Orthogonal Polynomials

1. A little history & introduction.

Murphy [1835] first defined orthogonal functions.

The field of orthogonal polynomials was developed by (great contributions) P.L. Chebyshev. Pursued by A.A. Markov & T.J. Stieltjes. Many mathematicians made their contributions to the area.

Mainly in the late 19^{th} century.

Fourier Series
continued fractions $\xrightarrow{\text{Motivate}}$ Orthogonal Polynomials
approximation theory

$\Big\downarrow$ related to

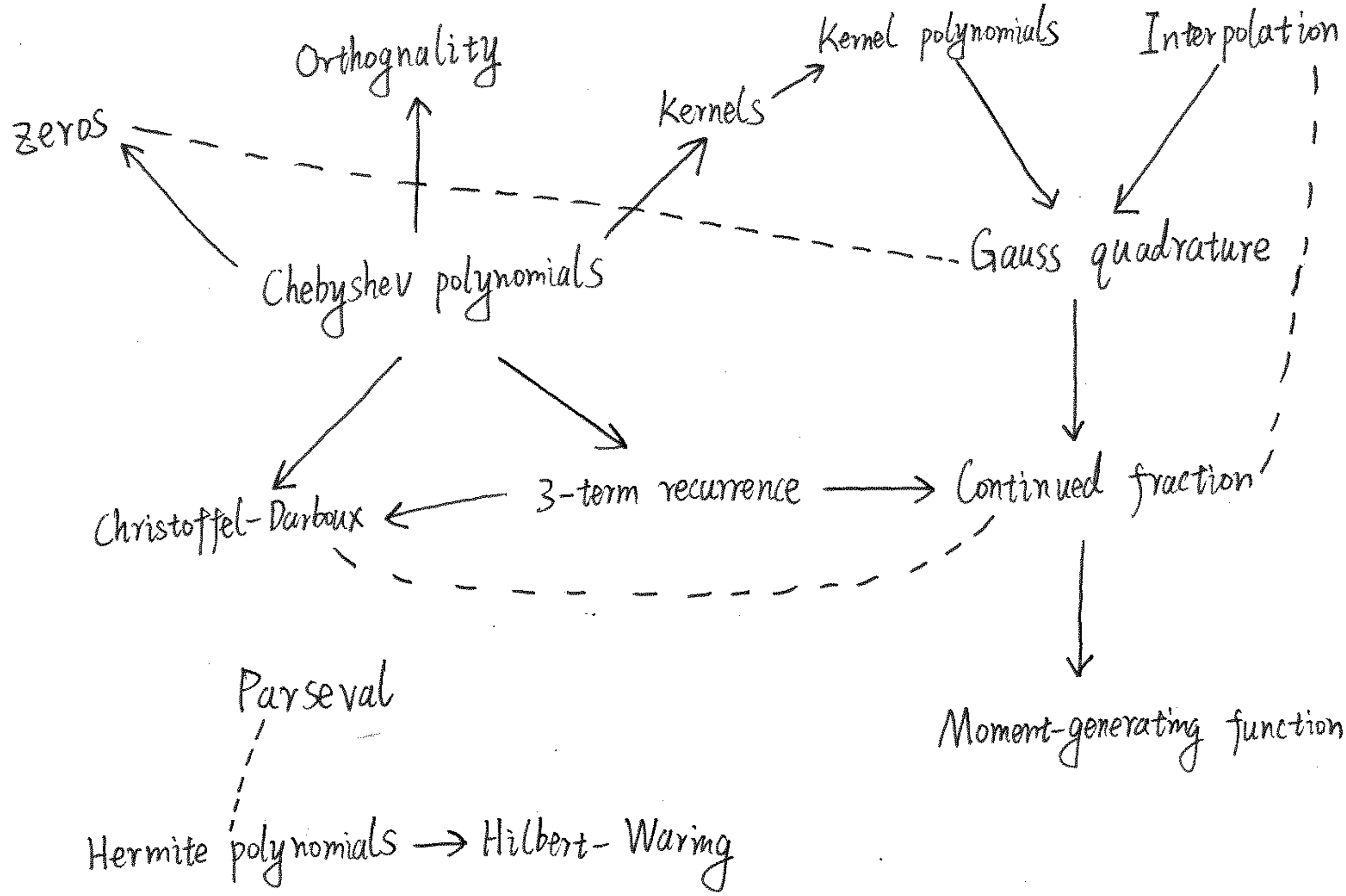
trigonometric, hypergeometric, Bessel, elliptic

continued fractions, interpolation, mechanical quadrature

quantum mechanics, mathematical statistics

And much more ...

2. The general structure of the presentation



5.1 Chebyshev Polynomials

Def. The Chebyshev Polynomials of the first and second kinds, denoted respectively by $T_n(x)$ and $U_n(x)$, are defined by the formulas

$$P_n^{(-\frac{1}{2},-\frac{1}{2})}(x) = \frac{(2n)!}{2^{2n}(n!)^2} T_n(x) = \frac{(2n)!}{2^{2n}(n!)^2} \cos n\theta \qquad (1.1)$$

and

$$P_n^{(\frac{1}{2},\frac{1}{2})}(x) = \frac{(2n+2)!}{2^{2n+1}[(n+1)!]^2} U_n(x) = \frac{(2n+2)!}{2^{2n+1}[(n+1)!]^2} \frac{\sin(n+1)\theta}{\sin\theta} \qquad (1.2)$$

where $x = \cos\theta$.

Recall: The Jacobi polynomial of degree n is defined by

$$P_n^{(\alpha,\beta)}(x) := \frac{(\alpha+1)^{(n)}}{n!} \, {}_2F_1\left(\begin{matrix} -n, \; n+\alpha+\beta+1 \\ \alpha+1 \end{matrix} ; \frac{1-x}{2}\right)$$

And we have $\displaystyle\int_{-1}^{+1} P_n^{(\alpha,\beta)}(x) P_m^{(\alpha,\beta)}(x) (1-x)^{\alpha}(1-x)^{\beta} dx =$

$$\frac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)\, n!} \delta_{mn},$$

so Jacobi Polynomials themselves are also orthogonal.

We have the elementary results:

$$\int_0^{\pi} \cos m\theta \cos n\theta \, d\theta = 0 \quad (m \neq n) \quad \overset{x=\cos\theta}{\Longleftrightarrow} \quad \int_{-1}^{1} T_n(x) T_m(x) \frac{dx}{\sqrt{1-x^2}} = 0$$

$$\int_0^{\pi} \sin(n+1)\theta \sin(m+1)\theta \, d\theta = 0 \quad (m \neq n) \quad \overset{x=\cos\theta}{\Longleftrightarrow} \quad \int_{-1}^{1} U_n(x) U_m(x) \sqrt{1-x^2} \, dx = 0$$

The orthogonality of $T_n(x)$, $U_n(x)$ are contained in trigometric functions.

3-term recurrence relation of Chebyshev polynomials

$$2xT_m(x) = T_{m+1}(x) + T_{m-1}(x) \quad (1.3) \Leftarrow T_m(x)T_n(x) = \frac{1}{2}(T_{m+n}(x) + T_{m-n}(x)) \quad (1.6)$$

$$\Updownarrow \qquad\qquad\qquad\qquad \Updownarrow$$

$$2\cos\theta\cos m\theta = \cos(m+1)\theta + \cos(m-1)\theta \ (1.4) \Leftarrow 2\cos m\theta\cos n\theta = \cos(m+n)\theta + \cos(m-n)\theta \quad (1.5)$$

More generally, we want to determine coefficients $a(k,m,n)$ in

$$P_m(x)P_n(x) = \sum_{k=0}^{m+n} a(k,m,n)P_k(x) \qquad (1.7)$$

where $\{P_n(x)\}$ is a sequence of polynomials with $\deg(P_n(x)) = n$. (For example, $x^m x^n = x^{m+n}$). In general this is very hard.

For $U_n(x)$:
$$\frac{\sin(m+1)\theta}{\sin\theta}\,\frac{\sin(n+1)\theta}{\sin\theta} = \sum_{k=0}^{m\wedge n} \frac{\sin(m+n+1-2k)\theta}{\sin\theta} \qquad (1.8)$$

dual $\updownarrow$
$$\sin(n+1)\theta\,\sin(n+1)\varphi = \frac{n+1}{2}\int_{\theta-\varphi}^{\theta+\varphi} \sin(n+1)\psi\, d\psi$$

where as dual of (1.5) is essentially the same:

$$(1.5) \xleftrightarrow{\text{dual}} \cos n\theta\cos n\varphi = \frac{1}{2}(\cos n(\theta+\varphi) + \cos n(\theta-\varphi)) \qquad (1.9)$$

The kernels, and a preview of Christoffel-Darboux formula

In Fourier analysis, we analyze partial sums of the form

$$\frac{1}{2}a_0 + \sum_{i=1}^{n}(a_i\cos i\theta + b_i\sin i\theta),$$ where

$$a_i = \frac{1}{\pi}\int_0^{2\pi} f(\varphi)\cos i\varphi\, d\varphi \quad \text{and} \quad b_i = \frac{1}{\pi}\int_0^{2\pi} f(\varphi)\sin i\varphi\, d\varphi.$$

Therefore
$$\frac{1}{2}a_0 + \sum_{i=1}^{n}(a_i\cos i\theta + b_i\sin i\theta)$$
$$= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{1}{2} + \sum_{i=1}^{n}(\cos i\theta\cos i\varphi + \sin i\theta\sin i\varphi)\right] f(\varphi)\,d\varphi$$
$$= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{1}{2} + \sum_{i=1}^{n}\cos i(\varphi-\theta)\right] f(\varphi)\,d\varphi$$

it can be verified that:

$$\frac{1}{2}+\sum_{i=1}^{n}\cos i\theta = \frac{\sin(n+\frac{1}{2})\theta}{2\sin\frac{\theta}{2}} =: D_n(\theta) \qquad (1.10)$$

Thus, the sum inside the last integral is $D_n(\theta-\varphi)$.

The function $D_n(\theta)$ is called the dirichlet kernel.

Define Chebyshev polynomials of the third kind by

$$V_n(x) = 2D_n(\theta), \text{ where } x=\cos\theta \qquad (1.11)$$

we have:

$$\int_{-1}^{1} V_n(x)V_m(x)\sqrt{\frac{1-x}{1+x}}\,dx = 0.$$

A generalization of (1.10) is:

$$1+\sum_{i=1}^{n} 2\cos i\theta\cos i\varphi = \frac{\cos(n+1)\theta\cos n\varphi - \cos n\theta\cos(n+1)\varphi}{\cos\theta-\cos\varphi} \qquad (1.12)$$

This is a special case of Christoffel-Darboux formula.

To see why (1.12) holds, we see

$$(1.4) \Rightarrow 2\cos m\theta\cos m\varphi(\cos\theta-\cos\varphi) = [\cos(m+1)\theta+\cos(m-1)\theta]\cos m\varphi$$
$$-[\cos(m+1)\varphi+\cos(m-1)\varphi]\cos m\theta. \qquad (1.13).$$

Sum this over m $(1\le m\le n)$.

To understand the factor 2 in (1.12), note that

$$\int_0^\pi \cos^2 m\theta\,d\theta = \begin{cases} \frac{\pi}{2}, & m\neq 0 \\ \pi, & m=0 \end{cases}$$

Thus, the normalized function is $\sqrt{\frac{2}{\pi}}\cos m\theta$ $(m\neq 0,\ m=0 \Rightarrow \frac{1}{\sqrt{\pi}})$.

The Poisson kernel for the Chebyshev polynomials defined by $\cos nx$ is given by:

$$1+\sum_{i=1}^{\infty}(2\cos i\theta\cos i\varphi)r^i =: P_r(\cos\theta,\cos\varphi).$$

$$\varphi=0 \Rightarrow 1+\sum_{i=1}^{\infty} 2\cos i\theta\, r^i = \frac{1-r^2}{1-2r\cos\theta+r^2} \overset{|r|<1}{>} 0.$$

$$\left.\begin{array}{l} P_r(\cos\theta,1)>0,\ \forall |r|<1. \\ 2\cos m\theta\cos m\varphi = \cos m(\theta+\varphi)+\cos m(\theta-\varphi) \end{array}\right\} \Rightarrow P_r(\cos\theta,\cos\varphi)>0,\ \forall |r|<1$$

Zeros of $T_m(x)$

$\cos m\theta = 0$ when $\theta = \frac{2n+1}{2m}\pi \Rightarrow T_m(x)=0$ when $x = \cos\frac{2n+1}{2m}\pi$, for $n \in [0, m-1]$, $n \in \mathbb{Z}$. Moreover, the zeros of $T_m(x)$ and $T_{m+1}(x)$ mutally separate each other. Also between two successive zeros of $T_m(x)$, $\cos\frac{2k+3}{2m}\pi$ and $\cos\frac{2k+1}{2m}\pi$, there is a zero of $T_n(x)$ for $n>m$. To see why, notice we can always find a integer $0 \le l$, $l \le n-1$ s.t. $\frac{2k+1}{2m} < \frac{2l+1}{2n} < \frac{2k+3}{2m}$.

All the properties above about Chebyshev polynomials have their generalizations to general orthogonal polynomials.

5.2 Recurrence

Let $\alpha(x)$ denote a nondecreasing function with an infinite number of points of increase in the interval $[a,b]$. The latter interval may be infinite. Assume moments of all orders exist, i.e.

$$\int_a^b x^n \, d\alpha(x) \text{ exists } \forall n \ge 0, n \in \mathbb{Z}.$$

Def. We say that a sequence of polynomials $\{P_n(x)\}_0^\infty$, where $\deg(P_n(x)) = n$, is orthogonal with respect to the distribution $d\alpha(x)$ if

$$\int_a^b P_n(x) P_m(x) \, d\alpha(x) = h_n \delta_{mn} \quad \forall m,n. \qquad (2.1)$$

Thm 2.2 (3-term recurrence relation): A sequence of orthogonal polynomials $\{P_n(x)\}$ satisfies:

$$P_{n+1}(x) = (A_n x + B_n) P_n(x) - C_n P_{n-1}(x), \quad n \ge 0 \qquad (2.2)$$

where $P_{-1}(x) = 0$. Here $\forall n \ge 0$, $A_n, B_n, C_n \in \mathbb{R}$; $A_{n-1} A_n C_n > 0 \; \forall n \ge 1$.

If $[x^n] P_n(x) = k_n > 0$, then $A_n = \frac{k_{n+1}}{k_n}$, $C_{n+1} = \frac{A_{n+1}}{A_n} \frac{h_{n+1}}{h_n}$ (h_n is given by (2.1))

Remark: In general, $A_n > 0 \wedge C_n > 0$. Also, $C_{n+1} = \frac{A_{n+1}}{A_n} = \frac{k_{n+1}k_{n-1}}{k_n^2}$ if $\{P_n(x)\}$ are normalized.

Proof: First determine A_n so that $P_{n+1}(x) - A_n x P_n(x)$ is a polynomial of degree n. Then we can write

$$P_{n+1}(x) - A_n x P_n(x) = \sum_{k=0}^{n} b_k P_k(x) \quad \text{for some } b_k \in \mathbb{R}. \quad (2.3)$$

$Q(x) = P_m(x)$, $m<n$, by (2.1), we have

$$\int_a^b P_n(x) Q(x)\, d\alpha(x) = 0. \quad \underset{\deg(Q(x))=m<n}{\overset{\forall Q(x)}{\Longrightarrow}} \quad \int_a^b P_n(x) Q(x)\, d\alpha(x) = 0.$$

Since we can write $Q(x)$ in terms of $P_k(x$

$\forall k < n-1$, multiply (2.3) by $P_k(x)$ and integrate, we get:

$$b_k \int_a^b P_k^2(x)\, d\alpha(x) = 0 \Rightarrow b_k = 0$$

This proves (2.2). It is clear that $A_n = \frac{k_{n+1}}{k_n}$.

To derive C_n, multiply (2.2) by $P_{n-1}(x)$ and integrate,

we get $\quad 0 = A_n \int_a^b P_n(x)\, x P_{n-1}(x)\, d\alpha(x) - C_n \int_a^b P_{n-1}^2(x)\, d\alpha(x)$

write $\quad x P_{n-1}(x) = \frac{k_{n-1}}{k_n} P_n(x) + \sum_{k=0}^{n-1} d_k P_k(x)$, we get

$$\frac{A_n}{A_{n-1}} h_n - C_n h_{n-1} = 0 \Leftrightarrow C_n = \frac{A_n}{A_{n-1}} \frac{h_n}{h_{n-1}} \qquad \blacksquare$$

Remark. The converse of Thm 2.2 still holds. (i.e. If a sequence of polynomials satisfy a 3-term recurrence relation as described, then they are orthogonal with respect to a positive measure) This is usually called Favard's theorem (though it is used by Stieltjes in the theory of continued fractions many years before Favard's paper).

Coro: $h_n = \frac{A_{n-1} C_n h_{n-1}}{A_n} = \frac{A_{n-1} C_n}{A_n} \frac{A_{n-2} C_{n-1} h_{n-2}}{A_{n-1}} = \cdots = \frac{A_0}{A_n} C_1 C_2 \cdots C_n h_0$

This shows that the L^2 norm (i.e. h_n) can be computed from the recurrence.

An important consequence of the recurrence relation in Thm 2.2 is the Christoffel–Darboux formula.

Thm 2.4 (Christoffel–Darboux formula): Suppose that the $P_n(x)$ are normalized so that $h_n = \int_a^b P_n^2(x)\, d\alpha(x) = 1$. Then

(recall $k_n = [x^n]P_n(x)$)
$$\sum_{m=0}^{n} P_m(y)P_m(x) = \frac{k_n}{k_{n+1}} \frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{x-y} \quad (2.4)$$

Proof: Use Thm 2.2, we obtain:
$$P_n(y)P_{n+1}(x) = (A_n x + B_n)P_n(x)P_n(y) - C_n P_{n-1}(x)P_n(y)$$
and
$$P_n(x)P_{n+1}(y) = (A_n y + B_n)P_n(y)P_n(x) - C_n P_{n-1}(y)P_n(x)$$
subtract and divide by $A_n(x-y)$, we get:
$$\frac{1}{A_n}\frac{P_n(y)P_{n+1}(x) - P_n(x)P_{n+1}(y)}{x-y} = P_n(x)P_n(y) + \frac{1}{A_{n-1}}\frac{P_{n-1}(y)P_n(x) - P_{n-1}(x)P_n(y)}{x-y} \quad (2.5)$$
(Notice we used $C_n = \frac{A_n}{A_{n-1}}$ as our polynomials are normalized)

Now sum (2.5) over n ($0 \le n \le m$), we get (2.4) ▨

Remark: If h_n are not normalized, then we have
$$\sum_{m=0}^{n} \frac{P_m(y)P_m(x)}{h_m} = \frac{k_n}{k_{n+1}} \frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{(x-y)h_n}$$
And if we write RHS of (2.4) as
$$\frac{k_n}{k_{n+1}} \frac{(P_{n+1}(x) - P_{n+1}(y))P_n(y) - (P_n(x) - P_n(y))P_{n+1}(y)}{x-y}$$
let $x-y \to 0$, we get the confluent form of (2.4).

Thm 2.5: if $h_n = 1$, then
$$\sum_{k=0}^{n} P_k^2(x) = \frac{k_n}{k_{n+1}} (P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x))$$

Coro: $P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x) > 0, \forall x$.

The three-term recurrence relation for Jacobi polynomials.

Consider polynomials $P_n(x) = \frac{n!}{(\alpha+1)^{(n)}} P_n^{(\alpha,\beta)}(x) = {}_2F_1\left(\begin{matrix} -n, n+\alpha+\beta+1 \\ \alpha+1 \end{matrix}; \frac{1-x}{2}\right)$ for $n \geq 0$, Write the recurrence relation as

$$(1-x)P_n(x) = A_n P_{n+1}(x) + B_n P_n(x) + C_n P_{n-1}(x), \quad n \geq 0, \text{ where } P_{-1}(x)=0$$

Equate the coefficients of $(1-x)^{n+1}$ to obtain A_n.

Take $x=1$, we get $0 = A_n + B_n + C_n$.

And the L^2 norm of Jacobi polynomials are given before.

So $C_n = \frac{A_n}{A_{n-1}} \frac{h_n}{h_{n-1}}$ is found, also $B_n = -(A_n + C_n)$.

Note that we have also proved the orthognality of Jacobi polynomials by Favard's theorem.

5.3 Gauss Quadrature

Suppose $x_1 < x_2 < \cdots < x_n$ is a set of n numbers in an increasing sequence and $y_1, y_2, \cdots, y_n$ is an arbitrary set of numbers.

Def. The Lagrange interpolation polynomial is a polynomial of degree $n-1$ that takes the value y_i at x_i for $i=1,\cdots,n$. This polynomial is given by

$$L_n(x) = \sum_{j=1}^{n} \frac{P(x)\, y_j}{P'(x_j)(x-x_j)} \tag{3.1}$$

where $P(x) = (x-x_1)\cdots(x-x_n)$. We write

$$l_j(x) := \frac{P(x)}{P'(x_j)(x-x_j)}, \quad \text{for } j \in [n] \tag{3.2}$$

It is clear that $l_j(x_k) = \delta_{jk}$

Thus if $f(x)$ is a continuous function whose values $f(x_i)$ are known at the points $x_i, i \in [n]$ in an interval $[a,b]$, then

$$L_n(x) = \sum_{j=1}^{n} l_j(x) f(x_j) \tag{3.3}$$

is a polynomial of degree $\leq n-1$ which interpolates the function f in $[a,b]$. Formula (3.3) can be applied to approximate integration.

We have
$$\int_a^b f(x)\,d\alpha(x) \approx \sum_{j=1}^{n} f(x_j)\int_a^b l_j(x)\,d\alpha(x) = \sum_{j=1}^{n} \lambda_j f(x_j), \qquad (3.4)$$
where $\lambda_j := \int_a^b l_j(x)\,d\alpha(x)$. (3.5) Notice that if $f(x)$ is a polynomial of degree $\le n-1$, $L_n(x) = f(x)$.

There are a number of ways to measure how well the quadrature method approximates the integral.

The most obvious way is to see how large the difference is.

Another way that is very fruitful: Require that the quadrature method to be exact for as large a class of functions as possible.

For Lagrange's interpolation, when x_k are given in advance, two functions are identical when f is a polynomial of degree at most $n-1$.

However, if we do not require the points x_k to be fixed, there is a possibility of increasing the degree of the polynomial by one for each x_k, which is allowed to vary. In this case, the maximum degree can be increased to $2n-1$.

This seems to be hard, for we seek the solution of the system
$$\sum_{k=1}^{n} \lambda_k x_k^j = \int_a^b x^j\,d\alpha(x), \quad j = 0, 1, \cdots, 2n-1.$$

The solution is contained in the Gauss quadrature formula.

Before stating it, we introduce some notation.

Suppose $\{P_n(x)\}$ is a sequence of orthogonal polynomials, associated with distribution $d\alpha(x)$ and interval $[a,b]$, then we denote the zeros of $P_n(x)$ by $x_j = x_{j,n} = x_{jn}$, $j \in [n]$. We shall prove later that these zeros are all simple and lie in the interval $[a,b]$.

Thm 3.2 (Gauss quadrature formula): There are positive numbers $\lambda_1, \lambda_2, \cdots, \lambda_n$ such that for every polynomial $f(x)$ of degree at most $2n-1$, we have

$$\int_a^b f(x)\,d\alpha(x)=\sum_{j=1}^{n}\lambda_j f(x_j) \qquad (3.7)$$

where x_j, $j\in[n]$ are defined above, and $\lambda_j=\lambda_{j,n}=\lambda_{jn}$.

Proof: Let $f(x)$ be an polynomial of any degree, $\{P_n(x)\}$ be a sequence of orthogonal polynomials, then

$$f(x)=P_n(x)Q(x)+R(x) \text{ for some } Q, R \text{ and } \deg R\le n-1$$

x_j are zeros of $P_n(x) \Rightarrow f(x_j)=R(x_j)$, $\forall j\in[n]$, and

$$\int_a^b f(x)\,d\alpha(x)=\int_a^b P_n(x)Q(x)\,d\alpha(x)+\sum_{j=1}^{n}\lambda_j f(x_j) \qquad (3.8)$$

Now (3.7) is exact if $\int_a^b f(x)\,d\alpha(x)=\sum_{j=1}^{n}\lambda_j f(x_j) \Leftrightarrow \int_a^b P_n(x)Q(x)\,d\alpha(x)=0$ (3.9)

Since (3.9) is true for $\deg Q(x)\le n-1 \Rightarrow$ (3.7) is exact if $\deg f(x)\le 2n-1$. Now we show the positivity of λ_j.

Note that $l_j^2-l_j$ is a polynomial of degree $2n-2$ which vanishes at x_k for $k\in[n]$. So $l_j^2-l_j=P_n(x)Q(x)$, where $\deg Q\le n-2$. Thus

$$\int_a^b (l_j^2-l_j)\,d\alpha(x)=\int_a^b P_n(x)Q(x)\,d\alpha(x)=0$$

And $\lambda_j=\int_a^b l_j(x)\,d\alpha(x)=\int_a^b l_j^2(x)\,d\alpha(x)>0.$ ∎

If $f(x)$ is not a polynomial of degree $\le 2n-1$, then (3.7) is not exact, but we can still use RHS as an approximation. Here the error involved is of great importance. We don't talk about this in depth but merely give the following result:

Thm 3.3. If $f(x)$ is continuous on a finite interval $[a,b]$, then

$$\lim_{n\to\infty}\sum_{j=1}^{n}\lambda_{jn}f(x_{jn})=\int_a^b f(x)\,d\alpha(x).$$

Remark: In Thm 3.2 (Gauss quadrature formula), special distributions give us specials cases.

$d\alpha(x)=dx \to P_n(x)$ are Legendre polynomials given by

$$P_n(x)=P_n^{(0,0)}(x)=\frac{(-1)^n}{2^n n!}\frac{d^n}{dx^n}(1-x^2)^n, \quad \text{in interval } [-1,1].$$

$d\alpha(x)=\frac{dx}{\sqrt{1-x^2}} \to P_n(x)$ are $T_n(x)$, in this case, it can be shown that $\lambda_1=\lambda_2=\cdots=\lambda_n$, and

$$\int_{-1}^{1} f(x)\frac{dx}{\sqrt{1-x^2}}=\frac{\pi}{n}\sum_{j=1}^{n} f\left(\cos\frac{2j-1}{2n}\pi\right).$$

when f is a polynomial of degree $\le 2n-1$. A converse of this result is also true.

5.4 Zeros of Orthogonal Polynomials

We have seen that the Chebyshev polynomials $T_n(x)$ have n simple roots in $[-1,1]$, more generally, one can prove the same about Jacobi polynomials. The next theorem shows that a similiar result is true for all orthogonal polynomials.

Thm 4.1: Suppose that $\{P_n(x)\}$ is a sequence of orthogonal polynomials with respect to the distribution $d\alpha(x)$ on the interval $[a,b]$. Then $P_n(x)$ has n simple zeros in $[a,b]$.

Proof: Suppose $P_n(x)$ has m distinct zeros $x_1,x_2,\cdots,x_m$ in $[a,b]$ that are of odd order. In that case

$$Q(x)=P_n(x)(x-x_1)(x-x_2)\cdots(x-x_m)\ge 0,\ \forall x\in[a,b] \tag{4.1}$$

If $m<n$, then by orthogonality

$$\int_a^b Q(x)\,d\alpha(x)=0 \tag{4.2}$$

However, $(4.1)\Rightarrow \int_a^b Q(x)\,d\alpha(x)>0$. Contradiction.

Thus $m=n$ and all zeros are simple. ▨

Thm 4.2. The zeros of $P_n(x)$ and $P_{n+1}(x)$ separate each other.

Proof: By Corollary of Thm 2.5,

$$P_{n+1}(x)P_n'(x) - P_n(x)P_{n+1}'(x) < 0$$

Since $x_{k,n+1}$ is a zero of $P_{n+1}(x)$, we get

$$P_n(x_{k,n+1})P_{n+1}'(x_{k,n+1}) > 0$$

The simplicity of the zeros implies that $P_{n+1}'(x_{k,n+1})$ and $P_{n+1}'(x_{k+1,n+1})$ have different signs. It follows that $P_n(x_{k,n+1})$ and $P_n(x_{k+1,n+1})$ have different signs.

By continuity of P_n, we know it has a zero between $x_{k,n+1}$ and $x_{k+1,n+1}$ for $k=1,2,\cdots,n$, and our result follows. $\blacksquare$

We can obtain an extension of Thm 4.2 by Gauss quadrature formula.

Thm 4.3 Let $m<n$. Between any two zeros of $P_m(x)$ there is a zero of $P_n(x)$

Proof: Suppose there is no zero of $P_n(x)$ between x_{km} and $x_{k+1,m}$. Consider the polynomial $g(x) = \dfrac{P_m(x)}{(x-x_{km})(x-x_{k+1,m})}$

It is clear that $g(x)P_m(x) \geq 0$ for $x \notin (x_{km}, x_{k+1,m})$.

By Gauss quadrature formula

$$\int_a^b g(x)P_m(x)\,d\alpha(x) = \sum_{j=1}^{n} \lambda_j g(x_{jn})P_m(x_{jn}).$$

By orthogonality, LHS $=0$.

On the other hand, $g(x_{jn})P_m(x_{jn}) \geq 0$ cannot vanish for all $j \in [n]$, and $\lambda_j > 0$ $\forall j$, so RHS >0, contradiction. $\blacksquare$

We conclude this part with stating (without proof) the Markov-Stieltjes inequalities for the sums $\sum_{k=1}^{j} \lambda_k$, where $j \leq n$. Once again we let x_j, $j \in [n]$ denote the zeros of $P_n(x)$ in increasing order.

Thm 4.4: The Markov-Stieltjes inequalities

$$\sum_{k=1}^{j-1} \lambda_k \le \int_a^{x_j} d\alpha(x) \le \sum_{k=1}^{j} \lambda_k \quad \text{hold for } j \in [n].$$

5.5 Continued Fractions

Continued Fractions of a certain type are closely connected with orthogonal polynomials, but here we shall merely touch on this topic.

Suppose $\{a_n\}_1^\infty$ and $\{b_n\}_0^\infty$ are sequences of complex numbers. One notation for an infinite continued fraction is

$$b_0 + \frac{a_1}{b_1+}\frac{a_2}{b_2+}\frac{a_3}{b_3+}\cdots \qquad (5.1)$$

We shall denote the nth convergent of this continued fraction by C_n. So $C_0 = b_0 =: \frac{A_0}{B_0}$, $\quad C_1 = b_0 + \frac{a_1}{b_1} = \frac{b_0 b_1 + a_1}{b_1} =: \frac{A_1}{B_1}$,

$$C_2 = b_0 + \frac{a_1}{b_1+}\frac{a_2}{b_2} = b_0 + \frac{a_1 b_2}{b_1 b_2 + a_2} = \frac{b_0 b_1 b_2 + a_2 b_0 + a_1 b_2}{b_1 b_2 + a_2} =: \frac{A_2}{B_2}.$$

Def: We say that the continued fraction (5.1) converges, if at most a finite number of C_n are undefined and $\lim_{n\to\infty} C_n$ exists (C_2 is undefined above if $B_2 = b_1 b_2 + a_2 = 0$).

The sequences $\{A_n\}$ and $\{B_n\}$ defined above satisfy a three-term recurrence relation.

Lemma 5.2: For $n \ge 1$, $\quad A_n = b_n A_{n-1} + a_n A_{n-2}$, $A_{-1} = 1 \quad (5.2)$

and $\quad B_n = b_n B_{n-1} + a_n B_{n-2}$, $B_{-1} = 0 \quad (5.3)$.

This can be shown by induction. Also, by a method similiar to the proof of Christoffel-Darboux formula, one can obtain the following result:

Lemma 5.3 $\quad A_n B_{n-1} - B_n A_{n-1} = (-1)^{n+1} a_1 a_2 \cdots a_n, \ n \geq 1.$

Recall: The 3-term recurrence relation of orthogonal polynomials

$$P_{n+1}(x) = (A_n x + B_n) P_n(x) - C_n P_{n-1}(x)$$

compare it with $A_n = b_n A_{n-1} + a_n A_{n-2}$, which suggests us to study the continued fraction of the form

$$\frac{A_0}{A_0 x + B_0 -} \ \frac{C_1}{A_1 x + B_1 -} \ \frac{C_2}{A_2 x + B_2 -} \cdots \qquad (5.4)$$

In this case, the nth convergent is a rational function whose denominator is $P_n(x)$, and we denote the numerator by $P_n^*(x)$.

The sequence $\{P_n^*(x)\}$ satisfy the same recursion, namely

$$P_{n+1}^*(x) = (A_n x + B_n) P_n^*(x) - C_n P_{n-1}^*(x), \ n \geq 1 \qquad (5.5)$$

with different initials $P_0^*(x) = 0$, $P_1^*(x) = A_0$.

Suppose that the sequence $\{P_n(x)\}$ is orthogonal with respect to the distribution $d\alpha(x)$ on $[a,b]$.

The next theorem relates $P_n^*(x)$ to $P_n(x)$.

Thm 5.4: With $P_n(x)$ and $P_n^*(x)$ as defined above, we have

$$P_n^*(x) = \delta \int_a^b \frac{P_n(x) - P_n(t)}{x - t} \, d\alpha(t), \ n \geq 0, \qquad (5.6)$$

where δ is a constant.

Next result is an application of the Gauss quadrature formula.

Recall $x_k, k \in [n]$ are zeros of $P_n(x)$.

Theorem 5.5: Using the notation of Thm 3.2 and Thm 3.3, we have

$$\frac{P_n^*(x)}{P_n(x)} = \delta \sum_{k=1}^{n} \frac{\lambda_k}{x - x_k} \qquad (5.7),$$

where δ is the constant in (5.6).

Proof: Expressed $\boxed{\frac{P_n^*(x)}{P_n(x)}}$ as a partial fraction, we get (Using interpolation)

$$\frac{P_n^*(x)}{P_n(x)} = \sum_{k=1}^{n} \frac{P_n^*(x_k)}{P_n'(x_k)(x-x_k)} \quad (\deg P_n^*(x) < \deg P_n(x))$$

By Thm 5.4, we get

$$\frac{P_n^*(x_k)}{P_n'(x_k)} = \delta \int_a^b \frac{P_n(t)}{P_n'(x_k)(t-x_k)} d\alpha(t) = \delta \lambda_k \quad (5.8)$$

The last step follows from the Gauss quadrature formula. ▨

Thm 5.6: Let $[a,b]$ be a finite integral. For any $x \notin [a,b]$

$$\lim_{n\to\infty} \frac{P_n^*(x)}{P_n(x)} = \delta \int_a^b \frac{d\alpha(t)}{x-t} \quad (5.9)$$

Proof: $\forall x \notin [a,b]$, $\frac{1}{x-t}$ is a continuous function of t in $[a,b]$. Use Thm 5.5 and Thm 3.3 give us the result. ▨

Remark: $\lambda_k > 0$, so (5.8) $\Rightarrow$ zeros of P_n^* and P_n alternate.

Also, In Thm 5.6, let $x \in \mathbb{C}$, $x \notin [a,b]$. If we denote RHS by $F(x)$, the inversion formula of stieltjes is given by

$$\alpha(c) - \alpha(d) = -\frac{1}{\pi} \lim_{v\to 0^+} \int_c^d \operatorname{Im}\{F(u+iv)\} du.$$

Which means if we know F, we can recover the distribution.

5.6 Kernel Polynomials

We have seen the partial sum of the Fourier series of a function when expressed as an integral gave us the Chebyshev polynomials of the third kind.

More generally, we get the kernel polynomials when we study partial sums involving orthogonal polynomials.

Let $\{P_n(x)\}$ be a sequence of polynomials orthogonal with respect to the distribution $d\alpha(t)$ on an interval $[a,b]$. Here $-\infty \le a < b \le \infty$. Let f be a function such that $\int_a^b f(t)P_n(t)\,d\alpha(t)$ exists for all n.

The series corresponding to the Fourier series is given by

$$a_0P_0(x) + a_1P_1(x) + \cdots + a_nP_n(x) + \cdots, \text{ where} \qquad (6.1)$$

$$a_n = \frac{\int_a^b f(t)P_n(t)\,d\alpha(t)}{\int_a^b (P_n(t))^2\,d\alpha(t)} \qquad (6.2)$$

In this section we assume that the denominator of a_n is 1. i.e. the sequence $\{P_n(x)\}$ is orthonormal.

The nth partial sum $S_n(x)$ is given by

$$S_n(x) = \sum_{k=0}^{n} P_k(x)\int_a^b f(t)P_k(t)\,d\alpha(t) = \int_a^b f(t)K_n(t,x)\,d\alpha(t) \qquad (6.3)$$

where $$K_n(t,x) = \sum_{k=0}^{n} P_k(t)P_k(x) \qquad (6.4).$$

Def. For a sequence of orthonormal polynomials $\{P_n(x)\}$, the sequence $\{K_n(x_0,x)\}$, where $K_n(x_0,x) = \sum_{k=0}^{n} P_k(x_0)P_k(x)$, is called the kernel polynomial sequence.

Lemma 6.2: If $Q(x)$ is a polynomial of degree $\le n$, then

$$Q(x) = \int_a^b K_n(t,x)Q(t)\,d\alpha(t).$$

Proof: Clearly, $Q(x) = \sum_{k=0}^{n} a_kP_k(x)$ for some a_k.

Multiply both sides by $P_j(x)$ and integrate, we get

$$\int_a^b Q(t)P_j(t)\,d\alpha(t) = a_j$$

compare with (6.3), the result follows. ▨

Thm 6.3 Suppose $x_0 \le a$ are both finite. The sequence $\{K_n(x_0,x)\}$ is orthogonal with respect to the distribution $(t-x_0)\,d\alpha(t)$.

Proof: In Lemma 6.2, let $Q(t)=(t-x_0)Q_{n-1}(t)$, where $Q_{n-1}(t)$ is an arbitrary polynomial with degree $n-1$. ∎

Note that a similiar result can be obtained for $b \le x_0$ are finite.

Remark: In the case of Chebyshev polynomials $T_n(x)$ with $x_0=a=-1$, we see for $x=\cos\theta$,

$$K_n(-1,\cos\theta)=\frac{1}{2}-\cos\theta+\cos 2\theta-\cdots+(-1)^n\cos n\theta=(-1)^n\frac{\cos(n+\frac{1}{2})\theta}{\cos\frac{\theta}{2}}$$

The polynomials $W_n(x)=\dfrac{\cos(n+\frac{1}{2})\theta}{\cos\frac{\theta}{2}}$ are the Chebyshev polynomials of the fourth kind and we can find the distribution via Thm 6.3.

$$\int_{-1}^{1} W_n(x)W_m(x)\sqrt{\frac{1+x}{1-x}}\,dx=0,\quad n\ne m.$$

The Christoffel-Darboux formula (Thm 2.4) gives a compact expression for the kernel polynomials:

$$K_n(x_0,x)=\frac{k_n}{k_{n+1}}\,\frac{P_{n+1}(x)P_n(x_0)-P_{n+1}(x_0)P_n(x)}{x-x_0} \qquad (6.5)$$

If we choose x_0 to be $x_k\,(=x_{kn})$, then

$$K_n(x_k,x)=-\frac{k_n}{k_{n+1}}\,\frac{P_{n+1}(x_k)P_n(x)}{x-x_k} \qquad (6.6)$$

This expression for K_n suggests a connection with the Gauss quadrature formula. In fact, we have the theorem which allows us to find all λ_k:

Thm 6.4: The numbers λ_k (or λ_{kn}) occuring in the Gauss quadrature formula are given by

$$\lambda_k = -\frac{k_{n+1}}{k_n} \frac{1}{P_{n+1}(x_k) P_n'(x_k)} ; \qquad (6.7)$$

their reciprocal is $\quad \dfrac{1}{\lambda_k} = K_n(x_k, x_k) = \sum_{k=0}^{n} (P_k(x_k))^2 \qquad (6.8).$

Proof: The expression for λ_k in Gauss's formula is

$$\lambda_k = \int_a^b \frac{P_n(t)\, d\alpha(t)}{P_n'(x_k)(t-x_k)}$$

By 6.6 and Lemma 6.2,

$$\lambda_k = -\frac{k_{n+1}}{k_n} \cdot \frac{1}{P_{n+1}(x_k) P_n'(x_k)} \int_a^b K_n(x_k, t)\, d\alpha(t)$$

$$= -\frac{k_{n+1}}{k_n} \frac{1}{P_{n+1}(x_k) P_n'(x_k)} .$$ This concludes (6.7).

To prove (6.8), let $x \to x_k$ in (6.6),

$$K(x_k, x_k) = -\frac{k_n}{k_{n+1}} P_{n+1}(x_k) P_n'(x_k)$$ ▨

The kernel polynomials also have a maximum property, but we shall only state it here.

Thm 6.5: Let $x_0 \in \mathbb{R}$ and $Q(x)$ be a polynomial with degree $\le n$, normalized by $\int_a^b (Q(t))^2 d\alpha(t) = 1$. Then the maximum value of $(Q(x_0))^2$ is given by the polynomial $Q(x) = \pm \dfrac{K_n(x_0, x)}{\sqrt{K_n(x_0, x_0)}}$, and the maximum itself is $K_n(x_0, x_0)$.

5.7 Parseval's Formula

Let $L_\alpha^p(a,b)$ denote the class of functions f such that

$$\int_a^b |f|^p d\alpha(x) < \infty .$$

As always, we assume moments of all orders exist.

In this section, we are interested in the space $L^2_\alpha(a,b)$. By the Cauchy-Schwartz inequality, we infer that $\int_a^b f(x)x^n d\alpha(x)$ exists for $n \geq 0$.

Thm 7.1: Suppose $f \in L^2_\alpha(a,b)$. Let $Q(x)$ be a polynomial of degree n, such that $Q(x) = \sum_{k=0}^{n} a_k P_k(x)$,

where $\{P_k(x)\}$ is the orthonormal sequence of polynomials for $d\alpha$.

The integral
$$\int_a^b [f(x) - Q(x)]^2 d\alpha(x) \qquad (7.1)$$
becomes a minimum when
$$a_k = \int_a^b f(x) P_k(x) d\alpha(x) \qquad (7.2)$$
Moreover, with a_k above, we have
$$\sum_{k=0}^{n} a_k^2 \leq \int_a^b [f(x)]^2 d\alpha(x) \qquad (7.3).$$

Proof: Let $c_k = \int_a^b f(x) P_k(x) d\alpha(x)$. By orthonormality of $\{P_n(x)\}$, we get $0 \leq \int_a^b [f(x) - Q(x)]^2 d\alpha(x) = \int_a^b \left[f(x) - \sum_{k=0}^{n} a_k P_k(x) \right]^2 d\alpha(x)$
$$= \int_a^b [f(x)]^2 d\alpha(x) - 2\sum_{k=0}^{n} a_k c_k + \sum_{k=0}^{n} a_k^2 = \int_a^b [f(x)]^2 d\alpha(x) - \sum_{k=0}^{n} c_k^2$$
$+ \sum_{k=0}^{n} (a_k - c_k)^2$. The last expression assumes its least value when $a_k = c_k$. This proves both parts of our theorem. ▨

Coro (Bessel's inequality): For $f \in L^2_\alpha(a,b)$ and a_k as above, we have
$$\sum_{n=0}^{\infty} a_n^2 \leq \int_a^b [f(x)]^2 d\alpha(x). \qquad (7.4)$$

Proof: The sequence of partial sums $S_n = \sum_{k=0}^{n} a_k^2$ is increasing and bounded. ▨

We now seek the situation where equality holds.

Assume that $[a,b]$ is a finite interval.

Lemma 7.3: For $f \in L^2_\alpha(a,b)$ and a given $\varepsilon>0$, then there exists a continuous function g such that

$$\int_a^b [f(x)-g(x)]^2 d\alpha(x) < \varepsilon .$$

Thm 7.4: Let $[a,b]$ be a finite interval. With the notation of Theorem (7.1), we have Parseval's formula:

$$\sum_{k=0}^{\infty} a_k^2 = \int_a^b [f(x)]^2 d\alpha(x) \qquad (7.5)$$

Proof: Suppose g is as in Lemma 7.3.

By Weierstrass's approximation theorem, for a given $\varepsilon>0$ there exists a polynomial $Q_n(x)$ s.t.

$$\int_a^b [g(x)-Q_n(x)]^2 d\alpha(x) < \varepsilon$$

$$\Rightarrow \int_a^b [f(x)-Q_n(x)]^2 d\alpha(x) < 4\varepsilon \qquad (7.6)$$

By Thm 7.1, we may choose $Q_n(x) = \sum_{k=0}^{n} a_k P_k(x)$ with a_k given in (7.2). As in the proof of Thm 7.1, we get from (7.6) that:

$$\int_a^b [f(x)]^2 d\alpha(x) - \sum_{k=0}^{n} a_k^2 < 4\varepsilon . \qquad \blacksquare$$

Coro: Suppose $f \in L^2_\alpha(a,b)$, where $-\infty<a<b<\infty$. If

$$\int_a^b f(x) x^n d\alpha(x) = 0 \text{ for all integers } n \geq 0 \qquad (7.7).$$

Then $f=0$ almost everywhere.

Proof: $\forall k$, $a_k=0$. The result follows from Parseval's formula. $\blacksquare$

Remark: Parseval's formula and its Corollary are in general False when $[a,b]$ is not finite. However there are important examples where infinite intervals do work. For example, Hermite polynomials on $(-\infty,\infty)$.

5.8 The moment-generating Function

In this section, we shall obtain a continued fraction expansion of the moment-generating function $\sum_{n\geq 0}\mu_n x^n$, where

$$\mu_n = (1, t^n) = \int_a^b t^n \, d\alpha(t) \qquad (8.1).$$

Consider graphs with weighted arcs so that the entry $(A)_{ij}$ is the weight of the arc (i,j). We denote this matrix by $A = A(G)$.

Let $\varphi(G,x) = \det(xI - A(G))$ and $W_{ij}(G,x) = \sum_{n\geq 0}(A^n)_{ij}\, x^n$.

$W_{ij}(G,x)$ is the generating function for the set of all walks in G from vertex i to vertex j. Let $W(G,x)$ be the matrix whose entries are $W_{ij}(G,x)$. Then

$$W(G,x) = \sum_{n\geq 0} A^n x^n \quad \text{since} \quad A\,\mathrm{adj}(A) = \det(A)I.$$

It follows that

$$W(G,x) = x^{-1}\varphi(G,x^{-1})^{-1}\,\mathrm{adj}(x^{-1}I - A) \qquad (8.2)$$

$$\Rightarrow W_{ii}(G,x) = x^{-1}\varphi(G\setminus i, x^{-1})/\varphi(G,x^{-1}) \qquad (8.3).$$

The connection of the above discussion with orthogonal polynomials is obtained as follows:

Suppose $\{P_n(x)\}$ is an orthogonal polynomial sequence with the 3-term recurrence relation

$$P_{n+1}(x) = (x - a_n)P_n(x) - b_n P_{n-1}(x), \quad n \geq 1 \qquad (8.4)$$

It is assumed that the polynomials are monic. Let A denote the matrix

$$\begin{bmatrix} a_0 & b_1 & & & \\ 1 & a_1 & b_2 & & \\ & 1 & a_2 & b_3 & \\ & & & \ddots & \ddots \end{bmatrix}$$

where the rows and columns are indexed by nonnegative integers.

Let A_n be the matrix obtained by A by taking the first n rows and columns.

Expand $\det(xI-A_n)$ about the last row,

$$\det(xI-A_n) = (x-a_{n-1})P_{n-1}(x) - b_{n-1}P_{n-2}(x) = P_n(x).$$

Observe that the matrix A is the adjacency matrix of a centain infinite weighted digraph whose vertices are indexed by nonnegative integer.

If we only take the first n vertices of G_∞, denote this (induced) subgraph by G_n, A_n is the adjacency matrix of G_n.

We shall give a continued fraction expansion for the moment-generating function $\sum_{n\geq 0}(1,t^n)x^n$ with convention $\mu_0=(1,1)=1$.

Lemma 8.1. For $n\geq 0$, $\mu_n=(1,x^n)=(A^n)_{00}$

Proof: First note $(A^k)_{00}=(A_n^k)_{00}$ for $k\leq 2n+1$, because no closed walk starting at 0 and of length $\leq 2n+1$ can include a vertex after the nth vertex. Thus $(P_n(A))_{00}=(P_n(A_n))_{00}$. We have seen that for $n\geq 1$, $P_n(x)$ is the characteristic polynomial of A_n. By Cayley-Hamilton theorem, $P_n(A_n)=0$, hence $(1,P_n)=(P_n(A_n))_{00}$ for $n\geq 0$. Now x^n is a linear combination of $P_0, P_1, \cdots, P_n$. ▨

Thm 8.2 With a_n and b_n in (8.4),

$$\sum_{n\geq 0}(1,t^n)x^n = \frac{1}{1-xa_0-}\,\frac{x^2b_1}{1-xa_1-}\,\frac{x^2b_2}{1-xa_2-}\cdots$$

Proof: Let $A_{n,k}$ be the matrix obtained from A_n by removing the first k rows and columns. Set

$$q_{n-k}(x) = \det(I - xA_{n,k})$$

Observe $\quad \varphi(G_n,x) = \det(xI-A_n) = x^n\det(I-x^{-1}A_{n,0}) = x^n q_n(\tfrac{1}{x})$

and $\quad \varphi(G_n\backslash 0, x) = \det(xI-A_{n,1}) = x^{n-1}\det(I-x^{-1}A_{n,1}) = x^{n-1}q_{n-1}(\tfrac{1}{x})$

By (8.3), we conclude $\quad x^{-1}W_{00}(G_n,x^{-1}) = \dfrac{\varphi(G_n\backslash 0,x)}{\varphi(G_n,x)} = x^{-1}\dfrac{q_{n-1}(x^{-1})}{q_n(x^{-1})}$ (8.5).

Expand $\det(I-xA_n)$ about the first row,

$$q_n(x) = (1-xa_0)q_{n-1}(x) - x^2b_1q_{n-2}(x)$$

$$\Longleftrightarrow \quad \frac{q_{n-1}(x)}{q_n(x)} = \frac{1}{1-xa_0-x^2b_1\dfrac{q_{n-2}(x)}{q_{n-1}(x)}} \qquad (8.6).$$

By Lemma 8.1,

$$\sum_{n\geq 0}(1,t^n)x^n = \sum_{n\geq 0}(A^n)_{00}x^n = W_{00}(G,x) = \lim_{n\to\infty}\frac{q_{n-1}(x)}{q_n(x)}.$$

Combine this with (8.6), our result follows. $\blacksquare$

6.1 Hermite polynomials.

The normal integral $\int_{-\infty}^{\infty} e^{-x^2}dx$ plays an important role in probability theory and other areas of mathematics.

This integral has several interesting properties.

For example, it is essentially its own Fourier transform.

$$e^{-x^2}=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}e^{2ixt}dt \qquad (9.1)$$

This integral is uniformly convergent in any disk $|x|\leq r$ and is majorized in that region by the convergent integral

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}e^{2rt}dt.$$

Thus the integral can be repeatedly differentiated by x, we have

$$\frac{d^n e^{-x^2}}{dx^n}=\frac{(2i)^n}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}t^n e^{2ixt}dt \qquad (9.2)$$

The polynomials orthogonal with respect to the normal distribution e^{-x^2} are the Hermite polynomials. They can be defined by the formula

$$H_n(x)=(-1)^n e^{x^2}\frac{d^n e^{-x^2}}{dx^n} \qquad (9.3)$$

The orthogonality is given by

$$\int_{-\infty}^{\infty} e^{-x^2}H_n(x)H_m(x)=2^n n!\sqrt{\pi}\,\delta_{mn}. \qquad (9.4).$$

Moreover, the Hermite polynomials have very nice generating function:

$$\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}r^n=e^{2xr-r^2} \qquad (9.5).$$

Instead of proving more properties for the Hermite polynomials, we put more number theoretic flavour into the end of this section.

Waring's problem for exponent k is to prove that the set of nonnegative integers is a basis of finite order, that is, to prove that every nonnegative integer can be written as the sum of a bounded number of kth powers.

We denote by $g(k)$ the smallest number s such that every nonnegative integer is the sum of s kth powers of nonnegative integers. $g(2)=4$, as we all know.

Other cases of Waring's problem can be deduced from these results by means of polynomial identities. We give some such examples. Here we use the notation

$$(x_1 \pm x_2 \pm \cdots + x_h)^k = \sum_{\varepsilon_2, \cdots, \varepsilon_h = \pm 1} (x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_h x_h)^k$$

(Liouville)

$$(x_1^2+x_2^2+x_3^2+x_4^2)^2 = \frac{1}{6} \sum_{1 \le i < j \le 4} (x_i+x_j)^4 + \frac{1}{6} \sum_{1 \le i < j \le 4} (x_i - x_j)^4$$

is a polynomial identity, thus $g(4) \le 53$.

(Fleck)

$$(x_1^2+x_2^2+x_3^2+x_4^2)^3 = \frac{1}{60} \sum_{1 \le i < j < k \le 4} (x_i \pm x_j \pm x_k)^6 + \frac{1}{30} \sum_{1 \le i < j \le 4} (x_i + x_j)^6 + \frac{3}{5} \sum_{1 \le i \le 4} x_i^6$$

so $g(6) < \infty$.

(Hurwitz)

$$(x_1^2+x_2^2+x_3^2+x_4^2)^4 = \frac{1}{840} (x_1 \pm x_2 \pm x_3 \pm x_4)^8 + \frac{1}{5040} \sum_{1 \le i < j < k \le 4} (2x_i \pm x_j \pm x_k)^8 +$$
$$\frac{1}{84} \sum_{1 \le i < j \le 4} (x_i \pm x_j)^8 + \frac{1}{840} \sum_{1 \le i \le 4} (2x_i)^6$$

, so $g(8) < \infty$

Suppose that

$$(x_1^2+x_2^2+x_3^2+x_4^2)^k=\sum_{i=1}^{M} a_i(b_{i1}x_1+b_{i2}x_2+b_{i3}x_3+b_{i4}x_4)^{2k} \qquad (9.6)$$

for some $M\geq 1$, $b_{i,j}\in\mathbb{Z}$, $a_i\in\mathbb{Q}^+$.

Hurwitz observed that if Waring's problem is true for exponent k, then it is immediate from (9.6) and Lagrange's theorem that it is also true for exponent $2k$.

Hilbert first proved that identities of the form (9.6) exists, using Hermite polynomials.

Here we shall use $H_n(x)=(\frac{-1}{2})^n e^{x^2}\frac{d^n}{dx^n}(e^{-x^2})$.

Lemma 1: The Hermite polynomial $H_n(x)$ has n simple zeros.

Lemma 2: Let $n\geq 1$, $f(x)$ be a polynomial of degree at most $n-1$.

Then $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}H_n(x)f(x)\,dx=0.$

Lemma 3: For $n\geq 0$,

$$C_n=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-x^2}x^n\,dx=\begin{cases}\dfrac{n!}{2^n(\frac{n}{2})!} & 2\mid n\\ 0 & 2\nmid n.\end{cases} \qquad (9.7)$$

Lemma 4: For $n\geq 1$, let $\beta_1,\cdots,\beta_n$ be n distinct real numbers, $C_0,C_1,\cdots,C_{n-1}$ as defined in Lemma 3. Then the system of linear equations

$$\sum_{j=1}^{n}\beta_j^k x_j=C_k \quad ,k=0,1,\cdots,n-1 \qquad (9.8)$$

has a unique solution $\rho_1,\cdots,\rho_n$. If $r(x)$ is a polynomial of degree at most $n-1$, then

$$\sum_{j=1}^{n} r(\beta_j)\rho_j=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-x^2}r(x)\,dx.$$

Lemma 5. Let $n \geq 1$, $\beta_1, \cdots, \beta_n$ be the n distinct real zeros of the Hermite polynomial $H_n(x)$, let $\rho_1, \cdots, \rho_n$ be defined as in lemma 4. Let $f(x)$ be a polynomial of degree at most $2n-1$. Then

$$\sum_{j=1}^{n} f(\beta_j)\rho_j = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} f(x)\,dx.$$

Lemma 6. Let $\rho_1, \cdots, \rho_n$ be defined above, then $\rho_i > 0, \forall i \in [n]$.

Lemma 7. Let $n \geq 1$, and let $C_0, C_1, \cdots, C_{n-1}$ be as defined in lemma 3. Then there exist pairwise distinct rational numbers $\beta_1^*, \cdots, \beta_n^*$ and positive rational numbers $\rho_1^*, \cdots, \rho_n^*$ such that

$$\sum_{j=1}^{n} (\beta_j^*)^k \rho_j^* = C_k, \text{ for } k = 0, 1, \cdots, n-1.$$

Proof: By Lemma 3.4, for any set of n distinct real numbers $\beta_1, \cdots, \beta_n$, the system of n linear equations

$$\sum_{j=1}^{n} \beta_j^k x_j = C_k, \text{ for } k = 0, 1, \cdots, n-1$$

has a unique solution $(\rho_1, \cdots, \rho_n)$.

Let $\mathcal{R}$ be the open subset of $\mathbb{R}^n$ consisting of all $(\beta_1, \cdots, \beta_n)$ with $\beta_i \neq \beta_j, \forall i \neq j$. And let $\Phi : \mathcal{R} \to \mathbb{R}^n$ be the function that sends $(\beta_1, \cdots, \beta_n) = (\rho_1, \cdots, \rho_n)$. This function is continuous, since by Cramer's rule, we can express each ρ_j as a rational function of $\beta_1, \cdots, \beta_n$.

Now since $\mathbb{R}_+^n$ is an open subset of $\mathbb{R}^n$, $\Phi^{-1}(\mathbb{R}_+^n)$ is an open neighbourhood of $(\beta_1, \cdots, \beta_n)$ in $\mathcal{R}$. Since the points with rational coordinates are dense in $\mathcal{R}$, it follows that this neighbourhood contains a rational point $(\beta_1, \cdots, \beta_n)$. Let $(\rho_1^*, \cdots, \rho_n^*) = \Phi(\beta_1^*, \cdots, \beta_n^*) \in \mathbb{R}_+^n$

Since each number ρ_i^* can be expressed as a rational function with rational coefficients of rational variables, our proof is complete. $\blacksquare$

Lemma 8: Let $n \geq 1$, $C_0, \cdots, C_{n-1}$ be as defined in Lemma 3, let $\beta_1, \cdots, \beta_n$ be n distinct real numbers, and let $\rho_1, \cdots, \rho_n$ be the solution of the linear system given in Lemma 4. $\forall r \in \mathbb{Z}^+$, $m = 1, 2, \cdots, n-1$,

$$C_m (x_1^2 + \cdots + x_r^2)^{m/2} = \sum_{j_1=1}^{n} \cdots \sum_{j_r=1}^{n} \rho_{j_1} \cdots \rho_{j_r} (\beta_{j_1} x_1 + \cdots + \beta_{j_r} x_r)^m$$

is a polynomial identity.

Proof:
$$\sum_{j_1=1}^{n} \cdots \sum_{j_r=1}^{n} \rho_{j_1} \cdots \rho_{j_r} (\beta_{j_1} x_1 + \cdots + \beta_{j_r} x_r)^m$$

$$= \sum_{j_1=1}^{n} \cdots \sum_{j_r=1}^{n} \rho_{j_1} \cdots \rho_{j_r} \sum_{\substack{\mu_1 + \cdots + \mu_r = m \\ \mu_i \geq 0, i \in [r]}} \frac{m!}{\mu_1! \cdots \mu_r!} (\beta_{j_1} x_1)^{\mu_1} \cdots (\beta_{j_r} x_r)^{\mu_r}$$

$$= m! \sum_{j_1=1}^{n} \cdots \sum_{j_r=1}^{n} \sum_{\substack{\mu_1 + \cdots + \mu_r = m \\ \mu_i \geq 0, i \in [r]}} \frac{x_1^{\mu_1}}{\mu_1!} (\beta_{j_1}^{\mu_1} \rho_{j_1}) \cdots \frac{x_r^{\mu_r}}{\mu_r!} (\beta_{j_r}^{\mu_r} \rho_{j_r})$$

$$= m! \sum_{\substack{\mu_1 + \cdots + \mu_r = m \\ \mu_i \geq 0, i \in [r]}} \sum_{j_1=1}^{n} \cdots \sum_{j_r=1}^{n} \prod_{i=1}^{r} \frac{x_i^{\mu_i}}{\mu_i!} (\beta_{j_i}^{\mu_i} \rho_{j_i})$$

$$= m! \sum_{\substack{\mu_1 + \cdots + \mu_r = m \\ \mu_i \geq 0, i \in [r]}} \prod_{i=1}^{r} \left(\frac{x_i^{\mu_i}}{\mu_i!} \sum_{j=1}^{n} \beta_j^{\mu_i} \rho_j \right)$$

$$= m! \sum_{\substack{\mu_1 + \cdots + \mu_r = m \\ \mu_i \geq 0}} \prod_{i=1}^{r} \frac{C_{\mu_i} x_i^{\mu_i}}{\mu_i!}$$

By Lemma 3, $C_m = 0$ if m is odd. If m is odd and $\mu_1 + \cdots + \mu_r = m$, the μ_i is odd for some i, so:

$$\sum_{j_1=1}^{n} \cdots \sum_{j_r=1}^{n} \rho_{j_1} \cdots \rho_{j_r} (\beta_{j_1} x_1 + \cdots + \beta_{j_r} x_r)^m = 0.$$

This proves the lemma for odd m. For even m, we only need to consider partitions of m into even parts $\mu_i = 2\nu_i$.

Inserting the expression for C_n in lemma 3:

$$\sum_{j_1=1}^{n} \cdots \sum_{j_r=1}^{n} \rho_{j_1} \cdots \rho_{j_r} (\beta_{j_1} x_1 + \cdots + \beta_{j_r} x_r)^m$$

$$= m! \sum_{\substack{2\nu_1 + \cdots + 2\nu_r = m \\ \nu_i \geq 0,\, i \in [r]}} \prod_{i=1}^{r} \frac{C_{2\nu_i} x_i^{2\nu_i}}{(2\nu_i)!}$$

$$= m! \sum_{\substack{\nu_1 + \cdots + \nu_r = \frac{m}{2} \\ \nu_i \geq 0,\, i \in [r]}} \prod_{i=1}^{r} \frac{(2\nu_i)!}{2^{2\nu_i}\, \nu_i!} \frac{x_i^{2\nu_i}}{(2\nu_i)!}$$

$$= \frac{m!}{2^m} \sum_{\substack{\nu_1 + \cdots + \nu_r = \frac{m}{2} \\ \nu_i \geq 0,\, i \in [r]}} \prod_{i=1}^{r} \frac{x_i^{2\nu_i}}{\nu_i!}$$

$$= \frac{m!}{2^m (\frac{m}{2})!} \sum_{\substack{\nu_1 + \cdots + \nu_r = \frac{m}{2} \\ \nu_i \geq 0,\, i \in [r]}} \frac{(\frac{m}{2})!}{\nu_1! \cdots \nu_r!} (x_1^2)^{\nu_1} \cdots (x_r^2)^{\nu_r}$$

$$= C_m (x_1^2 + \cdots + x_r^2)^{\frac{m}{2}}$$

▨

Thm 1 (Hilbert's identity).

$\forall k \geq 1\ \forall r \geq 1, \exists M \geq 1, a_i \in \mathbb{Q}^+, b_{i,j} \in \mathbb{Z}^+$ for $i = 1, \ldots, M$ and $j = 1, \ldots, r$ such that

$$(x_1^2 + \cdots + x_r^2)^k = \sum_{i=1}^{M} a_i (b_{i,1} x_1 + \cdots + b_{i,r} x_r)^{2k}$$

Proof: Choose $n > 2k$, let $\beta_1^*, \ldots, \beta_n^*$ and $\rho_1^*, \ldots, \rho_n^*$ be the rational numbers constructed in Lemma 7. We use these numbers in

Lemma 8: with $m=2k$ to obtain:

$$C_{2k}(x_1^2+\cdots+x_r^2)^k=\sum_{j_1=1}^{n}\cdots\sum_{j_r=1}^{n}\beta_{j_1}^*\cdots\beta_{j_r}^*\left(\beta_{j_1}^*x_1+\cdots+\beta_{j_r}^*x_r\right)^{2k}$$

Let q be a common denominator of $\beta_1^*,\cdots,\beta_n^*$, then $q\beta_j^*$ is an integer for all j, and we have

$$(x_1^2+\cdots+x_r^2)^k=\sum_{j_1=1}^{n}\cdots\sum_{j_r=1}^{n}\frac{\beta_{j_1}^*\cdots\beta_{j_r}^*}{C_{2k}q^{2k}}(q\beta_{j_1}^*x_1+\cdots+q\beta_{j_r}^*x_r)^{2k}. \quad \blacksquare$$

Lemma 9: Let $k\geq 1$. If there exists positive rational numbers $a_1,\cdots,a_M$ such that every sufficiently large integer n can be written in the form $n=\sum_{i=1}^{M}a_iy_i^k$ where $y_1,\cdots,y_M$ are nonegative integers, the Waring's problem is true for exponent k. $\blacksquare$

Proof: Choose n_0 s.t. $\forall n\in\mathbb{N}, n\geq n_0$, n can be represented in such a form. Let q be a common denominator of $a_1,\cdots,a_M$. Then $qa_i\in\mathbb{Z}$, $\forall i\in[M]$, and qn is a sum of $\sum_{i=1}^{M}qa_i$ nonegative kth powers for every $n\geq n_0$. Since every integer $N\geq qn_0$ can be written in the form $N=qn+r$, it follows that N can be written as the sum of $\sum_{i=1}^{M}qa_i+q-1$ nonegative kth powers. And clearly every integer $N\in\mathbb{N}, N<qn_0$ can be written as bounded number of kth powers. $\blacksquare$

Theorem 2: If Waring's problems holds for k, then it holds for $2k$.

Proof: We use Hilbert's identity for k with $r=4$:

$$(x_1^2+x_2^2+x_3^2+x_4^2)^k=\sum_{i=1}^{M}a_i(b_{i_1}x_1+\cdots+b_{i_4}x_4)^{2k}$$

Let y be a nonnegative integer, by Lagrange's theorem, there exist nonnegative integers x_1, x_2, x_3, x_4 s.t.

$$y = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

So $y^k = \sum_{i=1}^{M} a_i z_i^{2k}$, where $z_i = b_{i1}x_1 + \cdots + b_{i4}x_4$ are nonnegative integers.

If Waring's problem is true for k, then every nonnegative integer is the sum of a bounded number of kth powers, thus a integer is the sum of bounded number of $2k$th powers with rational coefficients.

By Lemma 9, Waring's problem is true for $2k$. $\blacksquare$