right] \le \frac{\varepsilon}{c},$$

where $c := 1 + (\theta + k - 1)k$. Let G be the orthogonal projection of $\widetilde{G}(\zeta)$ onto $\mathbb{F}_0 \oplus \ldots \oplus \mathbb{F}_k$. By orthogonality, we then have

$$\begin{aligned}\mathbb{E}[(F_0 - \widetilde{G}(\zeta))^2] &= \mathbb{E}[((F_0 - G(\zeta)) + (G(\zeta) - \widetilde{G}(\zeta)))^2] = \mathbb{E}[(F_0 - G(\zeta))^2] + \mathbb{E}[(G(\zeta) - \widetilde{G}(\zeta))^2] \\ &\ge \mathbb{E}[(F_0 - G(\zeta))^2].\end{aligned}$$

Let g_n, $n \in \mathbb{N}$, denote the kernel functions of G. By Proposition 4.15 (cf. (4.16)), it holds

$$g_n(x_1, \ldots, x_n) = \frac{\theta + 2n - 1}{n!} \sum_{j=0}^{n} (-1)^{n-j} (\theta + j)^{(n-1)} \sum_{1 \le i_1 < \ldots < i_j \le n} \mathbb{E}\left[\widetilde{G}(\zeta_{\rho + \delta_{x_{i_1}} + \ldots + \delta_{x_{i_j}}})\right], \quad n \in [k],$$

and $g_n \equiv 0$, $n > k$. Since $\widetilde{G} \in \mathbb{S}_0$, Corollary 3.12 yields that $G = G_0(\zeta)$ $\mathbb{P}$-a.s. for some $G_0 \in \mathbb{S}_0$. Moreover,

$$\mathcal{E}_1(F - G, F - G) = \mathbb{E}\left[(F(\zeta) - G(\zeta))^2\right] + \sum_{n=1}^{\infty} \frac{n! + (\theta + n - 1)nn!}{\theta^{(2n)}} \int_{\mathbb{X}^n} (f_n(x) - g_n(x))^2 \, \rho^{[n]}(\mathrm{d}x)$$

$$\leq \mathbb{E}\left[(F(\zeta)-G(\zeta))^2\right] + \sum_{n=1}^{k} \frac{n! + (\theta+n-1)nn!}{\theta^{(2n)}} \int_{\mathbb{X}^n} (f_n(x)-g_n(x))^2 \, \rho^{[n]}(\mathrm{d}x) + \varepsilon$$
$$\leq c \left(\mathbb{E}\left[(F(\zeta)-G(\zeta))^2\right] + \sum_{n=1}^{k} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} (f_n(x)-g_n(x))^2 \, \rho^{[n]}(\mathrm{d}x) \right) + \varepsilon$$
$$= c\,\mathbb{E}[(F_0 - G(\zeta))^2] + \varepsilon \leq 2\varepsilon. \qquad \square$$

6.4. The generator

In this section, the name Fleming–Viot operator is justified by establishing that the operator L from Chapter 5 is indeed the generator of the Fleming–Viot process with parent-independent mutation.

In Stannat (2000), a calculation on p. 678 shows that L_ρ from (2.6) with the mutation operator A given by (2.7) satisfies

$$\mathbb{E}\left[(-L_\rho(\zeta)(F))G\right] = \frac{1}{2}\mathbb{E}\left[\mathrm{Cov}_\zeta(\nabla^* F, \nabla^* G)\right], \quad F, G \in \mathbb{S}.$$

For completeness, we recall the computation using our shorthand notation $\mu(f) = \int_{\mathbb{X}} f(x)\,\mu(\mathrm{d}x)$ for a measure μ on $\mathbb{X}$ and an integrable function f (cf. Section 2.4). Let $F, G \in \mathbb{S}$ with representations

$$F(\mu) = \varphi\left(\mu(f_1), \dots, \mu(f_d)\right) \qquad \text{and} \qquad G(\mu) = \psi\left(\mu(g_1), \dots, \mu(g_e)\right), \quad \mu \in \mathbf{M}_1(\mathbb{X}),$$

where $d, e \in \mathbb{N}$, $\varphi \in C^\infty(\mathbb{R}^d)$, $\psi \in C^\infty(\mathbb{R}^e)$ and $f_i, g_j \colon \mathbb{X} \to \mathbb{R}$ are assumed to be measurable and bounded for $i \in [d], j \in [e]$. Given that with F and G the product FG also belongs to $\mathbb{S}$, we may exploit the explicit definition of L_ρ in order to derive

$$\begin{aligned} 2L_\rho(\mu)(FG) = &\sum_{i,j=1}^{d} (\partial_i\partial_j\varphi)(\mu(f_1),\dots,\mu(f_d))\,\mathrm{Cov}_\mu(f_i,f_j)G + \sum_{i,j=1}^{e} (\partial_i\partial_j\psi)(\mu(g_1),\dots,\mu(g_e))\,\mathrm{Cov}_\mu(g_i,g_j)F \\ &+ 2\sum_{i=1}^{d}\sum_{j=1}^{e} (\partial_i\varphi)(\mu(f_1),\dots,\mu(f_d))(\partial_j\psi)(\mu(g_1),\dots,\mu(g_e))\,\mathrm{Cov}_\mu(f_i,g_j) \\ &+ \sum_{i=1}^{d} (\partial_i\varphi)(\mu(f_1),\dots,\mu(f_d))\mu(Af_i)G + \sum_{i=1}^{e} (\partial_i\psi)(\mu(g_1),\dots,\mu(g_e))\mu(Ag_i)F, \quad \mu \in \mathbf{M}_1(\mathbb{X}). \end{aligned}$$

Given $\mu \in \mathbf{M}_1(\mathbb{X})$, this simplifies to the following expression upon collecting terms

$$2(L_\rho(\mu)(F))G + 2(L_\rho(\mu)(G))F + 2\sum_{i=1}^{d}\sum_{j=1}^{e} (\partial_i\varphi)(\mu(f_1),\dots,\mu(f_d))(\partial_j\psi)(\mu(g_1),\dots,\mu(g_e))\,\mathrm{Cov}_\mu(f_i,g_j).$$

Recalling the definition of ∇^* in (6.1), taking expectations and using that $L_\rho(\zeta)$ is symmetric, we infer

$$\begin{aligned} \mathbb{E}\left[(-L_\rho(\zeta)(F))G\right] &= \frac{1}{2}\mathbb{E}\left[(-L_\rho(\zeta)(F))G\right] + \frac{1}{2}\mathbb{E}\left[(-L_\rho(\zeta)(G))F\right] \\ &= \frac{1}{2}\mathbb{E}\left[(-L_\rho(\zeta)(FG))\right] + \frac{1}{2}\mathbb{E}\left[\mathrm{Cov}_\zeta(\nabla^* F, \nabla^* G)\right]. \end{aligned}$$

As the distribution of ζ is the stationary distribution of the corresponding Fleming–Viot process, i.e. it holds $\mathbb{E}[L_\rho(\zeta)(H)] = 0$ for every $H \in \mathbb{S}$, we thus have

$$\mathbb{E}\left[(-L_\rho(\zeta)(F))G\right] = \frac{1}{2}\mathbb{E}\left[\mathrm{Cov}_\zeta(\nabla^* F, \nabla^* G)\right].$$

Since $\mathbb{E}[\mathrm{Var}_\zeta(\nabla^* F)] = \mathbb{E}[\mathrm{Var}_\zeta(\nabla F)]$, $F \in \mathbb{S}$ by Lemma 6.2, we obtain

$$\mathbb{E}\left[(-L_\rho(\zeta)(F))G\right] = \frac{1}{2}\mathcal{E}(F,G) = \frac{1}{2}\mathbb{E}\left[(-L(F))G\right], \quad F, G \in \mathbb{S}.$$

The generator $\widetilde{L}$ associated with the bilinear form $\mathcal{E}$ from (6.10) is a linear mapping from a subset of $\mathrm{dom}(\nabla)$ into $L^2(\zeta)$ which is defined as follows. The domain $\mathrm{dom}(\widetilde{L})$ of $\widetilde{L}$ is the set of all $F \in \mathrm{dom}(\nabla)$ such that there exists $H \in L^2(\zeta)$ satisfying

$$\mathcal{E}(F,G) = \mathbb{E}[H(\zeta)G(\zeta)], \quad G \in \mathrm{dom}(\nabla).$$

In this case, one defines $\widetilde{L}F = H$. It turns out that $\widetilde{L}$ is the Fleming–Viot operator L from Chapter 5.

Theorem 6.6. *It holds* $\mathrm{dom}(L) = \mathrm{dom}(\widetilde{L})$ *and*

$$\widetilde{L}F = LF, \quad \mathbb{P}\text{-}a.s., \quad F \in \mathrm{dom}(L).$$

Proof. The inclusion $\mathrm{dom}(L) \subseteq \mathrm{dom}(\widetilde{L})$ and $\widetilde{L}F = LF$, $\mathbb{P}$-a.s. for $F \in \mathrm{dom}(L)$ follow from (6.13).

Conversely, let $F \in \mathrm{dom}(\widetilde{L})$, $G \in \mathrm{dom}(\nabla)$ and let $H \in L^2(\zeta)$ be such that $\mathcal{E}(F,G) = \mathbb{E}[H(\zeta)G(\zeta)]$ holds. Choosing $G \equiv 1$ shows $\mathbb{E}[H(\zeta)] = 0$. Let h_n, $n \in \mathbb{N}$, denote the kernel functions in the chaos expansion of H. We can proceed exactly as in the proof of Theorem 5.20 (cf. (5.21) and (5.22)) to show $h_n = (\theta + n - 1)f_n$, $n \in \mathbb{N}$: By (6.12), we have

$$\mathcal{E}(F,G) = \sum_{n=1}^{\infty} \frac{nn!}{\theta^{(2n)}}(\theta + n - 1) \int_{\mathbb{X}^n} f_n(x) g_n(x)\, \rho^{[n]}(\mathrm{d}x),$$

where the kernel functions of F and G are denoted by f_n and g_n, $n \in \mathbb{N}$, respectively. Since this expression is equal to

$$\mathbb{E}[H(\zeta)G(\zeta)] = \sum_{n=1}^{\infty} \frac{n!}{\theta^{[2n]}} \int_{\mathbb{X}^n} h_n(x) g_n(x)\, \rho^{[n]}(\mathrm{d}x)$$

and this equality holds for every $G \in \mathrm{dom}(\nabla)$, we can choose $G \in \mathrm{dom}(\nabla)$ such that its kernel functions satisfy $g_n \equiv 0$ for all except one $n \in \mathbb{N}$ to obtain $h_n = (\theta + n - 1)f_n$, $n \in \mathbb{N}$. Because of $H \in L^2(\zeta)$ and (4.23), the series

$$\sum_{n=1}^{\infty} \frac{n!n^2(\theta+n-1)^2}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x) = \sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} h_n(x)^2\, \rho^{[n]}(\mathrm{d}x) = \mathbb{E}[H(\zeta)^2]$$

converges. Thus, $F \in \mathrm{dom}(L)$. □

Since there is a one-to-one correspondence between the family of closed symmetric forms and the family of non-positive definite self-adjoint operators (cf. Theorem 1.3.1 in Fukushima, Oshima and Takeda (1994)) and $\mathcal{E}$ is generated by L, we obtain that L is the closure of $2L_\rho(\zeta)$.

The semigroup generated by L can be described in terms of the chaos expansion.

Definition 6.7. Let $t \geq 0$ and define

$$T_t(F) := \sum_{n=0}^{\infty} e^{-n(\theta+n-1)t} \int_{\mathbb{X}^n} f_n(x)\, \zeta^n(\mathrm{d}x),$$

for $F \in L^2(\zeta)$ with chaos expansion (4.22) and $f_0 = \mathbb{E}[F(\zeta)]$.

We note that the convergence of the series is guaranteed since for $F \in L^2(\zeta)$ with chaos expansion (4.22) and $m_0, n_0 \in \mathbb{N}$ with $m_0 \leq n_0$ it holds

$$\mathbb{E}\left[\left(\sum_{n=m_0}^{n_0} e^{-n(\theta+n-1)t} \int_{\mathbb{X}^n} f_n(x)\zeta^n(\mathrm{d}x)\right)^2\right] = \sum_{n=m_0}^{n_0} \frac{e^{-2n(\theta+n-1)t} n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x), \quad t \geq 0.$$

For every $t \geq 0$, an upper bound on this is given by $\mathbb{E}[F(\zeta)^2]$, which, according to (4.23), equals

$$\mathbb{E}[F(\zeta)]^2 + \sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x).$$

The subsequent lemma shows that the operator L generates the family $\{T_t, t \geq 0\}$, where we follow the definition of a strongly continuous semigroup from Fukushima, Oshima and Takeda (1994).

Lemma 6.8. *The family* $\{T_t, t \geq 0\}$ *forms a strongly continuous semigroup with generator* L.

Proof. Let $F, G \in L^2(\zeta)$ with kernel functions f_n and g_n, $n \in \mathbb{N}$, respectively. For ease of notation, we set $f_0 = \mathbb{E}[F(\zeta)]$ and $g_0 = \mathbb{E}[G(\zeta)]$. Let $m_0, n_0 \in \mathbb{N}$ and $s, t \geq 0$. By Corollary 4.11, we obtain

$$\mathbb{E}\left[\left(\sum_{m=0}^{m_0} \int_{\mathbb{X}^m} f_m(x)\, \zeta^m(\mathrm{d}x)\right)\left(\sum_{n=0}^{n_0} e^{-n(\theta+n-1)t} \int_{\mathbb{X}^n} g_n(x)\, \zeta^n(\mathrm{d}x)\right)\right]$$

$$= \sum_{n=0}^{n_0} \frac{e^{-n(\theta+n-1)t} n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x) g_n(x)\, \rho^{[n]}(\mathrm{d}x)$$
$$= \mathbb{E}\left[\left(\sum_{m=0}^{m_0} e^{-m(\theta+m-1)t} \int_{\mathbb{X}^m} f_m(x)\, \zeta^m(\mathrm{d}x) \right) \left(\sum_{n=0}^{n_0} \int_{\mathbb{X}^n} g_n(x)\, \zeta^n(\mathrm{d}x) \right) \right],$$

establishing the symmetry of T_t, $t \geq 0$. Moreover, by definition, it holds $T_0(F) = F$. Concerning the semigroup property, we note that $T_t(F)$ is given in its chaos expansion. Hence, we obtain

$$T_t(T_s(F)) = T_{t+s}(F), \quad F \in L^2(\zeta), \;\; s, t \geq 0.$$

Another consequence of the fact that $T_t(F)$ is expressed in terms of its chaos expansion is the contraction property

$$\mathbb{E}[(T_t(F))^2] = \sum_{n=0}^{\infty} \frac{e^{-2n(\theta+n-1)t} n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x) \leq \mathbb{E}[F(\zeta)^2]$$

following from (4.23). In order to show the strong continuity, we consider

$$\mathbb{E}\left[(T_t(F) - F)^2\right] = \mathbb{E}\left[\left(\sum_{n=0}^{\infty} (e^{-n(\theta+n-1)s} - 1) \int_{\mathbb{X}^n} f_n(x)\, \zeta^n(\mathrm{d}x) \right)^2 \right]$$
$$= \sum_{n=0}^{\infty} \frac{(e^{-n(\theta+n-1)t} - 1)^2 n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x).$$

Since the series $\sum_{n=0}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x)$ is convergent, we can apply dominated convergence and infer

$$\mathbb{E}\left[(T_t(F) - F)^2\right] \to 0, \quad t \downarrow 0.$$

Finally, we first assume $F \in \mathrm{dom}(L)$. Using orthogonality and dominated convergence again, which is applicable because of $|\frac{1}{t}(e^{-n(\theta+n-1)t} - 1)| \leq n(\theta + n - 1)$ for all $n \in \mathbb{N}$, $t \geq 0$, we compute

$$\mathbb{E}\left[\left(\frac{1}{t}(T_t(F) - F) - L(F) \right)^2 \right] = \mathbb{E}\left[\left(\sum_{n=0}^{\infty} \left(\frac{1}{t}\left(e^{-n(\theta+n-1)t} - 1 \right) + n(\theta + n - 1) \right) \int_{\mathbb{X}^n} f_n(x)\, \zeta^n(\mathrm{d}x) \right)^2 \right]$$
$$= \sum_{n=0}^{\infty} \left(\frac{e^{-n(\theta+n-1)t} - 1}{t} + n(\theta + n - 1) \right)^2 \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x) \to 0, \quad t \downarrow 0.$$

On the other hand, suppose $F \in L^2(\zeta)$ with chaos expansions (4.22) satisfies

$$\frac{1}{t}(T_t(F) - F) \to H, \quad t \downarrow 0,$$

in $L^2(\zeta)$ for some $H \in L^2(\zeta)$. We set $f_0 = \mathbb{E}[F(\zeta)]$ as well as $h_0 = \mathbb{E}[H(\zeta)]$ and denote the kernel functions of H by h_n, $n \in \mathbb{N}$. The isometry (4.23) gives

$$\mathbb{E}\left[\left(\frac{1}{t}(T_t(F) - F) - H \right) G \right] = \sum_{n=0}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} \left(\frac{e^{-n(\theta+n-1)t} - 1}{t} f_n(x) - h_n(x) \right) g_n(x) \rho^{[n]}(\mathrm{d}x), \quad t \geq 0,$$

for every $G \in L^2(\zeta)$ with kernel functions g_n, $n \in \mathbb{N}$, and $g_0 = \mathbb{E}[G(\zeta)]$. We conclude that F is an element of $\mathrm{dom}(L)$ and $H = L(F)$. □

CHAPTER 7

Variance bounds

In this chapter, we apply the concepts developed in this work to establish both a Poincaré inequality and a reverse Poincaré inequality in the spirit of Schulte and Trapp (2024). We note that the Poincaré inequality for Dirichlet processes is proven in Stannat (2000) by an approximation from the corresponding Poincaré inequality for the Dirichlet distribution from Shimakura (1977).

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space carrying a finite measure ρ with $\theta := \rho(\mathbb{X}) > 0$ and let ζ be a Dirichlet process with parameter measure ρ.

7.1. Poincaré inequality

We can state the result directly.

Theorem 7.1. *Let $F \in \operatorname{dom}(\nabla)$. Then*

$$\operatorname{Var}(F) \leq \frac{1}{\theta}\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)^2 \, \zeta(\mathrm{d}x)\right]. \tag{7.1}$$

Equality holds if and only if there exists a function $g \in L^2(\rho)$ such that $F(\zeta) = \int_{\mathbb{X}} g(y)\, \zeta(\mathrm{d}y)$, $\mathbb{P}$-a.s.

Proof. Let the chaos expansion of F be given by (4.22). According to (6.12), the right-hand side of (7.1) is

$$\frac{1}{\theta}\sum_{n=1}^{\infty} \frac{nn!}{\theta^{(2n)}}(\theta + n - 1)\int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x).$$

From (4.23), we obtain that the left-hand side of (7.1) equals

$$\mathbb{E}\left[(F(\zeta) - \mathbb{E}[F(\zeta)])^2\right] = \sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x).$$

Because of

$$1 \leq \frac{n(\theta + n - 1)}{\theta}$$

for all $n \in \mathbb{N}$, the first part of the claim follows. Note that the inequality $1 \leq \frac{n(\theta+n-1)}{\theta}$ with $n \in \mathbb{N}$ is strict for $n \geq 2$ and equality holds if and only if $n = 1$. (This can be seen, for example, by rewriting the inequality in the equivalent form $0 \leq n^2 + n\theta - n - \theta = (n + \theta)(n - 1)$.)

If equality in (7.1) holds,

$$\int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x)$$

has to vanish for $n \geq 2$. Hence, F is an element of $\mathbb{F}_0 \oplus \mathbb{F}_1$.

Conversely, let $g \in L^2(\rho)$ and $F \in \operatorname{dom}(\nabla)$ with $F(\zeta) = \int_{\mathbb{X}} g(y)\, \zeta(\mathrm{d}y)$, $\mathbb{P}$-a.s. In this case, $F \in \mathbb{F}_0 \oplus \mathbb{F}_1$, as shown in Example 4.17. Consequently, equality holds in (7.1). □

7.2. Reverse Poincaré inequality

In this section, we derive a reverse Poincaré inequality in the spirit of Schulte and Trapp (2024). We begin by defining a suitable higher-order derivative.

Definition 7.2. Let $k \in \mathbb{N}$. The set $\operatorname{dom}(\nabla^k)$ is the set of all $F \in L^2(\zeta)$ with chaos expansion (4.22) such that

$$\sum_{n=k}^{\infty} \frac{n!(n-k+1)^{(k)}}{\theta^{(2n)}}\left((\theta + 2n - k)^{(k)} - (n - k + 1)^{(k)}\right)\int_{\mathbb{X}^n} f_n(x)^2 \, \rho^{[n]}(\mathrm{d}x) < \infty. \tag{7.2}$$

In the case $k = 1$, we obtain $\mathrm{dom}(\nabla^1) = \mathrm{dom}(\nabla)$.

Following the same line of reasoning as in Lemma 5.2, we establish the convergence of a series that will serve to define a higher-order gradient. Of course, Lemma 5.2 is subsumed by the more general statement given below. The choice to defer the general formulation to this section was made to maintain clarity and readability in the earlier part of the thesis (as we have so far only worked with $\nabla^1 = \nabla$).

Let $k \in \mathbb{N}$ and C_{ζ^k} be the Campbell measure of ζ^k, i.e. the probability measure on $\Omega \times \mathbb{X}^k$ defined by

$$C_{\zeta^k}(A) := \int_\Omega \int_{\mathbb{X}^k} \mathbb{1}_A(\omega, x)\, \zeta^k(\omega, \mathrm{d}x)\, \mathbb{P}(\mathrm{d}\omega), \quad A \in \mathcal{A} \otimes \mathcal{X}^{\otimes k}.$$

If $k = 1$, we have $C_{\zeta^1} = C_\zeta$, the Campbell measure of ζ, which has already been frequently used in this thesis.

Lemma 7.3. *Let $k \in \mathbb{N}$ and $F \in \mathrm{dom}(\nabla^k)$ with chaos expansion* (4.22)*. Then*

$$(\omega, x_1, \dots, x_k) \mapsto \sum_{n=k}^{\infty} (n-k+1)^{(k)} \left(\int_{\mathbb{X}^{n-k}} f_n(x_1, \dots, x_k, y_1, \dots, y_{n-k})\, \zeta^{n-k}(\omega, \mathrm{d}(y_1, \dots, y_{n-k})) \right.$$
$$\left. - \int_{\mathbb{X}^n} f_n(y)\, \zeta^n(\omega, \mathrm{d}y) \right)$$

converges in $L^2(C_{\zeta^k})$.

Proof. For $n \in \mathbb{N}$ with $n \geq k$, we define $H_n \colon \Omega \times \mathbb{X}^k \to \mathbb{R}$ by

$$H_n(\omega, \mathbf{x_k}) := \int_{\mathbb{X}^{n-k}} f_n(\mathbf{x_k}, \mathbf{y_{n-k}})\, \zeta^{n-k}(\omega, \mathrm{d}\mathbf{y_{n-k}}) - \int_{\mathbb{X}^n} f_n(y)\, \zeta^n(\omega, \mathrm{d}y), \quad n > k,$$
$$H_k(\omega, \mathbf{x_k}) := f_k(\mathbf{x_k}) - \int_{\mathbb{X}^k} f_k(y)\, \zeta^k(\omega, \mathrm{d}y).$$

Let $m, n \in \mathbb{N}$ with $m, n \geq k$. By Jensen's inequality and Corollary 3.8, we have

$$\begin{aligned}
&\int_{\Omega \times \mathbb{X}^k} \left(\int_{\mathbb{X}^{n-k}} f_n(\mathbf{x_k}, \mathbf{y_{n-k}})\, \zeta^{n-k}(\omega, \mathrm{d}\mathbf{y_{n-k}}) \right)^2 C_{\zeta^k}(\mathrm{d}(\omega, \mathbf{x_k})) \\
&= \mathbb{E}\left[\int_{\mathbb{X}^k} \left(\int_{\mathbb{X}^{n-k}} f_n(\mathbf{x_k}, \mathbf{y_{n-k}})\, \zeta^{n-k}(\mathrm{d}\mathbf{y_{n-k}}) \right)^2 \zeta^k(\mathrm{d}\mathbf{x_k}) \right] \\
&\leq \mathbb{E}\left[\int_{\mathbb{X}^k} \int_{\mathbb{X}^{n-k}} f_n(\mathbf{x_k}, \mathbf{y_{n-k}})^2\, \zeta^{n-k}(\mathrm{d}\mathbf{y_{n-k}})\, \zeta^k(\mathrm{d}\mathbf{x_k}) \right] = \mathbb{E}\left[\int_{\mathbb{X}^n} f_n(y)^2\, \zeta^n(\mathrm{d}y) \right] \\
&= \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(y)^2\, \rho^{[n]}(\mathrm{d}y).
\end{aligned}$$

As the second term in the definition of H_n is also an element of $L^2(C_{\zeta^k})$, we obtain that H_n belongs to $L^2(C_{\zeta^k})$. Furthermore, it holds

$$\begin{aligned}
&\int_{\Omega \times \mathbb{X}^k} H_m(\omega, x) H_n(\omega, x)\, C_{\zeta^k}(\mathrm{d}(\omega, x)) \\
&= \mathbb{E}\left[\int_{\mathbb{X}} \left(\int_{\mathbb{X}^{m-k}} f_m(\mathbf{x_k}, \mathbf{y_{m-k}})\, \zeta^{m-k}(\mathrm{d}\mathbf{y_{m-k}}) - \int_{\mathbb{X}^m} f_m(y)\, \zeta^m(\mathrm{d}y) \right) \right. \\
&\qquad \left. \left(\int_{\mathbb{X}^{n-k}} f_n(\mathbf{x_k}, \mathbf{z_{n-k}})\, \zeta^{n-k}(\mathrm{d}\mathbf{z_{n-k}}) - \int_{\mathbb{X}^n} f_n(z)\, \zeta(\mathrm{d}z) \right) \zeta^k(\mathrm{d}\mathbf{x_k}) \right].
\end{aligned}$$

An application of Corollary 3.8 shows that this equals

$$\frac{1}{\theta^{(m+n-k)}} \int_{\mathbb{X}^{m+n-k}} f_m(\mathbf{x_k}, \mathbf{y_{m-k}}) f_n(\mathbf{x_k}, \mathbf{z_{n-k}})\, \rho^{[m+n-k]}(\mathrm{d}(\mathbf{x_k}, \mathbf{y_{m-k}}, \mathbf{z_{n-k}}))$$
$$- \frac{1}{\theta^{(m+n)}} \int_{\mathbb{X}^{m+n}} f_m(\mathbf{y_m}) f_n(\mathbf{z_n})\, \rho^{[m+n]}(\mathrm{d}(\mathbf{y_m}, \mathbf{z_n})).$$

By Corollary 4.5, this becomes

$$\mathbb{1}_{\{m=n\}}\left(\frac{(n-k)!}{\theta^{(2n-k)}}\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x)-\frac{n!}{\theta^{(2n)}}\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x)\right)$$
$$=\mathbb{1}_{\{m=n\}}\frac{(n-k)!}{\theta^{(2n)}}\left((\theta+2n-k)^{(k)}-(n-k+1)^{(k)}\right)\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x).$$

Hence, for $n_0\in\mathbb{N}$, $n_0\geq k$, we obtain

$$\begin{aligned}
&\int_{\Omega\times\mathbb{X}^k}\left(\sum_{n=k}^{n_0}(n-k+1)^{(k)}H_n(\omega,x)\right)^2 C_{\zeta^k}(\mathrm{d}(\omega,x))\\
&=\sum_{n=k}^{n_0}\sum_{m=k}^{n_0}\mathbb{E}\left[\int_{\mathbb{X}^k}(n-k+1)^{(k)}(m-k+1)^{(k)}H_n(\omega,x)H_m(\omega,x)\,\zeta^k(\mathrm{d}x)\right] \qquad (7.3)\\
&=\sum_{n=k}^{n_0}\frac{\left((n-k+1)^{(k)}\right)^2(n-k)!}{\theta^{(2n)}}\left((\theta+2n-k)^{(k)}-(n-k+1)^{(k)}\right)\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x).
\end{aligned}$$

Because of (7.2), the sequence under consideration is a Cauchy sequence in $L^2(C_{\zeta^k})$. □

Building on the previous Lemma, we define a higher-order derivative. The gradient from Chapter 5 is included in this framework as the case $k=1$.

Definition 7.4. Let $k\in\mathbb{N}$ and let $\nabla^k\colon \operatorname{dom}(\nabla^k)\to L^2(C_{\zeta^k})$ be defined by

$$(\nabla^k F)(\omega,x_1,\dots,x_k):=\sum_{n=k}^{\infty}(n-k+1)^{(k)}\left(\int_{\mathbb{X}^{n-k}} f_n(x_1,\dots,x_k,y_1,\dots,y_{n-k})\zeta^{n-k}(\omega,\mathrm{d}(y_1,\dots,y_{n-k})\right.$$
$$\left.-\int_{\mathbb{X}^n} f_n(x)\,\zeta^n(\omega,\mathrm{d}x)\right),\quad (\omega,x_1,\dots,x_k)\in\Omega\times\mathbb{X}^k,$$

for $F\in\operatorname{dom}(\nabla^k)$ with chaos expansion (4.22).

Note that for $k,n\in\mathbb{N}$ with $k\geq n$, we have

$$(n-k+1)^{(k)}=n(n-1)\dots(n-k+1)=(n)_{(k)},$$

where $(n)_{(k)}$ denotes the falling factorial, defined by $(n)_{(k)}:=n(n-1)\dots(n-k+1)$.

Again, we consider a measurable version and obtain the norm of the gradient directly.

Remark 7.5. Let $k\in\mathbb{N}$ and $F\in\operatorname{dom}(\nabla^k)$ with chaos expansion (4.22). Equation (7.3) yields

$$\mathbb{E}\left[\int_{\mathbb{X}^k}\left((\nabla^k F)(\mathbf{x_k})\right)^2\zeta^k(\mathrm{d}(\mathbf{x_k}))\right]$$
$$=\sum_{n=k}^{\infty}\frac{n!(n-k+1)^{(k)}}{\theta^{(2n)}}\left((\theta+2n-k)^{(k)}-(n-k+1)^{(k)}\right)\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x).$$

Furthermore, as in the case of ∇ (cf. Corollary 5.4), we obtain

$$\mathbb{E}\left[\int_{\mathbb{X}^k}(\nabla^k F)^2\,\zeta^k(\mathrm{d}x)\right]=\sum_{n=k}^{\infty}\frac{n!(n-k+1)^{(k)}}{\theta^{(2n)}}\left((\theta+2n-k)^{(k)}-(n-k+1)^{(k)}\right)\int_{\mathbb{X}^n} f_n(x)^2\,\rho^{[n]}(\mathrm{d}x) \qquad (7.4)$$

and, as with the lower-order gradient, the higher-order gradient is centred (with respect to ζ^k), meaning that

$$\int_{\mathbb{X}^k}(\nabla^k_x F)\,\zeta^k(\mathrm{d}x)=0,\quad \mathbb{P}\text{-a.s.}$$ ◇

We can now follow the approach of Schulte and Trapp (2024) and derive a reverse Poincaré inequality. Schulte and Trapp (2024) prove a lower variance bound for square-integrable functionals of a Poisson process. As noted in their Remark 2.1, their proof relies solely on the Fock space representation of the functional under consideration and the first two difference operators. Therefore, their approach is applicable to any setting where a Fock space representation for the functional and the first two difference operators (or gradients, in our terminology) is available. Since our formulas for the chaos expansions of F, ∇F and $\nabla^2 F$ for $F \in \operatorname{dom}(\nabla^2)$ are more complicated than the corresponding ones (for F, DF and D^2F, where D denotes the difference operator, cf. Chapter 18 in Last and Penrose (2017)) in the Poisson case, the calculations here become more involved, although the underlying approach remains the same.

Theorem 7.6. *Let $F \in \operatorname{dom}(\nabla^2)$. If the inequality*

$$\mathbb{E}\left[\int_{\mathbb{X}^2} (\nabla_x^2 F)\, \zeta^2(\mathrm{d}x)\right] \le C\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x)\right] \tag{7.5}$$

holds for some constant $C > 0$, then there exists $c \ge \frac{1}{256}(9 + 7\theta + C + \theta^2 + \theta C)^4$ such that

$$\operatorname{Var}(F) \ge \frac{1}{c}\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x)\right]. \tag{7.6}$$

Proof. Take $F \in \operatorname{dom}(\nabla^2)$ with chaos expansion (4.22) that satisfies (7.5) for a constant $C > 0$. By (7.4), we have

$$\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x)\right] = \sum_{n=1}^{\infty} \frac{n!n(\theta+n-1)}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x)$$

and

$$\mathbb{E}\left[\int_{\mathbb{X}^2} (\nabla_x^2 F)\, \zeta^2(\mathrm{d}x)\right] = \sum_{n=2}^{\infty} \frac{n!(n-1)n}{\theta^{(2n)}} \left((\theta+2n-2)(\theta+2n-1) - n(n-1)\right) \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x).$$

Using the absolute convergence of the series, we obtain from the assumption that

$$\sum_{n=1}^{\infty} \frac{nn!}{\theta^{(2n)}} \left[C(\theta+n-1) - (n-1)((\theta+2n-2)(\theta+2n-1) - n(n-1))\right] \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x) \ge 0.$$

On the other hand, for every $c > 0$, we have

$$c\operatorname{Var}(F) - \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x)\right] = c\sum_{n=1}^{\infty} \frac{n!}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x) - \sum_{n=1}^{\infty} \frac{nn!(\theta+n-1)}{\theta^{(2n)}} \int_{\mathbb{X}^n} f_n(x)^2\, \rho^{[n]}(\mathrm{d}x).$$

Hence, if we can choose a constant $c > 0$ satisfying

$$c - n(\theta+n-1) \ge n(C(\theta+n-1) - (n-1)((\theta+2n-2)(\theta+2n-1) - n(n-1))) \tag{7.7}$$

for all $n \in \mathbb{N}$, it follows that

$$c\operatorname{Var}(F) - \mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x)\right] \ge C\mathbb{E}\left[\int_{\mathbb{X}} (\nabla_x F)\, \zeta(\mathrm{d}x)\right] - \mathbb{E}\left[\int_{\mathbb{X}^2} (\nabla_x^2 F)\, \zeta^2(\mathrm{d}x)\right] \ge 0.$$

The inequality (7.7) can equivalently be written as

$$c \ge -3n^4 + (8-4\theta)n^3 + (7\theta - \theta^2 + C - 6)n^2 + (\theta^2 + \theta(C-2) - C + 1)n, \quad n \in \mathbb{N}.$$

An upper bound for the right-hand side is $-3n^4 + (9 + 7\theta + C + \theta^2 + \theta C)n^3$, $n \in \mathbb{N}$. We will now derive an upper bound for this expression. Let $\alpha := (9 + 7\theta + C + \theta^2 + \theta C) > 0$ and $f\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) := -3x^4 + \alpha x^3.$$

As a polynomial, f is a smooth function and its first three derivatives are given by

$$f'(x) = -12x^3 + 3\alpha x^2 = x^2(3\alpha - 12x), \quad f''(x) = -36x^2 + 6\alpha x \quad \text{and} \quad f''(x) = -72x + 6\alpha, \quad x \in \mathbb{R}.$$

Thus, f has a saddle point in 0 and a local maximum in $\frac{\alpha}{4}$. Since $f(x) \to -\infty$, $x \to \pm\infty$, this maximum is global and we obtain

$$f(x) \le f\left(\frac{\alpha}{4}\right) = -3\left(\frac{\alpha}{4}\right)^4 + \alpha\left(\frac{\alpha}{4}\right)^3 = \frac{\alpha^4}{256}, \quad x \in \mathbb{R}.$$

□

APPENDIX A

SIMPLE SUMMATION FORMULAS

The following section contains two summation identities used in various arguments in this thesis. We begin with a simple identity that is used in Example 4.17.

Lemma A.1. *For $\theta > 0$ and $m \in \mathbb{N}$ it holds*

$$\sum_{k=1}^{m} \frac{\theta + 2k - 1}{k!}(-1)^k \theta^{(k-1)} = \frac{(-1)^m}{m!}\theta^{(m)} - 1.$$

Proof. We proceed by induction on $m \in \mathbb{N}$. In the base case $m = 1$, both expressions evaluate to $-\theta - 1$.

Assuming the inductive hypothesis for some $m \in \mathbb{N}$, we obtain

$$\sum_{k=1}^{m+1} \frac{\theta + 2k - 1}{k!}(-1)^k \theta^{(k-1)} = \frac{(-1)^m}{m!}\theta^{(m)} - 1 + \frac{\theta + 2m + 1}{(m+1)!}(-1)^{m+1}\theta^{(m)}$$
$$= \frac{(-1)^m}{m!}\theta^{(m)}\left(1 - \frac{\theta + 2m - 1}{m+1}\right) - 1 = \frac{(-1)^m}{m!}\theta^{(m)}\left(-\frac{\theta + m}{m+1}\right) - 1 = \frac{(-1)^{m+1}}{(m+1)!}\theta^{(m+1)} - 1. \qquad \square$$

The subsequent lemma deals with a finite hypergeometric series.

Lemma A.2. *Let $\theta > 0$ and $m \in \mathbb{N}$, $m \geq 2$. For each $j \in [m-1]$ it holds*

$$\sum_{n=j}^{m}(-1)^{n-j}\frac{\theta + 2n - 1}{(n-j)!}(\theta + j)^{(n-1)} = (-1)^{m-j}\frac{(\theta + j)^{(m)}}{(m-j)!}.$$

Proof. The sum can be evaluated directly. In a first step, we obtain

$$\sum_{n=j}^{m}(-1)^{n-j}\frac{\theta + 2n - 1}{(n-j)!}(\theta + j)^{(n-1)}$$
$$= (\theta + 2j - 1)(\theta + j)^{(j-1)} - (\theta + 2j + 1)(\theta + j)^{(j)} + \sum_{n=j+2}^{m}(-1)^{n-j}\frac{\theta + 2n - 1}{(n-j)!}(\theta + j)^{(n-1)}.$$

Using $(\theta + 2j - 1)(\theta + j)^{(j-1)} = (\theta + j)^{(j)}$, this is equal to

$$(\theta + j)^{(j)}\left(-(\theta + 2j) + \sum_{n=j+2}^{m}(-1)^{n-j}\frac{\theta + 2n - 1}{(n-j)!}\prod_{k=j}^{n-2}(\theta + k + j)\right).$$

If $j \geq m - 2$, the remaining sum contains at least one additional term and the expression thus simplifies to

$$(\theta + j)^{(j+1)}\left(-1 + \frac{\theta + 2j + 3}{2!} + \sum_{n=j+3}^{m}(-1)^{n-j}\frac{\theta + 2n - 1}{(n-j)!}\prod_{k=j+1}^{n-2}(\theta + k + j)\right).$$

This, in turn, is equal to

$$\frac{(\theta + j)^{(j+2)}}{2}\left(1 + 2\sum_{n=j+3}^{m}(-1)^{n-j}\frac{\theta + 2n - 1}{(n-j)!}\prod_{k=j+2}^{n-2}(\theta + k + j)\right).$$

Iterating this procedure leads to

$$\frac{(\theta + j)^{(m-2)}}{(m-j-2)!}(-1)^{m-j}\left(1 - \frac{\theta + 2m - 3}{m - 1 - j} + \frac{(\theta + 2m - 1)(\theta + m - 2 + j)}{(m-j-1)(m-j)}\right)$$
$$= \frac{(\theta + j)^{(m-1)}}{(m-j-1)!}(-1)^{m-j}\left(-1 + \frac{\theta + 2m - 1}{m-j}\right) = \frac{(\theta + j)^{(m)}}{(m-j)!}(-1)^{m-j}. \qquad \square$$

Bibliography

Aldous, D. J. (1985). Exchangeability and related topics, *in* P. L. Hennequin (ed.), *École d'Été de Probabilités de Saint-Flour XIII - 1983*, Springer, Berlin, Heidelberg, pp. 1–198.

Bakry, D., Gentil, I. and Ledoux, M. (2014). *Markov Semigroups*, Springer International Publishing, Cham.

Billingsley, P. (1972). On the distribution of large prime divisors, *Periodica Mathematica Hungarica* **2**: 283–289.

Blackwell, D. and MacQueen, J. B. (1973). Ferguson Distributions Via Pólya Urn Schemes, *The Annals of Statistics* **1**(2): 353–355.

Burden, C. J. and Griffiths, R. C. (2019). The stationary distribution of a sample from the Wright–Fisher diffusion model with general small mutation rates, *Journal of Mathematical Biology* **78**(4): 1211–1224.

Crane, H. (2016). The Ubiquitous Ewens Sampling Formula, *Statistical Science* **31**(1): 1–19.

Dawson, D. A. (1993). Measure-valued Markov processes, *in* P.-L. Hennequin (ed.), *Ecole d'Eté de Probabilités de Saint-Flour XXI - 1991*, Springer, Berlin, Heidelberg, pp. 1–260.

Dello Schiavo, L. (2019). Characteristic functionals of Dirichlet measures, *Electronic Journal of Probability* **24**: 1–38.

Dello Schiavo, L. (2022). The Dirichlet–Ferguson diffusion on the space of probability measures over a closed Riemannian manifold, *The Annals of Probability* **50**(2): 591–648.

Dello Schiavo, L. and Lytvynov, E. (2023). A Mecke-type characterization of the Dirichlet–Ferguson measure, *Electronic Communications in Probability* **28**: 1–12.

Dello Schiavo, L. and Quattrocchi, F. (2023). Multivariate Dirichlet Moments and a Polychromatic Ewens Sampling Formula, *arXiv preprint:2309.11292* .

Di Nunno, G., Øksendal, B. and Proske, F. (2009). *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer, Berlin, Heidelberg.

Döring, M. and Stannat, W. (2009). The logarithmic Sobolev inequality for the Wasserstein diffusion, *Probability Theory and Related Fields* **145**(1–2): 189–209.

Engen, S. (1975). A Note on the Geometric Series as a Species Frequency Model, *Biometrika* **62**(3): 697–699.

Etheridge, A. (2011). *Some Mathematical Models from Population Genetics*, Vol. 2012 of *Lecture Notes in Mathematics*, Springer, Berlin, Heidelberg. École d'Été de Probabilités de Saint-Flour XXXIX - 2009.

Ethier, S. N. (1990). The infinitely-many-neutral-alleles diffusion model by ages, *Advances in Applied Probability* **22**(1): 1–24.

Ethier, S. N. and Griffiths, R. C. (1993). The Transition Function of a Fleming-Viot Process, *The Annals of Probability* **21**(3): 1571–1590.

Ethier, S. N. and Kurtz, T. G. (1981). The Infinitely-Many-Neutral-Alleles Diffusion Model, *Advances in Applied Probability* **13**(3): 429–452.

Ethier, S. N. and Kurtz, T. G. (1993). Fleming–Viot Processes in Population Genetics, *SIAM Journal on Control and Optimization* **31**(2): 345–386.

Ethier, S. N. and Kurtz, T. G. (2005). *Markov Processes: Characterization and Convergence*, Wiley Series in Probability and Statistics, reprint edn, John Wiley & Sons, Hoboken, NJ.

Ewens, W. J. (2004). *Mathematical population genetics 1: Theoretical introduction*, 2nd edn, Springer, New York.

Feng, S. (2010). *The Poisson-Dirichlet Distribution and Related Topics: Models and Asymptotic Behaviors*, Springer, Berlin, Heidelberg.

Feng, S., Miclo, L. and Wang, F.-Y. (2017). Poincaré inequality for Dirichlet distributions and infinite-dimensional generalizations, *ALEA. Latin American Journal of Probability and Mathematical Statistics* **14**(1): 361–380.

Feng, S. and Sun, W. (2010). Some diffusion processes associated with two parameter poisson–dirichlet distribution and dirichlet process, *Probability Theory and Related Fields* **148**(3–4): 501–525.

Feng, S., Sun, W., Wang, F.-Y. and Xu, F. (2011). Functional inequalities for the two-parameter extension of the infinitely-many-neutral-alleles diffusion, *Journal of Functional Analysis* **260**(2): 399–413.

Ferguson, T. S. (1973). A Bayesian Analysis of Some Nonparametric Problems, *The Annals of Statistics* **1**(2): 209–230.

Ferreira da Silva, A. R. (2007). A Dirichlet process mixture model for brain MRI tissue classification, *Medical Image Analysis* **11**(2): 169–182.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford.

Fleming, W. H. and Viot, M. (1979). Some Measure-Valued Markov Processes in Population Genetics Theory, *Indiana University Mathematics Journal* **28**(5): 817–843.

Flint, I. and Torrisi, G. L. (2023). An Integration by Parts Formula for Functionals of the Dirichlet–Ferguson Measure, and Applications, *Potential Analysis* **58**(4): 703–730.

Fukushima, M., Oshima, Y. and Takeda, M. (1994). *Dirichlet Forms and Symmetric Markov Processes*, De Gruyter, Berlin, New York.

Griffiths, R. C. (1980). unpublished notes.

Hofrichter, J., Jost, J. and Tran, T. D. (2017). *Information Geometry and Population Genetics: The Mathematical Structure of the Wright–Fisher Model*, Understanding Complex Systems, Springer International Publishing, Cham.

Hogg, R., McKean, J. and Craig, A. (2019). *Introduction to Mathematical Statistics*, 8th edn, Pearson, Boston.

Itô, K. (1951). Multiple Wiener Integral, *Journal of the Mathematical Society of Japan* **3**(1): 157–169.

Jost, J. (2013). *Partial Differential Equations*, Springer, New York.

Kallenberg, O. (2017). *Random Measures, Theory and Applications*, Springer International Publishing, Cham.

Kingman, J. F. C. (1975). Random Discrete Distributions, *Journal of the Royal Statistical Society. Series B (Methodological)* **37**(1): 1–22.

Kingman, J. F. C. (1993). *Poisson processes*, Oxford studies in probability, Clarendon Press, Oxford.

Last, G. (2016). Stochastic analysis for poisson processes, *in* G. Peccati and M. Reitzner (eds), *Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry*, Springer International Publishing, Cham, pp. 1–36.

Last, G. (2020). An Integral Characterization of the Dirichlet Process, *Journal of Theoretical Probability* **33**: 918–930.

Last, G. and Penrose, M. D. (2011). Poisson process Fock space representation, chaos expansion and covariance inequalities, *Probability Theory and Related Fields* **150**(3-4): 663–690.

Last, G. and Penrose, M. D. (2017). *Lectures on the Poisson Process*, Cambridge University Press.

Ma, Z.-M. and Röckner, M. (1992). *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms*, Springer, Berlin, Heidelberg.

Mahmoud, H. M. (2008). *Polya urn models*, Chapman and Hall texts in statistical science series, CRC Press, Boca Raton [u.a.].

Malliavin, P. (1976). Stochastic calculus of variation and hypoelliptic operators, *Proceedings of the International Symposium on Stochastic Differential Equations*, Wiley, Research Institute for Mathematical Sciences, Kyoto University, Kyoto, pp. 195–263.

McCloskey, J. W. (1965). *A model for the distribution of individuals by species in an environment*, Ph. d. thesis, Michigan State University.

Mecke, J. (1967). Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **9**: 36–58.

Narasimhan, R. (1985). *Analysis on real and complex manifolds*, North Holland mathematical library; 35, 3rd edn, North-Holland, Amsterdam.

Ng, K. W., Tian, G.-L. and Tang, M.-L. (2011). *Dirichlet and Related Distributions: Theory, Methods and Applications*, Wiley Series in Probability and Statistics, Wiley.

Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd edn, Springer, Berlin, Heidelberg.

Ohta, T. and Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population, *Genetical Research* **22**(2): 201 – 204.

Overbeck, L., Röckner, M. and Schmuland, B. (1995). An Analytic Approach to Fleming-Viot Processes with Interactive Selection, *The Annals of Probability* **23**(1): 1–36.

Patil, G. P. and Taillie, C. (1977). Diversity as a Concept and its Implications for Random Communities, *Bulletin of the International Statistical Institute* **41**: 497–515.

Peccati, G. (2008). Multiple Integral Representation for Functionals of Dirichlet Processes, *Bernoulli* **14**(1): 91–124.

Petrov, L. A. (2009). Two-parameter family of infinite-dimensional diffusions on the Kingman simplex, *Functional Analysis and Its Applications* **43**(4): 279–296.

Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme, *in* T. S. Ferguson, L. S. Shapley and J. B. MacQueen (eds), *Statistics, probability and game theory. Papers in honor of David Blackwell*, Vol. 30, Institute of Mathematical Statistics Lecture Notes-Monograph Series, Hayward, California, p. 245–267.

Pitman, J. (2006). Combinatorial stochastic processes, *in* J. Picard (ed.), *Ecole d'Eté de Probabilités de Saint-Flour XXXII - 2002*, Springer, Berlin, Heidelberg.

Pitman, J. and Yor, M. (1997). The Two-Parameter Poisson-Dirichlet Distribution Derived from a Stable Subordinator, *The Annals of Probability* **25**(2): 855–900.

Privault, N. and Schoutens, W. (2002). Discrete chaotic calculus and covariance identities, *Stochastics and Stochastic Reports* **72**(3-4): 289–316.

Ren, P. and Wang, F.-Y. (2020). Spectral gap for measure-valued diffusion processes, *Journal of Mathematical Analysis and Applications* **483**(2): 123624.

Rodríguez-Álvarez, M. X., Inácio, V. and Klein, N. (2025). Density regression via Dirichlet process mixtures of normal structured additive regression models, *Statistics and Computing* **35**(2).

Schulte, M. and Trapp, V. (2024). Lower bounds for variances of Poisson functionals, *Electronic Journal of Probability* **29**: 1–43.

Sethuraman, J. (1994). A constructive definition of Dirichlet priors, *Statistica Sinica* **4**(2): 639–650.

Shao, J. (2011). A New Probability Measure-Valued Stochastic Process with Ferguson-Dirichlet Process as Reversible Measure, *Electronic Journal of Probability* **16**: 271–292.

Shiga, T. (1990). A stochastic equation based on a poisson system for a class of measure-valued diffusion processes, *Journal of Mathematics of Kyoto University* **30**(2): 245–279.

Shimakura, N. (1977). Équations différentielles provenant de la génétique des populations, *Tohoku Mathematical Journal* **29**(2): 287–318.

Stannat, W. (2000). On the validity of the log-Sobolev inequality for symmetric Fleming-Viot operators, *The Annals of Probability* **28**(2): 667–684.

Teh, Y. W. (2010). Dirichlet process, *in* C. Sammut and G. I. Webb (eds), *Encyclopedia of Machine Learning*, Springer US, Boston, pp. 280–287.

Topkaya, I. S., Erdogan, H. and Porikli, F. (2013). Detecting and tracking unknown number of objects with dirichlet process mixture models and markov random fields, *in* G. Bebis, R. Boyle, B. Parvin, D. Koracin, B. Li, F. Porikli, V. Zordan, J. Klosowski, S. Coquillart, X. Luo, M. Chen and D. Gotz (eds), *Advances in Visual Computing*, Springer, Berlin, Heidelberg, pp. 178–188.

Tsilevich, N. V. and Vershik, A. M. (1999). Quasi-invariance of the gamma process and multiplicative properties of the Poisson–Dirichlet measures, *Comptes Rendus de l'Académie des Sciences. Série I. Mathématique* **329**(2): 163–168.

von Renesse, M.-K. and Sturm, K.-T. (2009). Entropic Measure and Wasserstein Diffusion, *The Annals of Probability* **37**(3): 1114–1191.

Watterson, G. A. (1976). The Stationary Distribution of the Infinitely-Many Neutral Alleles Diffusion Model, *Journal of Applied Probability* **13**(4): 639–651.

Wiener, N. (1938). The homogeneous chaos, *American Journal of Mathematics* **60**(4): 897–936.

Wright, S. (1931). Evolution in mendelian populations, *Genetics* **16**(2): 97–159.