PDMI PREPRINT — 1/1998

The Markov-Krein correspondence in several dimensions

Sergei KEROV and Natalia TSILEVICH

Russian Academy of Sciences
Steklov Mathematical Institute
St. Petersburg Department
Fontanka 27, St. Petersburg 191011, Russia
E-mail: kerov@pdmi.ras.ru, natalia@pdmi.ras.ru

January 1998

ABSTRACT

We compute the joint moments of several linear functionals with respect to a Dirichlet random measure and some of its generalizations.

Given a probability distribution τ on a space X, let $M = M_\tau$ denote the random probability measure on X known as Dirichlet random measure with the parameter distribution τ. We prove the formula

$$\left\langle \frac{1}{1 - z_1 F_1(M) - \ldots - z_m F_m(M)} \right\rangle = \exp \int \ln \frac{1}{1 - z_1 f_1(x) - \ldots - z_m f_m(x)} \tau(dx),$$

where $F_k(M) = \int_X f_k(x) M(dx)$, the angle brackets denote the average in M, and $f_1, \ldots, f_m$ are the coordinates of a map $f : X \to \mathbb{R}^m$. The formula describes implicitly the joint distribution of the random variables $F_k(M)$, $k = 1, \ldots, m$. Assuming that the joint moments $p_{k_1,\ldots,k_m} = \int f_1^{k_1}(x) \ldots f_m^{k_m}(x) d\tau(x)$ are all finite, we restate the above formula as an explicit description of the joint moments of the variables $F_1, \ldots, F_m$ in terms of $p_{k_1,\ldots,k_m}$.

In case of a finite space, $|X| = N+1$, the problem is to describe the image μ of a Dirichlet distribution

$$\frac{M_0^{\tau_0 - 1} M_1^{\tau_1 - 1} \ldots M_N^{\tau_N - 1}}{\Gamma(\tau_0)\Gamma(\tau_1)\ldots\Gamma(\tau_N)} dM_1 \ldots dM_N; \qquad M_0, \ldots, M_N \geq 0,\ M_0 + \ldots + M_N = 1$$

on the N-dimensional simplex Δ^N under a linear map $f : \Delta^N \to \mathbb{R}^m$. An explicit formula for the density of μ was already known in case of $m = 1$; here we find it in case of $m = N - 1$.

Partially supported by the RFBR Grant 96-15-96060.

1. Introduction.

We study in this paper the images of classical Dirichlet measures under linear or affine transformations. By *Dirichlet measure* we mean a probability distribution on an N-dimensional simplex Δ^N determined by the density

$$\frac{\mu(dM)}{dM_1\dots dM_N}=\frac{\Gamma(\tau_0+\dots+\tau_N)}{\Gamma(\tau_0)\dots\Gamma(\tau_N)}M_0^{\tau_0-1}\dots M_N^{\tau_N-1}. \tag{1.1}$$

Here $M_0,\dots,M_N\geq 0$, $M_0+\dots+M_N=1$ are the barycentric coordinates in Δ^N, and $\tau_0,\dots,\tau_N>0$, $\tau_0+\dots+\tau_N=1$ are arbitrary parameters.

Every measure μ in $\mathbb{R}^{N+1}$ is uniquely determined by its additive Cauchy-Stieltjes transform

$$R_\mu(z_0,\dots,z_N)=\int\frac{\mu(dx)}{1-z_0x_0-\dots-z_Nx_N}, \tag{1.2}$$

which is correctly defined at least for all $z=(z_0,\dots,z_N)\in i\mathbb{R}^{N+1}$. One can check that

$$R_\mu(z_0,\dots,z_N)=\prod_{j=0}^{N}(1-z_j)^{-\tau_j} \tag{1.3}$$

in case of the Dirichlet measure (1.1).

We define the *multiplicative* version of the Cauchy-Stieltjes transform of a probability measure τ as

$$Q_\tau(z_0,\dots,z_N)=\exp\int\ln\frac{1}{1-z_0x_0-\dots-z_Nx_N}\tau(dx). \tag{1.4}$$

For instance, consider a *free* discrete probability distribution τ in $\mathbb{R}^{N+1}$ determined by some positive weights $\tau_0,\dots,\tau_N$ at the basis vectors $e_0,\dots,e_N\in\mathbb{R}^{N+1}$. Then

$$Q_\mu(z_0,\dots,z_N)=\prod_{j=0}^{N}(1-z_j)^{-\tau_j}. \tag{1.5}$$

We say that a measure μ in $\mathbb{R}^n$ is the Markov-Krein transform of a distribution τ in $\mathbb{R}^n$, if

$$\int\frac{\mu(dx)}{1-z_1x_1-\dots-z_nx_n}=\exp\int\ln\frac{1}{1-z_1x_1-\dots-z_nx_n}\tau(dx), \tag{1.6}$$

for all $z=(z_1,\dots,z_n)\in i\mathbb{R}^n$. Comparing (1.3) and (1.5) we conclude that the Markov-Krein transform of a free measure τ is the corresponding Dirichlet measure (1.1).

The central observation of this paper is that the formula (1.6) behaves nicely upon affine transformations. Given an affine map $f:\mathbb{R}^n\to\mathbb{R}^m$, we denote by μ_f, τ_f the f-images of a pair of probability distributions μ, τ in $\mathbb{R}^n$. It is straightforward that if μ is the Markov-Krein transform of τ, then μ_f is the Markov-Krein transform of τ_f.

We derive from this observation that, for every probability distribution τ in $\mathbb{R}^n$ with the property that

$$\int_{\mathbb{R}^n}\ln(1+\|x\|)\tau(dx)<\infty, \tag{1.7}$$

there exists a unique probability distribution μ in $\mathbb{R}^n$ that satisfies the basic identity (1.6).

In case of dimension $n=1$, this fact was established in [CR], along with the explicit formula for the density,

$$\frac{\mu(da)}{da} = \frac{\sin \tau\{(a,\infty)\}\pi}{\pi} \exp \int \ln \frac{1}{|a-u|} \,\tau(du). \tag{1.8}$$

We provide here two simple direct proofs of this formula. We also find another particular case where the density of the measure μ in (1.6) can be written explicitly. Namely, we observe that there is an expression for this density in terms of Lauricella hypergeometric functions, if μ is a projection of some Dirichlet measure with one-dimensional kernel.

Assuming that the joint moments

$$p_{k_1,\dots,k_m} = \int f_1(x)^{k_1} \dots f_m(x)^{k_m} \;\tau(dx)$$

of a number of linear functionals $f_1, \dots, f_m$ with respect to a measure τ in $\mathbb{R}^n$ are all finite, we derive that the joint moments

$$h_{k_1,\dots,k_m} = \int f_1(x)^{k_1} \dots f_m(x)^{k_m} \;\mu(dx)$$

of the measure μ in (1.6) are also finite, and can be obtained as averages over symmetric groups,

$$h_{n_1,\dots,n_m} = \frac{1}{n!} \sum_{w\in\mathfrak{S}_n} \prod_{c\in C(w)} p_{k_1(c),\dots,k_m(c)}. \tag{1.9}$$

Here $C(w)$ is the set of cycles of a permutation w. The set B of n objects where the group $\mathfrak{S}_n$ acts is partitioned into the subsets $B_1, \dots, B_m$ of cardinalities $n_1, \dots, n_m$, and $k_i(c) = |c \bigcap B_i|$ is the number of elements of B_i in a cycle c.

The paper is organized as follows. We start in Section 2 by recalling known facts on the Markov-Krein transform on the real line. In Section 3 we define the additive and multiplicative Markov-Krein transforms in several dimensions, and prove the basic covariance property (Proposition 3.10). In Section 4 we check that the basic identity (1.6) holds for the classical Dirichlet density and the free measure in a finite-dimensional space. In Sections 5 and 6 we show that the density of a linear image of a Dirichlet distribution can be written explicitly in two particular cases: when the dimension of the image or that of the kernel of the linear map is 1. The basic notion we need here is that of Lauricella hypergeometric functions. In Section 7 we prove the formula relating the moments of the distribution τ and those of its Markov-Krein transform μ. In Section 8 we study the joint distributions of a finite number of linear functionals with respect to a Dirichlet random measure with a continuous parametric measure τ. In Section 9 we generalize the Markov–Krein correspondence to infinite-dimensional topological linear spaces. In the final Section 10 we consider exchangeable random sequences associated with random Dirichlet measures, and generalize the Markov-Krein transform to the case of two-parameter Pitman's partition structures.

In the conclusion, let us mention an interesting open problem. As is well known (cf. [K]), the measure μ in the basic identity (1.6) on the real line, $n=1$, can very well be positive even if the parametric measure τ is a signed measure. In fact, the proper condition on τ which ensures that μ be positive is the *interlacing property*

$$\tau\{x:\; x>a\} > 0; \qquad \tau\{x:\; x<a\} > 0, \qquad \text{for all } a \in \mathbb{R}.$$

It would be interesting to find a similar property of a measure τ in several dimensions which implies positivity of the distribution μ in (1.6). An obvious necessary condition is

$$\tau\{x:\ f(x)<a\}>0 \quad \text{for every } f:\mathbb{R}^n\to\mathbb{R} \text{ and } a\in\mathbb{R}. \tag{1.10}$$

We show in Example 6.5 that this condition is not sufficient.

2. The one-dimensional Markov-Krein transform.

The familiar Cauchy-Stieltjes transform R_μ of a distribution μ on the real line $\mathbb{R}$ is defined by the integral

$$R_\mu(z)=\int\frac{\mu(dx)}{1-zx}. \tag{2.1}$$

If $\operatorname{Im} z\neq 0$, the integral converges for all probability measures μ. Given a measure τ on $\mathbb{R}$ such that $\int\ln(1+|x|)\tau(dx)<\infty$ (we call such a measure *admissible*), a multiplicative analog of the Cauchy-Stieltjes transform can be naturally defined as the function

$$Q_\tau(z)=\exp\int\ln\frac{1}{1-zx}\tau(dx). \tag{2.2}$$

It is well known that every function Q_τ, where τ is an admissible probability distribution, also admits a representation (2.1) as an additive Cauchy-Stieltjes transform of a measure μ. Recall that the measure μ can be restored from its function R_μ by the Perron-Stieltjes inversion formula

$$\mu([a,b])=-\frac{1}{\pi}\lim_{v\to+0}\int_a^b \operatorname{Im}\,\widetilde{R}_\mu(u+iv), \tag{2.3}$$

where $\widetilde{R}_\mu(z)=R_\mu(1/z)/z=\int\mu(du)/(z-u)$. Therefore, an admissible probability measure τ uniquely determines the probability distribution μ via the identity

$$\int\frac{\mu(dx)}{1-zx}=\exp\int\ln\frac{1}{1-zx}\tau(dx). \tag{2.4}$$

(2.5) **Example.** One can check that if $\tau(dx)=\frac{1}{\pi}\frac{dx}{\sqrt{4-x^2}}$ is the arcsine distribution, and $\mu(dx)=\frac{1}{2\pi}\sqrt{4-x^2}$ is the semi-circle law, then (2.4) holds for all $z\notin[-2,2]$.

The formula (2.4) was studied, in particular, by A. A. Markov [Ma] and M. G. Krein and his school (see [KN]) in connection with the so called Markov moment problem. We shall say that μ is the *Markov-Krein transform* of the measure τ. See [K] for a survey of various applications of the Markov-Krein transform.

If the measure μ has finite moments

$$p_n=\int x^n\tau(dx),\qquad n=1,2,\dots,$$

then its Markov-Krein transform μ also has finite moments

$$h_n=\int x^n\mu(dx),\qquad n=1,2,\dots,$$

and the identity

$$h_n=\frac{1}{n!}\sum_{w\in S_n}p_1^{r_1(w)}p_2^{r_2(w)}\dots \tag{2.6}$$

holds, where $r_k(w)$ is the number of cycles of length k in a permutation $w\in S_n$ (see [K]).

3. Additive and multiplicative Cauchy-Stieltjes integrals in several dimensions.

Let z, x be two (column) vectors in $\mathbb{R}^m$, and denote by $z^T = (z_1, \dots, z_m)$, $x^T = (x_1, \dots, x_m)$ the corresponding row vectors. We write $A^T = (a_{ji})$ for the transpose of a matrix $A = (a_{ij})$. In particular, $z^T x = z_1 x_1 + \dots + z_m x_m$.

(3.1) **Definition.** The function

$$R_\mu(z) = \int_{\mathbb{R}^m} \frac{\mu(dx)}{1 - z^T x}, \qquad z \in \mathbb{C}^m \tag{3.2}$$

will be referred to as the *additive Cauchy-Stieltjes transform* of a probability distribution μ in $\mathbb{R}^m$. Note that the function is correctly defined for all vectors $z = iy$, $y \in \mathbb{R}^m$.

The function R_μ may be considered as a moment generating function of the measure μ. In fact, if the joint moments

$$h_{n_1, \dots, n_m} = \int_{\mathbb{R}^m} x_1^{n_1} \dots x_m^{n_m} \, \mu(dx) \tag{3.3}$$

of the measure μ are all finite, then $R_\mu(z)$ expands in a formal series

$$R_\mu(z) = \sum_{n_1, \dots, n_m \geq 0} n! h_{n_1, \dots, n_m} \frac{z_1^{n_1}}{n_1!} \dots \frac{z_m^{n_m}}{n_m!}. \tag{3.4}$$

Let us call a probability distribution τ in $\mathbb{R}^m$ *admissible* if

$$\int_{\mathbb{R}^m} \ln(1 + \|x\|)\tau(dx) < \infty,$$

where $\|x\| = \sqrt{x_1^2 + \dots + x_m^2}$. We define the *multiplicative Cauchy-Stieltjes transform* of an admissible probability measure τ in $\mathbb{R}^m$ as a natural generalization of the integral (2.2):

$$Q_\tau(z) = \exp \int \ln \frac{1}{1 - z^T x} \tau(dx); \qquad z \in \mathbb{C}^m. \tag{3.5}$$

The function is correctly defined for all vectors $z = iy$, $y \in \mathbb{R}^m$.

(3.6) **Example.** Assume that the measure τ is discrete, with some weights τ_j at the vectors $x(j)$, $j = 1, 2, \dots, N$. Then

$$Q_\tau(z) = \prod_j \left(1 - z^T x(j)\right)^{-\tau_j}. \tag{3.7}$$

(3.8) **Definition.** We say that a measure μ in $\mathbb{R}^m$ is the *Markov-Krein transform* of an admissible probability distribution τ, if the additive Cauchy-Stieltjes transform of μ coincides with the multiplicative Cauchy-Stieltjes transform of τ,

$$\int \frac{\mu(dx)}{1 - z^T x} = \exp \int \ln \frac{1}{1 - z^T x} \tau(dx), \tag{3.9}$$

for all $z \in i\mathbb{R}^m$.

Both additive and multiplicative Cauchy-Stieltjes transforms behave nicely upon affine changes of variables.

(3.10) **Proposition.** *Consider an affine transformation $f(x) = a + Ax$, where A is a $k \times n$ real matrix, and $a \in \mathbb{R}^k$ is a column vector. Let μ_f be the image in $\mathbb{R}^k$ of a probability measure μ in $\mathbb{R}^n$. Then*

$$R_{\mu_f}(z) = \frac{1}{1 - z^T a} R_\mu\Big(\frac{A^T z}{1 - z^T a}\Big). \tag{3.11}$$

Proof. By definition,

$$R_{\mu_f}(z) = \int \frac{\mu_f(dy)}{1 - z^T y} = \int \frac{\mu(dx)}{1 - z^T(a + Ax)} =$$
$$= \frac{1}{1 - z^T a} \int \frac{\mu(dx)}{1 - \frac{A^T z x}{1 - z^T a}} = \frac{1}{1 - z^T a} R_\mu\Big(\frac{A^T z}{1 - z^T a}\Big). \quad \square$$

Quite analogously, one can prove a similar proposition for the multiplicative Cauchy-Stieltjes transform.

(3.12) **Proposition.** *Let τ_f be the image of a probability measure τ under an affine transformation $f(x) = a + Ax$. Then*

$$Q_{\tau_f}(z) = \frac{1}{1 - z^T a} Q_\tau\Big(\frac{A^T z}{1 - z^T a}\Big). \tag{3.13}$$

(3.14) **Corollary.** *Let μ_f and τ_f be the images under an affine map $f(x) = a + Ax$ of some probability measures μ and τ. If μ is the Markov-Krein transform of the measure τ, then*

$$\int \frac{\mu_f(dy)}{1 - z^T y} = \exp \int \ln \frac{1}{1 - z^T y} \tau_f(dy); \qquad z \in i\mathbb{R}^k,$$

so that μ_f is the Markov-Krein transform of the measure τ_f.

4. The Dirichlet measures in $\mathbb{R}^m$.

Let Δ^N be an N-dimensional simplex with barycentric coordinates $M_0, M_1, \dots, M_N \geq 0$, $M_0 + M_1 + \ldots + M_N = 1$. The formula

$$\mu(dM) = \frac{\Gamma(\tau_0 + \tau_1 + \ldots + \tau_N)}{\Gamma(\tau_0)\Gamma(\tau_1)\ldots\Gamma(\tau_N)}\, M_0^{\tau_0 - 1} M_1^{\tau_1 - 1} \ldots M_N^{\tau_N - 1}\, dM_1 \ldots dM_N \tag{4.1}$$

determines a measure μ on Δ^N referred to as *Dirichlet distribution* with the parameters τ_0, τ_1, $\dots$, τ_N. The fact that μ is indeed a *probability* distribution follows easily from the well known *Dirichlet integral*

$$\begin{aligned} &\frac{\Gamma(\tau_1 + \ldots + \tau_N)}{\Gamma(\tau_1)\ldots\Gamma(\tau_N)} \int \cdots \int_{\Delta^N} \varphi(x_1 + \ldots + x_N)\, x_1^{\tau_1 - 1} \ldots x_N^{\tau_N - 1}\, dx_1 \ldots dx_N = \\ &\qquad = \int_0^1 \varphi(s)\, s^{\tau_1 + \ldots + \tau_N - 1}\, ds. \end{aligned} \tag{4.2}$$

One can also derive from the Dirichlet integral (4.2) the joint moments $h_{n_0, n_1, \dots, n_N}$ of the Dirichlet distribution μ:

$$h_{n_0, n_1, \dots, n_N} = \frac{1}{n!} (\tau_0)_{n_0} (\tau_1)_{n_1} \ldots (\tau_N)_{n_N}, \tag{4.3}$$

where $n = n_0 + n_1 + \ldots + n_N$ and $(a)_m = a(a+1)\ldots(a+m-1)$ is the Pochhammer symbol.

(4.4) **Proposition.** *Assume that* $\tau_0 + \tau_1 + \ldots + \tau_N = 1$. *Then the additive Cauchy-Stieltjes transform of the Dirichlet distribution (4.1) is*

$$\int_{\Delta^N}\cdots\int \frac{\mu(dx)}{1 - z_0x_0 - z_1x_1 - \ldots - z_Nx_N} = \prod_{j=0}^{N} \frac{1}{(1-z_j)^{\tau_j}}. \tag{4.5}$$

Proof. We shall employ the *Lauricella function* defined by the series (cf. [Ex, (2.1.4)])

$$\begin{aligned} &F_D^{(n)}(a,\, b_1,\, \ldots,\, b_n;\, c;\, z_1,\, \ldots,\, z_n) = \\ &= \sum_{m_1,\ldots,m_n \geq 0} \frac{(a)_{m_1+\ldots+m_n}(b_1)_{m_1}\ldots(b_n)_{m_n}}{(c)_{m_1+\ldots+m_n}} \frac{z_1^{m_1}}{m_1!}\cdots\frac{z_n^{m_n}}{m_n!}. \end{aligned} \tag{4.6}$$

This function admits the following Euler type integral representation (see [Ex, (2.3.5)]):

$$\begin{aligned} &\frac{\Gamma(b_0)\Gamma(b_1)\ldots\Gamma(b_n)}{\Gamma(b_0+b_1+\ldots+b_n)} F_D^{(n)}(a,\, b_1,\, \ldots,\, b_n;\, b_0+b_1+\ldots+b_n;\, z_1,\, \ldots,\, z_n) = \\ &= \int_{\Delta^N}\cdots\int \frac{x_1^{b_1-1}\ldots x_n^{b_n-1}(1-x_1-\ldots-x_n)^{b_0-1}}{(1-z_1x_1-\ldots-z_nx_n)^a}\, dx_1\ldots dx_n. \end{aligned} \tag{4.7}$$

Now put $a=1$ and $b_j = \tau_j$ in (4.7), so that $c = \tau_0 + \tau_1 + \ldots + \tau_N = 1$. Since $a = c$, it follows directly from (4.6) that the series simplifies to

$$\int_{\Delta^N}\cdots\int \frac{\mu(dx)}{1 - z_1x_1 - \ldots - z_Nx_N} = \prod_{j=1}^{N} \frac{1}{(1-z_j)^{\tau_j}}.$$

The formula (4.5) follows from this by a substitution $x_0 = 1 - x_1 - \ldots - x_N$ (see Section 5 for more details). $\square$

Comparing (4.5) with the Example 3.6 we arrive at the

(4.8) **Corollary.** *Assume that* μ *is the Dirichlet measure with the parameters* $\tau_0, \ldots, \tau_N$, *such that* $\tau_0 + \tau_1 + \ldots + \tau_N = 1$. *Then* μ *is the Markov-Krein transform of the* free *discrete measure* τ, *i.e., the measure with a weight* τ_j *at the* j*th basis vector,* $j = 0, 1, \ldots, N$.

(4.9) **Remark.** Let μ be the Dirichlet measure corresponding to a discrete distribution τ as in Corollary 4.8. Consider a linear map $f(x) = Ax$ induced by a map $f : B_n \to B_k$ of the standard linear basis B_n in $\mathbb{R}^n$ onto the basis B_k in $\mathbb{R}^k$, and let μ_f, τ_f be the images of the distributions μ, τ under this map. Then the measure μ_f coincides with the Dirichlet distribution corresponding to the parameter measure τ_f. This property is well known and allows one to extend the construction of the Dirichlet measure to the most general probability distributions τ. We return to this general definition in Section 8.

5. Rank 1 projections of Dirichlet measures.

Let μ_f be the image of the Dirichlet measure μ with positive parameters $\tau_0, \ldots, \tau_n$, $\tau_0 + \ldots + \tau_n = 1$, under a linear map $y_i = \sum_{j=0}^{n} a_{ij}x_j$, $i = 1, \ldots, k$. It follows from Corollary 3.14 that

$$R_{\mu_f}(z_1, \ldots, z_k) = \prod_{j=0}^{n} \Big(1 - \sum_{i=1}^{k} a_{ij}z_i\Big)^{-\tau_j}. \tag{5.1}$$

In this section we present an explicit formula for a linear image of a Dirichlet measure in a particular case when the image of the linear transformation is of dimension 1. In the next Section 6 we shall study linear projections with one-dimensional kernel.

Consider a Dirichlet measure (4.1) on a simplex Δ^N, and let μ_f be its image under a linear functional $f(x_0, \dots, x_N) = a_0x_0 + \dots + a_Nx_N$ that takes the value a_k at a vertex e_k of the simplex. We enumerate the vertices in the increasing order of the values of f, $a_0 < a_1 < \dots < a_N$.

(5.2) **Proposition** ([CR]). *The distribution μ_f of the functional f with respect to the Dirichlet measure μ has the density*

$$\frac{\mu_f(da)}{da} = \frac{\sin\pi(\tau_k + \dots + \tau_N)}{\pi} \prod_{j=0}^{N} |a - a_j|^{-\tau_j}; \qquad a \in (a_{k-1}, a_k), \tag{5.3}$$

for every $k = 1, \dots, N$.

First proof. By (5.1), the Cauchy-Stieltjes transform of the measure μ_f is

$$R_{\mu_f}(z) = \prod_{j=0}^{N} \frac{1}{(1 - a_j z)^{\tau_j}}.$$

Using the Perron-Stieltjes inversion formula (2.3) one can easily check that (5.3) provides the density of μ_f in the interval $a \in (a_{k-1}, a_k)$. $\square$

For the sake of clarity, we shall also compute the density (5.3) by a direct integration.

Second proof. Remark that the intersection of a simplex with a hyperplane is always linearly isomorphic to a prism, i.e., a direct product of two simplices. In fact, let $\Delta' = \Delta^{k-1}$ and $\Delta'' = \Delta^{N-k}$ be the faces of the simplex Δ^N, generated by the vertices $e_0, \dots, e_{k-1}$ (where the value $f(e_i) = a_i$ of the form f is smaller than a) and by the vertices $e_k, \dots, e_N$ (where $f(e_i) = a_i \geq a$). Every point $v \in \Delta^N$ can be uniquely represented as a barycenter

$$v = \frac{s}{s+t}v' + \frac{t}{s+t}v''; \qquad s, t \geq 0, \tag{5.4}$$

where $v' \in \Delta'$ and $v'' \in \Delta''$. Since $a_{k-1} < a < a_k$, the function f assumes the value $f(v) = a$ at exactly one point of every interval (5.4). For this point v one can set

$$t = \sum_{i=0}^{k-1}(a - a_i)v_i' = a - f(v'); \qquad s = \sum_{i=k}^{N}(a_i - a)v_i'' = f(v'') - a.$$

Note that $v' = (v_0', \dots, v_{k-1}')$ and $v'' = (v_k'', \dots, v_N'')$ are barycentric coordinates in the simplices Δ' and Δ''.

Consider a substitution

$$v_i = \frac{s}{s+t}v_i'; \quad i = 0, 1, \dots, k-1, \qquad v_i = \frac{t}{s+t}v_i''; \quad i = k, \dots, N,$$

so that we replace the initial coordinates $v_1, \dots, v_N$ with the new ones $v_1', \dots, v_{k-1}'$, a, and $v_{k+1}'', \dots, v_N''$. One can easily check that the Jacobian of this substitution is

$$J = \frac{s^{k-1}\, t^{N-k}}{(s+t)^N}.$$

It follows that the integral over the subset $\{v \in \Delta^N : a_{k-1} < f(v) < a_k\}$ with respect to the Dirichlet measure μ can be written, using new variables, as

$$\frac{\Gamma(\tau'+\tau'')}{\Gamma(\tau')\Gamma(\tau'')} \int_{a_{k-1}}^{a_k} \int\limits_{\Delta'}\int\limits_{\Delta''} \frac{s^{\tau'-1}\, t^{\tau''-1}}{(s+t)^{\tau'+\tau''-1}}\, \mu'(dv')\, \mu''(dv'')\, da, \tag{5.5}$$

where $\tau' = \tau_0 + \ldots + \tau_{k-1}$ and $\tau'' = \tau_k + \ldots + \tau_N$. We write μ' for the Dirichlet measure with parameters $\tau_0, \ldots, \tau_{k-1}$ on the simplex Δ', and μ'' for the Dirichlet measure with parameters $\tau_k, \ldots, \tau_N$ on the simplex Δ''.

Since $\tau' + \tau'' = 1$, we have $\Gamma(\tau')\Gamma(\tau'') = \pi/\sin \pi\tau'$, and the density of the measure μ_f factors as

$$\frac{\mu_f(da)}{da} = \frac{\sin \pi\tau'}{\pi} \int\limits_{\Delta'} t^{\tau''-1}\, \mu'(dv') \int\limits_{\Delta''} s^{\tau'-1}\, \mu''(dv'').$$

Using the Lauricella integral (4.7), we obtain

$$\int\limits_{\Delta'} t^{\tau''-1}\, \mu'(dv') =$$

$$= (a-a_0)^{\tau''-1}\, F_D^{(k-1)}\Big(1-\tau'', \tau_1, \ldots, \tau_{k-1}; \tau'; \frac{a_1-a_0}{a-a_0}, \ldots, \frac{a_{k-1}-a_0}{a-a_0}\Big) =$$

$$= \prod_{i=0}^{k-1} \frac{1}{(a-a_i)^{\tau_i}},$$

since the parameters $1-\tau'' = \tau'$ coincide. In a similar way,

$$\int\limits_{\Delta''} s^{\tau'-1}\, \mu''(dv'') = \prod_{i=k}^{N} \frac{1}{(a_i-a)^{\tau_i}},$$

and the formula (5.3) follows. $\square$

6. The images of Dirichlet measures under linear projections with one-dimensional kernel.

Now we consider linear maps $f : \Delta^{n+1} \to \mathbb{R}^n$ with a kernel of dimension one. We show that the image μ_f of a Dirichlet measure μ on the simplex Δ^{n+1} has a density p which is a piecewise Lauricella function.

Denote by $a_k \in \mathbb{R}^n$ the image $a_k = f(e_k)$ of the kth vertex e_k of the simplex Δ^{n+1}, $k = 0, 1, \ldots, n+1$. Let $a_k = (a_{1k}, \ldots, a_{nk})$ be the coordinates of the vector a_k. It will be convenient to use the notation

$$(b_0, b_1, \ldots, b_n) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ b_{10} & b_{11} & \ldots & b_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n0} & b_{n1} & \ldots & b_{nn} \end{vmatrix}$$

for the oriented volume of a simplex in $\mathbb{R}^n$ with the vertices $b_0, b_1, \ldots, b_n$ (more precicely, the Lebesgue volume multiplied by $n!$). Let S be the Lebesgue volume of the set $\delta = f(\Delta^{n+1})$. Since a generic point $x \in \delta$ belongs to exactly two simplices with the vertices a_k, one can write

$$S = \frac{1}{2n!} \sum_{j=0}^{n+1} |(a_0, \ldots, \widehat{a_j}, \ldots, a_{n+1})|.$$

Let us call a *diagonal* the convex hull of a subset of any $n-1$ points in the set $\{a_0, \dots, a_{n+1}\}$. The diagonals split the image δ into several *cells*. We shall show that on each cell the density $p(x) = S\mu_f(dx)/dx$ of μ_f (with respect to the normalized Lebesgue measure) is a continuous Lauricella function.

There is no loss of generality in the assumption that the point $x \in \delta$ belongs to the cell identified as the intersection of the simplices $\delta' = \{a_0, \dots, a_{n-1}, a_n\}$ and $\delta'' = \{a_0, \dots, a_{n-1}, a_{n+1}\}$. Let

$$x = \sum_{k=0}^{n-1} \alpha_k\, a_k + \alpha_n\, a_n; \qquad x = \sum_{k=0}^{n-1} \beta_k\, a_k + \beta_n\, a_{n+1} \tag{6.1}$$

be the barycentric decompositions of the vector x in the simplices δ', δ'' correspondingly.

(6.2) **Proposition.** *Assuming that $x \in \delta' \bigcap \delta''$, the density $p(x)$ of μ_f with respect to the normalized Lebesgue measure on δ can be written in the form*

$$\begin{aligned} p(x) = {} & \frac{\Gamma(\tau_0 + \ldots + \tau_{n+1})}{\Gamma(\tau_0)\ldots\Gamma(\tau_{n-1})\Gamma(\tau_n + \tau_{n+1})} \alpha_0^{\tau_0 - 1} \ldots \alpha_{n-1}^{\tau_{n-1}-1} \alpha_n^{\tau_n - 1} \beta_n^{\tau_{n+1}} \frac{S}{S'} \times \\ & \times F_D^{(n)}\Big(\tau_{n+1}, \tau_0, \dots, \tau_{n-1}; \tau_n + \tau_{n+1}; 1 - \frac{\beta_0}{\alpha_0}, \dots, 1 - \frac{\beta_{n-1}}{\alpha_{n-1}}\Big), \end{aligned} \tag{6.3}$$

where $S' = (a_0, \dots, a_n)$ is the volume of the simplex δ', and $F_D^{(n)}$ denotes the Lauricella function (4.6).

Proof. A generic barycenter representation of the vector x can be obtained as a convex combination of decompositions (6.1) with coefficients $1-u > 0$, $u > 0$ correspondingly. Note that

$$\alpha_k = \frac{(a_0, \dots, x, \dots, a_{n-1}, a_n)}{(a_0, \dots, a_k, \dots, a_{n-1}, a_n)}; \qquad \beta_k = \frac{(a_0, \dots, x, \dots, a_{n-1}, a_{n+1})}{(a_0, \dots, a_k, \dots, a_{n-1}, a_{n+1})}$$

for $k < n$, and $S'\alpha_n = (a_0, \dots, a_{n-1}, x) = S''\beta_n$, where $S'' = (a_0, \dots, a_{n-1}, a_{n+1})$. Another useful fact is that $S'\partial\alpha_n/\partial x_j = S''\partial\beta_n/\partial x_j$ for all $j = 1, \dots, n$. Using simple linear algebra one can check that the Jacobi determinant of the substitution

$$\begin{aligned} M_1 &= (1-u)\alpha_1 + u\,\beta_1 \\ &\dots \\ M_{n-1} &= (1-u)\alpha_{n-1} + u\,\beta_{n-1} \\ M_n &= (1-u)\alpha_n \\ M_{n+1} &= u\,\beta_n \end{aligned}$$

is actually

$$J = \frac{\beta_n}{S'} = \frac{\alpha_n}{S''} = \frac{(a_0, \dots, a_{n-1}, x)}{(a_0, \dots, a_{n-1}, a_n)(a_0, \dots, a_{n-1}, a_{n+1})},$$

and does not depend on u. Therefore,

$$\begin{aligned} p(x) = {} & \frac{\Gamma(\tau_0 + \ldots + \tau_{n+1})}{\Gamma(\tau_0)\ldots\Gamma(\tau_{n+1})} \alpha_n^{\tau_n - 1} \beta_n^{\tau_{n+1}-1} J S \times \\ & \times \int_0^1 u^{\tau_{n+1}-1}(1-u)^{\tau_n - 1} \prod_{j=0}^{n-1} \Big((1-u)\alpha_j + u\,\beta_j\Big)^{\tau_j - 1} du, \end{aligned}$$

and the formula (6.3) follows directly from the Lauricella integral (4.7). $\square$

(6.4) **Corollary.** *Let ε denote the distance of $x \in \delta$ from a generic point of a diagonal δ_0. Then*

$$p(x) = O(\varepsilon^{\tau_i+\tau_j-1}); \qquad \varepsilon \to 0,$$

where a_i and a_j are the only two vertices of δ not contained in the diagonal δ_0.

(6.5) **Example.** Consider a measure τ in $\mathbb{R}^2$ with equal weights $\varepsilon > 0$ at the points $a_0 = (-1/2, \sqrt{3}/2)$, $a_1 = (-1/2, -\sqrt{3}/2)$, $a_2 = (1, 0)$, and with the weight $1 - 3\varepsilon$ at the point $a_3 = (0, 0)$. If $\varepsilon \leq 1/3$, then all the weights are non-negative. Note that the condition (1.10) is satisfied in a wider range $0 < \varepsilon \leq 1/2$. We shall show that the density (6.3) is only positive when $0 < \varepsilon \leq 1/3$, so that the condition (1.10) does not imply the positivity of the density p.

The barycentric coordinates of a point $a = (x, 0)$, $-1/2 < x < 0$ in the triangle a_0, a_1, a_2 are

$$\alpha_0 = \frac{1-x}{3}; \qquad \alpha_1 = \frac{1-x}{3}; \qquad \alpha_2 = \frac{1+2x}{3};$$

the corresponding coordinates of a in the triangle a_0, a_1, a_3 are

$$\beta_0 = -x; \qquad \beta_1 = -x; \qquad \beta_2 = 1 + 2x.$$

Therefore, $z = 1 - \beta_0/\alpha_0 = 1 - \beta_1/\alpha_1 = (1 + 2x)/(1 - x)$ increases from 0 to 1, as far as x increases from $x = -1/2$ to $x = 0$.

It follows from the formula

$$\sum_{k=0}^{n} \frac{(\varepsilon)_k}{k!} \frac{(\varepsilon)_{n-k}}{(n-k)!} = \frac{(2\varepsilon)_n}{n!}$$

that

$$F_D^{(2)}(1 - 3\varepsilon, \varepsilon, \varepsilon; 1 - 2\varepsilon; z, z) = \sum_{n=0}^{\infty} \frac{(1-3\varepsilon)_n \, (2\varepsilon)_n}{(1-2\varepsilon)_n \, n!} z^n = \ {}_2F_1(1 - 3\varepsilon, 2\varepsilon; 1 - 2\varepsilon; z).$$

The density (6.3), up to a positive factor C, takes the form

$$\begin{aligned} C\,p(x) &= (1 - 2\varepsilon) \ \ {}_2F_1(1 - 3\varepsilon, 2\varepsilon; 1 - 2\varepsilon; z) = \\ &= (1 - 2\varepsilon) + (1 - 3\varepsilon) \sum_{n=1}^{\infty} \frac{(2-3\varepsilon)_{(n-1)}}{(2-2\varepsilon)_{(n-1)}} \frac{(2\varepsilon)_n}{n!} z^n, \end{aligned}$$

where $z = (1 + 2x)/(1 - x)$. Since the right hand side is negative in some interval $0 < z_0(\varepsilon) < z < 1$, for every $1/3 < \varepsilon < 1/2$, it follows that the density p is only positive for $0 < \varepsilon \leq 1/3$.

7. The moment formula.

Let $V = \{e_0, e_1, \ldots, e_N\}$ be the set of vertices of a simplex Δ^N, and τ a probability distribution on the set V with the weights $\tau_0, \tau_1, \ldots, \tau_N$ at the corresponding vertices. We denote by M a random point of the simplex Δ^N, subject to the Dirichlet distribution (4.1) with the parameter measure τ. The angle brackets will always denote the average in M with respect to this Dirichlet distribution μ.

Consider an affine map $f : \Delta^N \to \mathbb{R}^m$, and denote its values on the vertex set V as $f(v) = \big(f_1(v), \ldots, f_m(v)\big)$, $v \in V$. The value of f at a point $M \in \Delta^N$ with barycentric coordinates $(M_0, M_1, \ldots, M_N)$ can then be written in the form $f(M) = (X_1(M), \ldots, X_m(M))$, where

$$X_i(M) = \sum_{j=0}^{N} f_i(e_j)\, M_j; \qquad i = 1, \ldots, m. \tag{7.1}$$

Let τ_f and μ_f be the f-images in $\mathbb{R}^m$ of the measures τ and μ in the space $\mathbb{R}^N$. We use the following notations for the joint moments of the measures τ_f, μ_f:

$$p_{k_1,\dots,k_m} = \sum_{j=0}^{N} f_1^{k_1}(e_j)\dots f_m^{k_m}(e_j)\,\tau_j, \tag{7.2}$$

$$h_{n_1,\dots,n_m} = \left\langle X_1^{n_1}\dots X_m^{n_m}\right\rangle = \int\underset{\Delta^N}{\cdots}\int X_1^{n_1}(M)\dots X_m^{n_m}(M)\,\mu(dM). \tag{7.3}$$

In this section we compute the moments $h_{n_1,\dots,n_m}$ of the distribution μ_f in terms of the moments $p_{k_1,\dots,k_m}$ of the distribution τ_f. In order to make the formula look more combinatorial, we start with some notations.

Let $B = \bigcup_{i=1}^m B_i$ be a disjoint union of m subsets B_i of cardinalities $|B_i| = n_i$, and denote by $n = n_1 + \ldots + n_m$ the total number of elements in B. If $b \in B_i$, we write $i(b) = i$, and say that $i(b)$ is the *shape* of b.

A map $j : B \to \{0, 1, \dots, N\}$ will be referred to as a *coloring* of the set B, and we denote by $\mathcal{M}(B)$ the set of all such colorings.

Let $\mathfrak{S}(B)$ be the set of permutations of the set B. Given a coloring $j \in \mathcal{M}(B)$, we say that a permutation $w \in \mathfrak{S}(B)$ is *color-respecting* if b and $w(b)$ have the same color, for all $b \in B$. Let $\mathfrak{S}(B, j)$ denote the set of permutations of B respecting the coloring j. Note that all elements in a cycle c of a color-respecting permutation $w \in \mathfrak{S}(B, j)$ have the same color which we denote as $j(c)$.

(7.4) **Proposition.** *Define the moments of the distributions τ_f and μ_f by the formulae (7.2), (7.3) and let $n = n_1 + \ldots + n_m$. Then*

$$h_{n_1,\dots,n_m} = \frac{1}{n!}\sum_{w\in\mathfrak{S}_n}\prod_{c\in C(w)} p_{k_1(c),\dots,k_m(c)}, \tag{7.5}$$

where $C(w)$ is the set of cycles of a permutation w, and $k_i(c) = |c \bigcap B_i|$ is the number of elements of shape i in a cycle c.

(7.6) **Example.** If $m = 2$, then the formulas for the first few joint moments are as follows:

$$h_{10} = p_{10}, \qquad h_{01} = p_{01},$$
$$2\,h_{20} = p_{10}^2 + p_{20}, \qquad 2\,h_{11} = p_{10}p_{01} + p_{11}, \qquad 2\,h_{02} = p_{01}^2 + p_{02}.$$

The moments of order 3 have the form

$$\begin{aligned}
6\,h_{30} &= p_{10}^3 + 3p_{10}p_{20} + 2p_{30},\\
6\,h_{21} &= p_{10}^2p_{01} + p_{01}p_{20} + 2p_{10}p_{11} + 2p_{21},\\
6\,h_{12} &= p_{10}p_{01}^2 + p_{10}p_{02} + 2p_{01}p_{11} + 2p_{12},\\
6\,h_{03} &= p_{01}^3 + 3p_{01}p_{02} + 2p_{03}.
\end{aligned}$$

Proof. Using (7.2), one can record the right hand side of the moment formula (7.5) as

$$\frac{1}{n!}\sum_{j\in\mathcal{M}(B)}\sum_{w\in\mathfrak{S}(B,j)}\prod_{c\in C(w)} f_1^{k_1(c)}(e_{j(c)})\dots f_m^{k_m(c)}(e_{j(c)})\,\tau_{j(c)}. \tag{7.7}$$

In order to compute the left hand side of (7.5) note first that

$$X_i^{n_i} = \sum_{j \in \mathcal{M}(B_i)} \prod_{b \in B_i} f_i(e_{j(b)})\, M_{j(b)},$$

hence

$$h_{n_1, \dots, n_m} = \sum_{j \in \mathcal{M}(B)} \prod_{b \in B} f_{i(b)}(e_{j(b)}) \left\langle M_0^{k_0} M_1^{k_1} \dots M_N^{k_N} \right\rangle, \tag{7.8}$$

where $k_j = |\{b \in B : j(b) = j\}|$ is the number of elements of a color j.

Now recall the moment formula (4.3) and note that the Pochhammer symbol can be written in the form

$$(t)_k = \sum_{w \in \mathfrak{S}_k} t^{c(w)},$$

where $c(w)$ is the number of cycles in a permutation w. Therefore,

$$\left\langle M_0^{k_0} M_1^{k_1} \dots M_N^{k_N} \right\rangle = \sum_{(w_0, w_1, \dots, w_N)} \tau_0^{c(w_0)} \tau_1^{c(w_1)} \dots \tau_N^{c(w_N)},$$

where $(w_0, w_1, \dots, w_N)$ runs over a subgroup $\mathfrak{S}_{k_0} \times \mathfrak{S}_{k_1} \times \dots \times \mathfrak{S}_{k_N}$ in $\mathfrak{S}_n$. Hence one can state (7.8) in the form identical with (7.7), and the Proposition follows. □

(7.9) **Corollary.** *In case of a single functional, $m = 1$, (7.5) reduces to the familiar moment formula (2.6).*

In fact, the moment formula (7.5) is equivalent to the basic identity (3.9).

Assume for simplicity that a measure τ in $\mathbb{R}^m$ is finitely supported. Then its Markov–Krein transform μ is also finitely supported, and both distributions μ and τ have finite moments

$$p_{n_1, \dots, n_m} = \int \prod_{i=1}^m x_i^{n_i} \tau(dx), \qquad h_{n_1, \dots, n_m} = \int \prod_{i=1}^m x_i^{n_i} \mu(dx) \tag{7.10}$$

of all orders.

(7.11) **Proposition.** *If a measure τ is finitely supported, then the moment formula (7.5) is equivalent to the basic identity (3.9).*

Proof. Let us expand the functions $R_\mu(z)$ and $\ln Q_\tau(z)$ in a series in the variables $z_1, \dots, z_m$,

$$\begin{aligned} \ln Q_\tau(z_1, \dots, z_m) &= \sum_{n_1 + \dots + n_m \geq 1} \frac{(n_1 + \dots + n_m - 1)!}{n_1! \dots n_m!} p_{n_1, \dots, n_m} z_1^{n_1} \dots z_m^{n_m}, \\ R_\mu(z_1, \dots, z_m) &= 1 + \sum_{n_1 + \dots + n_m \geq 1} \frac{n!}{n_1! \dots n_m!} h_{n_1, \dots, n_m} z_1^{n_1} \dots z_m^{n_m}. \end{aligned} \tag{7.11}$$

We use the following combinatorial lemma which is a generalization of the well-known result for one variable ([M], Example I.2.11).

(7.13) **Lemma.** *Consider the formal power series*

$$A(x_1,...,x_m) = \sum_{n_1+...+n_m\geq 1} \frac{a_{n_1,...,n_m}}{n_1!...n_m!} x_1^{n_1}...x_m^{n_m},$$

$$B(x) = 1 + \sum_{n=1}^{\infty} \frac{b_n}{n!} x^n$$

and the composition $H(x_1,...,x_m) = B(A(x_1,...,x_m))$. Expand H into a power series,

$$H(x_1,...,x_m) = 1 + \sum_{n_1+...+n_m\geq 1} \frac{H_{n_1,...,n_m}}{n_1!...n_m!} x_1^{n_1}...x_m^{n_m}.$$

Then the coefficients $H_{n_1,...,n_m}$ take the form

$$H_{n_1,...,n_m} = \sum_{w\in\mathfrak{S}_n} b_k \prod_{c\in C(w)} \frac{a_{k_1(c),\ldots,k_m(c)}}{(k_1(c)+\ldots+k_m(c)-1)!},$$

where $k = k(w)$ is the number of cycles of a permutation $w \in \mathfrak{S}_n$.

Set $A(z_1,\ldots,z_m) = \ln Q_\tau(z)$, $B(x) = e^x$. Then we have $Q_\tau(z) = B(A(z_1,\ldots,z_m))$. By Lemma (4.11),

$$Q_\tau(z) = 1 + \sum_{n_1+\ldots+n_m\geq 1} \frac{z_1^{n_1}\ldots z_m^{n_m}}{n_1!\ldots n_m!} \sum_{w\in\mathfrak{S}_{n_1+\ldots+n_m}} \prod_{c\in C(w)} p_{k_1(c),\ldots,k_m(c)}. \tag{7.14}$$

By definition of the Markov–Krein correspondence, $R_\mu(z) = Q_\tau(z)$. The comparison of the coefficients for $R_\mu(z)$ in (7.12) with those in (7.14) concludes the proof of the Proposition. □

8. Dirichlet measures with continuous parametric distributions.

Consider an arbitrary finite space $X = \{s_0,\ldots,s_N\}$. We identify the set of all probability distributions on X with the standard simplex Δ^N: a point $M \in \Delta^N$ with barycentric coordinates $(M_0,\ldots,M_N)$ corresponds to the measure with weights $M_0,\ldots,M_N$ at the points $s_0,\ldots,s_N$. We use the same notation M for this distribution. Let τ be a probability measure on X with weights $\tau_0,\ldots,\tau_N$ at the points $s_0,\ldots,s_N$. Consider the Dirichlet distribution on the simplex Δ^N with parameters $\tau_0,\ldots,\tau_N$. Then a random point $M \in \Delta^N$ defines a random probability distribution $\sum_{i=0}^N M_i\delta_{s_i}$ on the space X. This distribution is said to be the *Dirichlet measure* on X with the parameter measure τ. By the Remark 4.9, this random measure is characterized by the following condition. If $X = A_0 \cup \ldots \cup A_k$ is a partition of X, the random vector $(M(A_0),\ldots,M(A_k))$ has the Dirichlet distribution on Δ^k with parameters $\tau(A_0),\ldots,\tau(A_k)$ (in Remark 4.9, consider the map $f : B_n \to B_k$ that sends a basis vector $e_i \in B_n$ to $e'_j \in B_k$ if $v_i \in A_j$). We take this property as a definition of Dirichlet measure for general probability distributions τ.

(8.1) **Definition** ([Fe]). Let X be an arbitrary measurable space, and τ be a probability measure on X. A random probability distribution M on X is called the *Dirichlet random measure* if, for each finite measurable partition $X = A_0 \cup \ldots \cup A_N$, the vector $(M(A_0),\ldots,M(A_N))$ has the Dirichlet distribution (4.1) on the simplex Δ^N with parameters $\tau(A_0),\ldots,\tau(A_N)$.

The Remark (4.9) guarantees the consistency of these conditions, and it was proved by Ferguson [Fe] that such a random measure does exist.

Let M be the random Dirichlet measure on X with a parameter measure τ. Denote by μ the distribution of this random measure. Consider a measurable function $f : X \to \mathbb{R}^m$. We shall call the map f *admissible* for a probability distribution τ, if

$$\int_X \ln(1+\|f(x)\|)\tau(dx) < \infty. \tag{8.2}$$

(8.3) Theorem. *The random mean value $\int f(x)M(dx)$ exists a.e. if and only if the function f is admissible for τ. Provided that this condition is satisfied,*

$$\int \frac{\mu(dM)}{1 - z^T\left(\int f(x)M(dx)\right)} = \exp\int \ln\frac{1}{1 - z^T f(x)}\tau(dx); \qquad z \in i\mathbb{R}^m. \tag{8.4}$$

Proof. The first part of the Theorem follows from [FT].

Let f be a "simple" function that takes a finite number of values $a_0, \dots, a_n \in \mathbb{R}^m$ on some sets $A_0, \dots, A_n$. Then the right-hand side of (8.4) equals

$$\prod_{j=0}^{n}(1 - z^T a_j)^{-\tau(A_j)}. \tag{8.5}$$

The left-hand side of (8.4) takes the form

$$\int \frac{\mu(dM)}{1 - z^T a_1 M(A_0) - \ldots - z^T a_n M(A_n)}. \tag{8.6}$$

By definition of μ, the vector $(M(A_0), \dots, M(A_n))$ has the Dirichlet distribution with parameters $\tau(A_0), \dots, \tau(A_N)$, thus (8.6) is just the additive Cauchy-Stieltjes transform of the Dirichlet distribution with parameters $\tau(A_0), \dots, \tau(A_n)$ which is equal to (8.5) by Proposition 4.4.

Now let f be an arbitrary admissible function. We can approximate it by a sequence of simple functions $f_1, f_2, \dots$, such that $f_n \to f$ and $\|f_n\| \le \|f\|$ a.e. Then

$$\int \frac{\mu(dM)}{1 - z^T\left(\int f_n(x)M(dx)\right)} = \exp\int \ln\frac{1}{1 - z^T f_n(x)}\tau(dx). \tag{8.7}$$

The right-hand side of this equality tends to the right-hand side of (8.4), since

$$\left|\ln\frac{1}{1 - z^T f_n(x)}\right| \le C\ln(1 + \|f(x)\|); \qquad \|f(x)\| \to \infty,$$

where C does not depend on x and n. By the first part of the Theorem, $\|\int f(x)M(dx)\| < \infty$ for almost all measures M with respect to the distribution μ. Then, by the Dominated Convergence Theorem, $\int f_n(x)M(dx) \to \int f(x)M(dx)$ a.e. Since the function $1/(1 - z^T x)$ is bounded, the left-hand side of (8.7) tends to the left-hand side of (8.4), and the Theorem follows.

(8.8) Corollary. *Let $f_1, \dots, f_m$ be admissible functions for τ, and consider the map $f : X \to \mathbb{R}^m$, where $f(x) = (f_1(x), \dots, f_m(x))$. Denote by τ_f the f-image of τ and by μ_f the distribution of the mean value $\int f(x)M(dx)$ in $\mathbb{R}^m$. Then μ_f is the Markov–Krein transform of τ_f.*

In case of $X = \mathbb{R}$ and $f(x) = x$, this follows from the results of [CR] (cf. a combinatorial proof in [DK]).

9. Generalized Dirichlet measures in linear spaces.

Let V be a real topological vector space, and $\mathcal{L}$ be a space of real linear functionals on V. We assume that the elements of $\mathcal{L}$ separate the points of V.

For the sake of simplicity, we shall assume in this section that all measures under consideration have compact support.

Given a probability measure μ in V, we define its *additive Cauchy-Stieltjes transform* as a functional

$$R_\mu(F) = \int \frac{\mu(dv)}{1 - F(v)}; \qquad F \in i\mathcal{L}. \tag{9.1}$$

Note that the Cauchy–Stieltjes transform uniquely determines the corresponding measure. Indeed, for each $F \in \mathcal{L}$ and $z \in i\mathbb{R}$, we have

$$R_\mu(zF) = \int \frac{\mu(dv)}{1 - zF(v)} = \int \frac{\mu_F(dx)}{1 - zx} = R_{\mu_F}(z),$$

where μ_F is the image of μ under the map $F : V \to \mathbb{R}$. Thus we know the Cauchy–Stieltjes transforms of all one-dimensional projections of the measure μ. These projections can be restored by (2.3), and they uniquely determine the original distribution μ, since they determine the Fourier transform $\langle e^{iF} \rangle = \int \exp(iF(v))\mu(dv)$ of μ.

We define the *multiplicative Cauchy-Stieltjes transform* of a probability measure τ in V as a functional

$$Q_\tau(F) = \exp \int \ln \frac{1}{1 - F(v)} \tau(dv); \qquad F \in i\mathcal{L}. \tag{9.2}$$

(9.3) **Definition.** We say that a measure μ in V is the *Markov-Krein transform* of a probability distribution τ, if the additive Cauchy-Stieltjes transform of μ coincides with the multiplicative Cauchy-Stieltjes transform of τ,

$$\int \frac{\mu(dv)}{1 - F(v)} = \exp \int \ln \frac{1}{1 - F(v)} \tau(dv); \qquad F \in i\mathcal{L}. \tag{9.4}$$

The function $R_\mu(F)$ uniquely determines the measure μ. Therefore, for each τ there is at most one distribution μ satisfying the basic identity (9.4).

One can easily generalize the proof of Corollary 3.14 to obtain that the Markov–Krein transform in a linear space is covariant under affine transformations.

(9.5) **Proposition.** *Let μ_g and τ_g be the images under an affine map $g(v) = a + Av$ of some probability measures μ and τ. If μ is the Markov-Krein transform of τ, then μ_g is the Markov-Krein transform of the measure τ_g.*

(9.6) **Example.** If $V = \mathbb{R}^m$ and $\mathcal{L} = V^*$ is the space of real linear forms on $\mathbb{R}^m$, we obtain the definition (3.8) of the Markov–Krein correspondence in $\mathbb{R}^m$.

(9.7) **Example.** Let $V = \mathcal{M}(X)$ be the space of finite Borel measures on a topological space X, and $\mathcal{L}$ be the space of bounded measurable real functions on X, where

$$F(v) = \int_X f(x) v(dx).$$

If we identify each point $x \in X$ with the δ-measure $\delta_x \in V$, then each probability measure τ on X can be considered as a distribution in V. Thus the basic identity (9.4) takes the form

$$\int \frac{\mu(dM)}{1 - z \int_X f(x) M(dx)} = \exp \int \ln \frac{1}{1 - zf(x)} \tau(dx); \qquad z \in i\mathbb{R}.$$

But this identity coincides with (8.7), thus the Markov–Krein transform of τ is just the distribution of the Dirichlet random measure with parameter τ.

We use this example to prove the existence of the Markov–Krein transform of a probability measure in a general linear space.

(9.8) **Proposition.** *For each probability measure τ with compact support in V there exists its Markov – Krein transform μ.*

Proof. Consider the Dirichlet measure M in V with the parameter measure τ. By Theorem 8.3, its distribution $\tilde\mu$ satisfies the identity

$$\int_{\mathcal{M}(V)} \frac{\tilde\mu(dM)}{1-F(M)} = \exp\int_V \ln\frac{1}{1-F(v)}\tau(dv); \qquad F\in i\mathcal{L}. \tag{9.9}$$

We consider the projection from $\mathcal{M}(V)$ onto V that sends a measure M to its mean value $v_M = \int vM(dv)$. Denote by μ the image of the distribution $\tilde\mu$ under this projection. Then for each linear functional F, the distribution of $F(M) = \int F(v)M(dv) = F(v_M)$ with respect to $\tilde\mu$ coincides with the distribution of $F(v)$ with respect to μ, thus

$$\int_{\mathcal{M}(V)} \frac{\tilde\mu(dM)}{1-F(M)} = \int_V \frac{\mu(dv)}{1-F(v)},$$

and the Proposition follows from (9.9).

10. Random measures and exchangeable sequences.

According to de Finetti celebrated result, there is a natural correspondence between random measures M on a space $\mathcal{X}$, and exchangeable sequences $(Z_1, Z_2, \dots, Z_n, \dots)$ of $\mathcal{X}$-valued random variables. In case of a Dirichlet random measure there is a simple special construction for the corresponding exchangeable sequence, the *Chinese restaurant process* (see [DK]). Consider a random sequence of permutations $w_1, w_2, \dots, w_n \in S_n$ constructed according to the following rule. Start with the identity permutation w_1. If we have already constructed $w_1, \dots w_n$, then $w_{n+1}\in S_{n+1}$ is obtained from w_n by inserting $n+1$ in a cycle of w_n to the immediate right of some element $i = 1, \dots n$, or by adding a new cycle consisting of the only element $n+1$. Each of these options has equal probabilities $1/(n+1)$. One can easily see that the distribution of w_n in this construction is the Haar measure on S_n, i.e. each permutation $w\in S_n$ has probability $1/n!$. Note that the sequence $\{w_n\}$ defines a partition of the set $\mathbb{N}$ into "infinite cycles" $C_1, C_2, \dots$, namely, $n\in C_i$ if n lies in the ith cycle of π_n. Associate with each cycle a random label taken independently from the distribution τ. Let X_n be the label of the cycle containing $n\in\mathbb{N}$. Then the sequence X is exchangeable and corresponds to the Dirichlet measure with parameter distribution τ.

As it was shown in [K], [P2], the constructions leading to random Dirichlet measures admit natural generalizations. We now describe a generalized Chinese restaurant model ([K]).

The parameters of this model consist of a probability measure τ and a family $\{\nu_n\}$ of coherent central measures on the symmetric groups S_n. This means that the weight $\nu_n(w)$ of a permutation $w\in S_n$ depends only on the cycle structure of w, and for each $w\in S_n$

$$\nu_n(w) = \sum_{\sigma:\sigma'=w} \nu_{n+1}(\sigma),$$

where σ' denotes the permutation obtained from σ by removing the element $n+1$ from its cycle. We choose a random sequence of permutations $\{w_n\}$ and the associated random sequence $\{X_n\}$ as follows. Start with the identity permutation $w_1\in S_1$. If we have already constructed $w_1, \dots w_n$, then choose $w_{n+1} = \sigma \in S_{n+1}$, such that $\sigma' = w_n$ with probability $\nu_{n+1}(\sigma)/\nu_n(w_n)$. As before, we associate with each (possibly infinite) cycle in $\mathbb{N}$ an independent random label with distribution τ and we set X_n to be the label of the cycle containing n. The conditions imposed on the family $\{\nu_n\}$ guarantee the correctness of this procedure and the exchangeability of the sequence $\{X_n\}$. Hence this sequence defines some random measure which we call a generalized Dirichlet measure and denote by $\mathcal{D}(\{\nu_n\},\tau)$.

One can easily generalize Proposition 7.4 to obtain the following result (cf. [K], 4.2.2.).

(10.1) **Proposition.** *Let τ be a finitely supported probability measure in $\mathbb{R}^m$, and let M be a generalized random Dirichlet measure $\mathcal{D}(\{\nu_n\},\tau)$. Denote by μ the distribution of the mean value $\int xM(dx)$. Then the moments (7.10) of measures μ and τ are related by the identity*

$$h_{n_1,\dots,n_m} = \sum_{w\in\mathfrak{S}_n} \nu_n(w) \prod_{c\in C(w)} p_{k_1(c),\dots,k_m(c)}, \tag{10.2}$$

where $C(w)$ is the set of cycles of a permutation w, and $k_i(c)$ is the number of elements of shape i in a cycle c (see Section 7 for the definition of 'shape').

(10.3) **Example.** Consider the *Ewens distributions* on symmetric groups,

$$M_n^\theta(w) = \frac{\theta^{k(w)}}{(\theta)_n}; \qquad \theta>0,\ w\in\mathfrak{S}_n,\ n=1,2,\dots, \tag{10.4}$$

and their two-parameter generalizations, the *Pitman distributions* ([P1], [PY])

$$M_n^{\alpha,\theta}(w) = \frac{(\theta+\alpha)(\theta+2\alpha)\dots(\theta+(k-1)\alpha)}{(\theta+1)_{n-1}} \prod_{j\ge1}\left(\frac{(1-\alpha)_{j-1}}{(j-1)!}\right)^{r_j}, \tag{10.5}$$

where $k=k(w)$ is the total number of cycles of w, r_j is the number of cycles of length j, and $(x)_m = x(x+1)\dots(x+m-1)$ is the Pochhammer symbol. The range of admissible parameters in (10.5) is

$$\{(\alpha,\theta): 0\le\alpha<1, \theta>-\alpha\} \quad \bigcup \quad \{(\alpha,-m\alpha),\ \alpha<0, m\in\mathbb{N}\}.$$

Let τ be a probability distribution in $\mathbb{R}^m$. Consider the generalized Dirichlet measure M with parameters $\{M_n^{\alpha,\theta}\}$, τ and denote by μ the distribution of the random mean value $\int xM(dx)$ in $\mathbb{R}^m$.

As a direct generalization of the main result in [Ts], one can show that the measures μ and τ satisfy one of the following identities generalizing the Markov–Krein correspondence (3.9):

$$\int(1-z^Tx)^{-\theta}\mu(dx) = \exp\int \ln(1-z^Tx)^{-\theta}\tau(dx), \qquad \text{if } \alpha=0 \tag{10.6}$$

in case of Ewens distributions, and

$$\begin{aligned} \left(\int(1-z^Tx)^{-\theta}\mu(dx)\right)^{-\frac{1}{\theta}} &= \left(\int(1-z^Tx)^{\alpha}\tau(dx)\right)^{\frac{1}{\alpha}}, \quad &\text{if } \alpha,\theta\ne0;\\ \exp\int\ln(1-z^Tx)^{\alpha}\mu(dx) &= \int(1-z^Tx)^{\alpha}\tau(dx), \quad &\text{if } \theta=0 \end{aligned} \tag{10.7}$$

if $\alpha\ne0$.

References

[CR] D. M. Cifarelli, E. Regazzini, *Some remarks on the distribution functions of means of a Dirichlet process*, Ann. Statist. **18** (1990), 429–442.

[DK] P Diaconis, J. Kemperman, *Some New Tools for Dirichlet Priors*, Bayesian Statistics 5 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.), Oxford University Press, 1996, pp. 97–106.

[Ex] H. Exton, *Multiple Hypergeometric Functions and Applications*, John Wiley & Sons Inc., NY, 1976.

[FT] P. D. Feigin, R. L. Tweedie, *Linear functionals and Markov chains associated with Dirichlet process*, Math. Proc. Camb. Phil. Soc. **105** (1989), 579–585.

[Fe] T. Ferguson, *A Bayesian analysis of some nonparametric problems*, Ann. Statist. **1** (1973), 209–230.

[K] S. Kerov, *Interlacing measures*, Amer. Math. Soc. Transl. (2) **181** (1998), 35–83.

[Ki] J. F. C. Kingman, *Poisson Processes*, Clarendon Press, Oxford, 1993.

[KN] M. G. Krein, A. A. Nudelman, *The Markov Moment Problem and Extremal Problems*, Translations of Mathematical Monographs, vol. 50, AMS, Providence, Rhode Island (1977).

[M] I. Macdonald, *Symmetric Functions and Hall Polynomials*, Clarendon Press, Oxford, 1979.

[Ma] A. A. Markov, *Nouvelles applications des fractions continues*, Math. Ann. **47** (1896), 579–597.

[P1] J. Pitman, *The Two–parameter Generalization of Ewens' Random Partition Structure*, Tech. Report **345** (1992), 1–23.

[P2] J. Pitman, *Some Developments of the Blackwell–MacQueen Urn Scheme*, Statistics, Probability and Game theory (Ferguson T. S., Shapley L. S., MacQueen J. B. eds.), IMS Lecture Notes – Monograph series, vol. 30, 1996, pp. 245–267.

[PY] J. Pitman, M. Yor, *The two – parameter Poisson – Dirichlet distribution derived from a stable subordinator*, Ann. Prob. **25** (1997), 855–900.

[Ts] N. V. Tsilevich, *Distribution of mean values of some random measures*, Zapiski Nauchnyh Seminarov POMI **240** (1997), 268–279.