4. Poisson-Dirichlet random measures

4.1 Distributions of mean values of random measures. Assuming that τ is a probability distribution, there exists a simple probabilistic construction providing its Markov transform μ (see [8, 10]). Let us start with an elementary case of a finitely supported measure τ.

(4.1.1) Example. Assume that a discrete probability distribution τ has positive weights $\tau_0, \tau_1, \dots, \tau_n$ (where $\sum \tau_k = 1$) at the points $x_0 < x_1 < \dots < x_n$ of the real line. Denote by Δ_n the simplex of all vectors $p = (p_0, p_1, \dots, p_n)$ with non-negative components, such that $\sum p_k = 1$. The formula

$$\frac{p_0^{\tau_0 - 1} p_1^{\tau_1 - 1} \dots p_n^{\tau_n - 1}}{\Gamma(\tau_0)\Gamma(\tau_1)\dots\Gamma(\tau_n)} dp_1 dp_2 \dots dp_n \tag{4.1.2}$$

determines the so called *Dirichlet distribution* with the parameters $\tau_0, \tau_1, \dots, \tau_n$ on the simplex Δ_n. Denote by M the discrete probability measure with a weight p_k at the point x_k, $k = 0, 1, \dots, n$, where $(p_0, p_1, \dots, p_n)$ is a random vector in Δ_n distributed according to (4.1.2). Then the distribution μ of the mean value $X = \sum_{k=0}^n x_k p_k$ of the random measure M is the Markov transform of τ. The measure μ is absolutely continuous, and it follows from the general formula (2.4.6) that its density equals

$$\frac{\mu(dx)}{dx} = \frac{1}{\pi} \sin\Big(\pi \sum_{j=0}^{k} \tau_j\Big) \prod_{k=0}^{n} \frac{1}{|x - x_k|^{\tau_k}} \qquad \text{if } x_k < x < x_{k+1}. \tag{4.1.3}$$

A more general construction, known as *Poisson-Dirichlet process* (see [27]), provides a random probability distribution M associated to every finite positive measure τ. The above example generalizes as follows.

(4.1.4) Theorem (see [10]). *Assume that τ is a probability distribution, and that the integral (1.3.11) converges. Let M denote a random probability measure obtained as a sample realization of the Dirichlet process with the parameter measure τ. Then the random integral*

$$X = \int_{-\infty}^{\infty} x M(dx) \tag{4.1.5}$$

converges absolutely for almost all M, and its distribution μ coincides with the Markov transform of τ:

$$\int_{-\infty}^{\infty} \frac{\mu(du)}{z - u} = \exp \int_{-\infty}^{\infty} \ln \frac{1}{z - u} \tau(du), \qquad \Im z \neq 0.$$

If the parameter measure of the Poisson-Dirichlet process is positive and has a finite total mass θ, we can write it in the form $\theta\tau(du)$, where τ is a probability measure. In this case the distribution μ of the mean value (4.1.5) of the random measure M is related to the parameters θ, $\tau(du)$ by the formula

$$\int_{-\infty}^{\infty} \frac{\mu(du)}{(z - u)^\theta} = \exp\Big(- \theta \int_{-\infty}^{\infty} \ln(z - u)\, \tau(du)\Big), \qquad \Im z \neq 0 \tag{4.1.6}$$

generalizing (3.7.9).

4.2 Exchangeable sequences. A useful tool in dealing with random measures is the notion of exchangeability. By definition, a sequence $X_1, X_2, \ldots, X_n, \ldots$ of real random variables is called *exchangeable* if its joint distributions are invariant upon every permutation of a finite number of its elements. Obviously, a sequence of independent random variables with a common distribution M is exchangeable. More generally, one can use a two-stage sampling: first choose a random measure M, then a sequence of independent variables with the distribution M. By the classical de Finetti theorem, every exchangeable sequence arises as a result of this procedure from appropriate random measure M which is determined by the sequence uniquely.

Every statement concerning a random measure M can be restated, in principle, in terms of the corresponding exchangeable sequence $X_1, X_2, \ldots$. For instance, the nth moment of the mean value $X = \int xM(dx)$ of a random measure M equals the average of the product of n distinct terms in the exchangeable sequence:

$$(4.2.1) \qquad < \Big(\int_{-\infty}^{\infty} xM(dx) \Big)^n > \quad = \quad < X_1 X_2 \ldots X_n > .$$

The exchangeable sequence $X_1, X_2, \ldots$ corresponding to the Poisson-Dirichlet random measure M with the parameter distribution τ can be obtained by a simple procedure known as *Chinese restaurant* construction (see [10]). One should imagine a series of large circular tables indexed by random numbers $x_1, x_2, \ldots$ taken independently from the distribution τ. The first guest takes a place at the first table, and we set $X_1 = x_1$. The nth guest takes a place either at the first empty table or to the immediate right of one of the $n-1$ persons already seated, with equal chances $1/n$. We assign to the nth variable X_n the label x_k of the table where the nth person took a seat, $X_n = x_k$.

There is an obvious generalization of the Chinese restaurant construction for arbitrary partition structures in place of the uniform distributions in the above description. Given a partition structure $\{M_n\}_{n=1}^{\infty}$, we associate an exchangeable sequence to every probability distribution τ in $\mathbf{R}$ as follows.

Let X_1 be a random number chosen from the distribution τ, and denote by $\pi_1 \in \mathfrak{S}_1$ the identity permutation. By induction argument, assume that we have already chosen the first n elements $X_1, X_2, \ldots, X_n$ of the sequence and a permutation $\pi_n \in \mathfrak{S}_n$. Independently of all previous choices, choose a permutation $\pi_{n+1} \in \mathfrak{S}_{n+1}$, such that $\pi'_{n+1} = \pi_n$ (i.e., π_n arises from π_{n+1} by deleting the element $n+1$ from its cycle), with probability $M_{n+1}(\pi_{n+1})/M_n(\pi_n)$. If the element $n+1$ is a fixed point of π_{n+1}, choose X_{n+1} independently from the distribution τ. Otherwise, if the cycle of $n+1$ contains an element $k \le n$, set $X_{n+1} = X_k$ (by the construction, the variables X_k indexed by the elements of the same cycle are equal).

The generalized Chinese restaurant construction provides a random variable with the moments h_n related to the moments p_n of the parameter measure τ by the formula (3.7.3).

(4.2.2) Theorem. *Let $X_1, \ldots, X_n, \ldots$ be the exchangeable sequence obtained via Chinese restaurant construction from a probability distribution τ and a partition structure $\{M_n\}_{n=1}^{\infty}$. Denote by μ the distribution of the mean value $X = \int xM(dx)$, where the random measure M is associated with the sequence by de Finetti theorem. Then the moments*

(3.1.1), (3.1.2) of the measures τ, μ are related by the formula

$$h_n = \sum_{\pi \in \mathfrak{S}_n} M_n(\pi) \prod_{k \geq 1} p_k^{r_k(\pi)}, \qquad n = 1, 2, \dots .$$

The proof coincides almost literally with that of Theorem 4.1.4 in Section 2 of [10].

5. Combinatorics of continued fractions

Assume that a positive measure μ has finite moments of all orders, and that μ is determined by its moments uniquely. Then the Cauchy-Stieltjes transform of the measure μ, $R(z) = \int (z-u)^{-1} \mu(du)$, can be represented by a continued fraction

$$R(z) = \cfrac{1}{z - b_0 - \cfrac{\lambda_1}{z - b_1 - \cfrac{\lambda_2}{z - b_2 - \dots}}}$$

which converges for all non-real $z \in \mathbf{C}$. The elements $b_0, b_1, \dots, \lambda_1, \lambda_2, \dots$ of the fraction are real and $\lambda_n > 0$ for all $n = 1, 2, \dots$ (such fractions are called *real J-type continued fractions*).

The moments $h_n = \int x^n \mu(dx)$ of the measure μ can be written as polynomials in the elements of the continued fraction. Flajolet [14] had suggested a simple combinatorial description of these polynomials (see [43] for a more detailed account of combinatorial theory of orthogonal polynomials). We provide a similar description for the moments p_n of the Rayleigh measure τ associated to μ by the master identity (2.4.2). By the equivalent form (2.4.3) of this identity, the moments p_n can be interpreted as the coefficients of the asymptotic series at infinity,

$$-\frac{R'(z)}{R(z)} = \sum_{n=0}^{\infty} \frac{p_n}{z^{n+1}},$$

of the logarithmic derivative of the above continued fraction.

5.1 Motzkin paths and the power series of a continued fraction. Consider a formal power series decomposition of a real J-type continued fraction,

$$\cfrac{1}{z - b_0 - \cfrac{\lambda_1}{z - b_1 - \cfrac{\lambda_2}{z - b_2 - \dots}}} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{h_n}{z^{n+1}}. \tag{5.1.1}$$

The coefficients h_n can be written as polynomials in the elements $b_0, b_1, b_2, \dots, \lambda_1, \lambda_2, \dots$ of the fraction in the left hand side. The first few examples are simple enough to be written explicitly:

$$h_1 = b_0,$$
$$h_2 = b_0^2 + \lambda_1, \tag{5.1.2}$$
$$h_3 = b_0^3 + 2b_0\lambda_1 + b_1\lambda_1,$$
$$h_4 = b_0^4 + 3b_0^2\lambda_1 + 2b_0b_1\lambda_1 + b_1^2\lambda_1 + \lambda_1^2 + \lambda_1\lambda_2.$$

For higher moments the equations are more cumbersome.

Since the coefficients of the polynomials (5.1.2) are positive integers, one can try to interpret these polynomials as generating functions of appropriate combinatorial objects. In order to identify these objects, consider the particular example of the continued fraction

$$\frac{1}{z-1-\cfrac{1}{z-1-\cfrac{1}{z-1-\ \dots}}} = \frac{1}{z} + \frac{1}{z^2} + \frac{2}{z^3} + \frac{4}{z^4} + \frac{9}{z^5} + \dots, \tag{5.1.3}$$

with all elements equal to one. Then the coefficients $h_n = M_n$ are the so called Motzkin numbers.

Recall that a *Motzkin path* can be defined as a sequence $s = (i_1, i_2, \dots, i_n)$ of non-negative integers, such that $i_n = 0$ and $|i_k - i_{k-1}| \leq 1$ for all $k = 1, 2, \dots, n$ (we assume that $i_0 = 0$). We denote the set of Motzkin paths of length n by $\mathcal{M}_n$. It is quite common to visualize Motzkin paths as the lattice paths in the upper half-plane, with the endpoints at zero level, see Fig. 9.

Define the *Motzkin number* $M_n = |\mathcal{M}_n|$ as the number of Motzkin paths. It is easy to see that these numbers satisfy the recurrence relations

$$M_{n+1} = M_n + (M_0M_{n-1} + M_1M_{n-2} + \dots + M_{n-1}M_0) \tag{5.1.4}$$

(for the properties and various interpretations of Motzkin numbers see [11]). It follows that the generating function

$$R(z) = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{M_n}{z^{n+1}} \tag{5.1.5}$$

satisfies a quadratic equation

$$R^2(z) - (z-1)R(z) + 1 = 0, \tag{5.1.6}$$

and one readily obtains the explicit formula

$$R(z) = \frac{z - 1 - \sqrt{z^2 - 2z - 3}}{2}. \tag{5.1.7}$$

Since the continued fraction in the left hand side of (5.1.3) also satisfies the equation (5.1.6), the identity (5.1.3) follows.

One can see from (5.1.3) that the coefficients of the polynomial representing h_n sum up to M_n, hence the monomials are likely to be associated with Motzkin paths. To this end, one labels the entries i_k of a Motzkin path $s \in \mathcal{M}_n$ by the appropriate elements m_k of the continued fraction (5.1.1), to obtain a monomial $m_1 m_2 \dots m_n$ associated with the path s.

The labeling rule is remarkably simple. Given a path $s = (i_1, i_2, \dots, i_n) \in \mathcal{M}_n$, define the factors m_k as

$$m_k = \begin{cases} \lambda_{i_k+1}, & \text{if } i_k = i_{k-1} - 1, \\ b_{i_k}, & \text{if } i_k = i_{k-1}, \\ 1, & \text{if } i_k = i_{k-1} + 1, \end{cases}$$

see Fig. 9, and set $m_s = m_1 m_2 \dots m_n$ for the monomial in the elements $b_0, b_1, \dots, \lambda_1, \lambda_2, \dots$ of the fraction.

(5.1.8) Theorem (Flajolet, [14]). *The nth coefficient h_n in the power series decomposition of the continued fraction (5.1.1) equals the sum of all monomials associated with Motzkin paths of length n,*

$$h_n = \sum_{s \in \mathcal{M}_n} m_s. \tag{5.1.9}$$

5.2 Cyclic Motzkin paths and the logarithm of a continued fraction. Let us write the formal power series of the logarithmic derivative of a J-type continued fraction as

$$-\frac{d}{dz} \ln \cfrac{1}{z - b_0 - \cfrac{\lambda_1}{z - b_1 - \cfrac{\lambda_2}{z - b_2 - \dots}}} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{p_n}{z^{n+1}}. \tag{5.2.1}$$

We aim at the combinatorial description of the coefficients p_n as polynomials in the elements of the fraction in the left hand side, similar to that of Flajolet theorem.

The first few examples of such polynomials are as follows:

$$\begin{aligned} p_1 &= b_0, \\ p_2 &= b_0^2 + 2\lambda_1, \\ p_3 &= b_0^3 + 3b_0\lambda_1 + 3b_1\lambda_1, \\ p_4 &= b_0^4 + 4b_0^2\lambda_1 + 4b_0 b_1 \lambda_1 + 4b_1^2\lambda_1 + 2\lambda_1^2 + 4\lambda_1\lambda_2. \end{aligned} \tag{5.2.2}$$

Note that if one derives these equations from (3.3.8) and (5.1.2), it is rather surprising that all the coefficients are positive.

The sums of coefficients can be obtained from the particular example of the fraction (5.1.3). To this end, set

$$R(z) \equiv \frac{1}{z} + \sum_{n=1}^{\infty} \frac{M_n}{z^{n+1}}$$

and denote by E_n the coefficients of the power series of the logarithmic derivative,

$$-\frac{R'(z)}{R(z)} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{E_n}{z^{n+1}}.$$

One readily derives from the equation (5.1.6) that $-R(z)/R'(z) = z - 1 - 2R(z)$, and it follows from (5.1.7) that the generating function for the numbers E_n equals

$$(5.2.3) \qquad -\frac{R'(z)}{R(z)} = \frac{1}{\sqrt{z^2-2z-3}} = \frac{1}{z} + \frac{1}{z^2} + \frac{3}{z^3} + \frac{7}{z^4} + \frac{19}{z^5} + \dots$$

The first coefficients $E_1 = 1$, $E_2 = 3$, $E_3 = 7$, $E_4 = 19$ conform to the formulae (5.2.2). These numbers arise in combinatorics in the so called "king walks problem" studied by L. Euler himself (see [35]). In Euler's problem, the number E_n represents the constant term in the decomposition of $(z + 1 + z^{-1})^n$.

Instead of the chess board king walks of Euler, we prefer to speak of shifted Motzkin paths.

(5.2.4) Definition. Call a sequence $s = (i_1, \dots, i_n)$ *shifted Motzkin path*, if it is conjugate to a true Motzkin path, i.e. can be obtained as a cyclic shift of an ordinary Motzkin path (see Fig. 9). If s is a cyclic shift of some $t \in \mathcal{M}_n$, we set $m_s = m_t$. The set of all shifted Motzkin n-paths is denoted by $\mathcal{S}_n$.

In accordance with the examples (5.2.2), the monomials for 19 cyclic Motzkin paths of length 4 are distributed as follows: there are four paths for each of the monomials $b_0^2\lambda_1$, $b_0b_1\lambda_1$, $b_1^2\lambda_1$, $\lambda_1\lambda_2$, two paths for the monomial λ_1^2, and one for the monomial b_0^4. The latter three are $(1,0,1,0)$, $(0,1,0,1)$ and $(0,0,0,0)$.

By the solution (5.2.3) of Euler's king walks problem, $E_n = |\mathcal{S}_n|$ is the number of shifted Motzkin paths.

(5.2.5) Theorem. *The nth coefficient p_n in the formal power series expansion of the logarithm of the continued fraction (6.2.1) equals the sum of all monomials associated with shifted Motzkin paths of length n,*

$$(5.2.6) \qquad p_n = \sum_{s \in \mathcal{S}_n} m_s.$$

Proof. The Theorem will follow as a direct corollary of a more general result in the next Section. □

(5.2.7) Example. Assume that $b_0 = b_1 = \dots = 0$ and $\lambda_1 = \lambda_2 = \dots = 1$. Then $h_{2n} = \binom{2n}{n}/(n+1)$ is the Catalan number. The number of shifted Dyck paths of length $2n$ (with the increments ± 1 only) equals $p_{2n} = \binom{2n}{n}$. Note that the measures μ and τ in this example are the semi-circle law and the arcsine law correspondingly,

$$\mu(du) = \frac{1}{2\pi}\sqrt{4-u^2}\,du, \qquad \tau(du) = \frac{1}{\pi}\frac{du}{\sqrt{4-u^2}}.$$

The identity (3.2.4) in this example is equivalent to the formula (9.6) in Chapter XII.9 of Feller's book [13], applied to the simple symmetric random walk over integers.

5.3 An abstract lemma. Consider an alphabet A and a set P of finite words in A^* called *prime*. Let W denote the set of all words representable as finite concatenations of prime words, and assume that such a representation is unique, for all words in W.

(5.3.1) Example. Let $A = \mathbf{N}$ be the set of non-negative integers, and let P be the set of *prime Motzkin paths* (a path $s = (i_1, i_2, \ldots, i_n)$ is said to be prime if $i_n = 0$ is its only zero entry). Every Motzkin path splits canonically into prime Motzkin subpaths.

Let $R = \mathbf{Q}[P]$ be the ring of polynomials in the variables m_p, $p \in P$. With every word $w \in W$ we associate a monomial $m_w = m_{p_1} m_{p_2} \ldots m_{p_k} \in R$, where $w = p_1 p_2 \ldots p_k$ is the unique representation of w as a concatenation of prime subwords. Write

$$h_n = \sum_{w \in W_n} m_w \in R \tag{5.3.2}$$

for the sum of monomials corresponding to the words of length n in W.

Denote by S the set of words conjugate to some words in W (i.e. the words representable as cyclic shifts of words in W). We call the elements of S *shifted* words. If a shifted word $s \in S$ is a cyclic shift of a word $w \in W$, we set $m_s = m_w$ for the monomial associated with s. Denote by

$$p_n = \sum_{s \in S_n} m_s \in R \tag{5.3.3}$$

the sum of monomials associated with all shifted words of length n.

(5.3.4) Lemma. *The polynomials (5.3.2), (5.3.3) satisfy the identities*

$$n\, h_n = \sum_{k=1}^{n} h_{n-k}\, p_k; \qquad n = 1, 2, \ldots . \tag{5.3.5}$$

Proof. Let $[n] \times W_n$ be the set of pairs (j, w) where $w \in W_n$ and the integer j *marks* the jth character of the word w. Let $\widetilde{W}_n = \bigcup_{k=1}^{n} W_{n-k} \times S_k$ be the set of pairs (v, s) where $v \in W_{n-k}$ and $s \in S_k$ for some $k = 1, 2, \ldots, n$. In order to prove (5.3.5), we establish a bijection $I : [n] \times W_n \to \widetilde{W}_n$, such that $I(j, w) = (v, s)$ implies $m_w = m_v m_s$.

Given $(j, w) \in [n] \times W_n$, denote by p the prime subword of w, containing the jth (marked) character. Let v be the concatenation of all prime subwords of w strictly to the left of p, and denote by $\widetilde{s}$ the concatenation of p and of other prime factors of w to the right of p. Let $s \in S$ be the cyclic shift of $\widetilde{s}$, such that the marked character comes first. By definition, $I(j, w) = (v, s)$.

The inverse bijection is obvious. Given a pair $(v, s) \in \widetilde{W}_n$, denote by $\widetilde{s} \in W$ the only cyclic shift of s, such that the first character a of s belongs to the first prime factor of $\widetilde{s}$. Set the marker j to the position of a in $\widetilde{s}$, and denote by $w = v\widetilde{s}$ the concatenation of v and $\widetilde{s}$. Then $I(j, w) = (v, s)$, and the Lemma follows. $\square$

(5.3.6) Example. Consider a Motzkin path $w = (1\ 2\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ \mathbf{0}\ 1\ 2\ 1\ 0) \in \mathcal{M}_{15}$ with the marked entry $i_{11} = 0$. The prime subpaths of w are $(1\ 2\ 1\ 1\ 0)$, (0), $(1\ 0)$, $(1\ 1\ 0)$ and $(1\ 2\ 1\ 0)$. Hence, $v = (1\ 2\ 1\ 1\ 0\ 0\ 1\ 0)$, $\widetilde{s} = (1\ 1\ \mathbf{0}\ 1\ 2\ 1\ 0)$ and $s = (\mathbf{0}\ 1\ 2\ 1\ 0\ 1\ 1)$.

5.4 A bijective proof of identity (3.3.2). The formula

$$n!\,h_n = \sum_{\pi\in\mathfrak{S}_n}\prod_{k\geq 1} p_k^{r_k(\pi)}, \tag{5.4.1}$$

is equivalent to the identity (5.3.5), hence is true under the assumptions of Lemma 5.3.4. Nevertheless, we present a direct bijective proof of (5.4.1). We use the notations and terminology of Section 5.3. In particular, A is an alphabet, P a set of *prime* words, W the set of concatenations of words in P, and S the set of words in A^* conjugate to some words in W (shifted words).

Consider a permutation $\pi\in\mathfrak{S}_n$ in the cycle notation,

$$\pi = (j_1^{(1)}, \ldots, j_{k_1}^{(1)})\,(j_1^{(2)}, \ldots, j_{k_2}^{(2)})\ldots(j_1^{(l)}, \ldots, j_{k_l}^{(l)}).$$

We denote by l the number of cycles in π, and by $j_2^{(t)} = \pi(j_1^{(t)}),\ \ldots,\ j_1^{(t)} = \pi(j_{k_t}^{(t)})$ the elements of the tth cycle. Let $\widetilde{\mathfrak{S}}_n$ be the set of pairs (π, s) where $\pi\in\mathfrak{S}_n$ and

$$s = (i_1^{(1)}, \ldots, i_{k_1}^{(1)})\,(i_1^{(2)}, \ldots, i_{k_2}^{(2)})\ldots(i_1^{(l)}, \ldots, i_{k_l}^{(l)})$$

is a family of shifted words $s_t = (i_1^{(t)}, \ldots, i_{k_t}^{(t)})$; $t = 1, \ldots, l$, matching the cycle decomposition of π. Denote by $m(\pi, s) = m_{s_1}m_{s_2}\ldots m_{s_l}\in R$ the product of monomials m_{s_t} associated to shifted words s_t in Section 5.3. It is obvious that

$$\sum_{\pi\in\mathfrak{S}_n}\prod_{k\geq 1} p_k^{r_k(\pi)} = \sum_{(\pi,s)\in\widetilde{\mathfrak{S}}_n} m(\pi, s),$$

and even more obvious that

$$n!\,h_n = \sum_{(\sigma,w)\in\mathfrak{S}_n\times W_n} m_w.$$

In order to prove (5.4.1) it suffices to establish a bijection $G:\mathfrak{S}_n\times W_n\to\widetilde{\mathfrak{S}}_n$, such that $G(\sigma, w) = (\pi, s)$ implies $m(\pi, s) = m_w$.

We proceed with the description of the bijection G. Given a pair $(\sigma, w)\in\mathfrak{S}_n\times W_n$, write σ as a sequence of its images,

$$\sigma = (\sigma(1), \sigma(2), \ldots, \sigma(n)). \tag{5.4.2}$$

Split the word $w = p_1p_2\ldots p_d$ into its prime subwords, and let $\sigma = \sigma_1\sigma_2\ldots\sigma_d$ be the factorization of the sequence (5.4.2) into subsequences $\sigma_1, \sigma_2, \ldots, \sigma_d$ of lengths matching the lengths of prime factors $p_1, p_2, \ldots, p_d$ of the word w.

Arrange the sequences $\sigma_1, \sigma_2, \ldots, \sigma_d$ in the increasing order of their first elements, and let $\rho\in\mathfrak{S}_d$ be the permutation which restores the original order of the pieces $\sigma_1, \ldots, \sigma_d$. Write the permutation ρ of the sequences $\sigma_1, \ldots, \sigma_d$ in the cycle notation to obtain a permutation $\pi \in \mathfrak{S}_n$. In a similar way, replace each σ_k in the cycle notation of ρ by the corresponding prime word p_k to obtain a family s of shifted words matching the cycles of π. Set $G(\sigma, w) = (\pi, s)$.

(5.4.3) Example. We apply here the map G to the pair

$$\begin{pmatrix} \sigma \\ w \end{pmatrix} = \begin{pmatrix} 7 & 16 & 3 & 4 & 13 & 15 & 10 & 8 & 2 & 12 & 9 & 1 & 5 & 11 & 14 & 6 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 2 & 1 & 0 & 1 & 2 & 1 & 1 & 0 \end{pmatrix}$$

in $\mathfrak{S}_{16} \times \mathcal{M}_{16}$. The path w factors as a product of prime Motzkin paths,

$$w = (1 \quad 1 \quad 0)\,(0)\,(1 \quad 0)\,(1 \quad 1 \quad 2 \quad 1 \quad 0)\,(1 \quad 2 \quad 1 \quad 1 \quad 0),$$

and the matching factorization of σ results in a sequence of words

$$(7 \quad 16 \quad 3)\,(4)\,(13 \quad 15)\,(10 \quad 8 \quad 2 \quad 12 \quad 9)\,(1 \quad 5 \quad 11 \quad 14 \quad 6).$$

We arrange these words in the increasing order of their first elements,

$$a(1 \quad 5 \quad 11 \quad 14 \quad 6)\; b(4)\; c(7 \quad 16 \quad 3)\; d(10 \quad 8 \quad 2 \quad 12 \quad 9)\; e(13 \quad 15).$$

To restore the original order, one should apply the permutation

$$\rho = \begin{pmatrix} a & b & c & d & e \\ c & b & e & d & a \end{pmatrix} = (a \quad c \quad e)\,(b)\,(d).$$

Replacing each letter a, b, c, d, e by the corresponding word, we derive the permutation

$$\pi = (1 \quad 5 \quad 11 \quad 14 \quad 6 \quad 7 \quad 16 \quad 3 \quad 13 \quad 15)\,(4)\,(2 \quad 12 \quad 9 \quad 10 \quad 8),$$

and the corresponding family of shifted Motzkin paths:

$$(1 \quad 2 \quad 1 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0)\,(0)\,(2 \quad 1 \quad 0 \quad 1 \quad 1).$$

Note that the word $d = (2\ 12\ 9\ 10\ 8)$ and the corresponding prime Motzkin path $(2\ 1\ 0\ 1\ 1)$ were shifted cyclically according to the convention that each cycle of a permutation should start with its least element.

To describe the inverse bijection G^{-1}, start with a pair $(\pi, s) \in \widetilde{\mathfrak{S}}_n$. Read shifted words $s_1, s_2, \ldots, s_l$ corresponding to the cycles of π and split each of these words into prime subwords in P. Split the cycles of π into subsequences $\sigma_1, \sigma_2, \ldots, \sigma_d$ matching the factorization of the words $s_1, \ldots, s_l$. Let $\rho \in \mathfrak{S}_d$ be the permutation induced on the set $\{\sigma_1, \sigma_2, \ldots, \sigma_d\}$ by the cycles of π. We denote by $\sigma^1, \ldots, \sigma^d$ the sequence of pieces $\sigma_1, \ldots, \sigma_d$ arranged in the increasing order of their first elements, and we consider the concatenation of sequences $\rho(\sigma^1), \rho(\sigma^2), \ldots, \rho(\sigma^d)$. This is a sequence of integers $1, 2, \ldots, n$, all entries being distinct, so that it determines a permutation $\sigma \in \mathfrak{S}_n$. Denote by w the concatenation of the prime words $p^{(1)}, p^{(2)}, \ldots, p^{(d)}$ matching the pieces $\rho(\sigma^1), \rho(\sigma^2), \ldots, \rho(\sigma^d)$. It is easy to see that $G(\sigma, w) = (\pi, s)$.

(5.4.4) Example. We apply here the map G^{-1} to the pair $(\pi, s) \in \widetilde{\mathfrak{S}}_{16}$ where

$$\begin{pmatrix} \pi \\ s \end{pmatrix} = \begin{matrix} (1 & 14 & 12 & 6 & 10 & 13 & 9) & (2 & 3 & 15 & 16 & 8 & 5 & 11 & 4) & (7) \\ (1 & 0 & 1 & 2 & 1 & 0 & 1) & (0 & 1 & 1 & 0 & 1 & 1 & 0 & 1) & (0) \end{matrix}.$$

The system s of shifted Motzkin paths splits into prime Motzkin subpaths (1 2 1 0), (1 1 0), (1 1 0), (1 1 0), (1 0), (0). The corresponding splitting of π results in subsequences (12 6 10 13), (9 1 14), (3 15 16), (8 5 11), (4 2), (7). We arrange these pieces in the order of their first elements, $a = (3\ 15\ 16)$, $b = (4\ 2)$, $c = (7)$, $d = (8\ 5\ 11)$, $e = (9\ 1\ 14)$, $f = (12\ 6\ 10\ 13)$. The cycles of π induce the permutation

$$\rho = (a\, d\, b)\ (c)\ (e\, f) = \begin{pmatrix} a & b & c & d & e & f \\ d & a & c & b & f & e \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} \sigma \\ w \end{pmatrix} = \left(\begin{array}{ccc|ccc|c|cc|cccc|ccc} 8 & 5 & 11 & 3 & 15 & 16 & 7 & 4 & 2 & 12 & 6 & 10 & 13 & 9 & 1 & 14 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 2 & 1 & 0 & 1 & 1 & 0 \end{array}\right).$$

6. Interlacing measures in operator theory

We want now to interpret a measure μ and a function F related by the identity (2.3.2) in terms of a self-adjoint operator A in a Hilbert space H and a cyclic unit vector $\xi \in H$. The contents of this section is closely related to the theory of the *spectral shift function*, see [14], [15].

6.1 The statement of the theorem. Let A denote a bounded self-adjoint operator in a separable Hilbert space H. Given a unit vector $\xi \in H$, $||\xi|| = 1$, one can define the *spectral measure of A at the vector ξ* by the identity

$$\big((zI - A)^{-1}\xi, \xi\big) = \int_{-\infty}^{\infty} \frac{\mu(du)}{z - u}; \qquad z \in \mathbf{C} \setminus \mathbf{R},$$

or, equivalently, by the equation

$$\big(Q(A)\xi, \xi\big) = \int_{-\infty}^{\infty} Q(u)\mu(du) \tag{6.1.1}$$

valid for every polynomial $Q \in \mathbf{C}[x]$. Since we assume (for simplicity) that the operator A is bounded, the measure μ is compactly supported and has the moments $h_n = \int u^n \mu(du)$ of all orders. Using (6.1.1), one can also write the moments directly in terms of A and ξ,

$$h_n = \big(A^n\xi, \xi\big); \qquad n = 0, 1, 2, \ldots . \tag{6.1.2}$$

Denote by H_ξ the hyperplane in H orthogonal to the vector ξ, and by P_ξ the orthogonal projection $P_\xi \eta = \eta - (\eta, \xi)\,\xi$, $\eta \in H$ on the hyperplain H_ξ. Given a bounded operator B in H, we write $B_\xi = P_\xi B P_\xi$ for the restriction of B to the hyperplane H_ξ.

For each polynomial $Q \in \mathbf{C}[x]$ the operator $Q(A) - Q(A_\xi)$ is of trace class, and we consider the linear functional $Q \mapsto \operatorname{tr}\big(Q(A) - Q(A_\xi)\big)$.

(6.1.3) Theorem. *There exists a Rayleigh function F, such that*

$$\operatorname{tr}\big(Q(A) - Q(A_\xi)\big) = \mathcal{M}_F(Q) \tag{6.1.4}$$

is the moment functional (3.4.3) of F. The function F and the spectral measure μ of the operator A with respect to a vector ξ determine each other via the master identity

$$\int_{-\infty}^{\infty} \frac{\mu(dt)}{z-t} = \frac{1}{z} \exp\Big(- \int_{-\infty}^{0} \frac{F(u)du}{z-u} + \int_{0}^{\infty} \frac{1-F(u)}{z-u} du \Big). \tag{6.1.5}$$

The idea of the *trace formula* (6.1.4) belongs to I. M. Lifshits [30] and was extensively developed by M. G. Krein (see [29]). The function F in this context is called the *spectral shift function*, see [7] for a recent survey of the theory.

A direct combinatorial proof of the Theorem based on the equation (3.2.4) for the moments of F and μ will be given in Section 6.3. Let us first consider the finite dimensional case.

6.2 The case of finite dimension. Let $H = \mathbf{C}^n$ have a finite dimension n. With no loss of generality we assume that the vector ξ is *cyclic* for the operator A, i.e., the vectors $\xi, A\xi, A^2\xi, \dots$ generate all of H. As a result, the eigenvalues of A are all distinct,

$$x_1 < x_2 < \ldots < x_n.$$

Choose a basis $\{\eta_k\}_{k=1}^n$ in H in which the matrix of A is diagonal, $A\,\eta_k = x_k\,\eta_k$; $k = 1, 2, \dots, n$. Since the vector $\xi = \sum_{k=1}^n a_k\,\eta_k$ is cyclic, all of its coordinates a_k are non-zero.

(6.2.1) Lemma. *The spectral measure μ of the operator A at the vector ξ is discrete; it has positive weights $\mu_k = |a_k|^2$ at the points x_k; $k = 1, 2, \dots, n$. The eigenvalues $y_1, \dots, y_{n-1}$ of the restriction A_ξ of A to the hyperplane H_ξ orthogonal to ξ are all distinct and interlace with the eigenvalues of A,*

$$x_1 < y_1 < x_2 < \ldots < x_{n-1} < y_{n-1} < x_n. \tag{6.2.2}$$

The eigenvalues of A_ξ and A are related to the spectral measure μ by the identity

$$\prod_{k=1}^{n-1} (z - y_k) \prod_{k=1}^{n} (z - x_k)^{-1} = \sum_{k=1}^{n} \frac{\mu_k}{z - x_k}. \tag{6.2.3}$$

Proof. It is immediate that

$$(Q(A)\xi, \xi) = \sum_{k=1}^{n} Q(x_k)|a_k|^2 = \sum_{k=1}^{n} Q(x_k)\mu_k$$

for any polynomial $Q \in \mathbf{C}[x]$, hence the spectral measure is as indicated in Lemma. *Define* the numbers $y_1, \dots, y_{n-1}$ by the formula (6.2.3). Since the coefficients of the partial fractions are positive, $\mu_k > 0$, the inequalities (6.2.2) hold true. It remains to show that the eigenvalues of A_ξ coincide with $y_1, \ldots, y_{n-1}$.

Define the vectors $\zeta_j \in H$; $j = 1, \ldots, n-1$ as

$$\zeta_j = \sum_{k=1}^{n} \frac{a_k \eta_k}{y_j - x_k}, \tag{6.2.4}$$

and note that $\zeta_j \in H_\xi$, since

$$(\zeta_j, \xi) = \sum_{k=1}^{n} \frac{\mu_k}{y_j - x_k} = 0.$$

It is also clear that

$$A\,\zeta_j = \sum_{k=1}^{n} \frac{a_k x_k \eta_k}{y_j - x_k} = y_j\,\zeta_j - \xi,$$

hence $A_\xi\,\zeta_j = y_j\,\zeta_j$, and we are done. $\square$

The equation (6.1.4) follows from the Lemma immediately, since $\mathrm{tr}Q(A) = \sum_{k=1}^{n} Q(x_k)$, $\mathrm{tr}\,Q(A_\xi) = \sum_{k=1}^{n-1} Q(y_k)$, and

$$\mathcal{M}(Q) = \sum_{k=1}^{n} Q(x_k) - \sum_{k=1}^{n-1} Q(y_k) = \mathrm{tr}\,\big(Q(A) - Q(A_\xi)\big), \tag{6.2.5}$$

see Example 3.4.5.

6.3 The proof of the Theorem 6.1.3. It suffices to prove that $p_n = \mathrm{tr}\,(A^n - A_\xi^n)$ for all $n = 1, 2, \ldots$, where the moments p_n of the Rayleigh function F are defined by the equation (3.4.1).

Choose an orthonormal basis $\{\xi_1, \xi_2, \ldots\}$ in H_ξ and set $\xi_0 \equiv \xi$. Then $\{\xi_0, \xi_1, \xi_2, \ldots\}$ is a basis in H, and we write $a_{ij} = (A\xi_i, \xi_j)$ for the matrix elements of the operator A in this basis, $i, j = 0, 1, 2, \ldots$. The matrix of A_ξ may be derived from that of A by replacing the entries of the first row and of the first column of A by zeroes.

Note that the moment h_n of the spectral measure μ is a polynomial of degree n in the matrix elements,

$$h_n = (A^n\,\xi, \xi) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_{n-1}=0}^{\infty} a_{0k_1}\; a_{k_1 k_2} \ldots a_{k_{n-1}0}.$$

This is true for the moments p_n, too. The trace $\mathrm{tr}\,A = \sum_{k=0}^{\infty} a_{kk}$ of the operator A may very well diverge, but the matrix

$$A - A_\xi = \begin{pmatrix} a_{00} & a_{01} & a_{02} & \ldots \\ a_{10} & 0 & 0 & \ldots \\ a_{20} & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

has a finite trace $p_1 = a_{00}$. The matrix elements of the squares are

$$(A^2)_{ij} = \sum_{k=0}^{\infty} a_{ik}\, a_{kj}; \qquad (A^2_\xi)_{ij} = \sum_{k=1}^{\infty} a_{ik}\, a_{kj},$$

so that

$$A^2 - A^2_\xi = \begin{pmatrix} \sum a_{0k}\, a_{k0} & \sum a_{0k}\, a_{k1} & \sum a_{0k}\, a_{k2} & \cdots \\ \sum a_{1k}\, a_{k0} & a_{10}\, a_{01} & a_{10}\, a_{02} & \cdots \\ \sum a_{2k}\, a_{k0} & a_{20}\, a_{01} & a_{20}\, a_{02} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

and the trace of the matrix $A^2 - A^2_\xi$ equals

$$p_2 = a_{00}^2 + 2\sum_{k=1}^{\infty} a_{0k}\, a_{k0}.$$

It is finite, since $h_2 = a_{00}^2 + \sum_{k=1}^{\infty} a_{0k}\, a_{k0}$ is a matrix element of A^2. One derives that $h_2 = (p_1^2 + p_2)/2$, which is the second one of the desired equations (3.3.5).

In order to settle the general case we introduce more notations. A finite sequence $w = (i_1, i_2, \ldots, i_n)$ of non-negative integers terminating with $i_n = 0$ will be referred to as a *chain*, or *n-chain*. Let $\mathcal{H}_n$ denote the set of n-chains. With every such chain w we associate a monomial

$$a_w = a_{i_n i_1}\, a_{i_1 i_2} \ldots a_{i_{n-1} i_n} \tag{6.3.1}$$

in the matrix elements a_{ij} of the operator A. A chain w is *prime* if $i_n = 0$ is its only zero entry. Every chain factors canonically as a concatenation of prime subchains.

Any sequence $s = (i_1, i_2, \ldots, i_n)$ of non-negative integers will be called a *shifted chain* if it has at least one zero entry, not necessarily the last one. Let $\mathcal{P}_n$ be the set of all shifted n-chains. Define the monomial m_s of a shifted chain by the equation (6.3.1). The monomial m_s does not change if one shifts the chain cyclically.

(6.3.2) Lemma. *The polynomials h_n, p_n coincide with the generating functions of the sets $\mathcal{H}_n$, $\mathcal{P}_n$,*

$$h_n \equiv (A^n\xi, \xi) = \sum_{s\in\mathcal{H}_n} a_s; \qquad p_n \equiv tr\,(A^n - A^n_\xi) = \sum_{s\in\mathcal{P}_n} a_s. \tag{6.3.3}$$

They satisfy the equations (3.3.2) and (3.3.3).

Proof. The ith diagonal matrix element of the matrix A^n equals

$$(A^n)_{ii} = \sum_{i_1=0}^{\infty} \cdots \sum_{i_{n-1}=0}^{\infty} a_{ii_1}\, a_{i_1 i_2} \ldots a_{i_{n-1} i}.$$

Since the operator A^n is bounded, the series converges absolutely. In particular, the moment $h_n = (A^n)_{00}$ is finite and coincides with the first generating function in (6.3.3). A similar equation holds for the matrix elements of the operator A_ξ^n,

$$(A_\xi^n)_{ii} = \sum_{i_1=1}^{\infty} \cdots \sum_{i_{n-1}=1}^{\infty} a_{ii_1}\, a_{i_1 i_2} \ldots a_{i_{n-1} i},$$

but the indices now run through positive integers only. Hence, p_n is the generating function of shifted chains in $\mathcal{P}_n$. The series converge absolutely, since

$$\sum_{s\in\mathcal{P}_n} |a_s| \quad \leq \quad n \sum_{s\in\mathcal{H}_n} |a_s| \quad < \quad \infty. \tag{6.3.4}$$

The last claim of the Lemma follows at once from the abstract Lemma 5.3.4. In fact, take for the alphabet $A = \mathbf{N}^2$ the set of all pairs (i,j) of non-negative integers. A chain $(i_1, i_2, \ldots, i_n)$ determines a word $w = \big((0,i_1),(i_1,i_2), \ldots, (i_{n-1},0)\big)$ in A^* and every such word factors uniquely as a product of *prime* words (corresponding to prime chains). The valuation $v : (i,j) \mapsto a_{ij}$ determines a ring homomorphism $v : R \to \mathbf{C}$, and the equations (3.3.2), (3.3.3) follow from (5.3.5) and (5.4.1) correspondingly. $\square$

(6.3.5) Remark. The Theorem 6.1.3 is actually equivalent to our Theorem 5.2.5. In fact, one can always choose the basis $\{\xi_n\}_{n=0}^{\infty}$ in such a way that the matrix of A is a three-diagonal Jacobi matrix. Then the set $\mathcal{M}_n$ of Motzkin paths is a subset in $\mathcal{H}_n$, and $a_w = 0$ for $w \in \mathcal{H}_n \setminus \mathcal{M}_n$. In a similar fashion, the set $\mathcal{S}_n$ of shifted Motzkin paths is a subset of $\mathcal{P}_n$. The generating functions (6.3.3) coincide with those of equations (5.1.9), (5.2.6), and the equation (6.1.4) is equivalent to (5.2.6).

7. Interlacing measures as generalized Young diagrams

7.1 The Plancherel growth of Young diagrams. Recall that to every Young diagram λ there corresponds a positive integer $d(\lambda)$ which admits several equivalent definitions[1]:

$d(\lambda)$ is the number of standard Young tableaux of shape λ;

$d(\lambda)$ is the dimension of the irreducible representation of the symmetric group $\mathfrak{S}_n$, associated with λ;

$d(\lambda)$ is given by the *hook formula* $d(\lambda) = n!/\prod h_{ij}$;

$d(\lambda)$ is the only solution of the recurrence relation $d(\nu) = \sum_{\lambda:\lambda\nearrow\nu} d(\lambda)$ satisfying the initial condition $d(\emptyset) = 1$ for the empty diagram $\lambda = \emptyset$. (We write $\lambda \nearrow \nu$ to show that a Young diagram ν differs from μ by exactly one extra box).

The numbers $d(\lambda)$ also satisfy the identity

$$\sum_{\nu:\lambda\nearrow\nu} d(\nu) = (N+1)\, d(\lambda),$$

[1] See [32] for the terminology and basic facts related to Young diagrams and tableaux.

where $N = |\lambda|$ denotes the number of boxes in λ. By virtue of this formula we associate the transition distribution

$$(7.1.1) \qquad p(\lambda, \nu) = \frac{d(\nu)}{(N+1)\, d(\lambda)}; \qquad \lambda \nearrow \nu$$

with every Young diagram λ. The corresponding Markov chain on the set $\mathbf{Y}$ of all Young diagrams (starting with the empty diagram) is called the *Plancherel growth process*. One can easily check that the probability to obtain a Young diagram λ in the process of Plancherel growth equals $d^2(\lambda)/|\lambda|!$, the probability of λ with respect to the Plancherel measure of the symmetric group $\mathfrak{S}_n$.

Recall that a Young diagram λ can be represented by a pair (1.1.1) of interlacing sequences, see Example 1.1.4. If there is a legal transition from λ to a Young diagram ν, $\lambda \nearrow \nu$, then the content $c(b)$ of the new box $b = \nu \setminus \lambda$ coincides with a point in the first sequence $x_1, x_2, \dots, x_n$. Conversely, to every point x_k there corresponds a unique Young diagram ν_k, such that $\lambda \nearrow \nu_k$ and $c(\nu_k \setminus \lambda) = x_k$. Using this remark, we consider the Plancherel transition distribution (7.1.1) as a discrete probability measure μ on the real line. The following observation was first made in [22].

(7.1.2) Theorem. *Let $x_1 < y_1 < x_2 < \dots < y_{n-1} < x_n$ be the interlacing sequences of contents of consecutive corner points at the border of a Young diagram λ (see Fig. 1). Denote by $\mu_k = d(\nu_k) \,/\, |\nu_k| d(\lambda)$ the Plancherel transition probability of attaching to λ a new box with the content $c(\nu_k \setminus \lambda) = x_k$; $k = 1, 2, \dots, n$. Then*

$$(7.1.3) \qquad \sum_{k=1}^{n} \frac{\mu_k}{z - x_k} = \frac{(z-y_1)\dots(z-y_{n-1})}{(z-x_1)(z-x_2)\dots(z-x_n)}$$

for all $z \in \mathbf{C} \setminus \{x_1, \dots, x_n\}$.

Since (7.1.3) is a particular case of the master identity (2.4.2), one can restate the Theorem as follows.

(7.1.4) Corollary. *The Plancherel transition distribution of a Young diagram λ coincides with the Markov transform of the interlacing sequences (1.1.1) associated with λ.*

Note that the set of Rayleigh functions corresponding to Young diagrams (see Fig. 2) is dense in the space $\mathcal{F}$ of all Rayleigh functions. The Markov transform of a Rayleigh function generalizes the correspondence between a Young diagram and its Plancherel transition distribution. More precisely, the Markov transform is simply the unique continuous extension of the above correspondence.

7.2 Representing Rayleigh functions by diagrams. Given a pair of interlacing sequences

$$(7.2.1) \qquad x_1 < y_1 < x_2 < \dots < x_{n-1} < y_{n-1} < x_n,$$

one can visualize it using a shape similar to that of a Young diagram.

(7.2.2) Definition. A continuous piecewise linear function $\omega : \mathbf{R} \to \mathbf{R}$ is called a *rectangular diagram*, if $\omega'(u) = \pm 1$, and there exists a constant c, such that $\omega(u) = |u - c|$ for sufficiently large $|u|$, see Fig. 10.

With a rectangular diagram we associate a pair (7.2.1) of its interlacing maxima and minima points. Conversely, every pair (7.2.1) determines a rectangular diagram, such that $c = \sum x_k - \sum y_k$ and $\omega'(u) = 1$ for $u \in (x_k, y_k)$; $\omega'(u) = -1$ for $u \in (y_k, x_{k+1})$; $k = 1, \ldots, n-1$. If two polynomials $Q(u) = \prod(x - y_k)$, $P(u) = \prod(x - x_k)$ have interlacing roots (7.2.1), then $\omega'(u) = \operatorname{sign}\frac{Q(u)}{P(u)}$. For instance, Fig. 10 represents the rectangular diagram of interlacing roots of Chebyshev polynomials of the second kind of degrees 15 and 16.

This correspondence between interlacing sequences and rectangular diagrams can be extended to general Rayleigh functions.

Assume that a function $\omega : \mathbf{R} \to \mathbf{R}$ satisfies the Lipshitz condition

$$|\omega(u_2) - \omega(u_1)| \leq |u_2 - u_1| \qquad \forall u_1, u_2 \in \mathbf{R}. \tag{7.2.3}$$

Then its derivative $\omega'(u)$ exists for a.a. $u \in \mathbf{R}$, and $|\omega'(u)| \leq 1$.

(7.2.4) Definition. We say that a function $\omega : \mathbf{R} \to \mathbf{R}$ *represents a diagram*, if it satisfies the condition (7.2.3) and the following integrals converge:

$$\int_{-\infty}^{-1} \big(1 + \omega'(u)\big)\frac{du}{|u|} < \infty, \qquad \int_{1}^{\infty} \big(1 - \omega'(u)\big)\frac{du}{u} < \infty. \tag{7.2.5}$$

Two such functions are *equivalent*, if they differ by a constant. A *diagram* is a class of equivalent functions satisfying the conditions (7.2.3), (7.2.5), see Fig. 11.

We endow the space $\mathcal{D}$ of all diagrams with the topology of uniform convergence on compact subsets in $\mathbf{R}$. More precisely, a family ω_n of diagrams *converges* to a diagram $\omega \in \mathcal{D}$, if one can choose the representing functions to be uniformly converging on finite intervals.

Obviously, the equations

$$F(x) = \frac{1}{2}\big(1 + \omega'(x)\big); \quad 1 - F(x) = \frac{1}{2}\big(1 - \omega'(x)\big); \quad \omega(x) = \int_0^x \big(2F(u) - 1\big)du \tag{7.2.6}$$

establish a homeomorphism between the space of diagrams $\mathcal{D}$ and the space of Rayleigh functions $\mathcal{F}$.

(7.2.7) Example. If $F(u) = \frac{1}{2} + \frac{1}{\pi}\arctan u$ is the Cauchy distribution function, then the corresponding diagram is represented by the function $\omega(u) = \frac{2}{\pi}\big(u \arctan u - \ln\sqrt{1 + u^2}\big)$, so that $\omega'(u) = \frac{2}{\pi}\arctan u$.

The existence of asymptotes for a diagram $\omega \in \mathcal{D}$,

$$\lim_{x \to -\infty} \big(\omega(x) + x\big) = c; \qquad \lim_{x \to +\infty} \big(\omega(x) - x\big) = d, \tag{7.2.8}$$

is equivalent to the existence of the first moment p_1 of the corresponding Rayleigh function F. In this case one can replace the last formula in equation (7.2.6) by

$$\omega(x) = \int_{-\infty}^{x} F(u)\,du + \int_{x}^{\infty} \big(1 - F(u)\big)du,$$

see Fig. 12. Then $-d = c = p_1$ in the equation (7.2.8), the asymptotes cross at zero level, and the region $\{(x,y) : |x-c| \le y \le \omega(x)\}$ outlined by the asymptotes and the graph of $y = \omega(x)$ generalizes the shape of a Young diagram. The area of this shape (i.e., "the number of boxes") equals one half of the variance $p_2 - p_1^2$ of the corresponding Rayleigh function F, and also equals the variance $h_2 - h_1^2$ of the measure μ associated with F by the equation (2.3.2).

(7.2.9) Example. Take for F the distribution function of the arcsine law $\tau(du) = \frac{1}{\pi}\frac{du}{\sqrt{4-u^2}}$; $|u| \le 2$. Then the function

$$\Omega(x) = \begin{cases} \frac{2}{\pi}\big(u \arcsin\frac{u}{2} + \sqrt{4-u^2}\big), & |u| \le 2 \\ |u|, & |u| \ge 2 \end{cases} \tag{7.2.10}$$

represents the corresponding diagram (note that $\Omega'(u) = \frac{2}{\pi}\arcsin\frac{u}{2}$ for $|u| \le 2$). The graph of the function $y = \Omega(x)$ is shown in Fig. 13, along with that of the rectangular diagram of Fig. 10. One can observe the convergence of diagrams ω_n corresponding to the interlacing zeroes of Chebyshev polynomials to the diagram (7.2.10). See [21, 23] for more details. The limiting shape (7.2.10) was first found in the papers [31, 41, 42] in connection with the asymptotics of large Young diagrams.

8. Exponential representations related to Gaussian measure

8.1 The exponential representation of the Gaussian distribution. Let τ be an absolutely continuous measure with the density

$$\frac{d\tau(u)}{du} = \frac{1}{\sqrt{2\pi}}\frac{1}{|D_{-1}(iu)|^2}; \qquad -\infty < u < \infty, \tag{8.1.1}$$

where

$$D_{-1}(z) = e^{z^2/4}\int_{z}^{\infty} e^{-x^2/2}dx = \sqrt{\frac{\pi}{2}}e^{z^2/4}\,\mathrm{Erfc}\Big(\frac{z}{\sqrt{2}}\Big)$$

is a parabolic cylinder function. It is shown in [5], that τ is a probability distribution, and that its Stieltjes transform $T(z)$ has a simple continued fraction decomposition,

$$T(z) = \int_{-\infty}^{\infty}\frac{d\tau(u)}{z-u} = \cfrac{1}{z - \cfrac{2}{z - \cfrac{3}{z - \dots}}}. \tag{8.1.2}$$

(8.1.3) Theorem. *The standard gaussian measure with the density*

$$\frac{d\mu(u)}{du} = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}; \qquad -\infty < u < \infty$$

admits the exponential representation

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-u^2/2}du}{z-u} = \exp\Big(- \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\ln(z-u)du}{|D_{-1}(iu)|^2} \Big). \tag{8.1.4}$$

Equivalently, μ is the Markov transform of the measure τ.

Proof. Recall that the Stieltjes transform of the gaussian measure may be written as

$$R(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-u^2/2}}{z-u} = \cfrac{1}{z - \cfrac{1}{z - \cfrac{2}{z - \cfrac{3}{z - \dots}}}} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{(2n-1)!!}{z^{2n-1}}. \tag{8.1.5}$$

Considering the master identity in its form (2.4.3), we have to show that $R'(z) + R(z)T(z) = 0$, i.e., that

$$-\frac{d}{dz} \ln \cfrac{1}{z - \cfrac{1}{z - \cfrac{2}{z - \dots}}} = \cfrac{1}{z - \cfrac{2}{z - \cfrac{3}{z - \dots}}}.$$

Taking the derivative of the formal power series in (8.1.5) one readily obtains that

$$-R'(z) = \frac{1}{z^2} + \frac{1\cdot 3}{z^4} + \frac{1 \cdot 3 \cdot 5}{z^6} + \ldots = z\,R(z) - 1. \tag{8.1.6}$$

On the other hand, it follows from the continued fraction representation of $R(z)$ that

$$(z - T(z))R(z) = 1. \tag{8.1.7}$$

The identity $R'(z) + R(z)T(z) = 0$ is an immediate consequence of the equations (8.1.6) and (8.1.7). $\square$

One can feel rather awkward about taking the derivative of a diverging series in (8.1.6), and desire for a more combinatorial proof. Then the argument can be as follows. Our aim is to prove that the moments $\{p_{2n}\}$, $\{h_{2n}\}$ of the distributions τ, μ are related by the formula (3.3.3) which can also be written as

$$(2n+1)h_{2n} = \sum_{k=0}^{n} h_{2k}\, p_{2n-2k}; \qquad n = 1, 2, \ldots . \tag{8.1.8}$$

The identity (8.1.7) implies that $h_0 = 1$ and

$$h_{2n+2} = \sum_{k=0}^{n} h_{2k}\, p_{2n-2k}; \qquad n = 1, 2, \dots . \tag{8.1.9}$$

Since $h_{2n+2} = (2n+1)!! = (2n+1)\, h_{2n}$, the equation (8.1.8) follows.

(8.1.10) Remark. The first few moments of the distribution (8.1.1) are

$$p_0 = 1, \quad p_2 = 2, \quad p_4 = 10, \quad p_6 = 74, \quad p_8 = 706, \quad p_{10} = 8162.$$

J. Touchard [38] suggests the following combinatorial interpretation of these numbers.

Consider a system s of non-intersecting semi-circular arcs in the upper half plane, with the endpoints in the set $I_n = \{1, 2, \dots, 2n\}$, see Fig. 14. If the set I_n can be split in two non-empty subintervals, $I_n = I_k \cup \{2k+1, \dots, 2n\}$, in such a way that no arc in the system s has the endpoints in different intervals, then Touchard calls the system *decomposable*. If no such decomposition exists, the system s is *indecomposable*, and we denote by $\tilde{p}_{2n}$ the number of indecomposable systems on the set I_{2n+2}.

Every system of arcs splits uniquely as a union of an indecomposable subsystem on a subinterval $I_{2k+2} \subset I_{2n+2}$, and the remaining system of arcs on the complementary interval. As a result, there is a recurrence relation

$$(2n+1)!! = \sum_{k=0}^{n} (2k-1)!!\, \tilde{p}_{2n-2k}. \tag{8.1.11}$$

Comparing (8.1.11) and (8.1.9) one derives that $\tilde{p}_{2n} = p_{2n}$, so that the number p_{2n} enumerates the indecomposable systems of $(n+1)$ arcs.

(8.1.12) Remark. Let $Q_n(x)$, $n = 0, 1, 2 \dots$ be the sequence of unital polynomials orthogonal with respect to the measure τ. Then $Q_0(x) = 1$, $Q_1(x) = x$, and other polynomials can be obtained from the recurrence relation

$$Q_{n+1}(x) = xQ_n(x) - (n+1)Q_{n-1}(x); \qquad n = 1, 2, \dots . \tag{8.1.13}$$

The coefficients of the recurrence relation coincide with the elements of the continued fraction (8.1.2). There is the following combinatorial description of these polynomials in the spirit of [43].

Let D and R be two finite sets with the total number of $|D| + |R| = n$ elements. Call an *arc* any pair of distinct points in $D \cup R$. By definition, an *n-chip* is a system s of arcs in $D \cup R$ with disjoint endpoints, such that

(1) every point in D belongs to some arc;
(2) every arc has at least one point in R.

Denote by C_n the set of all n-chips. For instance, all three 2-chips are shown in Fig. 15. Let $a(s)$ denote the number of arcs in an n-chip $s \in C_n$. Then

$$Q_n(x) = \sum_{s \in C_n} (-1)^{a(s)} x^{n-2a(s)}. \tag{8.1.14}$$

The formula follows easily from the equation (4.18) in [5].

8.2 A generalization of the Theorem 8.1.3. Askey and Wimp have found in [5] the explicit orthogonality relations for associated Hermite polynomials. Using their result we shall find the exponential representations (2.4.2) for the corresponding distributions.

Recall that the *parabolic cylinder function* $y = D_\nu(z)$ can be defined as the solution of the Weber equation

$$\frac{d^2y}{dz^2} + \Big(\nu + \frac{1}{2} - \frac{z^2}{4}\Big)y = 0$$

with the initial conditions

$$D_\nu(0) = \frac{\Gamma(\frac{1}{2})2^{\frac{\nu}{2}}}{\Gamma(\frac{1-\nu}{2})}; \qquad D'_\nu(0) = \frac{\Gamma(\frac{-1}{2})2^{\frac{\nu-1}{2}}}{\Gamma(\frac{-\nu}{2})}$$

(see [33], p.324). If $\nu = -c < 0$, there is an integral representation (see [33], eq.(8.1.4))

$$D_{-c}(z) = \frac{e^{-z^2/4}}{\Gamma(c)} \int_0^\infty e^{-zx} x^{c-1} e^{-x^2/2} dx,$$

so that up to a constant factor the function $e^{z^2/4} D_{-c}(z)$ is the Laplace transform of the so called χ-distribution with c degrees of freedom (see [13]).

Consider a measure μ_c with the density

(8.2.1) $$\frac{d\mu_c(u)}{du} = W(u,c) \equiv \frac{1}{\sqrt{2\pi}\Gamma(c+1)} \frac{1}{|D_{-c}(iu)|^2}; \qquad -\infty < u < \infty.$$

We use the following result of [5] (slightly adopted for our purposes).

(8.2.2) Theorem (Askey, Wimp). *For every $c > -1$, the equation (8.2.1) determines a probability distribution. The Cauchy-Stieltjes transform of the measure μ_c has a simple continued fraction expansion, namely*

(8.2.3) $$R_c(z) \equiv \int_{-\infty}^{\infty} \frac{d\mu_c(u)}{z-u} = \cfrac{1}{z - \cfrac{c+1}{z - \cfrac{c+2}{z - \dots}}}$$

(Note that the continued fraction decomposition carries the same information as the recurrence relation for orthogonal polynomials, see equation (1.7) in [5].)

We show that the exponential representation of the measure μ_c is provided by an absolutely continuous probability distribution τ_c with the density

(8.2.4) $$\frac{d\tau_c(u)}{du} = (c+1)\,W(u,c+1) - c\,W(u,c); \qquad -\infty < u < \infty.$$

(8.2.5) Theorem. *For each $c > -1$, the measures (8.2.1) and (8.2.4) are related by the identity*

$$-\frac{d}{dz}\ln\int_{-\infty}^{\infty}\frac{d\mu_c(u)}{z-u} = \int_{-\infty}^{\infty}\frac{d\tau_c(u)}{z-u}.$$

Proof. By definition (8.2.4), the Cauchy-Stieltjes transform of the measure τ_c equals

$$T_c(z) \equiv \int_{-\infty}^{\infty}\frac{d\tau_c(u)}{z-u} = (c+1)R_{c+1}(z) - cR_c(z).$$

It also follows directly from (8.2.3) that $zR_c(z) - 1 = (c+1)R_c(z)R_{c+1}(z)$. If we show that $y = R_c(z)$ satisfies the Riccati equation $R'(z) = cR_c^2(z) - zR_c(z) + 1$, this would yield the master identity in the form $R_c'(z) + R_c(z)T_c(z) = 0$.

It is known since Euler and Lagrange that the continued fraction (8.2.3) is a formal solution of the Riccati equation. In order to show this, note that a substitution

$$R_c(z) = -\frac{1}{c}\frac{\varphi'(z)}{\varphi(z)} \tag{8.2.6}$$

transforms the Riccati equation into a second order linear equation

$$\varphi''(z) + z\varphi'(z) + c\varphi(z) = 0. \tag{8.2.7}$$

Taking consecutive derivatives of (8.2.7), one obtains

$$\varphi^{(n+2)}(z) + z\varphi^{(n+1)}(z) + (c+n)\varphi^{(n)}(z) = 0. \tag{8.2.8}$$

Now write the equations (8.2.7), (8.2.8) in the form

$$-\frac{1}{c}\frac{\varphi'(z)}{\varphi(z)} = \frac{1}{z - \left(-\frac{\varphi''(z)}{\varphi'(z)}\right)}; \qquad -\frac{\varphi^{(n+1)}(z)}{\varphi^{(n)}(z)} = \frac{c+n}{z - \left(-\frac{\varphi^{(n+2)}(z)}{\varphi^{(n+1)}(z)}\right)}.$$

We get, at least formally, that

$$R_c(z) = \cfrac{1}{z - \cfrac{c+1}{z - \cfrac{c+2}{z - \dots}}}$$

Though there is a vast literature on solving Riccati equation in terms of continued fractions (see [19] and the references there), we couldn't find an appropriate reference. For the reader's convenience we prove the convergence, too. In the course of the proof we use associated Hermite polynomials introduced in [5].

8.3 The associated Hermite polynomials. Denote by $H_n(x;c)$ the polynomials defined by the recurrence relation

$$H_{n+1}(x;c) = xH_n(x;c) - (c+n)H_{n-1}(x;c) \tag{8.3.1}$$

and the usual initial conditions. The first examples are

$$\begin{aligned}
H_0(x;c) &= 1\\
H_1(x;c) &= x\\
H_2(x;c) &= x^2 - (c+1)\\
H_3(x;c) &= x^3 - (2c+3)x\\
H_4(x;c) &= x^4 - (3c+6)x^2 + (c^2+4c+3).
\end{aligned}$$

It follows from [5] that the polynomials $H_n(x;c)$ (called there the *associated Hermite polynomials*) are orthogonal with respect to the measure μ_c,

$$\int_{-\infty}^{\infty} H_m(x;c)H_n(x;c)d\mu_c(x) = \delta_{mn}.$$

In the particular case of $c = 0$ one obtains the ordinary Hermite polynomials, and the polynomials $H_n(x;1) = Q_n(x)$ coincide with those of Remark 8.1.12. Note that the combinatorial description (8.1.14) of the latter polynomials can be extended to the case of general $c > -1$:

$$H_n(x;c) = \sum_{s\in C_n} (-1)^{a(s)} c^{m(s)} x^{n-2a(s)}. \tag{8.3.2}$$

Here $m(s)$ denotes the number of cycles in a permutation generated by the vertical arcs of an n-chip $s \in C_n$, see Fig. 16. An arc is called *vertical*, if it has a point in D. The sets D and R are assumed to be linearly ordered, so that the vertical arcs determine a permutation in $\mathfrak{S}_d$, $d = |D|$.

One can easily derive from the recurrence relation (8.3.1) that

$$H_n(x;c) = H_n(x;c+1) + H'_{n-1}(x;c+1), \tag{8.3.3}$$

$$H_{n+1}(x;c-1) - xH_n(x;c) + cH_{n-1}(x;c+1) = 0. \tag{8.3.4}$$

Applying the elementary theory of continued fractions one can write

$$\frac{H_{n-1}(z;c+1)}{H_n(z;c)} = \cfrac{1}{z - \cfrac{c+1}{z - \cfrac{c+2}{\cdots - \cfrac{c+n-1}{z}}}} \longrightarrow R_c(z), \tag{8.3.5}$$

as far as $n \to \infty$, and

$$\frac{H_n(z;c)}{H_{n+1}(z;c)} = \cfrac{1}{z - \cfrac{c+n}{z - \cfrac{c+n-1}{\dots - \cfrac{c+1}{z}}}}; \qquad \frac{H_{n-1}(z;c+1)}{H_n(z;c+1)} = \cfrac{1}{z - \cfrac{c+n}{z - \cfrac{c+n-1}{\dots - \cfrac{c+2}{z}}}}.$$

It follows from Flajolet theorem 5.1.8 that the power series expansions of these functions at infinity have identical terms up to the order z^{-2n}.

Add the ratio $H_n'(z;c)/H_n(z;c)$ to the inverses of the latter fractions. Then the formulas (8.3.3), (8.3.4) yield

$$\frac{H_{n+1}(z;c)}{H_n(z;c)} + \frac{H_n'(z;c)}{H_n(z;c)} = \frac{H_{n+1}(z;c-1)}{H_n(z;c)} = z - c\,\frac{H_{n-1}(z;c+1)}{H_n(z;c)} \tag{8.3.6}$$

and it follows that

$$\begin{aligned} &\frac{H_n(z;c+1)}{H_{n-1}(z;c+1)} + \frac{H_n'(z;c)}{H_n(z;c)} = \\ =&\frac{H_n(z;c)}{H_{n-1}(z;c+1)} - \Big(\frac{H_{n-1}'(z;c+1)}{H_{n-1}(z;c+1)} - \frac{H_n'(z;c)}{H_n(z;c)}\Big) = \\ =&\frac{H_n(z;c)}{H_{n-1}(z;c+1)} - \Big(\frac{H_{n-1}(z;c+1)}{H_n(z;c)}\Big)'\Big(\frac{H_{n-1}(z;c+1)}{H_n(z;c)}\Big)^{-1}. \end{aligned} \tag{8.3.7}$$

By (8.3.5), the limit of (8.3.6), as $n \to \infty$, is $z - cR_c(z)$, and (8.3.7) converges to $\frac{1}{R_c(z)} - \frac{R_c'(z)}{R_c(z)}$. Since both expressions (8.3.6), (8.3.7) have n identical first terms in their power series expansions at infinity, the limiting asymptotic series coincide, and

$$z - cR_c(z) = \frac{1}{R_c(z)} - \frac{R_c'(z)}{R_c(z)}, \qquad \text{for } z \in \mathbf{C} \setminus \mathbf{R}.$$

This is exactly the Riccati equation for $R_c(z)$, and the proof of Theorem 8.2.5 is completed. $\square$

8.4 A deformation of the gaussian distribution to the semi-circle one. Using the affine invariance of the master identity one can slightly generalize Theorem 8.2.5. To this end, we denote by $\mu_{c,\beta}$ the distribution with the density

$$w_{c,\beta}(u) = \frac{1}{\sqrt{\beta}} W\Big(\frac{u}{\sqrt{\beta}}, \frac{c}{\beta}\Big); \qquad \beta > 0. \tag{8.4.1}$$

In words, $\mu_{c,\beta}$ is the image of the measure $\mu_{c/\beta}$ upon the dilation $u \mapsto u\sqrt{\beta}$. In a similar way, we introduce the measure $\tau_{c,\beta}$ with the density

$$\frac{d\tau_{c,\beta}(u)}{du} = \Big(1 + \frac{c}{\beta}\Big) w_{c+\beta,\beta}(u) - c w_{c,\beta}(u). \tag{8.4.2}$$

Once again, $\tau_{c,\beta}$ is the image of $\tau_{c/\beta}$ upon the dilation $v = u\sqrt{\beta}$.

(8.4.3) Corollary. *The measures $\mu_{c,\beta}$ and $\tau_{c,\beta}$ are related by the identity*

$$\int_{-\infty}^{\infty} \frac{d\mu_{c,\beta}(u)}{z-u} = \exp \int_{-\infty}^{\infty} \ln \frac{1}{z-u}\, d\tau_{c,\beta}(u),$$

for any choice of the parameters $c > -1$, $\beta > 0$.

Proof. This is a direct consequence of Theorem 8.2.5. $\square$

We complete this section with the description of moments $h_n(c,\beta)$, $p_n(c,\beta)$ of the measures $\mu_{c,\beta}$, $\tau_{c,\beta}$. The first few examples are as follows:

$h_0(c,\beta) = 1;$	$p_0(c,\beta) = 1;$
$h_2(c,\beta) = (c+\beta);$	$p_2(c,\beta) = 2(c+\beta);$
$h_4(c,\beta) = (c+\beta)(2c+3\beta);$	$p_4(c,\beta) = 2(c+\beta)(3c+5\beta);$
$h_6(c,\beta) = (c+\beta)(5c^2+17c\beta+15\beta^2);$	$p_6(c,\beta) = 2(c+\beta)(10c^2+38c\beta+37\beta^2).$

(8.4.4) Lemma. *The moments $h_{2n}(c,\beta)$, $p_{2n}(c,\beta)$ are homogeneous polynomials of degree n in the parameters c, β. The coefficient of c^n is the Catalan number $\binom{2n}{n}/(n+1)$ for $h_{2n}(c,\beta)$, and the choose symbol $\binom{2n}{n}$ for $p_{2n}(c,\beta)$.*

Proof. Taking into account the general relations (3.3.3) between the moments h_n and p_n, and the Example 5.2.7, it suffices to prove our claim for h_n. By the definition of the measure $\mu_{c,\beta}$, its Cauchy-Stieltjes transform $R_{c,\beta}(z)$ is related to that of the measure μ_c by the equation

$$R_{c,\beta}(z) = \frac{1}{\sqrt{\beta}} R_{c/\beta}\left(\frac{z}{\sqrt{\beta}}\right).$$

Using the continued fraction expansion (8.2.3) of $R_c(z)$, one derives that

$$R_{c,\beta}(z) = \cfrac{1}{z - \cfrac{c+\beta}{z - \cfrac{c+2\beta}{z - \cfrac{c+3\beta}{z - \dots}}}}, \tag{8.4.5}$$

so that the elements $\lambda_n = c + n\beta$ of the continued fraction are linear and homogeneous in c and β. The Lemma follows now from Flajolet theorem 5.1.8. $\square$

(8.4.6) Corollary. *If $\beta = 0$, then $\mu_{c,0}$ is the semi-circle distribution*

$$\frac{d\mu_{c,0}(u)}{du} = \frac{1}{2\pi c}\sqrt{4c-u^2}; \qquad -\infty < u < \infty$$

with the variance c. Hence, the family $\mu_{c,\beta}$ provides a deformation between the Gaussian measure $\mu_{0,1}$ and the semi-circle law $\mu_{1,0}$.