Functional Equations and Electrostatic Models for Orthogonal Polynomials

MOURAD E. H. ISMAIL

To my dear friend Walter K. Hayman on the occasion of his 75th birthday.

ABSTRACT. This article deals with connections between orthogonal polynomials, functional equations they satisfy, and some extremal problems. We state Stieltjes electrostatic models and Dyson’s Coulomb fluid method. We also mention the evaluation of the discriminant of Jacobi polynomials by Stieltjes and Hilbert. We show how these problems can be extended to general orthogonal polynomials with absolutely continuous measures or having purely discrete orthogonality measures whose masses are located at most two sequences of geometric progressions.

1. Introduction

This is a survey article dealing with connections between orthogonal polynomials, functional equations they satisfy, and some extremal problems. Although the results surveyed are not new we believe that we are putting together results from different sources which appear together for the first time, many of them are of recent vintage.

One question in the theory of orthogonal polynomials is how the zeros of a parameter dependent sequence of orthogonal polynomials change with the parameters involved. Stieltjes [1885a; 1885b] proved that the zeros of Jacobi polynomials $P_N^{(\alpha,\beta)}(x)$ increase with $\beta$ and decrease with $\alpha$ for $a > -1$ and $\beta > -1$. The Jacobi polynomials satisfy the following orthogonality relation [Szegö 1975, (4.3.3)]:

1991 Mathematics Subject Classification. Primary 42C05; Secondary 33C45.

Key words and phrases. Discriminants, orthogonal polynomials, zeros, interacting particles.

Research partially supported by NSF grant DMS-99-70665.

225
\[
\int_{-1}^{1} (1-x)^{\alpha}(1+x)^{\beta} P_{m}^{(\alpha,\beta)}(x) P_{n}^{(\alpha,\beta)}(x) \, dx = \frac{2^{\alpha+\beta+1}}{2n^{\alpha+\beta+1}} \Gamma(n+\alpha+1) \Gamma(n+\beta+1) \delta_{m,n}. \quad (1-1)
\]

In Section 2 we state Stieltjes’s results and describe the circle of ideas around them. Section 3 surveys the Coulomb Fluid method of Freeman Dyson and its potential theoretic set-up.

We shall use the shifted factorial notion [Andrews et al. 1999]

\[(\lambda)_{n} = \prod_{k=1}^{n} (\lambda + k - 1),\]

and the hypergeometric notation [Andrews et al. 1999]

\[rF_{s}\left( \begin{array}{c} a_{1}, \ldots, a_{r} \\ b_{1}, \ldots, b_{s} \end{array} \bigg| z \right) = rF_{s}(a_{1}, \ldots, a_{r}; b_{1}, \ldots, b_{s}; z) = \sum_{n=0}^{\infty} (a_{1})_{n} \cdots (a_{r})_{n} \frac{z^{n}}{(b_{1})_{n} \cdots (b_{s})_{n} n!}.
\]

The Jacobi polynomials have the explicit form

\[P_{n}^{(\alpha,\beta)}(x) = \frac{(\alpha+1)_{n}}{n!} 2F_{1}\left( \begin{array}{c} -n, n+\alpha+\beta+1 \\ \alpha+1 \end{array} \bigg| \frac{1-x}{2} \right). \quad (1-2)
\]

The discriminant of a polynomial \(g_{n}\),

\[g_{n}(x) := \gamma_{n}x^{n} + \text{lower order terms}, \quad \gamma_{n} \neq 0, \quad (1-3)
\]

is defined by

\[D(g_{n}) := \gamma_{n}^{2n-2} \prod_{1 \leq j < k \leq n} (x_{j} - x_{k})^{2}, \quad (1-4)
\]

where \(x_{1}, x_{2}, \ldots, x_{n}\) are the zeros of \(g_{n}(x)\); see [Dickson 1939]. The discriminant has the alternate representation

\[D(g_{n}) = (-1)^{n(n-1)/2} \gamma_{n}^{n-2} \prod_{j=1}^{n} g'_{n}(x_{j}). \quad (1-5)
\]

See, for example, Dickson [1939, § 100].

2. Stieltjes

Stieltjes [1885a; 1885b] considered the following electrostatic model. Fix two charges \((\alpha + 1)/2\) and \((\beta + 1)/2\) at \(x = 1\) and \(x = -1\), respectively, then put \(N\) movable unit charges at distinct points in \((-1,1)\). The potential here is a logarithmic potential so the potential energy of a system of two charges \(e_{1}\) and
\( e_2 \) located at \( x \) and \( y \) is \(-2e_1e_2 \ln |x - y| \). Let the position of the unit charges be at \( x_1, x_2, \ldots, x_N \). Thus the energy of this system is

\[
E_N(x) = -(\alpha + 1) \sum_{k=1}^{N} \ln |1 - x_k| - (\beta + 1) \sum_{k=1}^{N} \ln |1 + x_k| \\
- 2 \sum_{1 \leq j < k \leq N} \ln |x_j - x_k|,
\]

where

\[
x = (x_1, x_2, \ldots, x_N).
\]

(2-1)

For convenience we consider the function

\[
T_N(x) := \exp(-E_N(x)),
\]

that is,

\[
T_N(x) = \prod_{k=1}^{N} (1 - x_k)^{\alpha + 1}(1 + x_k)^{\beta + 1} \prod_{1 \leq i < j \leq N} [x_i - x_j]^2.
\]

(2-2)

The equilibrium position of the system occurs at the points which minimize \( E_N(x) \) or equivalently maximize \( T_N(x) \).

**Theorem 2.1 (Stieltjes).** The maximum of the function \( T_N(x) \) taken over \( x \in \mathbb{R}^N \) is attained when \( x \) is formed by the zeros of \( P_N^{(\alpha, \beta)}(x) \). In other words the equilibrium position of the movable charges in the electrostatic model described above is attained at the zeros of the Jacobi polynomial \( P_N^{(\alpha, \beta)}(x) \).

In Section 4, we will give a proof of a generalization of Theorem 2.1.

Stieltjes found a closed form expression for the maximum value of \( T_N(x) \) in Theorem 2.1. He observed that (1-2) implies

\[
P_N^{(\alpha, \beta)}(x) = (N + \alpha + \beta + 1)_N \frac{2^{-N}}{N!} x^N + \text{lower order terms},
\]

so that if \( x_{1, N} > x_{2, N} > \cdots > x_{N, N} \) are the zeros of \( P_N^{(\alpha, \beta)}(x) \) then

\[
\prod_{k=1}^{N} (\pm x - x_{k, N}) = \frac{2^N N!}{(N + \alpha + \beta + 1)_N} P_N^{(\alpha, \beta)}(\pm x),
\]

so that at equilibrium the first product in (2-2) can be found from

\[
\prod_{k=1}^{N} (1 - x_{j, N}) = \frac{2^N N!}{(N + \alpha + \beta + 1)_N} P_N^{(\alpha, \beta)}(1)
\]

\[
\prod_{k=1}^{N} (1 + x_{j, N}) = \frac{(-1)^N 2^N N!}{(N + \alpha + \beta + 1)_N} P_N^{(\alpha, \beta)}(-1).
\]
Clearly (1–2) implies
\[ P_N^{(\alpha,\beta)}(1) = (\alpha + 1)_N/N! . \]
Since the weight function and the right-hand side in (1–1) are symmetric under the exchange \((x, \alpha, \beta) \rightarrow (-x, \beta, \alpha)\) then
\[ P_N^{(\alpha,\beta)}(-1) = (-1)^N P_N^{(\beta,\alpha)}(1) = (-1)^N \frac{(\beta + 1)_N}{N!} . \]
The second product in (2–2) at equilibrium is the discriminant of Jacobi polynomials. Stieltjes then found explicit formulas for the discriminants of the Jacobi polynomials. His formula in our notation is
\[ D(P_N^{(\beta,\alpha)}) = 2^{-N(N-1)} \prod_{k=1}^{N} k^{k+2-2N} (k + \alpha)^{k-1}(k + \beta)^{k-1}(N + k + \alpha + \beta)^{N-k} . \]
(2–3)
Since the Hermite and Laguerre polynomials are limiting cases of Jacobi polynomials, (2–3) yields explicit evaluations for the discriminants of the Hermite and Laguerre polynomials. Shortly after Stieltjes work appeared, Hilbert [1888] gave another proof of (2–3). Schur [1931] also gave a very elegant proof. We reproduce the proof here to make the paper as self-contained as possible and also because Schur’s method proved to be central in the generalizations of Stieltjes results to general orthogonal polynomials.

**Lemma 2.2** [Schur 1931]. Assume that \( \{\rho_n(x)\} \) is a sequence orthogonal polynomials satisfying a three-term recurrence relation of the form
\[ \rho_{n+1}(x) = (\xi_{n+1} x + \eta_{n+1}) \rho_n(x) - \zeta_{n+1} \rho_{n-1}(x) , \]
(2–4)
and the initial conditions
\[ \rho_0(x) = 1, \quad \rho_1(x) = \xi_1 x + \eta_1 . \]
(2–5)
If
\[ x_{1,n} > x_{2,n} > \cdots > x_{n,n} \]
(2–6)
are the zeros of \( \rho_n(x) \) then
\[ \prod_{k=1}^{n} \rho_{n-1}(x_{k,n}) = (-1)^{n(n-1)/2} \prod_{k=1}^{n} \zeta_k^{n-2k+1} s_k^{k-1} , \]
(2–7)
with \( \zeta_1 := 1 \).
Proof. Let $\Delta_n$ denote the left-hand side of (2-7). The coefficient of $x^n$ in $\rho_n(x)$ is $\xi_1 \xi_2 \ldots \xi_n$. Thus by expressing $\rho_n$ in terms and $\rho_{n+1}$ of their zeros we find

$$\Delta_{n+1} = \left(\xi_1 \xi_2 \ldots \xi_n\right)^{n+1} \prod_{k=1}^{n+1} \prod_{j=1}^{n} (x_k, n+1 - x_j, n) \prod_{j=1}^{n+1} (x_j, n - x_k, n+1)\right) \left(\xi_1 \xi_2 \ldots \xi_n\right)^{n+1} \prod_{j=1}^{n} \rho_{n+1}(x_j, n).

On the other hand, the three-term recurrence relation (2-4) simplifies the extreme right-hand side in the above equation and we get

$$\Delta_{n+1} = \xi_1 \xi_2 \ldots \xi_n \sum_{i=1}^{n+1} \xi_i (-\zeta_{n+1})^n \Delta_n.$$

By iterating this relation we establish (2-7).

The relevance of Schur’s lemma to the evaluation of discriminants of Jacobi polynomials is the fact that the Jacobi polynomials satisfy a lowering (annihilation) relation of the type

$$\frac{d}{dx} P_N^{(\alpha, \beta)}(x) = A_N(x) P_N^{(\alpha, \beta)}(x) - B_N(x) P_{N-1}^{(\alpha, \beta)}(x),$$

with [Rainville 1960, (7), §136]

$$A_N(x) = \frac{2(N + \alpha)(N + \beta)}{(\alpha + \beta + 2N)(1 - x^2)}, \quad B_N(x) = \frac{N(\beta - \alpha + (\alpha + \beta + 2N)x)}{(\alpha + \beta + 2N)(1 - x^2)}.$$

Thus

$$\frac{d}{dx} P_N^{(\alpha, \beta)}(x_k, N) = A_N(x_k, N) P_N^{(\alpha, \beta)}(x_k, N)$$

and (1-5) and Schur’s lemma lead to the evaluation of the discriminant.

We now formulate this procedure as a general property of polynomials satisfying three term recurrence relations and possessing a lowering operator.

Theorem 2.3. Let a system of polynomials $\{p_n(x)\}$ be generated by (2-4) and (2-5) and assume that the zeros of $p_n(x)$ be arranged as in (2-6). If $\{p_n(x)\}$ satisfies the differential recurrence relation

$$\frac{d}{dx} p_n(x) = A_n(x) p_{n-1}(x) - B_n(x) p_n(x),$$

then the discriminant of $p_n(x)$ is given by

$$D(p_n) = \left(\prod_{j=1}^{n} A_n(x_j, n)\right) \prod_{k=1}^{n} \xi_k^{2n-2k-1} \zeta_k^{k-1}.$$
PROOF. The $\gamma_n$ in (1–3) is $\xi_1 \ldots \xi_n$. Thus (1–5) and Lemma 2.2 give

$$D(p_n) = (-1)^{n(n-1)/2}(\xi_1 \ldots \xi_n)^{n-2}(-1)^{n(n-1)/2} \prod_{k=1}^{n} A_n(x_{k,n}) \xi_n^{2k+1} \xi_k^{k-1},$$

which establishes the theorem. $\square$

It is clear that Theorem 2.3 implies the evaluation (2–3) of the discriminant of the Jacobi polynomials once we know the three-term recurrence relation satisfied by Jacobi polynomials [Rainville 1960, (1) §137]. From [Bauldry 1990; Bonan and Clark 1990; Chen and Ismail 1997], we now know that every polynomial sequence orthogonal with respect to a weight function, satisfying certain smoothness conditions (see §4), satisfies a differential recurrence relation of the type (2–8). Hence, Theorem 2.3 holds for orthogonal polynomials in this generality, a result from [Ismail 1998].

Selberg [1944] proved

$$\int_{[0,1]^n} \left( \prod_{j=1}^{n} t_j^{x-1}(1-t_j)^{y-1} \right) \prod_{1 \leq i < k \leq n} |t_i-t_k|^d dt_1 \ldots dt_n$$

$$= \prod_{j=1}^{n} \frac{\Gamma(x+(n-j)z) \Gamma(y+(n-j)z) \Gamma(jz+1)}{\Gamma(x+y+(2n-j-1)z) \Gamma(z+1)}, \quad (2-9)$$

for $\text{Re } x > 0$, $\text{Re } y > 0$, and $Re z > -\min\{1/n, \text{Re } x/n - 1, \text{Re } y/(n-1)\}$. Here $[0,1]^n$ is the unit cube in $\mathbb{R}^n$. The integral evaluation (2–9) is the multivariate generalization of the beta integral and is now called the Selberg integral [Andrews et al. 1999]. It is important to note that if we normalize the Jacobi polynomials to be orthogonal on $[0,1]$ then the Stieltjes–Hilbert results provide the $L_\infty$ norm of

$$\left( \prod_{j=1}^{n} t_j^{x-1}(1-t_j)^{y-1} \right) \prod_{1 \leq i < k \leq n} |t_i-t_k|^d, \quad t_j \in (-1,1) \text{ for } 1 \leq j \leq n. \quad (2-10)$$

On the other hand, the Selberg integral (2–9) essentially gives the $L_n$ norm of the expression in (2–10). One is then led to view the Stieltjes–Hilbert results as a limiting case of the Selberg integral.

3. Dyson and Potential Theory

As a generalization of the Stieltjes electrostatic problem we consider a system of $N$ logarithmically repelling particles obeying Boltzmann statistics subject to a common external potential $v(x)$ in one dimension. Let $\Phi(x)$ denote the total energy of the system, that is

$$\Phi(x) := \sum_{1 \leq j \leq N} v(x_j) - 2 \sum_{1 \leq j < k \leq N} \ln |x_j - x_k|,$$
and \( x \) is as in (2.1). Here the particles are assumed to be confined to a real interval \( K \), finite, semi-infinite, or infinite. Dyson’s idea [1962a; 1962b; 1962c] was that for large \( N \), one would expect that this collection of particle can be approximated by a continuous fluid where techniques of macroscopic physics such as thermodynamics and electrostatics can be applied. The Coulomb fluid approximation is described by an equilibrium density \( \sigma(x) \), supported on a set \( J \subset K \), which is obtained by minimizing the free energy functional, \( F[\sigma] \)

\[
F[\sigma] = \int_J v(x)\sigma(x)
- \int_J \int_J \sigma(x) \ln |x - y| \sigma(y)
\, dy
dx, \quad (3.1)
\]

subject to the side condition

\[
\int_J \sigma(x)\, dx = N. \quad (3.2)
\]

From the Euler-Lagrange equations for the system (3.1) it follows that the density \( \sigma(x) \) satisfies the singular integral equation

\[
v'(x) = 2P \int_J \frac{\sigma(y)}{x - y} dy, \quad x \in J. \quad (3.3)
\]

Therefore

\[
A = v(x) - 2 \int_J \sigma(y) \ln |x - y| \, dy, \quad x \in J, \quad (3.4)
\]

where \( A \) is the Lagrange multiplier for the constraint (3.2) and is recognized as the chemical potential for the fluid. Determining \( J \) is part of the solution of the variational problem.

Let \( v(x) \) is convex for \( x \in \mathbb{R} \), so that \( v''(x) \geq 0 \) almost everywhere. We shall assume \( v'(x) > 0 \) on a set of positive measures. At this stage, some physical considerations are invoked. With the condition \( v(x) > 0 \) it follows that \( J \) is a single interval denoted by \( (a, b) \). Intuitively, this physical principle can be understood by using an analogy from elasticity theory [Muskhelishvili 1953], where the fluid density \( \sigma(x) \) is identified with the pressure under a stamp pressing vertically downwards against an elastic half-plane. If the applied force is moderate, the end points of the interval, \( a \) and \( b \), are the points for which the elastic material come into contact with the rigid stamp. On the other hand if the force applied to the stamp is too great the end stamp will be fixed as the end points of the boundary of the stamp.

We seek a solution of (3.3) which is nonnegative on \( (a, b) \). If imposing the boundary conditions \( \sigma(a) = \sigma(b) = 0 \) lead to \( \sigma \) satisfying \( \sigma(x) \geq 0 \) on \( (a, b) \) then, according to the standard theory of singular integral equations [Gakhov 1990], the solution of (3.3) is

\[
\sigma(x) = \frac{\sqrt{(b - x)(x - a)}}{2\pi^2} \int_a^b \frac{v'(x)}{(y - x)\sqrt{(b - y)(y - a)}} dy, \quad (3.5)
\]
and $a$ and $b$ must satisfy the constraint (3–2) as well as the supplementary condition,

$$0 = \int_a^b \frac{v'(x)}{\sqrt{(b-x)(x-a)}} \, dx.$$  \hfill (3–6)

Using (3–6) the normalization condition becomes

$$N = \frac{1}{2\pi} \int_a^b \frac{xv'(x)}{\sqrt{(b-x)(x-a)}} \, dx.$$  \hfill (3–7)

The end points of the support of the density, $a$ and $b$, that are solutions of (3–6) and (3–7) are denoted by $a(N)$ and $b(N)$.

Sometimes the boundary conditions that $\sigma(x)$ vanishes at the endpoints of $J$ do not lead to a solution $\sigma(x)$ which is nonnegative on $J$. In this case other forms of solutions of (3–4) can be used. A good application of this physical approach is to the Freud weights

$$w(x, \alpha) = \exp(-|x|^{\alpha}), \quad \alpha > 0,$$  \hfill (3–8)

so that $v(x) = |x|^\alpha$. Since $w$ is even then $a(N) = -b(N)$. From (3–5) and (3–6) we find that the density, denoted by $\sigma(x, \alpha)$, to be

$$\sigma(x, \alpha) = \frac{\alpha}{\pi} \frac{2^{1-\alpha} \Gamma(\alpha)}{\Gamma(\alpha/2)^2} (b(N))^{\alpha-2} \sqrt{(b(N))^2 - x^2} \, _2F_1(1-\alpha/2, 1; 3/2; 1-(x/b(N))^2).$$

The form (3–7) of side condition (3–2) gives

$$N = \frac{2^{1-\alpha} \frac{1}{\pi} \frac{2^{1-\alpha} \Gamma(\alpha)}{\Gamma(\alpha/2)^2} (b(N))^{\alpha-2}}{\Gamma(\alpha/2) \Gamma(\alpha/2)} (b(N))^{\alpha}.$$  

Since the Coulomb fluid approximation works for large $N$ then we find the asymptotic result

$$b(N) \simeq \left( \frac{\Gamma^2(\alpha/2) z_0^{\alpha-1} N^{1/\alpha}}{\Gamma(\alpha)} \right).$$

This gives the large $N$ behavior of the position of the charges on the extreme right ($= b(N)$) and extreme left ($= -b(N)$). Furthermore, motivated by the Stieltjes result, this was believed to be the positions of the largest and smallest zeros of the polynomials orthogonal with respect to $\exp(-|x|^{\alpha})$.

What is outlined so far was started by Dyson [1962a; 1962b; 1962c] for the circular ensemble of random matrix theory. It was further developed and applied to orthogonal polynomials by many physicists, too many to be cited in this short article. Having said this, perhaps it is not too impertinent to mention a sample of the work of Yang Chen and his collaborators [Chen et al. 1995; Chen and Manning 1990], especially since my good friend, Yang Chen, is the one who introduced me to this subject, and one of the co-authors of the cited work was a co-organizer of the MSRI program on Random Matrices.
This approach can be made rigorous by using potential theory [Saff and Totik 1997] and Riemann–Hilbert problems starting from the pioneering work of Bleher and Its [1999] and cumulating in the recent series of monumental papers of Deift, Kriecherbauer, McLaughlin, Zhou, and others. First one starts with polynomials \( \{p_n(x) : 0 \leq n\} \) orthogonal with respect to a weight function \( w(x) \) on an interval \( K \) and associate with it an external field \( v(x) := -\ln w(x) \). The weight function \( w(x) \) is assumed to be positive on the interior of \( K \). Let \( x_{1,N} > x_{2,N} > \cdots > x_{N,N} \) be the zeros of \( p_N(x) \). Define a probability measure \( \nu_N \) to have masses \( 1/N \) at \( x_{k,N} \), for \( 1 \leq k \leq N \). When this sequence of measures converge in the weak-* topology to a probability measure \( \nu \) then \( \nu \) will be called the equilibrium measure associated with the external field \( v(x) \). In general the equilibrium position of \( N \) movable charges in the external field \( v(x) \) is not at the zeros of \( p_N(x) \) but at what is called the Fekete points, say \( y_{1,N} > y_{2,N} > \cdots > y_{N,N} \). One can think of the equilibrium measure as \( N^{-1}\sigma(x)dx \), where \( \sigma \) is Dyson’s fluid density. The absolute continuity of \( \nu \) was not proved till very recently by Deift, Kriecherbauer, and McLaughlin [Deift et al. 1998], for real analytic external fields, thus confirming Dyson’s intuition. Potential theory [Saff and Totik 1997] also confirms that the largest \( (y_{1,N}) \) and smallest \( (y_{N,N}) \) Fekete points are asymptotically equivalent to the largest \( (x_{1,N}) \) and smallest \( (x_{N,N}) \) zero of \( p_N(x) \), respectively, that is,

\[
\lim_{N \to \infty} \frac{x_{N,N}}{y_{N,N}} = \lim_{N \to \infty} \frac{x_{1,N}}{y_{1,N}} = 1.
\]

4. Differential Equations and Discriminants

In this section, we shall assume that \( \{p_n(x)\} \) is a sequence of orthonormal polynomials whose weight function is \( w(x) \), that is

\[
\int_a^b p_m(x)p_n(x)w(x)dx = \delta_{m,n}.
\]

Throughout this section we shall assume that \( w(x) > 0 \) on \( (a,b) \),

\[
w(x) = \exp(-v(x)), \quad x \in (a,b),
\]

and that \( v \) has a continuous derivative on \( (a,b) \). We shall normalize \( w \) by setting

\[
\int_a^b w(x)dx = 1.
\]

The initial values and three-term recurrence relation of \( \{p_n(x)\} \) take the form

\[
p_0(x) = 1, \quad p_1(x) = (x - b_0)/a_1,
\]

\[
xp_n(x) = a_{n+1}p_{n+1}(x) + b_np_n(x) + a_np_{n-1}(x), \quad n > 0.
\]
Assume \( v \) has a continuous first derivative on \((a, b)\). We define \( A_n(x) \) and \( B_n(x) \) via
\[
A_n(x) = \frac{a_n w(b^-) p_n^2(b)}{b-x} + \frac{a_n w(a^+) p_n^2(a)}{x-a} + a_n \int_a^b \frac{v'(x) - v'(y)}{x-y} p_n^2(y) w(y) \, dy
\]
and
\[
B_n(x) = \frac{a_n w(a^+) p_n(a) p_{n-1}(a)}{x-a} + \frac{a_n w(b^-) p_n(b) p_{n-1}(b)}{b-x} + a_n \int_a^b \frac{v'(x) - v'(y)}{x-y} p_n(y) p_{n-1}(y) \, w(y) \, dy.
\]

In (4-3) and (4-4) it is assumed that
\[
y^n \frac{v'(x) - v'(y)}{x-y} \, w(y), \quad n = 0, 1, \ldots,
\]
are integrable over \((a, b)\) and the boundary terms in (4-3) and (4-4) exist. Under the latter assumptions, the orthonormal polynomials \( p_n \)'s satisfy the differential recurrence relation [Bauldry 1990; Bonan and Clark 1990; Chen and Ismail 1997],
\[
p_n(x) = A_n(x) p_{n-1}(x) - B_n(x) p_n(x),
\]
and the second-order differential equation
\[
p_n''(x) + R_n(x) p_n'(x) + S_n(x) p_n(x) = 0,
\]
where
\[
R_n(x) := -\left( v'(x) + \frac{A_n'(x)}{A_n(x)} \right),
\]
and
\[
S_n(x) := B_n'(x) - B_n(x) \frac{A_n'(x)}{A_n(x)} - B_n(x) \left( v'(x) + B_n(x) \right) + \frac{a_n}{a_{n-1}} A_n(x) A_{n-1}(x).
\]

Bonan and Clark [1990] and Bauldry [1990] were the first to establish (4-5) and (4-6), with the boundary terms assumed to vanish. Chen and Ismail [1997] rediscovered these results, and proved several others. The form of \( R_n(x) \) of (4-7) in [Bonan and Clark 1990] and [Bauldry 1990] was more complicated, but Chen and Ismail [1997] observed that
\[
B_{n+1}(x) + B_n(x) = \frac{x - b_n}{a_n} A_n(x) - v'(x),
\]
which simplified \( R_n(x) \) to the form in (4-10). Ismail and Wimp [1998] further proved that
\[
B_{n+1}(x) - B_n(x) = \frac{a_{n+1}}{x - b_n} A_{n+1}(x) - \frac{a_n^2 A_{n-1}(x)}{a_{n-1}(x - b_n)} - \frac{1}{x - b_n}.
\]
The recurrence relations (4–9) and (4–10) are curious because by solving them for $B_n(x)$ or $B_{n+1}(x)$ we find expressions for $B_n(x)$ and $B_{n+1}(x)$ whose consistency lead to five term inhomogeneous recursion relations for $A_n(x)$. Furthermore (4–9) gives $A_n(x)$ in terms of $B_n(x)$ and $B_{n+1}(x)$ then (4–10) yields a five-term inhomogeneous recursion relations for $B_n(x)$. The details are in [Ismail and Wimp 1998]. They have another implication. It is clear from (4–9) and (4–10) that if for some $n$, $A_n(x)$, $B_n(x)$, and $B_{n+1}(x)$ are polynomial (or rational) functions then $v'(x)$ is a polynomial (or rational) function. Furthermore the relationships (4–3) and (4–4) show that if $v'(x)$ is a polynomial (or rational) function then both $A_n(x)$ and $B_n(x)$ are polynomial (or rational, respectively) functions for all $n$.

For completeness, we indicate how differential equation (4–6) follows from (4–5). Eliminate $p_{n-1}(x)$ between (4–5) and the second line in (4–2) to get

$$-p_n'(x) + \left( \frac{x-b_n}{a_n} A_n(x) - B_n(x) \right) p_n(x) = \frac{a_{n+1}}{a_n} A_n(x) p_{n+1}(x).$$

In view of (4–9), we see that the polynomials $\{p_n(x)\}$ have the raising and lowering operators $L_{n,1}$ and $L_{n,2}$

$$L_{1,n} = \frac{d}{dx} + B_n(x), \quad L_{2,n} = -\frac{d}{dx} + B_n(x) + v'(x).$$

Indeed

$$L_{1,n} p_n(x) = A_n(x) p_n(x), \quad L_{2,n} p_{n-1}(x) = \frac{a_n}{a_{n-1}} A_{n-1}(x) p_n(x).$$

The differential equation (4–6) is the expanded form of

$$L_{2,n} \left( \frac{1}{A_n(x)} (L_{1,n} p_n(x)) \right) = \frac{a_n}{a_{n-1}} A_{n-1}(x) p_n(x).$$

The differential equation (4–6) and the creation and annihilation operators $L_{2,n}$ and $L_{1,n}$ have many applications. First Chen and Ismail [1997] pointed out that $L_{2,n}$ and $L_{1,n}$ are adjoints in the Hilbert space $L_2(\mathbb{R}, w(x))$. Chen and Ismail also proved that the Lie algebras generated by $L_{1,n}$ and $L_{2,n}$ are finite dimensional when $v(x)$ is a polynomial. The rest of this section will cover two applications of (4–5) and (4–6), where we generalize Stieltjes’ electrostatic problem and the evaluation of discriminants of orthogonal polynomials.

We now indicate how (4–6) leads to a generalization of the Stieltjes problem. This material is from [Ismail 2000b]. We propose that a weight function $w(x)$ creates two external fields. One is an external field whose potential at a point $x$ is $v(x)$, $v(x)$ is as in (4–1). In addition in the presence of $N$ unit charges $w$ produces a second field whose potential is $\ln(A_N(x)/a_N)$. Thus the total external potential $V(x)$ is

$$V(x) = v(x) + \ln(A_N(x)/a_N). \quad (4-11)$$
Consider the system of \( N \) movable unit charges in \([a, b]\) in the presence of the external potential \( V(x) \) of (4-11). Let \( \mathbf{x} \) be as in (2-1) where \( x_1, \ldots, x_N \) are the positions of the particles arranged in decreasing order. The total energy of the system is

\[
E_N(\mathbf{x}) = \sum_{k=1}^{N} V(x_k) - 2 \sum_{1 \leq j < k \leq N} \ln x_j - x_k.
\]

Let

\[
T_N(\mathbf{x}) := \exp(-E_N(\mathbf{x})).
\]

**Theorem 4.1** [Ismail 2000b]. Assume \( w(x) > 0 \), \( x \in (a, b) \) and let \( v(x) \) of (4-1) and \( v(x) + \ln A_N(x) \) be twice continuously differentiable functions whose second derivative is nonnegative on \((a, b)\). Then the equilibrium position of \( N \) movable unit charges in \([a, b]\) in the presence of the external potential \( V(x) \) of (4-11) is unique and attained at the zeros of \( p_N(x) \), provided that the particle interaction obeys a logarithmic potential and that \( T_N(\mathbf{x}) \to 0 \) as \( \mathbf{x} \) tends to any boundary point of \([a, b]^N\), where

\[
T_N(\mathbf{x}) = \left( \prod_{j=1}^{N} \frac{\exp(-v(x_j))}{A_N(x_j)/A_N} \right) \prod_{1 \leq l < k \leq N} (x_l - x_k)^2.
\]

**Proof.** Since \( T_N \) is symmetric in \( x_1, \ldots, x_N \) and \( T_N(\mathbf{x}) \) vanishes when two of the \( x \)'s coincide, we may assume that

\[
x_1 > x_2 > \cdots > x_N.
\]  

(4-12)

The assumption \( v''(x) > 0 \) ensures the positivity of \( A_N(x) \). To find an equilibrium position we solve

\[
\frac{\partial}{\partial x_j} \ln T_N(\mathbf{x}) = 0, \quad j = 1, 2, \ldots, N.
\]

This system is

\[
-v'(x_j) - \frac{A_N'(x_j)}{A_N(x_j)} + 2 \sum_{1 \leq k \leq N} \frac{1}{x_j - x_k} = 0, \quad j = 1, 2, \ldots, N.
\]  

(4-13)

The system of equations (4-13) is nonlinear in the unknowns \( x_1, \ldots, x_N \). To change this to a linear system we set

\[
f(\mathbf{x}) := \prod_{j=1}^{N} (x - x_j),
\]
and turn the system \((4-13)\) to a differential equality in \(f(x)\) satisfied at the points \(x_j, 1 \leq j \leq N\). To see this, first observe that

\[
\sum_{1 \leq k \leq N \atop k \neq j} \frac{1}{x_j - x_k} = \lim_{x \to x_j} \left( \frac{f'(x)}{f(x)} - \frac{1}{x - x_j} \right) = \lim_{x \to x_j} \left( \frac{(x - x_j)f'(x) - f(x)}{(x - x_j)f(x)} \right),
\]

which implies, via L’Hôpital’s rule,

\[
2 \sum_{1 \leq k \leq N \atop k \neq j} \frac{1}{x_j - x_k} = \frac{f(x_j)}{f'(x_j)}.
\]

Now this and \((4-13)\)

\[-v'(x_j) - \frac{A'(x_j)}{A(x_j)} + \frac{f''(x_j)}{f'(x_j)} = 0,
\]

or equivalently

\[f''(x) + R_N(x)f'(x) = 0, \quad x = x_1, \ldots, x_N,
\]

with \(R_N\) as in \((4-7)\). In other words

\[f''(x) + R_N(x)f'(x) + S_N(x)f(x) = 0, \quad x = x_1, \ldots, x_N. \quad (4-14)
\]

To check for local maxima and minima consider the Hessian matrix

\[H = (h_{ij}), \quad h_{ij} = \frac{\partial^2 \ln T_N(x)}{\partial x_i \partial x_j}. \quad (4-15)
\]

It readily follows that

\[
h_{ij} = 2(x_i - x_j)^{-2}, \quad i \neq j,
\]

\[
h_{ii} = -v'(x_i) - \frac{\partial}{\partial x_i} \left( \frac{A'(x_i)}{A(x_i)} \right) - 2 \sum_{1 \leq j \leq N \atop j \neq i} \frac{1}{(x_i - x_j)^2}.
\]

This shows that the matrix \(-H\) is positive definite because it is real, symmetric, strictly diagonally dominant and its diagonal terms are positive [Horn and Johnson 1992, Cor. 7.2.2]. Therefore, \(\ln T_N(x)\) has no relative minima nor saddle points. Thus any solution of \((4-14)\) will provide a local maximum of \(\ln T_N(x)\) or \(T_N(x)\). There cannot be more than one local maximum since \(T_N(x) \to 0\) as \(x \to \) any boundary point along a path in the region defined in \((4-12)\). Thus the system \((4-14)\) has at most one solution. On the other hand, \((4-6)-(4-8)\) show that the zeros of

\[f(x) = a_1 a_2 \ldots a_N p_N(x),
\]

satisfy \((4-14)\), hence the zeros of \(p_N(x)\) solve \((4-14)\). This completes the proof of Theorem 4.1. \(\square\)
Observe that the convexity of \( v \) in Theorem 4.1 can be replaced by requiring that \( A_n(x) > 0 \) for \( a < x < b \).

The next result is a generalization of the Stieltjes–Hilbert evaluation of the discriminant of Jacobi polynomials (2–3). Observe that (4–5) is exactly the assumption (2–8) of Theorem 2.3. Thus we have established the following result.

**Theorem 4.2** [Ismail 1998]. Let \( \{p_N(x)\} \) be orthonormal polynomials and let \( \{a_N\} \) and \( \{b_N\} \) be the recursion coefficients in (4–2). Let

\[
x_{1,N} > x_{2,N} > \cdots > x_{N,N},
\]

be the zeros of \( p_N(x) \). Then the discriminant of \( \{p_N(x)\} \) is given by

\[
D(p_N) = \left( \prod_{j=1}^{N} \frac{A_N(x_j,N)}{a_N} \right) \left( \prod_{k=1}^{N} a_k^{2k-2N+2} \right).
\]

We next give a representation for the maximum value of \( T_N(x) \) or the minimum value of \( E_N(x) \) in terms of the recursion coefficients \( \{a_n\} \).

**Theorem 4.3.** Let \( T_N \) and \( E_N \) be the maximum value of \( T_N(x) \) and the equilibrium energy of the \( N \) particle system in Theorem 4.1. Then

\[
T_N = \left( \prod_{j=1}^{N} w(x_j,N) \right) \left( \prod_{k=1}^{N} a_k^{2k} \right), \quad E_N = \sum_{j=1}^{n} v(x_j,N) - 2 \sum_{j=1}^{N} j \ln a_j.
\]

This follows from Theorems 4.1 and 4.2.

Theorem 4.3 extends the Stieltjes results from Jacobi polynomials to general orthogonal polynomials, so it would be of interest to explore the analogue of the Selberg integral. This means replace the \( L_\infty \) norm in Theorem 4.1 by the \( L_p \) norm and evaluate the integral

\[
\int_{[a,b]^N} \left( \prod_{j=1}^{N} \frac{\exp(-v(t_j))}{A_N(t_j)/a_N} \right)^p \prod_{1 \leq i < k \leq N} (t_i - t_k)^{2p} dt_1 \cdots dt_N.
\]

Chihara [1985] studied orthogonal polynomials which result from modifying the orthogonality measure of a given set orthogonal by adding a one-point mass at the end of the spectral interval. He also considered specific cases of adding two point masses in certain special cases. Chihara’s construction uses the kernel polynomials. Recently, Kiesel and Wimp [1996] found a different approach which avoids the use of kernel polynomials and in their later work [Kiesel and Wimp 1996], applied their results to the Koornwinder polynomials [Koornwinder 1984]. They derived closed form expressions for the recursion coefficients and the coefficients in the differential equation satisfied by the Koornwinder polynomials.

Grünbaum [1998] described an electrostatic interpretation for the zeros of the Koornwinder polynomials [Koornwinder 1984], which are orthogonal on \([-1,1]\) with respect to measure with the absolutely continuous component \((1 - x)^a \times (1 + x)^b\) on \([-1,1]\) and two discrete masses at \( \pm 1 \). This motivated us to write
[Ismail 2000c] where we derived second order differential equations for general polynomials orthogonal with respect a measure with a non-trivial absolutely continuous part supported on an interval and a finite discrete part outside the interval. We also extended the electrostatic models of Stieltjes and Hilbert to polynomials orthogonal with respect to a weight function supported on an interval \([a, b]\) with at most two discrete mass points at the finite end points of the interval. So far the only interesting example of this type is the Koornwinder polynomials and their special cases. Kiesel and Wimp studied the Kornwinder polynomials extensively in [Kiesel and Wimp 1996; Wimp and Kiesel 1996], Grünbaum \([\geq 2001]\), however, continued his research employing the Darboux transformation and described electrostatic models where the movable charges are restricted to an interval but the external field is generated by fixed charges in the plane.

5. Generalized and Quantized Discriminants

Motivated by (1–5) we were led in [Ismail 2000a] to define discriminants associated with linear operators, which reduce to (1–3) when the linear operator is the derivative operator.

**Definition 5.1.** Let \(T\) be a linear degree reducing operator, that is \((Tf)(x)\) is a polynomial of exact degree \(n - 1\) whenever \(f\) has precise degree \(n\) and the leading terms in \(f\) and \(Tf\) have the same sign. We define the (generalized) discriminant relative to \(T\) by

\[
D(g_n, T) := (-1)^{n(n-1)/2} \gamma_n^{-1} R\{g_n, Tg_n\} = (-1)^{n(n-1)/2} \gamma_n^{-2} \prod_{j=1}^{n} (Tg_n)(x_j),
\]

for \(g_n\) as in (1–3), and \(R\{g_n, f_m\}\) is the resultant

\[
R\{g_n, f_m\} := \gamma_n^m \prod_{j=1}^{n} f_m(x_j), \quad \text{for} \quad f_m(x) = \alpha_m x^m + \text{lower order terms}.
\]

In this set up Theorem 2.3 becomes

**Theorem 5.2.** Let \(T\) be a linear degree reducing operator and assume that \(\{p_n(x)\}\) is a system of polynomials generated by (2–4) and (2–5) and assume that the zeros of \(p_n(x)\) be arranged as in (2–6). If \(\{p_n(x)\}\) satisfies the functional recurrence relation

\[
T p_n(x) = A_n(x) p_{n-1}(x) - B_n(x) p_n(x), \quad (5-1)
\]

then the discriminant of \(p_n(x)\) relative to \(T\) is given by

\[
D(p_n) = \left( \prod_{j=1}^{n} A_n(x, j, n) \right) \prod_{k=1}^{n} \xi_k^{2n-2k-1} \xi_k^{-1}.
\]
For particular $T$ the above definition provides a quantization of the concept of a discriminant and leads to what we call $q$-discriminants which correspond to the case $T = D_q$,

\[
(D_q f)(x) = \frac{f(x) - f(qx)}{x - qx}.
\]

The operator $D_q$ is called the $q$-difference operator [Andrews et al. 1999; Gasper and Rahman 1990]. Furthermore, $D_q f \to f'$ as $q \to 1$, for differentiable functions $f$. An easy calculation gives

\[
D(g_n; q) := D(g_n, D_q) = \gamma^{2n-2}q^n(n-1)/2 \prod_{1 \leq i < j \leq n} (q^{1/2}x_i - q^{1/2}x_j)(q^{1/2}x_i - q^{-1/2}x_j).
\]

In other words,

\[
D(g_n; q) := \gamma^{2n-2}q^n(n-1)/2 \prod_{1 \leq i < j \leq n} (x_i^2 + x_j^2 - (q^{-1} + q)x_i x_j).
\]

The representations (1–4) and (5–2) reaffirm the fact that $D(g_n; q) \to D(g_n)$ as $q \to 1$. We shall refer to $D(g_n; q)$ as the quantized discriminant. For $n = 2$, that is $g(x) = Ax^2 + Bx + C$, the quantized discriminant is $qB^2 - (1 + q)^2AC$, to be contrasted with the usual discriminant $B^2 - 4AC$.

In [Ismail 2000a] we extended many of the results of in §4 to polynomials orthogonal with respect to a discrete measure whose masses are at the union of at most two geometric progressions. To state these results we need the concept of a $q$-integral [Andrews et al. 1999; Gasper and Rahman 1990]:

\[
\int_a^b f(x)q(x) := b(1 - q) \sum_{n=0}^{\infty} q^n f(bq^n) - a(1 - q) \sum_{n=0}^{\infty} q^n f(aq^n),
\]

with

\[
\int_0^\infty f(x)q(x) := (1 - q) \sum_{n=0}^{\infty} q^n f(q^n).
\]

The orthogonality relation of discrete $q$-orthonormal polynomials is of the form

\[
\int_a^b p_m(x)p_n(x)w(x)q(x)dx = \delta_{m,n}.
\]

In [Ismail 2000a] we proved the following extension of (4–3)–(4–5) to discrete $q$-orthonormal polynomials.

**Theorem 5.3** [Ismail 2000a]. Let $\{p_n(x)\}$ be a sequence of discrete $q$-orthonormal polynomials satisfying equalities (4–2). Then they have a lowering (annihilation) operator of the form

\[
D_q p_n(x) = A_n(x)p_{n-1}(x) - B_n(x)p_n(x),
\]

(5–5)
where $A_n(x)$ and $B_n(x)$ are given by

\[ A_n(x) = a_n \frac{w(y/q)p_n(y)p_n(y/q)}{x - y/q} \int_a^b + a_n \int_a^b \frac{u(qy) - u(y)}{qy - y} p_n(y) p_n(y/q) w(y) \, dqy, \]

and

\[ B_n(x) = a_n \frac{w(y/q)p_n(y)p_{n-1}(y/q)}{x - y/q} \int_a^b + a_n \int_a^b \frac{u(qy) - u(y)}{qy - y} p_n(y)p_{n-1}(y/q) w(y) \, dqy, \]

where $u$ is defined by

\[ D_q w(x) = -u(qx) w(qx). \]

Furthermore the corresponding raising operator is

\[ \frac{x - b_n}{a_n} A_n(x) p_n(x) - B_n(x) p_n(x) - D_q p_n(x) = \frac{a_{n+1}}{a_n} A_n(x) p_{n+1}(x). \]

We used (5–4) and (5–5) to evaluate the functions the functions $A_n(x)$ and $B_n(x)$ explicitly in [Ismail 2000a]. In [Ismail 2000a] we also computed the quantized discriminants of the big $q$-Jacobi polynomials. It is interesting to note that by experimentation we found that the usual discriminants for these polynomials for degrees 5 and 6 do not factor nicely.

We now proceed to study the generalized discriminants when $T$ is the Askey–Wilson operator. Given a polynomial $f$ we set $f(e^{i\theta}) := f(x)$ for $x = \cos \theta$; that is,

\[ f(z) = f((z + 1/z)/2), \quad z = e^{i\theta}. \]

In other words, we think of $f(\cos \theta)$ as a function of $e^{i\theta}$. In this notation, the Askey–Wilson divided difference operator $D_q$ is defined by

\[ (D_q f)(x) := \frac{\tilde{f}(q^{1/2} e^{i\theta}) - \tilde{f}(q^{-1/2} e^{i\theta})}{e^{i(1/2) q e^{i\theta}) - e^{i(-1/2) q e^{i\theta})}, \quad x = \cos \theta, \quad (5-6) \]

with

\[ e(x) = x. \]

A calculation reduces (5–6) to

\[ (D_q f)(x) = \frac{\tilde{f}(q^{1/2} e^{i\theta}) - \tilde{f}(q^{-1/2} e^{i\theta})}{(q^{1/2} - q^{-1/2}) \sin \theta}, \quad x = \cos \theta. \]

It is not difficult to see that

\[ \lim_{q \to 1} (D_q f)(x) = f'(x). \]

It is not difficult to compute the action of Askey–Wilson operator on the Chebyshev polynomials of the first kind, $T_n(x)$,

\[ T_n(\cos \theta) = \cos(n\theta). \quad (5-7) \]
Indeed,
\[ D_q T_n(x) = \frac{(q^{n/2} - q^{-n/2})}{(q^{1/2} - q^{-1/2})} U_{n-1}(x). \]  

(5-8)

In [Ismail 2000a] we showed how that the appropriate generalized discriminant for the continuous \( q \)-Hermite and continuous \( q \)-Jacobi polynomials is \( D(f, D_q) \) and we applied Theorem 5.2 because (5-1) is known when \( T = D_q \) and \( p_n(x) \) is any of the above mentioned polynomials.

It may be of interest to compute \( D(f, D_q) \) for \( f(x) = Ax^2 + Bx + C \) and compare it with the familiar \( B^2 - 4AC \). Since, with \( x = \cos \theta \),
\[ A x^2 + B x + C = \frac{A}{2} \cos(2\theta) + B \cos \theta + C + \frac{A}{2}, \]
then (5-7) and (5-8) give
\[ (D_q f)(x) = \frac{A(q - q^{-1})}{q^{1/2} - q^{-1/2}} x + B, \]
so that \( D(f, D_q) \) is \((q^{1/2} + q^{-1/2} - 1) B^2 - (q^{1/2} + q^{-1/2})^2 AC \).

One advantage of visiting an institute like MSRI is getting the opportunity to meet very bright young mathematicians with different backgrounds. I had the good fortune of meeting Naihuan Jing and talking to him about the different discriminants I encountered. He found interpretation of \( D(f) = D(f, \frac{d}{dx}) \), \( D(f; D_q) \) and \( D(f; \Delta_h) \) as expectation values of vertex operators, where
\[ (\Delta_h f)(x) := f(x + h) - f(x). \]

Our conversations led to [Ismail and Jing \( \geq 2001 \)].

Acknowledgments

The author is very grateful to Drs. B. H. Yoo and G. J. Yoon for their careful proofreading of this paper. We gratefully acknowledge partial research support of MSRI and the Department of Mathematics at the University of California, Berkeley.

I thank Denise Marks for the excellent job she did in putting this paper together.

References


MOURAD E. H. ISMAIL
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH FLORIDA
Tampa, FL 33620-5700
UNITED STATES
ismail@math.usf.edu