# UNA APROXIMACIÓN A LOS POLINOMIOS ORTOGONALES EN EL CÍRCULO UNITARIO

YENIFER ANDREA GONZÁLEZ GARCÍA

Universidad Distrital Francisco José de Caldas
Bogotá D.C.
2019

# Una Aproximación a Los Polinomios Ortogonales en el Círculo Unitario

Yenifer Andrea González García

Monografía de grado

Presentada como requisito para optar por el título de Matemática

Director: Luis Oriol Mora Valbuena

Universidad Distrital Francisco José de Caldas

Facultad de Ciencias y Educación

Proyecto curricular de Matemáticas

Bogotá D.C

2019

Dedicado a mi mamá, mi papá y mi hermano.

# Agradecimientos

Principalmente, destaco mi agradecimiento hacia mis papás, mi hermano y mi familia en general quienes me han apoyado siempre durante toda mi vida en cada paso que doy. Sus valiosos consejos y su acompañamiento me permiten hoy cumplir con excelencia el desarrollo de este trabajo.

Por otro lado, quiero agradecer al doctor Luis Oriol Mora, no solo por transmitirme sus conocimientos y dedicar su tiempo a supervisar y dirigir éste trabajo, sino porque fué un apoyo sumamente importante para mi crecimiento personal. Al igual que el doctor Arturo Sanjuan, quien además fué un amigo incondicional todo el tiempo dentro y fuera de las aulas. Gracias a los dos por maximizar mi amor por la matemática.

También, agradezco a Lorena Sosa por estar constantemente en el camino y ser la mejor de las amigas, a Marcela Beltran por aportar, desde un principio tanto para el desarrollo de este trabajo como para mi vida personal. Por último, gracias a Martín Puentes, por que desde que llegó, su apoyo fué de mucha ayuda y contribuyo para la finalización de esta monografía a pesar de no pertenecer a esta área.

Finalmente, doy gracias a mis compañeros de universidad y a los docentes que también fueron partícipes en el transcurso de toda mi carrera que hoy culmina.

Muchas gracias a todos.

# Índice general

# Introducción

El estudio de polinomios ortogonales, se remonta al primer tercio del sigo $XX$ gracias al trabajo de Legendre sobre el movimiento planetario. Tiene importantes aplicaciones en la física, en la probabilidad, en los procesos estocásticos y en otras ramas de la matemática. Existen muchos tipos de polinomios ortogonales, como por ejemplo los de Legendre, los de Jacobi o los de Chebyshev, entre otros, éstos se encuentran en la recta real y tienen versión en el plano complejo. El desarrollo de la teoría de polinomios ortogonales en la circunferencia de radio uno, se desarrolla en la década de los sesenta, gracias a los trabajos de Gabor Szegö, y retomada más tarde por Mourad. E. Ismail. En los últimos tiempos, tales polinomios han sido de gran útilidad en campos modernos como el procesamiento de señales digitales, la teoría de operadores y el cálculo de probabilidades; estos han hecho del mismo, un importante campo de investigación matemática. Como se afirma en [5] y [8].

Este trabajo, es el desarrollo del capítulo $XI$ del libro "Polinomios Ortogonales"de Szegö [7], titulado "Polinomios ortogonales en el círculo unitario". Con esto en mente, se estudiaron los conocimientos necesarios para entender, tanto la teoria general de polinomios ortogonales, como el caso particular de los polinomios ortogonales en el círculo unitario. Con el fin de hacer el desarrollo completo del capítulo, se emplearon las propiedades de los polinomios ortogonales en la recta y de los polinomios ortogonales en el círculo unitario; para entender como se relacionan éstos polinomios entre sí.

Para el estudio de dicho capítulo, se propuso lo siguiente:

- Detallar los conceptos necesarios para entender la teoría general de polinomios ortogonales.

- Estudiar las propiedades generales de los polinomios ortogonales en la recta real.
- Explicar la relación entre polinomios ortogonales en el círculo unitario y polinomios ortogonales en la recta real.

Se hizo un seminario en el que se estudiaron polinomios, de manera general, algunas de sus propiedades y en donde se conocieron algunos ejemplos de polinomios ortogonales junto con sus funciones peso y demas. Luego, se eligió trabajar con base en el libro [7], ya que durante todo el seminario, el estudio se basó en ese y en [5], pero me llamó la atención el capítulo $XI$ del primer libro, pues nombraba polinomios que no se habian estudiado hasta el momento. Esta información se presenta por capítulos organizados así: En el primer capítulo, se introducen los conceptos necesarios del álgebra lineal y de polinomios, del análisis funcional y de teoría de la medida; cuya intención es facilitar el entendimiento del segundo capítulo en el que se encuentra el ejemplo de los polinomios de Chebyshev (los cuales son de suma importancia en el tercer capítulo) también están los conceptos básicos y las propiedades generales de polinomios ortogonales en la recta real. Finalmente, en el tercer capítulo se encuentran las definiciones necesarias para mostrar la teoria general de polinomios ortogonales, aterrizada a los polinomios en el círculo unitario; se hace uso de lo visto en el capítulo dos para hacer un contraste de la teoría de los dos casos y por último, a través de los polinomios de Chebyshev, se explica la relación que tienen entre sí.

Para terminar, se encuentran las conclusiones a las que se llegaron, teniendo en cuenta los objetivos planteados, y la bibliografía empleada durante todo el trabajo.

# CAPÍTULO 1

## Preliminares

En este capítulo, se presentan los conceptos, definiciones, teoremas y ejemplos necesarios para comprender los temas previos al estudio de los polinomios ortogonales en el circulo unitario. Los temas base que se encontraran en éste primer capitulo son: Álgebra lineal, polinomios, análisis funcional y teoría de la medida.

Una vez terminado éste capitulo, se tienen los conocimientos adecuados para la comprensión de toda la teoría sobre polinomipos ortogonales y su debida relevancia.

## 1.1. Álgebra Lineal y Polinomios

Además de contener definiciones y teoremas básicos e importantes para toda la teoría de polinomios y la matemática en general; esta sección contiene ejemplos que permiten introductir a los *polinomios* y al espacio de funciones en el que están junto con sus operaciones básicas. PEl estudio de ésta sección, se hizo con base en el libro [1] y todo lo que se muestra acontinuación, fue tomado del mismo.

**Definición 1.** Un *espacio vectorial* consta de lo siguiente:

1. Un campo K de escalares;

2. un conjunto V de objetos llamados vectores;

3. una operación llamada *adición*, que asocia a cada par de vectores $x, y$ de $V$ un vector $x + y$ de $V$, tal que

   a) la adición es conmutativa, $x + y = y + x$;

   b) la adición es asociativa $x + (y + z) = (x + y) + z$;

   c) existe un único vector 0 de V, llamado *vector nulo*, tal que $x + 0 = x$, para todo $x\ inV$;

   d) para cada vector $x \in V$, existe un único vector $-x \in V$, tal que $x + (-x) = 0$;

4. una operación llamada *multiplicación escalar*, que asocia a cada escalar de $\alpha \in F$ y cada vector $x \in V$ a un vector $\alpha x \in V$, de tal modo que

   a) $1x = x$ para tdo $x \in V$;

   b) $(\alpha\beta)x = \alpha(\beta x)$;

   c) $\alpha(x + y) = \alpha x + \alpha y$;

   d) $(\alpha + \beta)x = \alpha x + \beta x$.

En esta definición se observa como se establece un espacio vectorial como un objeto compuesto que consta de un campo, de un conjunto de *vectores* y de dos operaciones con propiedades especiales. De ahora en adelante, la notación $(V, F)$ representa al espacio vectorial $V$ sobre el campo $F$. Tomado de [1].

*Ejemplo* 1. (El espacio de funciones de un conjunto sobre un cuerpo)

Sean $F$ un campo y $X$ un conjunto no vacio. Sea $V$ el conjunto de todas las funciones de $X$ en $F$. La suma de dos vectores $f$ y $g$ de $V$ es el vector $f + g$; es decir, la función de $X$ en $F$ definida por:

$$(f + g)(x) = f(x) + g(x).$$

El producto del escalar $\alpha \in F$ y la función $f \in V$ es la función $\alpha f$ definida por

$$(\alpha f)(x) = \alpha f(x).$$

**Definición 2.** Sea $V$ un espacio vectorial sobre el campo $F$. Un *subespacio* de $V$ es un subconjunto no vacio $W$ de $V$ que, con las operaciones de adición vectorial y multiplicación escalar sobre $V$, es el mismo un espacio vectorial sobre $F$.

**Teorema 1.1.** *Un subconjunto no vacio $W$ de $V$ es un subespacio de $V$ si, y solo si, para todo par de vectores $x$, $y$ de $W$ y todo escalar $\alpha$ de $F$, el vector $\alpha x + y$ está en $W$.*

*Demostración.* (Tomada de [1] pp. 35) Sea $W$ un subconjunto no vacio de $V$ tal que $\alpha x + y$ pertenece a $W$ para todos los vectores $x$ y $y$ de $W$ y todos los escalares $\alpha$ de $F$. Como $W$ es no vacio, existe un vector $z$ en $W$, y, por lo tanto, $(-1)z + z = 0$ está en $W$. Ahora bien, si $x$ es cualquier vector de $W$ y $\alpha$ cualquier escalar, el vector $\alpha x = \alpha x + 0$ está en $W$. En particular, $(-1)x = -x$ está en $W$. Finalmente, si $x$ y $y$ están en $W$, entonces $x + y = 1x + y$ está en $W$. Así, $W$ esun subespacio de $V$. Recíprocamente, si $W$ es un subespacio de $V$, $x$ y $y$ están en $W$ y $\alpha$ es un escalar, claramente $\alpha x + y$ está en $W$.

□

*Ejemplo* 2. (El espacio de las funciones polinomios sobre el campo $F$) Sea $F$ un campo y sea $V$ el conjunto de todas las funciones $f$ de $F$ en $F$ definidas de la forma

$$f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n \tag{1.1}$$

donde $\alpha_0, \alpha_1, \cdots, \alpha_n$son escalares fijos de $F$. Una función de este tipo se llama *función polinomio sobre $F$*. Sean la adición y la multiplicación escalar definidas como en el ejemplo 1. se observa que si $f$ y $g$ sn funciones polinomios y $\alpha$ está en $F$, entonces $f + g$ y $\alpha f$ también son funciones polinomiales.

El espacio de las funciones polinomios sobre el campo $F$ es un subespacio del espacio de todas las funciones de $F$ en $F$.

**Definición 3.** Un vector $x$ del espacio vectorial $V$ es una *combinación lineal* de los vectores $\alpha_1, \cdots, \alpha_n$ en $V$, si existen escalares $\alpha_1, \cdots, \alpha_n$ de $F$ tales que

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n = \sum_{i=1}^{n} \alpha_i x_i.$$

**Definición 4.** El *subespacio generado* por un subconjunto $S$ no vacio de un espacio vectorial $V$ es el conjunto de todas las combinaciones lineales de los vectores de $S$. El subcojunto generado, es notado por

$$span\ S \qquad \text{o} \qquad gen\ S$$

**Definición 5.** Sea $V$ un espacio vectorial sobre $F$. Un subconjunto $S$ de $V$ es *linealmente dependiente* si existen vectores distintos $x_1, x_2, \cdots, x_n$ de $S$ y escalares $\alpha_1, \alpha_2, \cdots, \alpha_n$ de $F$, no todos nulos, tales que

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0.$$

Un conjunto que no es linealmente dependiente es *linealmente independiente*.

Esta definición, y la siguiente, se extienden a espacion vectoriales de dimensión infinita, lo cual se estudiara en la siguiente sección de análisis funcional.

**Definición 6.** Sea $V$ un espacio vectorial. Una *base* de $V$ es un conjunto de vectores linealmente independiente de $V$ que genera el espacio $V$. El espacio $V$ es de *dimensión finita* si tienen una base finita.

*Ejemplo* 3. (En [1]) Sea $F$ un subconjunto de los números complejos, y sea $V$ el espacio de las funciones polinomios sobre $F$ como se planteó en el ejemplo 2. Sea $f_k(x) = x^k$ con $k = 0, 1, 2, \cdots$. El conjunto $\{f_0, f_1, \ldots\}$ es una base: Inicialmente, se demuestra que el conjunto $\{f_0, f_1, f_2, \ldots\}$ es linealmente independiente. Reescribiendo la función polinomio (1.1) en términos del conjunto dado, e igualando a cero, se obtiene

$$\begin{aligned} \alpha_0 f_0 + \alpha_1 f_1 + \ldots + \alpha_n f_n &= 0 \\ \alpha_0 + \alpha_1 x^1 + \ldots + \alpha_n x^n &= 0. \qquad \text{para todo} \quad x \in F \end{aligned} \tag{1.2}$$

En otras palabras, cada $x$ de $F$ es raíz del polinomio (1.2). Por el teorema fundamental del álgebra, se sigue que $\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0$. Finalmente, gracias a que la función polinomio (1.1) se puede escribir como (1.2), por medio del conjunt $\{f_0, f_1, f_2, \cdots\}$, es claro que dicho conjunto genera a $V$. Por lo tanto, como $V$ tiene una base infinita, se concluye que el espacop de las funciones polinmios sobre $F$ es de dimensión infinita.

**Definición 7.** Sean $V$y $W$ dos espacios vectoriales sobre el campo $F$. Una *transformación lineal* de $V$ en $W$, es una función $T$ de $V$ en $W$ tal que

$$T(\alpha x + y) = cT(x) + T(y),$$

para todos los vectores $x$ e $y$ de $V$ y todos los escalares $\alpha$ de $F$.

**Definición 8.** Si $V$ es un espacio vectorial sobre el cuerpo $F$, un *operador lineal* sobre $V$ es una transformación lineal de $V$ en $V$

### 1.1.1. Polinomios

Cómo se puede apreciar hasta el momento, el trabajo está encaminado hacia el estudio de los polinomios, en particular, el de los polinomios ortogonales en el círculo unitario; por lo tanto, es importante resaltar aspectos importantes de la teoría general de polinomios que permiten el desarrollo de conceptos y teoremas más elaborados en cuanto a éste tema. A continuación se estudian algunas definiciones y propiedades importantes de los polinomios que serán usadas más adelante. En [1].

**Definición 9.** Sea $F[x]$ el subespacio del espacio de funciones de un conjunto en un campo, $F$, generado por los vectores $1,\ x,\ x^2, \cdots$. Un elemento de $F[x]$ se llama *polinomio sobre F*

Cómo $F$ consta de todas las combinaciones lineales (finitas) de $x$ y sus potencias, un vector no nulo $x$ de $F$ es un polinomio si, y solo si, existe un entero $n \geq 0$ tal que $\alpha_n \neq 0$ y tal que $\alpha_k = 0$ para todos los enteros $k > n$ cuando este entero existe y es único, se llama el *grado* de $f$.

Un polinomio no nulo $f$ de grado $n$ tal que $\alpha_n = 1$ se dice que es un *polinomio mónico*.

**Teorema 1.2.** *Sean $f$ y $g$ polinomios no nulos sobre $F$. Entonces*

*(i)* $fg$ *es un polinomio no nulo;*

*(ii)* $grad\ (fg) = grad\ (f) + grad\ (g)$;

*(iii)* $fg$ *es un polinomio mónico si ambos,* $f$ *y* $g$*, son polinomios mónicos;*

*(iv)* $fg$ *es un polinomio escalar si, y solo si, ambos* $f$ *y* $g$ *son polinomios escalares;*

*(v)* *si* $f + g \neq 0$,
$$grad(f+g) \leq max(grad\ f,\ grad\ g).$$

**Teorema 1.3.** *Sean $F$ un campo y $X$ un espacio vectorial sobre $F$. Se supone que $f$ y $g$ son polinomios sobre $F$, que $x$ es un elemento de $X$ y que $\alpha$ pertenece a $F$. Entonces*

*(i)* $(\alpha f + g) = \alpha f(x) + g(x)$;

*(ii)* $(fg) = f(x)g(x)$.

Para detalles de las demostraciones de los teoremas 1.2 y 1.3 ver [1].

**Definición 10.** Sea $F$ un campo. Un elemento $\alpha$ de $F$ se dice *raíz*, o un *cero*, de un polinomio dado $f$ sobre $F$ si $f(\alpha) = 0$. Si $\alpha$ es una raíz del polinomio $f$, la *multiplicidad* de la raíz $\alpha$ de $f$ es el mayor entero positivo $r$ tal que $(x - \alpha)^r$ divide a $f$.

**Teorema 1.4.** *Un polinomio $f$ de grado $n$ sobre un cuerpo $F$ tiene a lo más $n$ raíces en $F$.*

**Teorema 1.5.** *Si $F$ es un campo, un polinomio mónico no escalar en $F[x]$ puede descomponerse en el producto de polinomios mónicos irreducibles en $F[x]$ de manera única, salvo en lo que respecta al orden.*

En la factorización de un polinomio no escalar $f$ dado, algunos de los factores mónicos pueden estar repetidos. Si $p_1, p_2, \cdots, p_r$ son los polinomios irreducibles mónicos distintos que aparecen en ésta factorización de $f$, entonces

$$f = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r},$$

siendo el exponente $n_j$ el número de veces que el primo $p_j$ aparece en la factorización. Ésta descomposióción es también única y se llama *descomposición prima* de $f$.

**Polinomios Trigonométricos**

**Definición 11.** Un *polinomio trigonométrico* es una suma finita de la forma

$$f(\theta) = a_0 + \sum_{m=1}^{N} (a_m \cos m\theta + b_m \operatorname{sen} m\theta) \qquad \theta \in \mathbb{R} \tag{1.3}$$

con coeficientes arbitrarios complejos. Éste polinomio es de grado $m$ si $|a_m| + |b_m| > 0$.

Teniendo en cuenta la igualdad

$$e^{im\theta} = \cos m\theta + i \operatorname{sen} m\theta$$

el polinmio (1.3) puede escribirse de la forma

$$f(\theta) = \sum_{m=0}^{N} c_n e^{im\theta} \qquad \theta \in \mathbb{R}$$

que, en la mayoría de casos, es mas conveniente.

**Definición 12.** Sea $\rho(x)$ un polinomio de grado $m$, el polinomio recíproco de $\rho(x)$ está definido por

$$\rho^*(x) = x^m \overline{\rho}(x^{-1}). \tag{1.4}$$

Si los ceros de $\rho(x)$ son $x_1, x_2, \cdots, x_m$, los de $\rho^*(x)$ son $x_1^*, x_2^*, \cdots, x_m^*$, donde $x_\mu^* = \overline{x_\mu}^{-1}$ es el punto obtenido de $x_\mu$ por inversión con respecto al círculo unitario $|x| = 1$ en el plano complejo. Los ceros pueden ser contados de acuerdo a su multiplicidad, y $0^* = \infty, \infty^* = 0$; $\infty$ como un cero de orden $k$ significa que los coeficientes de potencias de rden $k$ desaparecen.Ésta y la definición 11 fue tomado de [7].

## 1.2. Análisis Funcional

Acontinuación, se definen algunos conceptos importantes tanto para este trabajo como para la matemática en general, ya que la *norma* es muy utilizada para operaciones algebraicas de espacios vectoriales y los *espacios normados*, son probablemente el tipo de espacio más importante en el análisis funcional (al menos desde el punto de vista de las últimas aplicaciones). Adicionalmente se introduce el espacio en el que se desarrolla todo el trabajo de los polinomios ortogonales, junto con sus propiedades, su producto interno y finalmente se define la ortogonalidad, se dan algunos ejemplos y desigualdades importantes para el resto del documento. Para el estudio de esta sección, se consultó [2], [3] y [9]. La traducción de los textos bibliográficos fue hecha por la autora de este trabajo.

**Definición 13.** Una *norma* en un espacio vectorial (real o complejo) $X$ es una función a valor real en $X$ cuyo valor en un $x \in X$ está denotado por

$$||x||.$$

Un *espacio normado* $X$ es un espacio vectorial con una norma definida en él. Las propiedades de la norma, son:

1. $||x|| \geq 0$
2. $||x|| = 0 \Longleftrightarrow x = 0$

3. $||\alpha x|| = |\alpha|||x||$

4. $||x + y|| \leq ||x|| + ||y||$

aqui, $x$ y $y$ son vectores arbitrarios en $X$ y $\alpha$ es cualquier escalar. Una norma en $X$ define una métrica $d$ en $X$ la cual está dada por

$$d(x, y) = ||x - y||$$

y es llamada *la métrica inducida por la norma*. El espacio normado es notado por $(X, ||\cdot||)$.

Un *espacio de Banach* es un espacio normado completo (completo en la métrica definida por la norma, como se explicó anteriormente).

*Ejemplo* 4. El espacio vectorial de todas las funciones contínuas a valor real en $[a, b]$ forma un espacio normado $X$ con norma definida por

$$||x|| = \left( \int_a^b x(t)^2 \, dt \right)^{1/2}. \tag{1.5}$$

Como este espacio no es completo, por medio del teorema de completacion (Teorema 1.6-2 en [2] pp. 41), este espacio queda completado y es notado por $L^2[a,\ b]$; este es un espacio de Banach; de hecho, la norma en $x$ y las operaciones de espacio vectorial pueden ser extendidas a la completación de $X$ gracias al teorema de completación para espacios de Banach (Teorema 2.3-2 en [2] pp. 69). Entonces, para un número real fijo $p \geq 1$, el espacio de Banach

$$L^p[a,\ b]$$

es la completación del espacio normado de todas las funciones contínuas a valor real en $[a,\ b]$, como antes, y la norma definida por

$$||x||_p = \left( \int_a^b |x(t)|^p \, dt \right)^{1/p}. \tag{1.6}$$

El subíndice $p$ es para indicar que esa norma depende del $p$ que se escoja, que se mantiene fijo. Observese que para $p = 2$, (1.6) es igual a (1.5).

El espacio $L^p[a,\ b]$ también puede ser obtenido de forma directa con el uso de la integral de Lebesgue y las funciones medibles en el sentido de Lebesgue con $x$ en $[a,\ b]$, tal que la integral de Lebesgue de $|x|^p$ sobre $[a,\ b]$ existe y es finita. Los elementos de $L^p[a,\ b]$ son clases de equivalencia de esas funciones, donde $x$ es equivalente a $y$ si la integral de Lebesgue de $|x-y|^p$ sobre $[a,\ b]$ es cero; lo cual garantiza que el axioma 2 de espacios normados; pero ésto se eviencia con más detalle en la siguiente sección.

Para ver éste ejemplo con más detalle, remitase a [2] pp. 62.

**Definición 14.** Un *funcional* es un operador cuyo rango se encuentra en la linea real $\mathbb{R}$ o en el plano complejo $\mathbb{C}$.

**Definición 15.** Un *funcional lineal* $f$ es un operador lineal con dominio en un espacio vectorial $X$ y rango en un campo escalar $K$ de $X$; así,

$$f : X \to K$$

donde $K = \mathbb{R}$ si $X$ es real, y $K = \mathbb{C}$ si $X$ es complejo.

*Ejemplo* 5. (La integral es un funcional lineal) La *integral definida* es un número si se considera para una sola función, como se hace en el cálculo todo el tiempo. Sin embargo, la situación cambia completamente cuando se considera la integral, para todas las funciones en cierto espacio de funciones. Entonces, la integral es un funcional $f$ en ese espacio. Como el espacio escogido es $C[a,\ b]$; entonces $f$ está definido por

$$\begin{aligned} f : C[a,b] &\longrightarrow \mathbb{R} && (1.7)\\ x &\longrightarrow f(t) = \int_a^b x(t)\,dt \qquad x \in C[a,\ b] && (1.8) \end{aligned}$$

la cual es lineal.

**Definición 16.** Un *producto interno* en $X$ es una función de $X \times X$ al campo escalar $K$ de $X$; es decir, que para cada par de vectores $x$ y $y$ existe un escalar asociado notado como

$$\langle x, y \rangle,$$

llamado el *producto interno* entre $x$ e $y$ . Un *espacio con producto interno* es un espacio vectorial $(X, K)$ con un producto interno definido en él. Para todos los vectores $x$, $y$, $z$ y el escalar $\alpha$; el producto interno cumple las siguientes propiedades:

1. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

2. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$

3. $\langle x, y\rangle = \overline{\langle y, x\rangle}$

4. $\langle x, x\rangle \geq 0$
   $\langle x, x\rangle = 0 \Longleftrightarrow x = 0.$

Un producto interno en $X$ define una norma en $X$ dada por

$$||x|| = \sqrt{\langle x, x\rangle}$$

y una métrica en $X$ dada por

$$d(x,y) = ||x-y|| = \sqrt{\langle x-y, x-y\rangle}.$$

Un *espacio de Hilbert* es un espacio con producto interno completo (completo en la métrica definida por el producto interno).

**Teorema 1.6.** *Una norma, en un espacio con producto interno, satisface la importante igualdad del paralelogramo*

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2)$$

**Definición 17.** Un elemento $x$ de un espacio con producto interno $X$ se dice que es *ortogonal* a un elemento $y \in X$ si

$$\langle x, y\rangle = 0$$

en este caso, se dice que $x$ y $y$ son ortogonales y se escribe $x \perp y$.

*Ejemplo* 6. La norma en el ejemplo 4, definida por (1.5), puede ser obtenida del producto interno definido por

$$\langle x, y\rangle = \int_a^b x(t)y(t)\, dt.$$

En el ejemplo 4, por simplicidad, se asume que las funciones eran a valor real; pero en este caso, es mejor remover esa restricción y considerar funciones a valor complejo (conservando $t \in [a, b]$ real, como antes). Esas funciones forman un espacio vectorial complejo, el cual es un espacio con producto interno si se define

$$\langle x, y \rangle = \int_a^b x(t)\overline{y(t)}\, dt.$$

Con $\overline{y(t)}$ como la conjugada compleja de la función $y(t)$. De aqui, la propiedad 3 del producto interno, se satisface y se puede concluir que $\langle x, x \rangle$ también es real. Esta propiedad es necesaria en conexión con la norma definida por

$$||x|| = \left( \int_a^b |x(t)|^2\, dt \right)^{1/2},$$

ya que $x(t)\overline{x(t)} = |x(t)|^2$.

Gracias a que se puede externder el producto interno a un espacio con producto interno, por el teorema de completación para el producto interno (Teorema 3.2-3 en [2] pp. 139) se concluye que $L^p[a, b]$ es un espacio de Hilbert.

**Lema 1.1.** *Un espacio con producto interno y su respectiva norma, satisfacen la desigualdad de Schwarz y la desigualdad triangular como sigue:*

1. *Se tiene*

   $$|(x, y)| \leq ||x||\; ||y|| \qquad \textit{Desigualdad de Schwarz}$$

   *donde la desigualdad se cumple si y solo si { x, y } es un conjunto linealmente dependiente.*

2. *La norma también satisface*

   $$||x + y|| \leq ||x|| + ||y|| \qquad \textit{Desigualdad triangular}$$

   *donde la desigualdad se cumple si y solo si $y = 0$ o $x = cy$ (c real y mayor o igual a cero).*

**Definición 18.** Un *conjunto ortogonal* $M$ en un espacio con producto interno $X$ es un subconjunto $M \subset X$ cuyos elemento son ortogonales dos a dos. Un *conjunto ortonormal* $M \subset X$ es un conjunto ortogonal en $X$ cuyos elementos son de norma 1, esto es, para todo $x, y \in M$

$$\langle x, y \rangle = \begin{cases} 0 & \text{si} \quad x \neq y \\ 1 & \text{si} \quad x = y. \end{cases}$$

Si un conjunto ortogonal u ortonormal $M$ es contable, se puede organizar una sucesión $(x_n)$ y es llamada *sucesión ortogonal* u *ortonormal* respectivamente.

Generalmente un conjunto indexado, o una *familia*, $(x_\alpha)$, $\alpha \in I$, es llamada *ortogonal* si $x_\alpha \perp x_\beta$ para todo $\alpha, \beta \in I$, $\alpha \neq \beta$. La familia es llamada *ortonormal* y es ortogonal y cada $x_\alpha$ tiene norma 1, entonces, para todo $\alpha, \beta \in I$, se tiene

$$\langle x_\alpha, x_\beta \rangle = \delta_{\alpha\beta} = \begin{cases} 0 & \text{si} \quad \alpha \neq \beta \\ 1 & \text{si} \quad \alpha = \beta. \end{cases}$$

Aqui, $\delta_{\alpha\beta}$ es la *delta de Kronecker*.

**Teorema 1.7.** *Si $x, y$ son ortogonales en un espacio con producto interno $X$ entonces,*

$$||x + y||^2 = ||x||^2 + ||y||^2 \tag{1.9}$$

*llamada relación pitagórica.*

Ésto se extiende de manera general con más de dos elementos, como será utilizado más adelante.

**Lema 1.2.** *Un conjunto ortonormal es linealmente independiente.*

*Ejemplo* 7. Sea $X$ el espacio con producto interno de todas las funciones contínuas a valor real en $[0, 2\pi]$ con el producto interno definido por

$$\langle x, y \rangle = \int_0^{2\pi} x(t)y(t)\, dt.$$

Dos sucesiones ortogonales en $X$ son $(u_n)$ y $(v_n)$, donde

$$\begin{array}{rcll} u_n(t) & = & cos\ nt & n = 0, 1, 2, \cdots \\ v_n(t) & = & sen\ nt & n = 1, 2, \cdots , \end{array}$$

de hecho; por integración, se tiene

$$\langle u_m, u_n \rangle = \int_0^{2\pi} cos\ mt\ cos\ nt\ dt = \begin{cases} 0 & \text{si } m \neq n \\ \pi & \text{si } m = n = 1, 2, \cdots \\ 2\pi & \text{si } m = n = 0. \end{cases}$$

Análogamente para $\langle v_n \rangle$. Por lo tanto, una sucesión ortonormal es $(e_n)$, donde

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \qquad e_n(t) = \frac{u_n(t)}{||u_n||} = \frac{cos\ nt}{\sqrt{\pi}} \qquad n = 1, 2, \cdots .$$

Para $(v_n)$ se obtiene la sucesión ortnormal $(\tilde{e}_n)$, donde

$$\tilde{e}_n(t) = \frac{u_n(t)}{||u_n||} = \frac{cos\ nt}{\sqrt{\pi}} \qquad n = 1, 2, \cdots .$$

Inclusive, también se tiene que $u_m \perp v_n$ para todo $m$ y $n$. (Tomado de [2] pp. 154).

Una gran ventaja de las sucesiones ortonormales sobre sucesiones arbitrarias linealmente independientes, es la siguiente:

Si se sabe que un $x$ dado, puede ser representado como una combinación lineal de algunos elementos de una sucesión ortonormal, entonces la ortonormalidad hace que la determinación real de los coeficientes sea muy facil. De hecho, si $(e_1, e_2, \cdots)$ es una sucesión ortonormal en un espacio con producto interno $X$ y se tiene que $x \in span\{e_1, \cdots, e_n\}$ donde $n$ es fijo, entonces por la definición del generado,

$$x = \sum_{k=1}^{n} \alpha_k\ e_k,$$

y si se toma el producto interno con un $e_j$ fijo, se obtiene

$$\langle x, e_j \rangle = \left( \sum_{k=1}^{n} \alpha_k \, e_k, e_j \right) = \sum_{k=1}^{n} \alpha_k \, \langle e_j, e_k \rangle = \alpha_j.$$

Por lo tanto, los coeficientes de la primera sumatoria se vuelven

$$x = \sum_{k=1}^{n} \langle x, e_k \rangle e_k.$$

Lo cual muestra que la determinación de los coeficientes de la primera sumatoria es sencilla.

Otra ventaja de la ortonormalidad se vuelve aparente si, en las anteriores sumatorias, se adiciona otro término $\alpha_{n+1} \, e_{n+1}$; para cuidar de un

$$\tilde{x} = x + \alpha_{n+1} \, e_{n+1} \in span\{e_1, \cdots, e_{n+1}\};$$

es necesario calcular solo un coeficiente más, ya que los otros permanecen iguales.

**Teorema 1.8.** *Sea $(e_k)$ una sucesión ortornomal en un espacio con producto interno $X$. Entonces para todo $x \in X$*

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq ||x||^2 \qquad \textit{Desigualdad de Bessel.} \tag{1.10}$$

*Donde $\langle x, e_k \rangle$ son llamados los coeficientes de Fourier de $x$ con respecto a la sucesión ortonormal $(e_k)$.*

### Proceso de ortogonalización de Gram-Schmidt

Se ha visto que las sucesiones ortogonales son muy convenientes para trabajar; pero el "problema"practico restante es como obtener una sucesión ortogonal si tenemos una sucesion arbitraria linealmente independiente dada. Esto fué solucionado con un procedimiento constructivo, el *proceso de ortogonalización de* $Gram - Schmidt$ para ortogonalizar sucesiones $(x_j)$ lienalmente independientes en un espacio con producto interno. La sucesión ortogonal resultante $(e_j)$ tiene la propiedad de, que para todo $n$,

$$span\{e_1, \cdots, e_n\} = span\{x_1, \cdots, x_n\}$$

este proceso se desarrolla de la siguiente manera:

El primer elemento de $(e_k)$ es

$$e_1 = \frac{1}{||x_1||} x_1.$$

El segundo elemento de la base $x_2$, puede ser escribirse como

$$x_2 = \langle x_2, e_1 \rangle e_1 + v_2$$

así

$$v_2 = x_2 - \langle x_2, e_1 \rangle e_1$$

no es un vector cero ya que $(x_j)$ es linealmente independiente; también $v_2 \perp e_1$ ya que $(v_2, e_1) = 0$, entonces, el segundo elemento de $(e_k)$ es

$$e_2 = \frac{1}{||v_2||} v_2.$$

De manera general, para el $n-$ésimo paso, se escribe

$$v_n = x_n \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k$$

así, el $n-$ésimo término de $(e_k)$ es

$$e_n = \frac{1}{||v_n||} v_n.$$

**Definición 19.** Si la desigualdad de Bessel (1.10) se reduce a la igualdad, para todo $x \in X$, es decir

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 = ||x||^2; \tag{1.11}$$

se dice que la *relación de Parseval* se satisface.

**Teorema 1.9.** *(Teorema de Riesz) Sean $H$ un espacio de Hilbert y $K$ el campo, $f : H \longrightarrow K$ un funcional lineal acotado; entonces, para todo $x \in H$ existe un $z \in H$ tal que el funcional puede ser representado en terminos del producto interno como*

$$f(x) = \langle x, z \rangle$$

*donde $z$ depende de $f$, está univocamente determinado por $f$ y tiene norma*

$$||z|| = ||f||.$$

**Definición 20.** Sean $X$ y $Y$ espacios vectoriales sobre el mismo campo $K$ ($\mathbb{R}$ o $\mathbb{C}$). Entonces una *forma sesquilineal* (o un *funcional sesquilineal*) $h$ en $X \times Y$ es una función

$$h : X \times Y \longrightarrow K$$

tal que para todo $x$, $x_1$, $x_2 \in X$ y $y$, $y_1$, $y_2 \in Y$ y todos los escalares $\alpha$, $\beta$

$$\begin{aligned} a)\ h(x_1 + x_2, y) &= h(x_1, y) + h(x_2, y) \\ b)\ h(x, y_1 + y_2) &= h(x, y_1) + h(x, y_2) \\ c)\ h(\alpha x, y) &= \alpha h(x, y) \\ d)\ h(x, \beta y) &= \overline{\beta} h(x, y). \end{aligned}$$

**Teorema 1.10.** *(Teorema de representacion de Riesz) Sean $H_1$ y $H_2$ espacios de Hilbert y*

$$h : H_1 \times H_2 \longrightarrow K$$

*una forma sesquilineal acotada. Entonces h tiene representación*

$$h(x, y) = (Sx, y)$$

*donde $S : H_1 \longrightarrow h_2$ es un operador lineal acotado. S está univocamente determinado por h y tiene norma*

$$||S|| = ||h||$$

Esta y la anterior demostración, se encuentran en [2].

**Definición 21.** Sea $X$ un espacio vectorial sobre un campo $K$. Una *forma Hermitiana* $h$ en $X \times X$ es una función

$$h : X \times X \longrightarrow K$$

tal que para todo $x, y, z \in X$ y $\alpha in K$,

$$\begin{aligned} h(x+y,z) &= h(x,z)+h(y,z) && (1.12) \\ h(\alpha x,y) &= \alpha h(x,y) && (1.13) \\ h(x,y) &= \overline{h(y,x)} && (1.14) \end{aligned}$$

Un caso particular que se puede deducir de la definición anterior es cuando la matríz $A = (a_{j,k})$, $1 \leq j, k \leq n$ cumple la igualdad

$$\overline{a_{j,k}} = a_{k,j},$$

en éste caso, se dice que la matríz $A$ es *Hermitiana*.

En adelante, se usa el producto interno

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j \overline{y_j}, \quad x = (x_1, \ldots)^T, \quad y = (y_1, \ldots)^T,$$

en el espacio complejo $\mathbb{C}$ de dimensión $n$; donde $A^T$ es la matríz traspuesta de $A$.

Claramente,

$$\langle x, y \rangle = \overline{\langle y, x \rangle}, \qquad \langle ax, y \rangle = a \langle x, y \rangle, \quad a \in \mathbb{C}.$$

La matríz que satisface

$$\langle Ax, y \rangle = \langle x, A^* y \rangle \tag{1.15}$$

es la *matríz adjunta* de $A$; de aca se concluye que si $A = (a_{j,k})$ entonces $A^* = (\overline{a_{k,j}})$. Así, $A$ es Hermitiana sí, y solo si $A = A^*$.

Cualquier matríz Hermitiana genera una forma cuadrática

$$\sum_{j,k=1}^{n} a_{j,k}\overline{x_j}x_k; \tag{1.16}$$

recíprocamente, cualquier forma cuadrática con $a_{j,k} = a_{k,j}$ determina una matríz hermitiana por medio de

$$\sum_{j,k=1}^{n} a_{j,k}\overline{x_j}x_k = x^*Ax = \langle Ax, x\rangle.$$

En los espacios de Hilber de dimensión infinita, el adjunto es definido por (1.15), siempre que se satisfaga para todo $x, y$ en dicho espacio.

Una forma cuadrática (1.16) es *definidapositiva* si $\langle Ax, x\rangle > 0$ para cualquier $x \neq 0$. Tomado de [9] pp. 1.

**Teorema 1.11.** *(Teorema de Aproximación de Weierstrass) Sea $f(x)$ contínua sobre el intervalo $[a, b]$. Dado $\epsilon > 0$, existe un polinomio $P(x)$ tal que*

$$|f(x) - P(x)| < \epsilon \qquad \textit{para todo} x \in [a, b].$$

## 1.3. Aproximación por Polinomios

El objetivo de esta sección, es mostrar la forma más óptima para aproximar polinomios. La ventaja es que éstas aproximaciones puedes llevarse a polinomios ortogonales y, mejor aún, a polinomios ortogonales en el círculo unitario. Tomado de [3].

**Definición 22.** Sea $S$ un subespacio de dimensión finita de un espacio con producto interno $X$, y sea $\{e_1, e_2, \cdots, e_n\}$ una base ortonormal para $S$. Si $x \in X$; el elemento $s \in S$ definido por

$$s = \sum_{k=1}^{n} (x, e_k)e_k$$

se denomina la *proyección* de $x$ sobre el subespacio $S$.

Acontinuación se demuestra un teorema que indica que la proyección de $x$ sobre $S$ es el elemento más cercano de $S$ a $x$ que cualquier otro:

**Teorema 1.12.** *Sea $S$ un subespacio de dimensión finita de un espacio con producto interno $X$, y sea $x$ un elemento de $X$. la proyección de $x$ sobre $S$ es más próxima a $x$ que cualquier otro elemento de $S$. Esto es, si $s$ es la proyección de $x$ sobre $S$, tenemos*

$$||x-s|| \leq ||x-t||$$

*para todo $t \in S$*

*Demostración.* Gracias al teorema de descomposición ortogonal, (Teorema 1.15 [3] pp. 33 ) se puede escribir a $x$ como

$$x = s + s^{\perp} \qquad s \in S \text{ } y \text{ } s^{\perp} \in S^{\perp}.$$

Entonces, para cualquier $t \in S$, se tiene

$$\begin{aligned} x - t &= x - s + s - t \\ &= (x-s) + (s-t) \end{aligned} \tag{1.17}$$

como $(s-t) \in S$ y $(x-s) = s^{\perp} \in S^{\perp}$; (1.17) es una descomposición ortogonal de $x-t$, así que su norma viene de la relación pitagórica (1.9)

$$||x-t||^2 = ||x-s||^2 + ||s-t||^2$$

pero como $||s-t||^2 \geq 0$ por estar en $S$ entonces

$$||x-t||^2 \geq ||x-s||^2$$

por lo tanto

$$||x-s|| \leq ||x-t||.$$

□

Una idea básica del análisis numérico, es la de utilizar funciones sencillas de polinomios ordinarios para aproximar una función dada $f$.
Un tipo de aproximación por polinomios, es la llamada *aproximación por mínimos cuadrados*. En este caso, la función dada $f$ está definida y es integrable en el intervalo $[a, b]$ y se busca un polinomio $p(x)$ de grado menor o igual a $n$ tal que el *error cuadrático medio*

$$\int_a^b |f(x) - p(x)|^2 \, dx$$

sea lo menor posible.

Sea $C$ el espacio de funciones contínuas en el intervalo $[a, b]$ y sea $S$ un subespacio lineal que contiene todos los polinomios de grado menor o igual a $n$. Supongase también, que en $C$ se ha definido una norma. Se elije una función $f$ de $C$. Si existe un polinomio $P(x) \in S$ tal que

$$||f(x) - P(x)|| \leq ||f(x) - Q(x)||$$

para todo $Q \in S$, se dice que $P(x)$ es el *polinomio de aproximación óptima* para $f$ con cierto grado. La palabra "óptima" es relativa a la norma y se debe tener en cuenta que el polinomio de aproximación para una norma elegida, no necesariamente sigue siendo el polinomio 'optimo para otra norma.

Para encontrar dicho $P(x)$, aparecen tres problemas:

1. Existencia.
2. Unicidad.
3. Construcción.

Esto, pueden resolverse totalmente para cualquier norma inducida por un producto interno. En ese caso, por el teorema 1.12, existe un solo polinomio en $S$ para el que la norma $||f(x) - P(x)||$ es tan pequeña como se quiera. En efecto, este $P(x)$ es la pryección de $f$ sobre $S$ y viene dado por

$$P(x) = \sum_{k=0}^{n} \langle f, e_k \rangle e_k(x)$$

donde $\{e_0, e_1, \cdots, e_n\}$ es una base ortonormal, de funciones, para $S$.

**Teorema 1.13.** *Sea C el espacio lineal de las funciones contínuas en $[a, b]$ con producto interno definido por*

$$\int_a^b f(x)g(x)\,dx.$$

*Sea $\{e_0, e_1, \cdots, e_n\}$ una base ortonormal para el subespacio S de los polinomios de grado menor o igual a n. Sea $P(x)$ el polinomio en S de aproximación óptima para $f$ en C relativo a la norma cuadrática. Entonces el cuadrado de la normal del error, está dado por*

$$||f - P||^2 = ||f||^2 - \sum_{k=0}^{n} \langle f, e_k \rangle^2 \tag{1.18}$$

Tomado de [3] pp. 704.

## 1.4. Teoría de la Medida

Éste capítulo se termina con una parte importante y muy útil en los capítulos posteriores, ya que en esta sección se explica la medida que se utilizará más adelante para el producto interno entre polinomios ortogonales tanto en $\mathbb{R}$ como en el círculo unitario; también se explica la integral de Lebesgue, que es la que se utiliza todo el tiempo para los productos internos ya mencionados, junto con sus propiedades. Todo lo que se estudió en este capítulo fué tomado de [4] y la traducción estuvo a cargo de la autora de este trabajo.

**Definición 23.** Una familia $\mathbb{X}$ de subconjuntos de un conjunto $X$ se dice que es una *σ-álgebra* si:

1. y $X$ pertenecen a $\mathbb{X}$.
2. Si $A$ pertenece a $\mathbb{X}$, entonces, el complemento $A^c$ pertenece a $\mathbb{X}$.
3. Si $(A_n)$ es una sucesión de conjuntos en $\mathbb{X}$, entonces la unión $\cup_{n=1}^{\infty} A_n$ pertenece a $\mathbb{X}$.

**Definición 24.** Un par ordenado $(X, \mathbb{X})$, que consiste de un conjunto X y una *σ-álgebra* $\mathbb{X}$ de subconjuntos de $X$, es llamado un *espacio medible*. Y cualquier conjunto en $\mathbb{X}$ es llamado *conjunto medible*

**Definición 25.** Sea $A$ una colección, no vacia, de subconuntos de un conjunto $X$. observemos que existe la *σ-álgebra* más pequeña de subconjuntos de $X$ que contiene a $A$. Para ver esto, observe que la familia e todos los subconjuntos de $X$ es una *σ-álgebra* que contiene a $A$ y a la intersección de todas las *σ-álgebras* que contienen a $A$ es también una *σ-álgebra* que contiene a $A$. Esta *σ-álgebra* más pequeña, es llamada la *σ-álgebra generada por $A$*

**Definición 26.** Sea $X$ el conjunto $\mathbb{R}$ de todos los números reales. El *álgebra de Borel* es la $\sigma$-álgebra $B$ generada por todos los intervalos abiertos $(a, b) \in \mathbb{R}$. Observese que el álgebra de Borel $B$ también es la $\sigma$-álgebra generada por todos los intervalos cerrados $[a, b] \in \mathbb{R}$. Cualquier conjunto en $B$ es llamado *conjunto de Borel*.

**Definición 27.** Una función $f$ de $x$ a $\mathbb{R}$ se dice que es *medible* si para todo número real $\alpha$, el conjunto

$$\{x \in X : f(x) > \alpha\}$$

pertenece a $\mathbb{X}$.

**Lema 1.3.** *Las siguientes proposiciones son equivalentes entre si para una función $f$ de $X$ a $\mathbb{R}$:*

*a) Para todo $\alpha \in \mathbb{R}$, el conjunto $A_\alpha = \{x \in X : f(x) > \alpha\}$ pertenece a $\mathbb{X}$.*
*b) Para todo $\alpha \in \mathbb{R}$, el conjunto $B_\alpha = \{x \in X : f(x) \leq \alpha\}$ pertenece a $\mathbb{X}$.*
*c) Para todo $\alpha \in \mathbb{R}$, el conjunto $C_\alpha = \{x \in X : f(x) \geq \alpha\}$ pertenece a $\mathbb{X}$.*
*d) Para todo $\alpha \in \mathbb{R}$, el conjunto $D_\alpha = \{x \in X : f(x) < \alpha\}$ pertenece a $\mathbb{X}$.*

**Definición 28.** Sean $X$ un conjunto y $E$ un subconjunto de $X$.La *función característica* $\chi_E$ está definida por

$$\chi_E(x) = \begin{cases} 1 & \text{si} \quad x \in E \\ 0 & \text{si} \quad x \notin E, \end{cases}$$

es medible. De hecho, $\{x \in X : \chi_E(x) > \alpha\}$ puede ser $X, E$ o .

**Definición 29.** Si $f$ es cualquier función de un conjunto $X$ a $\mathbb{R}$, sean $f^-$ y $f^+$ funciones no negativas definidas en X por

$$f^+(x) = sup\{f(x), 0\}, \qquad f^-(x) = sup\{-f(x), 0\}$$

La función $f^+$ es la *parte positiva* de $f$ y $f^-$ es la *parte negativa* de $f$.Es claro que

$$f = f^+ - f^- \quad \text{y} \quad |f| = f^+ + f^-$$

**Definición 30.** Una *medida* es una función a valor real extendido $\mu$ definida en una $\sigma$-álgebra $\mathbb{X}$ de subconjuntos de $X$ tal que

1. $\mu(\emptyset) = 0$,
2. $\mu(E) \leq 0$ para todo $E \in \mathbb{X}$,
3. Si $(E_n)$ es una sucesión disjunta en $\mathbb{X}$, entonces

$$\mu(\cup_{n=1}^{\infty}) = \sum_{n=1}^{\infty} \mu(E)$$

*Ejemplo* 8. Sea $X = R$ y $\mathbb{X} = B$, el álgebra de Borel, y sea $\lambda(E) = ba$ donde $E = (a, b)$, intervalo no vacío, entonces $\lambda$ es una medida, usualmente llamada la *medida de Lebesgue*.

**Definición 31.** Un *espacio de medida* es una tripla $(X, \mathbb{X}, \mu)$ que consiste de un arreglo $X$, una $\sigma$-álgebra $\mathbb{X}$ de subconjutos de $X$, y una medida $\mu$ definida en $X$

**Definición 32.** Una función a valor real es *simple* si tiene solo un numero de valores finito.

Una función medible simple $\varphi$ puede ser representada como:

$$\varphi = \sum_{j=1}^{n} a_j x_{E_j} \tag{1.19}$$

donde $a_j \in \mathbb{R}$ y $\chi_{E_j}$ es la función característica de un conjunto $E_j$ en $\mathbb{X}$

**Definición 33.** Si $\varphi$ es una función simple en $M^+(X, \mathbb{X})$ con la representación estandar (1.19), se define la integral de $\varphi$ con respecto a $\mu$ para ser el número real extenido

$$\int \varphi \, d\mu = \sum_{j=1}^{n} a_j \, \mu(E_j) \tag{1.20}$$

**Definición 34.** Si $f$ pertenece a $M^+(X, \mathbb{X})$, se define la integral de $f$ con respecto a $\mu$ para ser el número real extenido

$$\int f \, d\mu = sup \int \varphi \, d\mu, \tag{1.21}$$

donde el supremo es extendido sobre todas las funciones simples $\varphi$ en $M^+(X, X)$ satisfaciendo $0 \leq \varphi(x) \leq f(x)$ para toda $x \in \mathbb{X}$. Si $f$ pertenece a $M^+(X, X)$ y $E$ pertenece a $\mathbb{X}$, entonces $f(\chi_E)$ pertenece a $M^+(X, \mathbb{X})$ y se define la integral de $f$ sobre $E$ con respecto a $\mu$ para ser el número real extendido

$$\int_E f \, d\mu = \int f(x_E) \, d\mu. \tag{1.22}$$

Hasta el momento, queda clara la forma de calcular la Integral de Lebesgue de funciones medibles $f$ definidas positivas. A pesar de ello, en éste capítulo se planea presentar una definición todavía más general. Donde se presente la Integral de Lebesgue para funciones medibles no necesariamente definidas positivas, por tal motivo, se presenta primero la colección $L$ de funciones integrables, para luego, introducir la definición de la integral de Lebesgue de funciones $f \in L$ sin ningún probloema.

**Definición 35.** La coleción $L = L(X, \mathbb{X}, \mu)$ de funciones integrables (o sumables) consiste en todas las funciones medibles a valor real $f$ deinidas en $X$, tal que ambas, la parte positiva $f^+$ y negativa $f^-$ de $f$, tienen integrales finitas con respecto a $\mu$. En este caso, se define la integral de $f$ con respecto a $\mu$ para ser

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu. \tag{1.23}$$

Si $E$ pretenece a $\mathbb{X}$, se define

$$\int_E f \, d\mu = \int_E f^+ \, d\mu - \int_E f^- \, d\mu. \tag{1.24}$$

El siguiente resultado es conocido como la *propiedad de absoluta integrabilidad* de la integral de Lebesgue. Notese que, aunque el valor absoluto de una funcion integrable de Riemann (propia) es integrable en el sentido de Riemann, esto no puede externderse para el caso de una funcion con una integral impropia de Riemann.

**Teorema 1.14.** *Un multiplo constante $\alpha f$ y una sum $f+g$ de funciones en L pertenece a L y*

$$\int \alpha f \, d\mu \;=\; \alpha \int f \, d\mu, \qquad \int (f+g)\, d\mu \;=\; \int f \, d\mu + \int g \, d\mu \tag{1.25}$$

Pära detalles de esta demostración, ver [4] pp. 43.

### Los espacios $L_p$, $1 \leq p < \infty$

Ahora, se desea considerar una familia de espacios lineales normados relacionados con clases de equivalencia de funciones medibles

**Definición 36.** Si $1 \leq p < \infty$, el espacio $L_p = L_p(X, \mathbb{X}, \mu)$ consiste de todas las funciones de valor real medibles de clases $\mu$-equivalentes $f$ para la cual $|f|^p$ tiene finitas integrales con respecto a $\mu$ sobre X. Dos funciones son $\mu$-equivalentes si son iguales $\mu$-casi siempre. Se hace

$$||f||_p = \left( \int |f|^p \, d\mu \right)^{1/p} \tag{1.26}$$

### Medida de Lebesgue

Ahora, se describe la generación de una medida en la linea real $\mathbb{R}$. En el lema 9.3, en [4] pp. 97, se vio que el conjunto $F$ de todas las uniones finitas de conjuntos de la forma

$$(a,\, b], \quad (-\infty,\, b], \quad (a,\, +\infty), \quad (-\infty,\, +\infty), \tag{1.27}$$

eran un subconjunto de $R$ y que la función longitud $l$ daba una medida en esta algebra $F$. Si se aplica el procedimiento de extensión a $l$ y a $F$, se genera una espacio de medida $(R,\ F^*,\ l^*)$. La $\sigma$-algebra $F^*$ obtenida es esta construcción es llamada la colección de conjuntos medibles de Lebesgue y la medida $l^*$ en $F^*$ es llamada medida de Lebesgue. Aunque aveces se desea trabajar con $(R,\ F^*,\ l^*)$, es más conveniente tratar con la $\sigma$-algebra más pequeña que contiene a $F^*$ en lugar de trabajar con todos los que contienen a $F^*$. Es notable que esta $\sigma$-algebra más pequeña es igual a la colección de conjuntos de Borel. La restricción de medida de Lebesgue a los conjuntos de Borel es llamada *medida de Borel* o *medida de Lebesgue*.

# CAPÍTULO 2

## Polinomios Ortogonales

Teniendo en cuenta toda la teoría vista hasta el momento, ya es posible entender las propiedades con las que los polinomios ortogonales están dotados como pór ejemplo, su producto interno, la medida con la que se opera dicho producto y la definición como tal de esa operacion, entre otras cosas con las que cuentan los polinomios ortogonales. Para el desarrollo de este capitulo, se usaron somo base [5], [6], [7], [8] y [9].

## 2.1. Polinomios de Chebyshev

Los polinomios de Chebyshev o polinomios trigonométricos, nacen a partir de un ejercicio elemental del cálculo, por medio del uso de identidades trigonométricas; estos polinomios, reciben su nombre gracias al trabajo hecho por Pafnuti Chebyshev, quien logró escribir los polinomios trigonométricos de manera recursiva. El estudio de ésta sección, se basó en [5] y la traducción estuvo a cargo de la autora de este trabajo.

Con ésto en mente, al sumar las identidades

$$\begin{aligned} cos(m+n)\theta = cos(m\theta+n\theta) &= cos\ m\theta\ cos\ n\theta - sen\ m\theta\ sen\ n\theta \\ cos(m-n)\theta = cos(m\theta-n\theta) &= cos\ m\theta\ cos\ n\theta + sen\ m\theta\ sen\ n\theta \end{aligned}$$

se obtiene

$$2\cos m\theta \cos n\theta = \cos(m+n)\theta + \cos(m-n)\theta \tag{2.1}$$

al integrar, se tiene

$$\begin{aligned} \int_0^\pi (2\ cos\ m\theta\ cos\ n\theta)\ d\theta &= \int_0^\pi (cos(m+n)\theta + cos(m-n)\theta)\ d\theta \\ &= \int_0^\pi cos(m+n)\theta\ d\theta + \int_0^\pi cos(m-n)\theta\ d\theta \end{aligned}$$

haciendo las sustituciones

$$\begin{aligned} u &= (m+n)\theta & \qquad v &= (m+n)\theta \\ du &= (m-n)\ d\theta & \qquad dv &= (m-n)\ d\theta \end{aligned}$$

se tiene

$$\begin{aligned} \int_0^\pi (2\ cos\ m\theta\ cos\ n\theta)\ d\theta &= \int_0^\pi \cos u\ \frac{du}{m+n} + \int_0^\pi \cos v\ \frac{dv}{m-n} \\ &= \left[\frac{sen\ u}{m+n}\right]_0^\pi + \left[\frac{sen\ v}{m-n}\right]_0^\pi \\ &= \left[\frac{sen\ (m+n)\theta}{m+n}\right]_0^\pi + \left[\frac{sen\ (m-n)\theta}{m-n}\right]_0^\pi \\ &= \frac{sen\ (m+n)(\pi)}{m+n} - \frac{sen\ (m+n)(0)}{m+n} + \frac{sen\ (m-n)(\pi)}{m-n} - \frac{sen\ (m-n)(0)}{m-n} \\ &= 0 \end{aligned}$$

así

$$\begin{aligned}
\int_0^{\pi} (2\ cos\ m\theta\ cos\ n\theta)\ d\theta &= 0 \\
2\int_0^{\pi} (cos\ m\theta\ cos\ n\theta)\ d\theta &= 0 \\
\int_0^{\pi} (cos\ m\theta\ cos\ n\theta)\ d\theta &= 0 \qquad m \neq n, \quad m, n = 0,\ 1,\ 2, \ldots
\end{aligned} \tag{2.2}$$

De aqui, se puede decir que cos $m\theta$ y cos $n\theta$ son ortogonales sobre el intervalo $[0,\ \pi]$ para $m \neq n$

Entonces, de manera general, se concluye que

$$\{1,\ cos\ \theta,\ cos\ 2\theta, \ldots,\ cos\ n\theta\}$$

es una sucesión ortogonal sobre el intervalo $[0,\ \pi]$.

Observese que con el cambio de variable $x = \cos\theta$ se tiene:

$$\begin{aligned}
cos\ n\theta &= cos\ (n\ cos^{-1}\ (cos\ \theta)) \\
&= cos\ (n\ cos^{-1}\ x) \\
&= T_n(x)
\end{aligned} \tag{2.3}$$

$$\begin{aligned}
cos\ m\theta &= cos\ (m\ cos^{-1}\ (cos\ \theta)) \\
&= cos\ (m\ cos^{-1}\ x) \\
&= T_m(x),
\end{aligned} \tag{2.4}$$

y si se deriva a ambos lados de $x =$ cos $\theta$, se eleva al cuadrado y se usan identidades trigonométricas, se obtiene

$$
\begin{aligned}
x &= cos\ \theta \\
dx &= -sen\ \theta\ d\theta \\
(dx)^2 &= (-sen\ \theta\ d\theta)^2 \\
(dx)^2 &= sen^2\ \theta\ (d\theta)^2 \\
(dx)^2 &= (1 - cos^2\ \theta)\ (d\theta)^2 \\
(dx)^2 &= (1 - x^2)\ (d\theta)^2 \\
\sqrt{(dx)^2} &= \sqrt{(1 - x^2)\ (d\theta)^2} \\
dx &= \sqrt{(1 - x^2)}\ d\theta \\
\frac{(dx)^2}{\sqrt{(1 - x^2)}} &= d\theta \qquad\qquad -1 \leq x \leq 1 \qquad (2.5)
\end{aligned}
$$

Ahora, reemplazando (2.3), (2.4) y (2.5) en (2.2), queda

$$\int_{-1}^{1} T_m(x)\ T_n(x)(1 - x^2)^{-1/2} = 0 \qquad m \neq n, \quad m, n = 0,\ 1,\ 2, \ldots$$

Dándole valores fijos a $n$, haciendo uso de $x = \cos\ \theta$ y de algunas identidades trigonométricas, se obtiene:
si se hace $n = 0$:

$$
\begin{aligned}
T_0(x) &= cos\ 0\theta \\
&= 1
\end{aligned}
$$

si se hace $n = 1$

$$
\begin{aligned}
T_1(x) &= cos\ 1\theta \\
&= x
\end{aligned}
$$

si se hace $n = 2$

$$\begin{aligned}
T_2(x) &= cos\ 2\theta \\
&= cos\ (\theta + \theta) \\
&= cos\ \theta\ cos\ \theta - sen\ \theta\ sen\ \theta \\
&= cos^2\ \theta - sen^2\ \theta \\
&= cos^2\ \theta - (1 - cos\ \theta) \\
&= 2\ cos^2\ \theta - 1 \\
&= 2x^2 - 1
\end{aligned}$$

si se hace $n = 3$

$$\begin{aligned}
T_3(x) &= cos\ 3\theta \\
&= cos\ (2\theta + \theta) \\
&= cos\ 2\theta\ cos\ \theta - sen\ 2\theta\ sen\ \theta \\
&= (2\ cos^2\ \theta - 1)\ cos\ \theta - sen\ (\theta + \theta)\ sen\ \theta \\
&= (2\ cos^2\ \theta - 1)\ cos\ \theta - 2\ sen\ \theta\ \cos\ \theta\ sen\ \theta \\
&= (2\ cos^2\ \theta - 1)\ cos\ \theta - 2\ sen^2\ \theta\ \cos\ \theta \\
&= 2\ cos^3\ \theta - cos\ \theta - 2\ (1 - cos^2\ \theta)\ \cos\ \theta \\
&= 2\ cos^3\ \theta - cos\ \theta - 2\ cos\ \theta + 2\ cos^3\ \theta \\
&= 4\ cos^3\ \theta - 3\ cos\ \theta \\
&= 4x^3 - 3x
\end{aligned}$$

Con los ejemplos anteriores, se puede notar que $T_0(x)$, $T_1(x)$, $T_2(x)$ y $T_3(x)$ son de grados 0, 1, 2 y 3 respectivamente; de manera inductiva, se demuestra que $T_n(x)$ es de grado $n$ para todo $n = 0, 1, 2, 3, ...$

Ya se evidenció que se cumple para los primeros números naturales. Cómo hipótesis inductiva, se supone que el polinomio $T_k(x) = \cos\ k\theta$ es de grado $k$. De manera inductiva se demuestra que se cumple para $n = k + 1$:

$$T_{k+1}(x) = cos\ (k+1)\theta$$

usando la identidad (2.1) con $n = 1$ y $m = k$, se tiene:

$$\begin{aligned} 2\ cos\ k\theta\ cos\ \theta &= cos\ (k+1)\theta + cos\ (k-1)\theta \\ 2\ cos\ k\theta\ cos\ \theta - cos\ (k-1)\theta &= cos\ (k+1)\theta \end{aligned}$$

así

$$\begin{aligned} T_{k+1}(x) &= 2\ cos\ k\theta\ cos\ \theta - cos\ (k-1)\theta \\ &= 2\ T_k(x)\ cos\ \theta - T_{k-1}(x) \end{aligned}$$

Por hipótesis de inducción cos $k\theta$ es de grado $k$, al ser multiplicado por cos $\theta$, queda un polinomio de grado $k+1$, y como cos $(k-1)\theta$ es de grado $k-1$, si se le resta al polinomio que ya es de grado $k+1$ es irrelevante para el grado de $T_{k+1}$; por lo tanto $T_{k+1}$ es de grado $k+1$.

Estos polinomios son llamados los *polinomios de Chebyshev de primer orden*.

Otro ejemplo de polinomios ortogonales, son los *polinomios de Chebyshev de segundo orden*; los cuales se representan como:

$$U_n(x) = \frac{sin\ (n+1)\theta}{sin\ \theta} \qquad \text{con} \quad x = cos\ \theta \qquad n = 0,\ 1,\ 2, \ldots \tag{2.6}$$

Por inducción, se muestra que $U_n$ también es de grado $n$, así como con los $T_n$: Inicialmente, se demuestra que se cumple para $n = 1$

$$\begin{aligned} U_1(x) &= \frac{sin\ (1+1)\theta}{sin\ \theta} \\ &= \frac{sin\ 2\theta}{sin\ \theta} \\ &= \frac{2\ cos\ \theta\ sin\ \theta}{sin\ \theta} \\ &= 2\ cos\ \theta \\ &= 2\ T_1(x) \\ &= 2x \end{aligned}$$

luego, se cumple para $n = 1$. Ahora, se supone que se cumple para $n = k$, es decir

$$U_k(x) = \frac{sin\ (k+1)\theta}{sin\ \theta}$$

es de grado $k$. Inductivaménte, se demuestra que se cumple para $n = k+1$

$$U_{k+1}(x) = \frac{sin\ (k+2)\theta}{sin\ \theta}$$

Usando la identidad

$$\begin{aligned} sen\ (m+n)\theta + sen\ (m-n)\theta &= sen\ (m\theta + n\theta) + sen\ (m\theta - n\theta) \\ &= sen\ m\theta\ cos\ n\theta + sen\ n\theta\ cos\ m\theta + sen\ m\theta\ cos\ n\theta - sen\ n\theta\ cos\ m\theta \\ &= 2\ sen\ m\theta\ cos\ n\theta \end{aligned}$$

con $m = k$ y $n = 2$, se obtiene

$$\begin{aligned} sen\ (k+2)\theta + sen\ (k-2)\theta &= 2\ sen\ k\theta\ cos\ 2\theta \\ sen\ (k+2)\theta &= 2\ sen\ k\theta\ cos\ 2\theta - sen\ (k-2)\theta \end{aligned}$$

Así

$$\begin{aligned} U_{k+1}(x) &= \frac{2\ sen\ k\theta\ cos\ 2\theta - sen\ (k-2)\theta}{sin\ \theta} \\ &= \frac{2\ sen\ k\theta\ cos\ 2\theta - sen\ (k\theta - 2\theta)}{sin\ \theta} \\ &= \frac{2\ sen\ k\theta\ cos\ 2\theta - (sen\ k\theta\ cos\ 2\theta - sen\ 2\theta\ cos\ k\theta)}{sin\ \theta} \\ &= \frac{sen\ k\theta\ cos\ 2\theta}{sen\ \theta} + \frac{sen\ 2\theta\ cos\ k\theta}{sen\ \theta} \\ &= \frac{sen\ k\theta}{sen\ \theta}\ cos\ 2\theta + \frac{sen\ 2\theta}{sen\ \theta}\ cos\ k\theta \\ &= U_{k-1}(x)\ T_2(x) + U_1(x)\ T_k(x) \end{aligned}$$

Por hipótesis de inducción, se sabe que $\dfrac{\text{sen}\ k\theta}{\text{sen}\ \theta} = U_{k-1}(x)$ es de grado $k-1$ y por lo anterior, se tiene que $\cos\ 2\theta = T_2(x)$ es de grado 2, luego $\dfrac{\text{sen}\ k\theta}{\text{sen}\ \theta}\ \cos\ 2\theta$ es de grado $k+1$, análogamente, por el paso base $\dfrac{\text{sen}\ 2\theta}{\text{sen}\ \theta} = U_1(x)$ es de grado 1 y $\cos\ k\theta = T_k(x)$ es de grado $k$, así $\dfrac{\text{sen}\ 2\theta}{\text{sen}\ \theta}\ \cos\ k\theta$ es de grado $k+1$, entonces, al sumar esta con la fracción anterior se concluye que $U_{k-1}(x)$ es de grado $k+1$

Por lo tanto, (2.6) es de grado $n$.

## 2.2. Teoria elemental de polinomios ortogonales

En esta sección, se verá, de manera general, las propiedades que cumplen los polinomios ortogonales; como por ejemplo la operación con la que están dotados, es decir, el funcional de momentos, las relaciones de recurrencia de tres términos, la identidad de Christoffel-Darboux (Tomadas de [5] y [8]), los determinantes a los que están asociados y las propiedades de los mismos, los ceros de esos polinomios, el núcleo de polinomios, entre otros. Todo esto, basado en [5], [6], [7], [8] y [9].

Ya se tiene un breve acercamiento a los polinomios ortogonales y a sus operaciones. En la sección anterior, se mostró un importante ejemplo de polinomios ortogonales en $\mathbb{R}$; si se detalla bien la integral con la que se construyen estos polinomios, es facil evidenciar que a parte de los dos polinomios que se operan en la integral, la función $(1-x^2)^{-1/2}$ hace parte de la misma. A ésta función se le denomina *función peso*.

De manera general; sea $[a,b]$ un subconjunto de $\mathbb{R}$, fnito o infinito y sea $w(x)$ una función peso positiva e integrable en ese intervalo; de modo que

$$\int_a^b w(x)\,dx > 0, \tag{2.7}$$

dado el caso en el que $[a,b]$ sea acotado, es necesario que los *momentos*

$$\mu_n = \int_a^b x^n w(x)\,dx \qquad n = 0,1,2,\ldots \tag{2.8}$$

Aplicando la integral al polinomio

$$P_n(x) = \sum_{k=0}^{n} C_k\, x^k$$

se tiene

$$\begin{aligned}
\int_a^b [P_n(x)] w(x)\, dx &= \int_a^b \left[\sum_{k=0}^{n} C_k\, x^k\right] w(x)\, dx \\
&= \sum_{k=0}^{n} C_k \int_a^b x^k w(x)\, dx \\
&= \sum_{k=0}^{n} C_k\, \mu_k,
\end{aligned}$$

luego, se define una medida positiva $\alpha$ en $[a, b]$, cuyos momentos (2.8) existen para todo $n = 0, 1, \ldots$. Dichos momentos, generan una forma cuadrática

$$\sum_{j,k=0}^{n} \alpha_{k+j}\, \overline{x_j}\, x_k. \tag{2.9}$$

Cómo $\alpha$ tiene soporte infinito, la forma (2.9) es definida positiva; así, la forma cuadrática tiene la forma:

$$\int_a^b \left| \sum_{j=0}^{n} x_j t^j \right|^2 d\alpha(t)$$

**Definición 37.** El *producto escalar* de dos funciones reales $f(x)$ y $g(x)$, donde $x$ se extiende sobre el intervalo real $[a, b]$, esta definido por la integral de Lebesgue

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, d\alpha(x) \tag{2.10}$$

en donde se asume que $f(x)g(x)$ es de clase $L_\alpha(a, b)$.

Este es el caso en el que si $f(x)$ y $g(x)$ son contínuas y $[a, b]$ es un intervalo finito, entonces, para una función fija $\alpha(x)$, la ortogonalidad con respecto a la "distribución"$d\alpha$ está definida por la relación

$$\langle f, g \rangle = 0$$

lo que quiere decir que $f(x)$ es ortogonal a $g(x)$.

Si $f(x)$ y $g(x)$ son funciones complejas; en general, la definición (2.10) puede ser modificada por

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\, d\alpha(x)$$

con este cambio, si $f(x)$ y $g(x)$ son ortogonales, $(f, g) = 0$ se mantiene.

Si $\alpha(x)$ es absolutamente contínua, el producto escalar (2.10) se reduce a

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)\, dx$$

donde se asume que la integral existe en el sentido de Lebesgue, para la cual (2.7) se satisface.

Como anteriormente, $w(x)$ es llamada la *función peso*, referente a la función peso del intervalo dado.

Sea $d\alpha(x)$ o $w(x)\, dx$ ($a \leq x \leq b$) una distribución fija,y se considera un espacio de vectores definidos por el conunto de las funciones reales $f(x)$ que pertenecen a la clase $L^2_\alpha(a, b)$. El producto escalar de los dos vectores (funciones), $f(x)$ y $g(x)$, está definido por (2.10) y la longitud (magnitud, norma) de un vector $f(x)$ por

$$||f|| = \langle f, f \rangle^{\frac{1}{2}}.$$

Los vectores (funciones) con $||f|| = 0$ son llamados *vectores cero* (*funciones cero*) y los vectores con $||f|| = 1$ son *normalizados*.

En el segundo caso, $f(x)$ es una función cero sí, y solo si $\{f(x)\}^2 w(x)$ o, lo que equivale a lo mismo, $f(x)w(x)$ desaparece en todo el intervalo $[a, b]$ excepto en un conjunto de medida cero. Si $w(x)$ y $f(x)$ son funciones integrables en el sentido de Riemann, $f(x)$ es una función cero, facilitando que, $f(x)w(x)$ desaparezca en todo punto de continuidad.

*Observación.* No toda sucesión de momentos $\{\mu_n\}$ da lugar a una sucesión de polinomios ortogonales.

Entonces, de manera más formal:

**Definición 38.** Un *funcional de momentos* $\mathscr{L}$, es una transformación $\mathbb{C}$-Lineal del espacio $\mathbb{C}[x]$ de los polinomios con coeficientes complejos en el campo $\mathbb{C}$ de los números complejos, es decir

$$\begin{aligned} \mathscr{L}: \mathbb{C}[x] &\longrightarrow \mathbb{C} && (2.11) \\ f(x) &\longrightarrow \mathscr{L}[f(x)] = \int_a^b f(x)w(x)\,dx && (2.12) \end{aligned}$$

tal que, para todo $a, b \in \mathbb{C}$ y $p, q \in \mathbb{C}[x]$

$$\mathscr{L}(ap + bq) = a\mathscr{L}(p) + b\mathscr{L}(q) \tag{2.13}$$

**Definición 39.** Un sistema $\{P_n(x)/n \geq 0\}$ de polinomios monicos (es decir, el coeficiente del máximo grado $P_n(x)$ es 1) es un *sistema ortogonal con respecto a un funcional de momentos* $\mathscr{L}$ si

$$P_n(x) \text{ es de grado } n$$

y

$$\mathscr{L}(P_n(x)P_m(x)) = \lambda_n \delta_{mn} \text{ con } n, m \geq 0$$

$$\textit{donde} \quad \lambda_0 = 1,\ \lambda_n \neq 0 \textit{ para todo } n \leq 1$$

**Teorema 2.15.** *Sea $\mathscr{L}$ un funcional de momentos, y sea $\{P_n(x)\}$ una sucesión de polinomios. Entonces, las siguientes afirmaciones son equivalentes:*

*(a)* $\{P_n(x)\}$ *es un sistema de polinomios ortogonales con respecto a* $\mathscr{L}$

*(b)* $\mathscr{L}[P_m(x)P_n(x)] = k_n\delta_{mn}$ *donde* $k_n \neq 0$, $m = 0, 1, \ldots, n$.

*Demostración.* Sea $\{P_n(x)\}$ un sistema de polinomios ortogonales para $\mathscr{L}$. Cómo cada $P_k(x)$ es de grado $k$ entonces

$$\{P_0(x), P_1(x), \ldots, P_m(x)\}$$

es una base para el subespacio vectrial de los polinomios de grado a lo sumo $m$. Así, para $P_m(x)$, existen constantes $c_k$ tales que

$$
\begin{aligned}
P_m(x) &= \sum_{k=0}^{m} c_k P_k(x) \\
P_m(x)P_n(x) &= \sum_{k=0}^{m} c_k P_k(x)P_n(x) \\
\mathscr{L}[P_m(x)P_n(x)] &= \mathscr{L}\left[\sum_{k=0}^{m} c_k P_k(x)P_n(x)\right] \\
&= \sum_{k=0}^{m} c_k \mathscr{L}\left[P_k(x)P_n(x)\right] = 0 \\
&= c_m \mathscr{L}[P_n^2(x)] \qquad \text{si} \quad m = n.
\end{aligned}
$$

□

**Definición 40.** Si un funcional de momentos $\mathscr{L}$ admite un sistema monico ortogonal de polinomios, se dice que $\mathscr{L}$ es *regular*.

**Teorema 2.16.** *Sea $\{P_n(x)\}$ un sistema de polinomios ortogonales con respecto a $\mathscr{L}$. entonces para todo polinomio $P_n(x)$*

$$P_n(x) = \sum_{k=0}^{n} c_k P_k(x)$$

*donde*

$$c_k = \frac{\mathscr{L}[P_n(x)P_k(x)]}{P_k^2(x)} \qquad k = 0, 1, \ldots, n \tag{2.14}$$

*Demostración.* Cómo $P(x)$ es un polinomio de grado $n$, entonces existen constantes $c_k$ tales que

$$
\begin{aligned}
P_m(x) &= \sum_{k=0}^{m} c_k P_k(x) && (2.15)\\
P_m(x)P_n(x) &= \sum_{k=0}^{m} c_k P_k(x) P_n(x) && (2.16)\\
\mathscr{L}[P_m(x)P_n(x)] &= \mathscr{L}\left[\sum_{k=0}^{m} c_k P_k(x) P_n(x)\right] && (2.17)\\
&= \sum_{k=0}^{m} c_k \mathscr{L}\left[P_k(x)P_n(x)\right] && (2.18)\\
&= c_m \mathscr{L}[P_n^2(x)] \qquad \text{si} \quad m=n, && (2.19)
\end{aligned}
$$

despejando

$$\frac{\mathscr{L}[P_m(x)P_n(x)]}{\mathscr{L}[P_n^2(x)]} = c_m$$

cambiando $m$ por $k$, se deduce (2.14); para todo $k = 0, 1, 2, \ldots, n$. □

De ahora en adelante, el funcional de momentos $\mathscr{L}$, será notado por su definición (con la integral 2.34). para continuar viendo sus propiedades, se define el siguiente determinante:

**Definición 41.** Sea $D_n$ el *determinante de Hankel* asociado a la sucesión de momentos $\{\mu_n\}_{n=0}^{\infty}$ definido como

$$D_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} \tag{2.20}$$

**Teorema 2.17.** *$D_n \neq 0$ para todo $n = 0, 1, 2, \ldots$*

*Demostración.* Se procede por contradicción: Se supone que $D_n = 0$ para todo $n = 1, 2, \ldots$. Se construye el sistema homogéneo de $n+1$ ecuaciones con $a_0, \ldots, a_n$ incognitas, así:

$$
\begin{cases}
a_0\mu_0 + a_1\mu_1 + \ldots + a_n\mu_n = & 0 \\
a_0\mu_1 + a_1\mu_2 + \ldots + a_n\mu_{n+1} = & 0 \\
\vdots \\
a_0\mu_n + a_1\mu_{n+1} + \ldots + a_n\mu_{2n} = & 0
\end{cases} \tag{2.21}
$$

Cómo $D_n = 0$ el sistema de ecuaciones (2.21) admite soluciones diferentes de cero. Por la definición de los momentos (2.8) el sistema (2.21) puede escribirse como

$$
\begin{cases}
a_0 \int_a^b x^0 w(x)\, dx + a_1 \int_a^b x^1 w(x)\, dx + \ldots + a_n \int_a^b x^n w(x)\, dx = & 0 \\
a_0 \int_a^b x^1 w(x)\, dx + a_1 \int_a^b x^2 w(x)\, dx + \ldots + a_n \int_a^b x^{n+1} w(x)\, dx = & 0 \\
\vdots \\
a_0 \int_a^b x^n w(x)\, dx + a_1 \int_a^b x^{n+1} w(x)\, dx + \ldots + a_n \int_a^b x^{2n} w(x)\, dx = & 0
\end{cases}
$$

así

$$
\begin{cases}
\int_a^b [a_0 + a_1 x^1 + \ldots + a_n x^n] w(x)\, dx = & 0 \\
\int_a^b [a_0 x + a_1 x^2 + \ldots + a_n x^{n+1}] w(x)\, dx = & 0 \\
\vdots \\
\int_a^b [a_n x^n + a_1 x^{n+1} + \ldots + a_n x^{2n}] w(x)\, dx = & 0
\end{cases} \tag{2.22}
$$

factorizando $x$ de manera conveniente en cada ecuación, se tiene

$$
\begin{cases}
\int_a^b [a_0 + a_1 x^1 + \ldots + a_n x^n] w(x)\, dx = & 0 \\
\int_a^b [a_0 + a_1 x^1 + \ldots + a_n x^n] x w(x)\, dx = & 0 \\
\vdots \\
\int_a^b [a_n + a_1 x^1 + \ldots + a_n x^n] x^n w(x)\, dx = & 0
\end{cases} \tag{2.23}
$$

finalmente, multiplicando cada ecuación, dependiendo el exponente de la $x$ por el que se haya factrizado, por $a_0, \ldots, a_n$ respectivamente y sumandolas, resulta

$$\int_a^b [a_0 + a_1 x^1 + \ldots + a_n x^n]^2 w(x)\, dx = 0,$$

pero esta igualdad es una contradicción puesto que algunas de las incognitas $a_i$ son distintas de cero; por lo tanto $D_n \neq 0$ para todo $n = 1, 2, \ldots$. □

### 2.2.1. Relaciones de Recurrencia

Acontinuación, se Muestran dos relaciones muy importantes entre polinomios ortogonales, la primera, tomada de [5], fue demostrada con base en la prueba que se hace en el texto así como la segunda, tomada de [8].

**Teorema 2.18.** *Si $\{P_n(x) | n \geq 0\}$ es un sistema de polinomios mónicos ortogonales con respecto a un funcional de momentos $\mathscr{L}$, existen números complejos $B_n$, $C_n$, $n \geq 0$ tales que*

$$C_n \neq 0 \quad n \geq 1 \tag{2.24}$$

*y que*

$$xP_n(x) = P_{n+1}(x) + B_n P_n(x) + C_n P_{n-1}(x) \quad n \geq 0 \tag{2.25}$$

*con*

$$P_{-1}(x) = 0 \quad P_0(x) = 1 \tag{2.26}$$

*Además $\{P_n(x)\}$ queda univocamente determinado por las tres igualdades anteriores.*

*Recíprocamente, si $P_n(x) | n \geq 0$ es un sistema de polinomios mónicos, determinado por la relación de recurrencia de tres términos (2.25) y las condiciones iniciales (2.26) y si (2.24) se verifica; entonces que existe un funcional $\mathscr{L}$ para el cual $\{P_n(x) | n \geq 0\}$ es un sistema de polinomios mónicos ortogonales.*
*Más aún*

$$\lambda_n = \mathscr{L}(P_n^2(x)) = C_1 \cdots C_n$$

$$y \quad \mathscr{L}(1) = 1; \quad \mathscr{L}(P_n(x)) = 0 \quad n \geq 1 \tag{2.27}$$

*y $\mathscr{L}$ queda unívocamente determinado por las dos igualdades anteriores.*

*Demostración.* ($\Rightarrow$) Como $P_n(x)$ es de grado $n$, se tiene que

$$a_0P_0(x) + a_1P_1(x) + \cdots + a_nP_n(x) = 0$$

lo cual obliga que todos los $a_i = 0$ y $\{P_n(x)|n \geq 0\}$ genera a cada polinomio de $\mathbb{C}[x]$; luego $\{P_n(x)|n \geq 0\}$ es una base de $\mathbb{C}[x]$.
Entonces

$$\begin{aligned} P_n &= \sum_{i=0}^{n} C_{i,n}P_i \qquad n \geq 0 \\ xP_n &= \sum_{i=0}^{n+1} C_{i,n}P_i \qquad C_{i,n} \in \mathbb{C}, C_{n+1,n} = 1 \end{aligned} \tag{2.28}$$

Segun (2.28), para $n = 0$, la relación de recurrencia de tres términos (2.25) es

$$\begin{aligned} xP_0 &= \sum_{i=0}^{1} C_{i,n}P_i(x) \\ &= C_{0,n}P_0(x) + C_{1,n}P_1(x) \\ &= C_{0,n}(1) + C_{1,n}P_1(x) \qquad con \;\; C_{1,n} = 1 \\ &= P_1(x) + C_{0,n}(1) + C_n(0) \end{aligned} \tag{2.29}$$

Entonces, para $n = 0$, la relación de recurrencia de tres términos se cumple.

Segun (2.28), para $n = 1$ tenemosse tiene:

$$
\begin{aligned}
xP_1(x) &= \sum_{i=0}^{1+1} C_{i,n}P_i(x) \\
&= C_{0,n}P_0(x) + C_{1,n}P_1(x) + C_{2,n}P_2(x) \\
&= C_{0,n}(1) + C_{1,n}P_1(x) + (1)P_2(x) \\
&= C_{0,n} + B_1P_1(x) + P_2(x) \qquad (2.30)
\end{aligned}
$$

Luego, para $n = 1$, la relación de recurrencia de tres términos, se cumple.

Se demuestra si (2.28) se cumple para $n \geq 2$:

Multiplicando a ambos lados de (2.28) por $P_k(x)$ con $0 \leq k \leq n-2$, se tiene

$$
xP_n(x)P_k(x) = \sum_{i=0}^{n+1} C_{i,n}P_i(x)P_k(x) \qquad (2.31)
$$

y como $P_k(x)$ es un polinomio, se sabe que

$$
xP_k(x) = \sum_{i=0}^{k+1} C_{i,k}P_i(x)
$$

Entonces

$$
xP_k(x)P_n(x) = \sum_{i=0}^{k+1} C_{i,k}P_i(x)P_n(x)
$$

Así, al aplicar el funcional a ambos lados de la igualdad y usando la propiedad (2.13) de $\mathscr{L}$, se tiene que

$$
\begin{aligned}
\mathscr{L}(xP_n(x)P_k(x)) &= \mathscr{L}\left(\sum_{i=0}^{k+1} C_{i,k}P_i(x)P_n(x)\right) \\
&= \sum_{i=0}^{k+1} C_{i,k}\mathscr{L}(P_i(x)P_n(x))
\end{aligned}
$$

Ahora, como $k \leq n-2$ entonces $k+1 \leq n-1 < n$ por tanto, por la propiedad de ortogonalidad (2.14), se tiene que

$$\mathscr{L}(P_i(x)P_n(x)) = 0 \qquad para \quad i \leq n$$

Entonces

$$\begin{aligned}
\mathscr{L}(xP_k(x)P_n(x)) &= \sum_{i=0}^{k+1} C_{i,k}(0) \\
&= 0.
\end{aligned}$$

Pero a su vez, las propiedades de ortogonalidad (2.14) y (2.31) implican que

$$\begin{aligned}
\mathscr{L}(xP_k(x)P_n(x)) &= \mathscr{L}(\sum_{i=0}^{n+1} C_{i,n}P_i(x)P_k(x)) \\
&= \sum_{i=0}^{n+1} C_{i,n}\mathscr{L}(P_i(x)P_k(x)) \\
&= \sum_{i=0}^{n+1} C_{i,n}\lambda_i\delta_{i,k} \qquad cuando\ i = k \\
&= C_{k,n}\lambda_k,
\end{aligned}$$

y como $\lambda_k \neq 0$ entonces $C_{k,n} = 0$ con $0 \leq k \leq n-2$

Luego, extendiendo (2.31) se tiene:

$$\begin{aligned}
xP_n(x) &= \sum_{i=0}^{n+1} C_{i,n}P_i(x) \\
x &= C_{0,n}P_0(x) + C_{1,n}P_1(x) + \cdots + C_{n-2,n}P_{n-2} + C_{n-1,n}P_{n-1}(x) \\
&\quad +C_{n,n}P_n(x) + C_{n+1,n}P_{n+1}(x) \\
&= (0)P_0(x) + (0)P_1(x) + \cdots + (0)P_{n-2}(x) + C_{n-1,n}P_{n-1}(x) \\
&\quad +C_{n,n}P_n(x) + (1)P_{n+1}(x) \\
&= C_{n-1,n}P_{n-1}(x) + C_{n,n}P_n(x) + P_{n+1}(x).
\end{aligned}$$

Haciendo $C_{n-1,n} = C_n$, $C_{n,n} = B_n$ y reordenando, se tiene

$$xP_n(x) = P_{n+1}(x) + B_nP_n(x) + C_nP_{n-1}(x). \tag{2.32}$$

Con lo que se concluye que, para $n \geq 2$, se cumple la relación de recurrencia de tres términos (2.25).

Ahora, Multiplindo por $P_{n-1}(x)$, con $n \geq 1$, a ambos lados de (2.32) así:

$$\begin{aligned} xP_n(x)P_{n-1}(x) &= (P_{n+1}(x) + B_nP_n(x) + C_nP_{n-1}(x))P_{n-1}(x) \\ &= P_{n+1}(x)P_{n-1}(x) + B_nP_n(x)P_{n-1}(x) + C_nP_{n-1}(x)P_{n-1}(x) \\ &= P_n^2(x) + B_nP_n(x)P_{n-1}(x) + C_nP_{n-2}(x)P_n(x). \end{aligned}$$

Al aplicar el funcional a ambos lados

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = \mathscr{L}(P_n^2(x) + B_nP_n(x)P_{n-1}(x) + C_nP_{n-2}(x)P_n(x))$$

por la propiedad (2.13) de $\mathscr{L}$

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = \mathscr{L}(P_n^2(x)) + B_n\mathscr{L}(P_n(x)P_{n-1}(x)) + C_n\mathscr{L}(P_{n-2}(x)P_n(x)),$$

y por la propiedad de ortogonalidad (2.14) de $\mathscr{L}$

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = \mathscr{L}(P_n^2(x)) + B_n(0) + C_n(0)$$

así

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = \mathscr{L}(P_n^2(x)), \tag{2.33}$$

ahora, si se multiplica nuevaménte a (2.32) por $P_{n-1}(x)$, se obtiene

$$
\begin{aligned}
xP_n(x)P_{n-1}(x) &= (P_{n+1}(x) + B_nP_n(x) + C_nP_{n-1}(x))P_{n-1}(x) \\
&= P_{n+1}(x)P_{n-1}(x) + B_nP_n(x)P_{n-1}(x) + C_nP_{n-1}(x)P_{n-1}(x) \\
&= P_{n+1}(x)P_{n-1}(x) + B_nP_n(x)P_{n-1}(x) + C_nP_{n-1}^2(x),
\end{aligned}
$$

al aplicar el funcional a ambos lados de esta igualdad, se tiene

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = \mathscr{L}(P_{n+1}(x)P_{n-1}(x) + B_nP_n(x)P_{n-1}(x) + C_nP_{n-1}^2(x)),$$

por la propiedad (2.13) de $\mathscr{L}$

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = \mathscr{L}(P_{n+1}(x)P_{n-1}(x) + B_n\mathscr{L}(P_n(x)P_{n-1}(x)) + C_n\mathscr{L}(P_{n-1}^2(x))$$

y, por la propiedad de ortogonalidad (2.14) de $\mathscr{L}$

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = (0) + B_n(0) + C_n\mathscr{L}(P_{n-1}^2(x)),$$

así

$$\mathscr{L}(xP_n(x)P_{n-1}(x)) = C_n\mathscr{L}(P_{n-1}^2(x)),$$

entonces, por (2.33) y por (2.14), se tiene que

$$
\begin{aligned}
\mathscr{L}(P_n^2(x)) &= C_n\mathscr{L}(P_{n-1}^2(x)) \\
\lambda_n &= C_n\lambda_{n-1}
\end{aligned}
$$

Ahora, para $n \geq 1$, $\lambda_n$ y $\lambda_{n-1}$ son no nulos, por lo tanto $C_n \neq 0$ para $n \geq 1$.

Como a partir de (2.32) se puede hallar cualquier polinomio, con las condiciones (2.26), se puede concluir que $\{P_n(x)|n \geq 0\}$ queda univocamente determinado por (2.25) y (2.26).

$(\Leftarrow)$ Se supone ahora, que $\{P_n(x)|n \geq 0\}$ satiface que existan constantes $B_n,\ C_n,\ n \geq 0$ tales que se cumplen las propiedades (2.24), (2.25) y (2.26).

Como ya se vió, $P_n(x)$ es de grado $n$, para todo $n \geq 0$, entonces $\{P_n(x)|n \geq 0\}$ es una base de $\mathbb{C}[x]$. Entonces, se define $\mathscr{L}$ por

$$\mathscr{L}(1) = 1 \quad y \quad \mathscr{L}(P_n(x)) = 0 \qquad n \geq 1, \tag{2.34}$$

si lo se extiende por linealidad a todo $\mathbb{C}[x]$, $\mathscr{L}$ es una transformación lineal de $\mathbb{C}[x]$ en $C$

$$\mathscr{L} : \mathbb{C}[x] \longrightarrow C.$$

Se debe ver que $\mathscr{L}$ satisface:

$$P_n(x) \text{ sea de grado } n$$

$$\mathscr{L}(P_n(x)P_m(x)) = \lambda_n \delta_{m,n} \quad n, m \geq 0 \tag{2.35}$$

$$\lambda_0 = 1,\ \lambda_n = 0. \tag{2.36}$$

Primero, se prueba que para $m \geq 0$ fijo

$$\mathscr{L}(x^m P_n(x)) = 0 \qquad para\ todo\ n, m \geq 0 \tag{2.37}$$

Para esto, se hace inducción sobre $m$: Si m=0:

$$\mathscr{L}(x^m P_n(x)) = \mathscr{L}(x^0 P_n(x)) = \mathscr{L}(P_n(x)) = 0 \quad \text{por (2.34).}$$

Como hipótesis inductiva, se supone que se cumple para $m = k$, es decir

$$\mathscr{L}(x^m P_n(x)) = \mathscr{L}(x^k P_n(x)) = 0.$$

De manera inductiva, se muestra que se cumple para $m = k + 1$: Como (2.25) se cumple, Multiplicando a ambos lados de esta igualdad por $x^m$, aplicando $\mathscr{L}$ y usando las propiedades de linealidad (2.13),las de ortogonalidad (2.14) y la hipótesis de inducción:

$$\begin{aligned}
x^m(xP_n(x)) &= x^m(P_{n+1}(x) + B_n P_n(x) + C_n P_{n-1}(x)) \\
x^{m+1}P_n(x) &= x^m P_{n+1}(x) + x^m B_n P_n(x) + x^m C_n P_{n-1}(x) \\
\mathscr{L}(x^{m+1}P_n(x)) &= \mathscr{L}(x^m P_{n+1}(x) + x^m B_n P_n(x) + x^m C_n P_{n-1}(x)) \\
&= \mathscr{L}(x^m P_{n+1}(x)) + B_n \mathscr{L}(x^m P_n(x)) + C_n \mathscr{L}(x^m P_{n-1}(x)) \\
&= 0
\end{aligned}$$

Por tanto $\mathscr{L}(x^{m+1}P_n(x)) = 0$. Así (2.37) queda establecida.

Ahora, por (2.37) y (2.34), se asegura que

$$\mathscr{L}(P_n(x)P_m(x)) = 0 \qquad m \neq n \tag{2.38}$$

se cumple, así que solo queda por ver que $\mathscr{L}(P_n^2(x)) \neq 0$: Al multiplicar, a ambos lados de (2.25), por $P_{n-1}$ y aplicar $\mathscr{L}$ a la igualdad, junto con las propiedades de linealidad (2.13), de ortogonalidad (2.14) y la anterior (2.38), se obtiene:

$$\begin{aligned}
(xP_n(x))P_{n-1} &= (P_{n+1}(x) + B_n P_n(x) + C_n P_{n-1}(x))P_{n-1} \\
xP_n(x)P_{n-1} &= P_{n+1}(x)P_{n-1} + B_n P_n(x)P_{n-1} + C_n P_{n-1}(x)P_{n-1} \\
&= P_n^2 + B_n P_n(x)P_{n-1} + C_n P_{n-2}(x)P_n \\
\mathscr{L}(xP_n(x)P_{n-1}) &= \mathscr{L}(P_n^2(x) + B_n P_n(x)P_{n-1} + C_n P_{n-2}(x)P_n) \\
&= \mathscr{L}(P_n^2(x)) + B_n \mathscr{L}(P_n(x)P_{n-1}) + C_n \mathscr{L}(P_{n-2}(x)P_n) \\
&= \mathscr{L}(P_n^2(x)),
\end{aligned}$$

luego

$$\mathscr{L}(xP_n(x)P_{n-1}) \quad = \quad \mathscr{L}(P_n^2(x)). \tag{2.39}$$

También, al multiplicar a ambos lados de (2.25), por $P_{n-1}$ y aplicar $\mathscr{L}$ y las propiedades de linealidad (2.13), ortogonalidad (2.14) y (2.38), se obtiene

$$\begin{aligned}
(xP_n(x))P_{n-1} &= (P_{n+1}(x) + B_nP_n(x) + C_nP_{n-1}(x))P_{n-1} \\
xP_n(x)P_{n-1} &= P_{n+1}(x)P_{n-1} + B_nP_n(x)P_{n-1} + C_nP_{n-1}(x)P_{n-1} \\
&= P_{n+1}(x)P_{n-1} + B_nP_n(x)P_{n-1} + C_nP_{n-1}(x)^2 \\
\mathscr{L}(xP_n(x)P_{n-1}) &= \mathscr{L}(P_{n+1}(x)P_{n-1} + B_nP_n(x)P_{n-1} + C_nP_{n-1}(x)^2) \\
&= \mathscr{L}(P_{n+1}(x)P_{n-1}) + B_n\mathscr{L}(P_n(x)P_{n-1}) + C_n\mathscr{L}(P_{n-1}(x)^2) \\
&= C_n\mathscr{L}(P_{n-1}(x)^2),
\end{aligned}$$

así

$$\mathscr{L}(xP_n(x)P_{n-1}) \quad = \quad C_n\mathscr{L}(P_{n-1}(x)^2). \tag{2.40}$$

Entonces, por (2.39), se tiene que

$$\mathscr{L}(P_n^2(x)) = C_n\mathscr{L}(P_{n-1}(x)^2) \qquad n \geq 1,$$

y por la propiedad de ortogonalidad (2.14)

$$\lambda_n = \mathscr{L}(P_n^2(x)).$$

Ahora, por inducción sbre $n$, se prueba que

$$\mathscr{L}(P_n^2(x)) = C_1, \ldots, C_n. \tag{2.41}$$

Inicialmente, se muestra que se cumple para $n = 1$:

$$\begin{aligned}
\mathscr{L}(P_1^2(x)) &= C_1\mathscr{L}(P_{1-1}(x)^2) \\
&= C_1\mathscr{L}(P_0(x)^2) \\
&= C_1(1) \qquad \text{por (2.26)} \\
&= C_1
\end{aligned}$$

Por tanto, para $n = 1$ se cumple. Como hipótesis inductiva, Se supone que se cumple para $n = k$:

$$\mathscr{L}(P_k^2(x)) = C_1, \ldots, C_k$$

Finalmente, de manera inductiva, Se demuestra que se cumple para $n = k + 1$:

$$\begin{aligned}
\mathscr{L}(P_{k+1}2(x)) &= C_{k+1}\mathscr{L}(P_{k-1}(x)^2) \\
&= C_{k+1}\mathscr{L}(P_k(x)^2) \\
&= C_{k+1}(C_1, \ldots, C_k) \qquad \text{por hipótesis de inducción} \\
&= C_1, \ldots, C_k C_{k+1}
\end{aligned}$$

Así, $\lambda_n = \mathscr{L}(P_n^2(x)) = C_1, \ldots, C_k C_n$

Esto asegura los items (2.35) y (2.36). Por lo tanto, el teorema ha quedado demostrado. $\square$

En la sección anterior, se vio la construcción y el desarrollo de los polinomios de Chebyshev y se concluyó que

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

*Ejemplo* 9. Recuerdese que $T_0(x) = 1$ y que $T_1(x) = x$; de manera general, se puede notar que el coeficiente principal de $T_n(x)$ está dado por $2^{n-1}$ para todo $n = 0, 1, \ldots$. Así, los polinomios mónicos correspondietes son

$$\hat{T}_0(x) = T_0 \qquad \hat{T}_n = 2^{1-n}T_n(x),$$

reescribiendo (2.25) como

$$xP_n(x) = (x + B_n)P_n(x) + C_nP_{n-1}(x),$$

se tiene

$$\begin{aligned}
x\hat{T}_{n-1}(x) - \frac{1}{4}\hat{T}_{n-2}(x) &= x(2^{1-(n-1)}T_{n-1}(x)) - \frac{1}{4}(2^{1-(n-2)}T_{n-2}(x)) \\
&= x(2^{2-n}T_{n-1}(x)) - \frac{1}{4}(2^{3-n}T_{n-2}(x)) \\
&= 2^{1-n}2xT_{n-1}(x)) - \frac{1}{2^2}(2^{3-n}T_{n-2}(x)) \\
&= 2^{1-n}2xT_{n-1}(x)) - 2^{3-n-2}T_{n-2}(x) \\
&= 2^{1-n}2xT_{n-1}(x)) - 2^{1-n}T_{n-2}(x) \\
&= 2^{1-n}(2xT_{n-1}(x) - T_{n-2}(x)) \\
&= 2^{1-n}(T_n(x)) \\
&= \hat{T}_n(x),
\end{aligned}$$

por tanto

$$\hat{T}_n(x) = x\hat{T}_{n-1}(x) - \frac{1}{4}\hat{T}_{n-2}(x).$$

**Teorema 2.19.** *(Identidad de Christoffel-Darboux) Sea $\{P_n(x)\}_{n=0}^{\infty}$ el sistema de polinomios ortogonales con respecto a $w(x)$. Entonces, para todo $x, y \in \mathbb{R}$, con $x \neq y$ y $n \geq 1$*

$$\sum_{k=0}^{n-1} P_k(x)P_k(y) = \alpha_n \left[\frac{P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)}{x-y}\right] \tag{2.42}$$

*con $p_n(x) = \alpha_n x^n + \ldots$, $\alpha_n > 0$ y $\alpha_n = \frac{\alpha_{n-1}}{\alpha_n}$.*

*Demostración.* Sí en el teorema anterior, $P_n(x)$ no fuera mónico, podria representarse como

$$xP_n(x) = \alpha_{n+1}P_{n+1}(x) + \beta_{n+1}P_n(x) + \alpha_n P_{n-1}(x) \quad n \geq 0, \tag{2.43}$$

multiplicando a ambos lados de ésta igualdad por $P_n(y)$, se tiene

$$\begin{aligned}
xP_n(x)P_n(y) &= [\alpha_{n+1}P_{n+1}(x) + \beta_{n+1}P_n(x) + \alpha_n P_{n-1}(x)]P_n(y) \\
&= \alpha_{n+1}P_{n+1}(x)P_n(y) + \beta_{n+1}P_n(x)P_n(y) + \alpha_n P_{n-1}(x)P_n(y),
\end{aligned}$$

cambiando $x$ por $y$ en la última igualdad, se tiene

$$yP_n(x)P_n(y) = \alpha_{n+1}P_{n+1}(y)P_n(x) + \beta_{n+1}P_n(x)P_n(y) + \alpha_n P_{n-1}(y)P_n(x), \tag{2.44}$$

Al restar (2.43) con (2.44), se obtiene

$$\begin{aligned}
(x-y)P_n(x)P_n(y) &= \alpha_{n+1}P_{n+1}(x)P_n(y) - \alpha_{n+1}P_{n+1}(y)P_n(x) + \alpha_n P_{n-1}(x)P_n(y) - \alpha_n P_{n-1}(y)P_n(x) \\
&= \alpha_{n+1}[P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)] - \alpha_n[P_{n-1}(y)P_n(x) - P_{n-1}(x)P_n(y)].
\end{aligned}$$

Reorganizando, factorizando $\alpha_n$, acomodando los coeficientes y sumando sobre $n$, resulta

$$\begin{aligned}
\sum_{k=0}^{n-1}(x-y)P_k(x)P_k(y) &= \alpha_n\left[P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)\right] \\
(x-y)\sum_{k=0}^{n-1}P_k(x)P_k(y) &= \alpha_n\left[P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)\right]
\end{aligned}$$

Dividiendo entre $(x-y)$ a ambos lados de la igualdad

$$\sum_{k=0}^{n-1}P_k(x)P_k(y) = \alpha_n\left[\frac{P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)}{x-y}\right]$$

Recordando que $P_{-1}(x) = 0$.

$\square$

### 2.2.2. Más propiedades de polinomios ortogonales

En esta sección, se muestra que la sucesión de polinomios mónicos ortogonales $\{P_n(x)\}$ puede escribirse en términos del determinante $D_n$ que es llamado el *determinante de momentos*; la ventaja de todo esto, es que, por medio de esas matrices, se define el coeficiente de la variable de mayor grado del polinomio y gracias él, se pueden dar propiedades del máximo y del mínimo, que se verán en las siguientes secciones, para dichas sucesiones. Este estudio fue hecho con base en [7], [8] y [9]. Traducción a cargo de la autora de este trabajo.

**Teorema 2.20.** *Sea $\alpha(x)$ una función no decreciente en $[a,b]$ que no es constante, existe una única sucesión de polinomios mónicos ortogonales $\{P_n(x)\}$ tales que*

$$P_n(x) = x^n + términos\ de\ orden\ inferior, \qquad n = 0, 1, \ldots$$

*y una sucesión de números positivos $\{a_n\}_0^\infty$, con $a_0 = 1$ tales que*

$$\int_a^b P_m(x)P_n(x)\, d\alpha(x) = a_n \delta_{n,m}. \tag{2.45}$$

*Demostración.* Se procede por inducción sobre $N$ cn $m, n = 0, \ldots, N$ y $N = 0, \ldots$:
Se define $P_0(x) = 1$; se asume que para $P_0(x), \ldots, P_N(x)$ (2.45) esta definido y se satisface. Se establece $P_{N+1}(x) = x^{N+1} + \sum_{j=0}^{N} c_j x^j$. Para $m < N+1$, se tiene

$$\begin{aligned}
\int_a^b x^m P_{N+1}(x)\, d\alpha(x) &= \int_a^b x^m \left( x^{N+1} + \sum_{j=0}^{N} c_j x^j \right) d\alpha(x) \\
&= \int_a^b \left( x^{N+1+m} + \sum_{j=0}^{N} c_j x^{j+m} \right) d\alpha(x) \\
&= \int_a^b x^{N+1+m}\, d\alpha(x) + \int_{\mathbb{R}} \sum_{j=0}^{N} c_j x^{j+m}\, d\alpha(x) \\
&= \mu_{N+1+m} + \sum_{j=0}^{N} c_j \int_a^b x^{j+m}\, d\alpha(x) \\
0 &= \alpha_{N+1+m} + \sum_{j=0}^{N} c_j \alpha_{j+m} \\
-\alpha_{N+1+m} &= \sum_{j=0}^{N} c_j \alpha_{j+m},
\end{aligned} \tag{2.46}$$

por medio de la solucion de (2.46) para $c_j$, puede verse que su determinante asociado $D_N$ es mayor que cero. Así, el polinomio $P_{N+1}(x)$ fué encontrado y $a_n = \int_{\mathbb{R}} P_{N+1}^2(x)\, d\alpha(x)$.(En [9])

□

Como consecuencia del teorema anterior, es posible escribir los polinomios de la sucesión y los $a_n$, como

$$P_n(x) = \frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} \qquad a_n = \frac{D_n}{D_{n-1}}, \tag{2.47}$$

como la matríz satisface los requerimientos del teorema anterior puesto que para $m < n$, $\int_{\mathbb{R}} x^m P_n(x)\, d\mu(x)$ las filas $n+1$ y $m+1$ del determinante (2.20) son iguales. Para ver la representación de $a_n$, note que

$$\begin{aligned} a_n &= \int_{\mathbb{R}} P_n^2(x)\, d\mu(x) \\ &= \int_{\mathbb{R}} x^n P_n(x)\, dx \end{aligned}$$

$$= \int_{\mathbb{R}} x^n \frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} dx$$

$$= \frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ \int_{\mathbb{R}} x^n 1\, d\mu(x) & \int_{\mathbb{R}} x^n x\, d\mu(x) & \cdots & \int_{\mathbb{R}} x^n x^n\, d\mu(x) \end{vmatrix}$$

$$= \frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ \int_{\mathbb{R}} x^n\, d\mu(x) & \int_{\mathbb{R}} x^{n+1}\, d\mu(x) & \cdots & \int_{\mathbb{R}} x^{2n}\, d\mu(x) \end{vmatrix},$$

$$= \frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}$$

por tanto

$$a_n = \frac{D_n}{D_{n-1}}. \tag{2.48}$$

A raíz de esto, se encuentra un importante resultado de Heine (en [9]), que es utilizado para representar las funciones ortonormales $\phi_n(x)$ que se definen acontinuación:

**Definición 42.** Un *conjunto ortonormal* de funciones $\phi_0(x), \phi_1(x), \phi_2(x), \cdots \phi_l(x)$, $l$ finito o infinito, está definida por las relaciones

$$(\phi_n, \phi_m) = \int_a^b \phi_n(x)\phi_m(x)\, d\alpha(x) = \delta_n m \qquad n, m = 0, 1, 2, \cdots, l.$$

Donde $\phi_n(x)$ es a valor real y pertenece a $L^2_\alpha(a, b)$.

Las funciones de esos ordenes, son necesariamente linealmente independientes. Si $\alpha(x)$ tiene solo un número finito $N$ de puntos de crecimiento (es decir, puntos en la vecindad en los cuales $\alpha(x)$ no es constante), $l$ es necesariamente finito y $l < N$.

Sean $f_0(x), f_1(x), f_2(x), \ldots, f_l(x)$ funciones a valor real de clase $L^2_\alpha(a, b)$ linealmente independientes. Al aplicar el proceso de ortogonalización de Gram-Schmidt, explicado en el capítulo anterior, se obtiene el conjunto ortonormal

$$\phi_0(x), \phi_1(x), \phi_2(x), \ldots, \phi_l(x). \tag{2.49}$$

Esas funciones pueden ser escritas como (en [7])

$$\phi_n(x) = \lambda_{n0} f_0(x) + \lambda_{n1} f_1(x) + \ldots + \lambda_{nn} f_n(x) \qquad \lambda_{nn} > 0.$$

Como consecuencia de las secciones anteriores, se tiene el siguiente teorema, que es uno de los resultados más importantes hechos por Heine, en lo que respecta a polinomios ortogonales:

**Teorema 2.21.** *El sistema de polinomios mónicos ortogonales $\{P_n(x)\}$ tiene la siguiente representación integral de Heine*

$$P_n(x) = \frac{1}{n!D_{n-1}} \underbrace{\int_a^b \int_a^b \cdots \int_a^b}_{n} \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_n(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{n-1}) & f_1(x_{n-1}) & \cdots & f_n(x_{n-1}) \\ f_0(x) & f_1(x) & \cdots & f_n(x) \end{vmatrix} \tag{2.50}$$

$$\cdot \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_{n-1}(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_{n-1}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{n-1}) & f_1(x_{n-1}) & \cdots & f_{n-1}(x_{n-1}) \end{vmatrix} d\alpha(x_0)\, d\alpha(x_1) \cdots d\alpha(x_{n-1}), \quad n \geq 1. \tag{2.51}$$

$$= \frac{1}{(n+1)!D_{n-1}} \underbrace{\int_a^b \int_a^b \cdots \int_a^b}_{n+1} \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_n(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_n) & f_1(x_n) & \cdots & f_n(x_n) \end{vmatrix}^2 d\alpha(x_0)\, d\alpha(x_1) \cdots d\alpha(x_n), \quad n \geq 1 \tag{2.52}$$

Los detalles de ésta demostración se encuentran en [9].

Si se reemplaza (2.47) en (2.51), se obtiene

$$\frac{1}{D_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} = \frac{1}{n!D_{n-1}} \underbrace{\int_a^b \int_a^b \cdots \int_a^b}_{n} \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_n(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{n-1}) & f_1(x_{n-1}) & \cdots & f_n(x_{n-1}) \\ f_0(x) & f_1(x) & \cdots & f_n(x) \end{vmatrix}$$

$$\cdot \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_{n-1}(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_{n-1}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{n-1}) & f_1(x_{n-1}) & \cdots & f_{n-1}(x_{n-1}) \end{vmatrix} d\alpha(x_0)\, d\alpha(x_1) \cdots d\alpha(x_{n-1})$$

$$D_n(x) = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} = \frac{1}{n!} \underbrace{\int_a^b \int_a^b \cdots \int_a^b}_{n} \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_n(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{n-1}) & f_1(x_{n-1}) & \cdots & f_n(x_{n-1}) \\ f_0(x) & f_1(x) & \cdots & f_n(x) \end{vmatrix} \tag{2.53}$$

$$\cdot \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_{n-1}(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_{n-1}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_{n-1}) & f_1(x_{n-1}) & \cdots & f_{n-1}(x_{n-1}) \end{vmatrix} d\alpha(x_0)\, d\alpha(x_1) \cdots d\alpha(x_{n-1}) \tag{2.54}$$

De ahora en adelante, el determinante del lado derecho de la igualdad es notado por $D_n(x)$.

Análogamente, si se reemplaza (2.47) en (2.52) se obtiene

$$D_n(x) = \frac{1}{(n+1)!} \underbrace{\int_a^b \int_a^b \cdots \int_a^b}_{n+1} \begin{vmatrix} f_0(x_0) & f_1(x_0) & \cdots & f_n(x_0) \\ f_0(x_1) & f_1(x_2) & \cdots & f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_0(x_n) & f_1(x_n) & \cdots & f_n(x_n) \end{vmatrix}^2 d\alpha(x_0)\, d\alpha(x_1) \cdots d\alpha(x_n), \quad n \geq 1 \tag{2.55}$$

Éstas igualdades, son usadas más adelante para representar los polinomios ortonormales.

**Definición 43.** Sea $\{\phi_n(x)\}$ un conjunto ortonormal dado, finito o infinito. Una función arbitraria $f(x)$ a valor real, le corresponde la espansión de Fourier formal

$$f(x) \sim f_0\phi_0(x) + f_1\phi_1(x) + \cdots + f_n\phi_n(x) + \cdots . \tag{2.56}$$

Los coeficientes $f_n$, llamados los coeficientes de Fourier de $f(x)$ con respecto al sistema dado, estan definidos por

$$f_n = (f, \phi_n) = \int_a^b f(x)\phi_n(x)\, d\alpha \qquad n = 0, 1, 2, \cdots . \tag{2.57}$$

Toda sección finita de la serie (2.56) tiene la siguiente propiedad del mínimo:

**Teorema 2.22.** *Sean $\phi_n(x)$ un conjunto ortonormal, $f_n$ los coeficientes de Fourier y $f(x)$ la expansión formal de Fourier. Sea $l \geq 0$ un entero fijo y $a_0, a_1, \cdots, a_l$ constantes reales arbitrarias. Si se escribe*

$$g(x) = a_0\phi_0(x) + a_1\phi_1(x) + \cdots + a_l\phi_l(x)$$

*y los coeficientes $a_\nu$ son variables, la integral*

$$\int_a^b \{f(x) - g(x)\}^2 \, d\alpha(x) \tag{2.58}$$

*se convierte en un mínimo si y solo si $a_\nu = \int_a^b f_\nu \phi_l(x) \, d\alpha(x)$ con $\nu = 0, 1, 2, \cdots, l$.*

*El mínimo en sí mismo es*

$$\int_a^b |f(x)|^2 d\alpha(x) - \sum_{j=0}^{l} |f_j|^2.$$

*Demostración.*

$$\begin{aligned}
\int_a^b (f(x) - g(x))^2 \, d\alpha(x) &= \int_a^b \left( f(x) - \sum_{j=1}^{l} a_j\phi_j \right)^2 \, d\alpha(x) \\
&= \int_a^b |f(x)|^2 d\alpha(x) - 2Re\left( \sum_{j=0}^{l} f_j \int_a^b p_j(x)\overline{f(x)} d\alpha(x) \right) + \sum_{j=0}^{l} |f_j|^2,
\end{aligned}$$

haciendo $a_j = \int_a^b p_j(x) f(x) \, d\mu(x)$

$$\begin{aligned}
\int_a^b |f(x)|^2 d\alpha(x) &- 2Re\left( \sum_{j=0}^{l} f_j \int_a^b p_j(x)\overline{f(x)} d\alpha(x) \right) + \sum_{j=0}^{l} |f_j|^2 \\
&= \int_a^b |f(x)|^2 \, d\mu(x) - \sum_{j=0}^{l} |a_j|^2 + \sum_{j=0}^{l} |f_j - a_j|^2,
\end{aligned}$$

lo cual se minimiza sí $f_j = a_j$ para $j = 0, 1, \ldots, l$.

□

La serie ordinaria de Fourier en términos de las funciones trigonométricas $1, \cos nx, \sin nx, n = 1, 2, 3, \cdots ; -\pi \leq x \leq \pi$, es un clásico ejemplo de la expansión de Fourier de este tipo.

Otra importante caracterización del conjunto ortonormal $\phi_n(x)$, que puede estar basada en la anterior propiedad del mínimo para sumas parciales, basada en la propiedad del mínimo del teorema anterior, en efecto, es que para variables a valores reales $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$, la expresión

$$||\lambda_0 f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_{n-1} f_{n-1}(x) + f_n(x)||$$

se vuelve un mínimo si y solo si

$$\lambda_0 f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_{n-1} f_{n-1}(x) + f_n(x) = \lambda_{nn}^{-1} \phi_n(x).$$

**Polinomios Ortogonales**

**Definición 44.** Sea $\alpha(x)$ una función fija no decreciente con infinitos puntos decrecimiento en el intervalo $[a, b]$ finito o infinito, entonces los *momentos*

$$c_n = \int_a^b x^n \, d\alpha(x) \qquad n = 0, 1, 2, \cdots$$

existen.

Si se ortogonaliza el conjunto de potencias no negativas de x;

$$1,\ x,\ x^2, \cdots, x^n, \cdots$$

por medio del proceso de Gram-Scmidt explicado en el capítulo 1, se obtiene el conjunto de polinomios ortonormales

$$p_0(x), p_1(x), p_2(x), \cdots, p_n(x), \cdots \tag{2.59}$$

univocamente determinado por las siguientes condiciones:

(a) $p_n(x)$ es un polinomio de, precisamente, grado $n$; en el cual, el coeficiente de $x^n$ es positivo;

(b) el sistema $\{p_n(x)\}$ es ortonormal, esto es

$$\int_a^b p_n(x)p_m(x)\, d\alpha(x) = \delta_{nm} \qquad n,m = 0,1,2\cdots$$

La existencia de los *momentos*, es equivalente al hecho de que las funciones $x^n$ sean de clase $L_\alpha(a,b)$.

Los polinomios ortonormales (2.49), notados por $\{p_n(x)\}$ pueden ser expresados en términos de la matríz de Gram

$$G_n = \begin{pmatrix} \langle 1,1\rangle & \langle x,1\rangle & \cdots & \langle x^n,1\rangle \\ \langle 1,x\rangle & \langle x,x\rangle & \cdots & \langle x^n,x\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle 1,x^n\rangle & \langle x,x^n\rangle & \cdots & \langle x^n,x^n\rangle \end{pmatrix}, \tag{2.60}$$

pero como el producto interno que se esta utilizando es hermitiano, es decir que $\langle f,g\rangle = \overline{\langle g,f\rangle}$, entonces la matríz $G_n$ siempre es Hermitiana y definida positiva; por lo tanto, el teorema 2,19 se satisface.

**Teorema 2.23.** *Las funciones ortonormales (2.59) pueden ser expresadas de la forma*

$$p_n(x) = \frac{1}{\sqrt{D_nD_{n-1}}} \begin{vmatrix} \langle 1,1\rangle & \langle x,1\rangle & \cdots & \langle x^n,1\rangle \\ \langle 1,x\rangle & \langle x,x\rangle & \cdots & \langle x^n,x\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle 1,x^{n-1}\rangle & \langle x,x^{n-1}\rangle & \cdots & \langle x^n,x^{n-1}\rangle \\ 1 & x & \cdots & x^n \end{vmatrix}. \tag{2.61}$$

Para detalles de ésta demostracion, ver [8]

Por el teorema anterior, los polinomios $\{p_n(x)\}$ se escriben como

$$p_n(x) = \frac{1}{\sqrt{D_nD_{n-1}}} \begin{vmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} \tag{2.62}$$

y, como ya se vio que el determinante de la matríz de Gram cumple el teorema 2.20, junto con las matrices $D_n$ y $D_{n-1}$; entonces

$$D_n = \begin{vmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n} \end{vmatrix},$$

con $p_0(x) = D_0^{\frac{-1}{2}} = c_0^{\frac{-1}{2}}$, esta asociado con la forma cuadrática definida positiva (1.16) de la forma

$$\sum_{\nu,\mu=0}^{n} c_{\nu+\mu} u_\nu u_\mu = \int_a^b (u_0 + u_1 x + u_2 x^2 + \ldots + u_n x^n)^2 d\alpha(x) \tag{2.63}$$

la cual es llamada *forma de Hankel* o de tipo *recurrente*.

Al hacer operaciones elementales entre columnas en (2.62), se obtiene una nueva forma de escribir $p_n(x)$

$$p_n(x) = \frac{1}{\sqrt{D_n D_{n-1}}} \begin{vmatrix} c_0 x - c_1 & c_1 x - c_2 & \cdots & c_{n-1} x - c_n \\ c_1 x - c_2 & c_2 x - c_3 & \cdots & c_n x - c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} x - c_n & c_n x - c_{n+1} & \cdots & c_{2n-2} x - c_{2n-1} \end{vmatrix}. \tag{2.64}$$

Finalmente, por teorema 2,23, (2.54) y (2.55), se tienen las siguientes representaciones integrales para $p_n(x)$

$$p_n(x) = \frac{1}{n!\sqrt{D_{n-1} D_n}} \quad \underbrace{\int_a^b \int_a^b \cdots \int_a^b}_{n} (x - x_0)(x - x_1) \cdots (x - x_{n-1})$$
$$\cdot \prod_{\nu,\mu=0,1,\ldots n-1; \nu<\mu} (x_\nu - x_\mu)^2 d\alpha(x_0) d\alpha(x_1) \cdots d\alpha(x_{n-1})$$

y

$$D_n = \frac{1}{(1+n)!} \underbrace{\int_a^b \int_a^b \cdots \int_a^b}_{n+1} \prod_{\nu,\mu=0,1,\ldots n-1;\nu<\mu} (x_\nu - x_\mu)^2 d\alpha(x_0) d\alpha(x_1) \cdots d\alpha(x_n).$$

La expansión de Fuerier de una función arbitraria $f(x)$ en términos de los polinomios $\{p_n(x)\}$ tiene la forma

$$f(x) \sim f_0 p_0(x) + f_1 p_1(x) + \cdots + f_n p_n(x) + \cdots$$

con

$$f_n = \int_a^b f(x) p_n(x)\, d\alpha \qquad n = 0, 1, \cdots .$$

**Teorema 2.24.** *Sea n un entero positivo y sea*

$$\rho_n(x) = x^n + \sum_{j=0}^{n-1} c_j x^j.$$

*Entonces, la integral*

$$\int_a^b |\rho_n(x)|^2\, d\alpha(x)$$

*se vuelve un mínimo sí, y solo si $\rho(x) = p_n(x)/k_n$, donde $k_n$ se obtiene al normalizar $\rho_n(x)$.*

*Demostración.* De manera directa;

$$\int_a^b |\rho_n(x)|^2\, d\alpha(x) = \int_a^b \left| x^n + \sum_{j=0}^{n-1} c_j x^j \right|$$

reescribiendo $\rho_n(x)$ como $\sum_{j=0}^{n} d_j p_j(x)$, se tiene

$$\begin{aligned} \int_a^b |\rho_n(x)|^2 \, d\alpha(x) &= \int_a^b \left| \sum_{j=0}^{n} d_j p_j(x) \right|^2 \\ &= |d_0 p_0(x) + d_1 p_1(x) + \ldots + d_n p_n(x)|^2, \end{aligned}$$

si se hace $d_0 p_0(x) + d_1 p_1(x) + \ldots + d_n p_n(x) = (k^{-1}) p_n$, por ortogonalidad se sigue que la integral se vuelve un mínimo.

□

Si $k_n$ denota el coeficiente director de $p_n(x)$, el mínimo es $k_n^{-2}$. De (2.62) y de (2.48), se concluye que

$$k_n = (D_{n-1}/D_n^{-1})^{\frac{-1}{2}}. \tag{2.65}$$

### 2.2.3. Ceros y Núcleo de Polinomios Ortogonales

En esta sección, se describen las propiedades del mínimo y del máximo ( y todo lo que eso conlleva), de los polinomios ortogonales; siemore teniendo en cuenta que la distribución de tipo Stieltjes $\alpha(x)$ es una función fija no decreciente con muchos puntos de crecimiento en el intervalo $[a, b]$ finito o infinito. Con base en el estudio de [7], [8] y [9] y traducido por la autora de este trabajo.

Sea $f(x)$ una función de clase $L_\alpha(a,b)^2$ y sea $x^n \in L_\alpha(a,b)$ para $n = 0, 1, 2, \cdots$. Entonces, las integrales

$$\int_a^b |x|^2 \, d\alpha(x) \quad , \quad \int_a^b f(x) x^n \, d\alpha(x) \quad n = 0, 1, 2, 3 \cdots$$

existen en el sentido de Lebesgue.

Generalizando el teorema 2.24 de la sección 2.2.2 y definiendo $\{p_n(x)\}$ como el conjunto ortonormal de polinomios asociado con la distribución $d\alpha(x)$ en $[a, b]$, se establece el siguiente teorema:

**Teorema 2.25.** *La desviación cuadrática ponderada*

$$\int_a^b |f(x) - \rho(x)|^2 \, d\alpha(x),$$

*donde $\rho(x)$ se extiende sobre el conjunto de todos los polinomios de grado menor o igual a $n$, se vuelve un mínimo si y solo si $\rho(x)$ es la $n$-ésima suma parcial de la expansión de Fourier*

$$f(x) \sim f_0 p_0(x) + f_1 p_1(x) + f_2 p_2(x) + \cdots \tag{2.66}$$

*donde*

$$f_n = \int_a^b f(x) p_n(x) \, d\alpha(x) \qquad n = 0, 1, 2, \cdots$$

La demostración de éste teorema se tiene por la sección de aproximacion por polinomios del capítulo anterior. Despues, se concluyó que el mínimo en sí mismo es

$$\int_a^b |f(x)|^2 \, d\alpha(x) - \sum_{\nu=0}^{n} |f_\nu|^2$$

lo cual implica la desgualdad de Bessel

$$|f_0|^2 + |f_1|^2 + |f_2|^2 + \cdots + |f_n|^2 + \cdots \leq \int_a^b |f(x)|^2 \, d\alpha(x). \tag{2.67}$$

Análogamente, se puede hacer una generalización en las hipótesis del teorema 2,26; ya que en las condiciones, se consideró el conjunto de todos los polinomios $\rho_n(x)$ con coeficiente $x^n$ unitario. Ahora, el polinomio $\rho_n(x)$ se extiende sobre todo el conjunto de los polinomios de grado $n$ donde el coeficiente más grande de $x^n$ también se vuelve un mínimo sí, y solo si $\rho(x) = p_n(x)/k_n$.

**Definición 45.** El *núcleo de polinomios* $\{K_n(x,y)\}$ de una función $F_\mu$ está definido por

$$K_n(x,y) = \sum_{k=0}^{n} \overline{p_k(x)} p_k(y) \tag{2.68}$$

$$= \sum_{k=0}^{n} \overline{P_k(x)} P_k(y)/\epsilon, \tag{2.69}$$

para $n = 0, 1, \ldots$

**Teorema 2.26.** *Sea $x_0$ una constante arbitraria compleja, $\rho_n(x)$ un polinomio con coeficientes complejos, normalizados por la condición*

$$\int_a^b |\rho_n(x)|^2 \, d\alpha(x) = 1. \tag{2.70}$$

*El máximo de $|\rho_n(x_0)|$ está dado por los polinomios*

$$\rho(x) = \epsilon K_n(x_0, x)(K_n(x_0, x_0))^{\frac{-1}{2}} \tag{2.71}$$

*donde $|\epsilon| = 1$. El máximo en sí mismo es $K_n(x_0, x_0)$.*

*Demostración.* Sea

$$\rho_n(x) = \sum_{k=0}^{n} c_k p_k(x),$$

tal que cumple (2.70). Entonces

$$\begin{aligned} |\rho_n(x_0)|^2 &= \left| \sum_{k=0}^{n} c_k p_k(x_0) \right|^2 \\ &\leq \left( \sum_{k=0}^{r} |c_k|^2 \right) \left( \sum_{k=0}^{n} |p_k(x_0)|^2 \right) \\ &= K_n(x_0, x_0). \end{aligned}$$

□

El núcleo de polinomios

$$\begin{aligned}
K_n(x_0, x) &= \sum_{k=0}^{n} \overline{p_k(x_0)} p_k(x) \\
&= \overline{\sum_{k=0}^{n} \overline{p_k(x)} p_k(x_0)} \\
&= \sum_{k=0}^{n} \overline{p_k(\overline{x})} p_k(\overline{x_0}) \\
&= K_n(\overline{x}, \overline{x_0}),
\end{aligned}$$

puede ser usado para la representación de la $n$-ésima suma parcial $s_n(x)$ de la expansión de Fuourier (2.66) en una integral, así:

$$\begin{aligned}
s_n(x) &= f_0 p_0(x) + f_1 p_1(x) + \ldots + f_n p_n(x) \\
&= \sum_{\nu}^{n} p_\nu(x) \int_a^b f(t) p_\nu(t)\, d\alpha(t) \\
&= \int_a^b f(t) K_n(t, x)\, d\alpha(t). \qquad (2.72)
\end{aligned}$$

El análogo del teorema 2,26, en términos del núcleo es:

**Teorema 2.27.** *Sea $\rho(x)$ un polinomio que se extiende sobre el conjunto de todos polinomios de grado n; la integral*

$$\int_a^b |\rho(x)|^2\, d\alpha(x)$$

*se vuelve un mínimo sí, y solo si $\rho(x) = \dfrac{K_n(t,x)}{K_n(t,t)}$. El mínimo en sí mismo es $1/K_n(t,t)$*

La demostración de éste teorema es análoga a la del teorema 2,26 pero en [8] pp. 67, se encuentra con más detalle.

**Teorema 2.28.** *Si se asume que $B_n$ es real y $C_n > 0$ para todo $n = 1, 2, \cdots$. Entonces, los ceros de los polinomios generados por (2.25) y (2.26) son reales y simples.*

**Teorema 2.29.** *Sea $\{P_n\}$ una sucesión de polinomios ortogonales que satisfacen (2.45), y sea $[a, b]$ el intervalo cerrado más pequeño que contiene el soporte de $\alpha(x)$. Entonces todos los ceros de $P_n$ estan en $[a, b]$*

Para ver la demostracion de éste y del teorema anterior, ver [9].

# CAPÍTULO 3

## Polinomios ortogonales en el círculo unitario

Se presenta el capítulo para el que se estuvo encaminado todo el trabajo. Para terinar, se planea desarrollar la misma teoría vista en el capítulo dos, pero aterrizada a polinomios ortogonales en el círculo unitario. Se presenta el análogo de loa problemas del máximo y del mínimo, del núcleo de polinomios y el de la identidad de Christoffel-Darboux, entre otros; se hace uso del teorema de Gram-Schmidt y para el teorema más importante de todo el trabajo, que es el que muestra la relación entre polinomios ortogonales en $\mathbb{R}$ y en el círculo unitario, se utilizan los porlinomios de Chebyshev vistos en la primera sección del segundo capítulo. Todo esto con base en [7] y [9], con la traducción hecha por la autora de este trabajo.

Para empezar, se definen los momentos de los polinomios ortogonales en el círculo unitario, tomando $f(\theta)$, una función no negativa de periodo $2\pi$, integrable en el sentido de Lebesgue en el intervalo $[-\pi, \pi]$ y se asume que

$$\int_{-\pi}^{\pi} f(\theta)\, d\theta > 0.$$

Sean las constantes de Fourier

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta}\, d\theta, \qquad n \in \mathbb{Z}.$$

Notese que $c_{-n} = \bar{c}_n$, ya que:

$$\begin{aligned} c_{-n} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{-i(-n)\theta}\, d\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{in\theta}\, d\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\overline{e^{-in\theta}}\, d\theta \\ &= \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\overline{e^{-in\theta}}\, d\theta} \\ &= \bar{c}_n. \end{aligned}$$

Para este caso, la matríz de Gram (2.60) es:

$$T_n = \begin{pmatrix} c_0 & c_1 & \cdots & c_n \\ c_{-1} & c_0 & \cdots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{-n} & c_{-n+1} & \cdots & c_0 \end{pmatrix}, \tag{3.1}$$

como se vio en el capítulo anterior, esta matríz es hermitiana. Observese que tiene valor constante en cada diagonal; a una matríz de este estilo se le denomina *matríz de tipo Toeplitz*.

La correspondiente forma Hermitiana (2.63) asociada con $T_n$ es:

$$
\begin{aligned}
H_n &= \sum_{\nu,\mu}^{n} c_{\nu-\mu} u_\mu \overline{u}_\nu \\
&= \sum_{\nu}^{n}\sum_{\mu}^{n} c_{\nu-\mu} u_\mu \overline{u}_\nu \\
&= \sum_{\nu}^{n} c_{\nu-0} u_0 \overline{u}_\nu + c_{\nu-1} u_1 \overline{u}_\nu + c_{\nu-2} u_2 \overline{u}_\nu + \cdots + c_{\nu-n} u_n \overline{u}_\nu \\
&= c_0 u_0 \overline{u}_0 + c_{-1} u_1 \overline{u}_0 + c_{-2} u_2 \overline{u}_0 + \cdots + c_{-n} u_n \overline{u}_0 + c_1 u_0 \overline{u}_1 + c_0 u_1 \overline{u}_1 + c_{-1} u_2 \overline{u}_1 + \\
&\quad \cdots + c_{1-n} u_n \overline{u}_1 + \ldots + c_n u_0 \overline{u}_n + c_{n-1} u_1 \overline{u}_n + c_{n-2} u_2 \overline{u}_n + \cdots + c_0 u_n \overline{u}_n \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(0)\theta}\, d\theta\; u_0 \overline{u}_0 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(-1)\theta}\, d\theta\; u_1 \overline{u}_0 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(-2)\theta}\, d\theta\; u_2 \overline{u}_0 + \\
&\quad \cdots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(-n)\theta}\, d\theta\; u_n \overline{u}_0 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(1)\theta}\, d\theta\; u_0 \overline{u}_1 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(0)\theta}\, d\theta\; u_1 \overline{u}_1 + \\
&\quad \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(-1)\theta}\, d\theta\; u_2 \overline{u}_1 + \cdots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(1-n)\theta}\, d\theta\; u_n \overline{u}_1 + \ldots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(n)\theta}\, d\theta\; u_0 \overline{u}_n + \\
&\quad \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(n-1)\theta}\, d\theta\; u_1 \overline{u}_n + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(n-2)\theta}\, d\theta\; u_2 \overline{u}_n + \cdots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i(0)\theta}\, d\theta\; u_n \overline{u}_n \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\, d\theta\; u_0 \overline{u}_0 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{i\theta}\, d\theta\; u_1 \overline{u}_0 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{2i\theta}\, d\theta\; u_2 \overline{u}_0 + \\
&\quad \cdots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{in\theta}\, d\theta\; u_n \overline{u}_0 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-i\theta}\, d\theta\; u_0 \overline{u}_1 + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\, d\theta\; u_1 \overline{u}_1 + \\
&\quad \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{i\theta}\, d\theta\; u_2 \overline{u}_1 + \cdots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{i(n-1)\theta}\, d\theta\; u_n \overline{u}_1 + \ldots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-in\theta}\, d\theta\; u_0 \overline{u}_n + \\
&\quad \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{i(1-n)\theta}\, d\theta\; u_1 \overline{u}_n + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{i(2-n)\theta}\, d\theta\; u_2 \overline{u}_n + \cdots + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\, d\theta\; u_n \overline{u}_n \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) |u_0 + u_1 z + u_2 z^2 + \cdots + u_n z^n|\, d\theta,
\end{aligned}
$$

donde $z = e^{i\theta}$, es definido positivo y $T_n$ tiene el determinante positivo.

Se pueden construir polinomios ortonormales con respecto a $\alpha(x)$, por medio del proceso de Gram-Schmidt como se define acontinuación:

**Definición 46.** Al ortogonalizar el sistema

$$
\{(f(\theta))^{\frac{1}{2}} z^n\}, \qquad z = e^{i\theta}; \quad n, m = 0, 1, 2, \cdots,
$$

se obtiene el sistema de polinomios

$$\phi_0(z), \phi_1(z), \phi_2(z), \cdots, \phi_n(z)$$

con las siguientes propiedades:

(a) $\phi_n(z)$ es un polinomio de grado $n$ en el que el coeficiente de $z^n$ es real y positivo.

(b) El sistema $\{\phi_n(z)\}$ es ortonormal; esto es

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\phi_n(z)\overline{\phi_m(z)}\, d\theta = \delta_{nm} \qquad z = e^{i\theta}; \quad n, m = 0, 1, 2, \ldots.$$

Además el sistema de polinomios ortonormales $\{\phi_n(z)\}$ está univocamente determinado por las condiciones $a)$ y $b)$.

Si $f(\theta)$ es una función par, los coeficientes de $\phi_n(z)$ son reales.

Por teorema 2.23, $\phi_n(z)$ puede ser expresado como

$$\phi_n(z) = \frac{1}{\sqrt{D_n D_{n-1}}} \begin{vmatrix} c_0 & c_{-1} & \cdots & c_{-n} \\ c_1 & c_0 & \cdots & c_{-n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \cdots & c_{-1} \\ 1 & z & \cdots & z^n \end{vmatrix}, \tag{3.2}$$

así, por (2.64), se tiene

$$\phi_n(x) = \frac{1}{\sqrt{D_n D_{n-1}}} \begin{vmatrix} c_0 x - c_{-1} & c_{-1}x - c_{-2} & \cdots & c_{-n+1}x - c_{-n} \\ c_1 x - c_0 & c_0 x - c_{-1} & \cdots & c_{-n+2}x - c_{-n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1}z - c_{n-2} & c_{n-2}x - c_{n-3} & \cdots & c_0 x - c_{-1} \end{vmatrix}. \tag{3.3}$$

También, como se vio en el capítulo anterior en el Teorema 2.24

$$\phi_n(z) = \kappa_n z^n + \sum_{k=0}^{n-1} l_k z^k \tag{3.4}$$

y por (2.65)

$$\kappa_n = \sqrt{\frac{D_{n-1}}{D_n}} \tag{3.5}$$

Note que

$$\phi_n(0) = \frac{(-1)^n}{\sqrt{D_n D_{n-1}}} \begin{vmatrix} c_{-1} & c_{-2} & \cdots & c_{-n} \\ c_0 & c_{-1} & \cdots & c_{-n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-2} & c_{n-3} & \cdots & c_{-1} \end{vmatrix} \tag{3.6}$$

De la definición (1.4), si $f$ es un polinomio de grado $n$ de la forma

$$f(z) = \sum_{k=0}^{n} a_k z^k \qquad a_n \neq 0$$

entonces su polinomio recíproco $f^*(z)$ es de la forma

$$f^*(z) = \sum_{k=0}^{n} \overline{a_k} z^{n-k}.$$

**Teorema 3.30.** *Sea $F(e^{i\theta})$ una función medible para la cual*

$$\int_{-\pi}^{\pi} |F(e^{i\theta})|^2 \, d\theta$$

*existe. La desviación cuadrática ponderada*

$$\int_{-\pi}^{\pi} |F(z) - \rho(z)|^2 \, d\theta,$$

*donde $\rho(z)$ se extiende sobre todo el conjunto de los polinomios de grado $n$, es un mínimo si $\rho(z)$ es la $n$-ésima suma parcial de fourier*

$$
\begin{aligned}
F(z) &\sim F_0\phi_0(z) + F_1\phi_1(z) + \ldots + F_n\phi_n(z) + \ldots, \\
F_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} F(z)\overline{\phi(z)}\, d\theta, \qquad z = e^{i\theta}; n = 0, 1, 2, \ldots.
\end{aligned} \tag{3.7}
$$

*Como consecuencia se tiene la desigualdad de bessel (2.67) y, adicionalmente, la igualdad de Parseval (1.11) se satisface si $F(z)$ es regular y acotado para $|z| < 1$.*

Como consecuecia de este teorema se tiene

**Teorema 3.31.** *Sí $\rho_n(z)$ se extiende sobre todo el conjunto de los polinomios de grado $n$, el polinomio $\phi_n(z)/\kappa_n$ minimiza la integral*

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)|\rho_n(z)|^2\, d\theta.
$$

*El valor mínimo de la integral es $1/\kappa_n^2$.*

*Demostración.* Sea

$$
\rho_n(z) = \sum_{k=0}^{n} a_k\overline{\phi_k(z)}.
$$

entonces

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)|\rho_n(z)|^2\, d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\left|\sum_{k=0}^{n} a_k\overline{\phi_k(z)}\right|^2 d\theta, \tag{3.8}
$$

por el terema 2,26, la integral (3.8) es un mínimo sí, y solo si $\rho_n(z) = \phi_n(z)/\kappa_n$; así

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\left|\rho_n(z)\right|^2\, d\theta = k^{-2} + \sum_{k=0}^{n-1} |a_k|^2; \tag{3.9}
$$

por lo tanto, el mínimo en sí mismo es $1/\kappa_n^2$.

□

La mayoria de expresiones vistas en los capítulos a pasados para polinomios ortogonales, tienen su análogo en términos de los polinomios ortonormales $\phi_n(z)$; por ejemplo, el núcleo de polinomios (2.68) se reescribe, de ahora en adelante, como

$$s_n(a,z) = \sum_{k=0}^{n} \overline{\phi_k(a)}\phi_k(z), \qquad n = 0, 1, \ldots. \tag{3.10}$$

**Teorema 3.32.** *Sea a una constante compleja fija y sea $\rho_n(z)$ un polinomio que satisface*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)|g(z)|^2 d\theta = 1 \qquad z = e^{i\theta}. \tag{3.11}$$

*el máximo de $|\rho_n(a)|^2$ es alcanzado cuando*

$$\rho(z) = \epsilon\, s_n(a,z)(s_n(a,a))^{\frac{-1}{2}} \tag{3.12}$$

*donde $|\epsilon| = 1$. El máximo valor de $|\rho_n(a)|$ es $s_n(a,a)$.*

*Demostración.* Sea $\rho_n(z) = \sum_{k=0}^{n} a_k \phi_k(z)$, entonces

$$\begin{aligned}
\rho_n(a) &= \sum_{k=0}^{n} a_k \phi_k(a) \\
|\rho_n(a)|^2 &= \left|\sum_{k=0}^{n} a_k \phi_k(a)\right|^2 \\
&\leq \left(\sum_{k=0}^{n} a_k\right)\left(\sum_{k=0}^{n} \phi_k(a)\right) \\
&= s(a,a)
\end{aligned}$$

ya que, por la condición del teorema $(\sum_{k=0}^{n} a_k) = 1$. La igualdad se satisface sí, y solo si, para algun $\epsilon$ en el cítculo unitario $a_k = \epsilon\phi_k(a)$ para todo $k$, $0 \leq k \leq n$. □

**Teorema 3.33.** *Para todo $a \neq 0$ y $\rho_n(z)$, el núcleo de polinomios sólo es el polinomio que tiene la propiedad reproductiva*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)s_n(a,z)\overline{\rho_n(z)}\, d\theta = \overline{\rho_n(a)}, \qquad z = e^{i\theta}. \tag{3.13}$$

*Demostración.* Sea $\rho_n(z) = \sum_{k=0}^{n} a_k\phi_k(z)$; entonces

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} s_n(a,z)\overline{\rho_n(z)}\, d\theta &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=0}^{n}\overline{\phi_k(a)}\phi_k(z)\overline{\sum_{k=0}^{n} a_k\phi_k(z)}\, f(\theta)d\theta && (3.14)\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=0}^{n}\overline{\phi_k(a)}\phi_k(z)\overline{a_k}\overline{\phi_k(z)}\, f(\theta)d\theta && (3.15)\\
&= \sum_{k=0}^{n}\overline{a_k}\overline{\phi_k(a)}\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi_k(z)\overline{\phi_k(z)}f(\theta)d\theta && (3.16)\\
&= \sum_{k=0}^{n}\overline{a_k}\overline{\phi_k(a)} && (3.17)\\
&= \overline{\rho_n(a)}. && (3.18)
\end{aligned}$$

□

El núcleo de polinomios $s_n(a,z)$,puede ser usado para la representación de la suma parcial de la expansión (3.7) en forma de integrales (2.72).

Una consecuencia inmediata del teorema anterior, es que el núcleo de polinomios puede expresarse como

$$s_n(a,z) = (\bar{a}z)^n s_n(1/\bar{z}, 1/\bar{a});$$

el caso en el que $a = 0$

$$s(0,z) = \sum_{k=0}^{n}\overline{\phi_k(0)}\phi_k(z) = \kappa_n z^n\phi_n^*(z).$$

Además

$$\begin{aligned}
s(0,0) &= \sum_{k=0}^{n} \overline{\phi_k(0)}\phi_k(0) \\
&= \sum_{k=0}^{n} |\phi_k(0)|^2 \\
&= \kappa^2 \\
&= \frac{D_{n-1}}{D_n}
\end{aligned}$$

consecuentemene

$$|\phi_n(0)|^2 = \kappa_n^2 - \kappa_{n-1}^2$$

En particular, esto demuestra que $\kappa_n$ no decrece con respecto a $n$.

## 3.1. Relaciones de recurrencia

Finalizando, en esta sección se muestra la analogía que tienen las relaciones de recurrencia vistas en la sección 2.2.1 del capítulo 2, con respecto a los polinomios ortogonales en el círculo unitario con las propiedades vistas hasta ahora en éste capítulo. Todo, estudiado de [7] y [9] y traducido por la autora de este trabajo.

**Teorema 3.34.** *El análogo a la identidad de Christoffel-Darboux (2.42) es*

$$\begin{aligned}
s_n(a,z) &= \sum_{k=0}^{n} \overline{\phi_k(a)}\phi_k(z) \\
&= \frac{\overline{\phi_{n+1}^*(a)}\phi_{n+1}^*(z) - \overline{\phi_{n+1}(a)}\phi_{n+1}(z)}{1-\bar{a}z}.
\end{aligned}$$

*Además los polinomios $\{\phi_n(z)\}$ satisfacen las relaciones de recurrencia*

$$\kappa_n z\phi_n(z) = \kappa_{n+1}\phi_{n+1}(z) - \phi_{n+1}(0)\phi_{n+1}^*(z), \tag{3.19}$$

$$\kappa_n\phi_{n+1}(z) = \kappa_{n+1}z\phi_n(z) + \phi_{n+1}(0)\phi_n^*(z). \tag{3.20}$$

*Demostración.* Sea $\rho_n(z)$ un polinomio de grado $n$, entonces, con $z = e^{i\theta}$, se encuentra que

$$\begin{aligned}\int_{-\pi}^{\pi} \frac{\overline{\phi_{n+1}^*(a)}\phi_{n+1}^*(z) - \overline{\phi_{n+1}(a)}\phi_{n+1}(z)}{1-\bar{a}z}\overline{\rho_n(z)}d\theta &= \overline{\rho_n(a)} \int_{-\pi}^{\pi} \frac{\overline{\phi_{n+1}^*(a)}\phi_{n+1}^*(z) - \overline{\phi_{n+1}(a)}\phi_{n+1}(z)}{1-\bar{a}z}d\theta \\ &+ \int_{-pi}^{\pi} \left[\phi_{n+1}^*(a)\phi_{n+1}^*(z) - \overline{\phi_{n+1}(a)\phi_{n+1}(z)}\right] \frac{\overline{\rho_n(z) - \rho_n(a)}}{1-\bar{a}z}\ \theta,\end{aligned}$$

pero como $\rho_n(z)$ es de grado $n$, $\rho_n(z) - \rho_n(a) = (z-a)g(z)$ con $g(z)$ de grado menor o igual a $n-1$, entonces

$$\int_{-\pi}^{\pi} \phi_{n+1}^*(z)\overline{zg(z)}\ d\theta = \int_{-\pi}^{\pi} \overline{\phi_{n+1}}z^n g(1/z)\ d\theta = 0$$

y

$$\int_{-\pi}^{\pi} \phi_{n+1}(z)\overline{zg(z)}\ d\theta = 0.$$

Por lo tanto

$$\frac{\overline{\phi_{n+1}^*(a)}\phi_{n+1}^*(z) - \overline{\phi_{n+1}(a)}\phi_{n+1}(z)}{1-\bar{a}z} = cs_n(a,z)$$

donde $c$ es una constante; y al cambiar la $z$ por la $a$ y viceversa, se obtiene la conjugada compleja del núcleo. Al hacer $a = z = 0$ se obtiene que $c = 1$; así queda demostrada la identidad.

Para probar la identidad (3.20) se parte del polinomio ortonrmal

$$\phi(z) = \kappa_n\phi_{n+1}(z) - \kappa_{n+1}z\phi_n(z)$$

que es de grado, a lo sumo, $n$. Al hacer $\phi(z) = 0$ se tiene

$$\begin{aligned} 0 &= \kappa_n\phi_{n+1}(z) - \kappa_{n+1}z\phi_n(z) \\ \kappa_n\phi_{n+1}(z) &= \kappa_{n+1}z\phi_n(z)\end{aligned}$$

lo cual muestra parcialmente (3.20); por otro lado, para $1 \leq k \leq n$, se tiene

$$
\begin{aligned}
\phi(z) &= \kappa_n \phi_{n+1}(z) - \kappa_{n+1} z \phi_n(z) \kappa_n \phi_{n+1}(z) - \kappa_{n+1} z \phi_n(z) \\
\phi(z)\overline{z^k} &= \kappa_n \phi_{n+1}(z)\overline{z^k} - \kappa_{n+1} z \phi_n(z)\overline{z^k} \\
\phi(z)\overline{z^k} &= \kappa_n \phi_{n+1}(z)\overline{z^k} - \kappa_{n+1} z \phi_n(z)\overline{z^k} \\
\int_{-\pi}^{\pi} \phi(z)\overline{z^k}\, d\theta &= \int_{-\pi}^{\pi} \kappa_n \phi_{n+1}(z)\overline{z^k}\, d\theta - \kappa_{n+1} z \phi_n(z)\overline{z^k}\, d\theta \\
\int_{-\pi}^{\pi} \phi(z)\overline{z^k}\, d\theta &= 0 - \kappa_{n+1} \int_{-\pi}^{\pi} \phi_n(z)\overline{z^{k-1}}\, d\theta = 0.
\end{aligned}
$$

En otras palabras

$$0 = \int_{-\pi}^{\pi} z^k \overline{\phi(z)}\, d\theta = \int_{-\pi}^{\pi} \overline{z^{n-k}} \phi^*(z)\, d\theta$$

para $0 \leq n - k < n$. Por lño tanto, $\phi(z) = c\phi^*(z)$, y análogo a los teoremas anteriores, $c = \phi(0)/\kappa_n$. Con ésto, se concluye (3.20).

La prueba de (3.19) se hace por medio de la relacion equivalente

$$\kappa_n \phi_n^*(z) = \kappa_{n+1} \phi_{n+1}^*(z) - \overline{\phi_{n+1}(0)} \phi_{n+1}(z),$$

es necesario escribir las relaciones de recurrencia en términos de los polinomios mónicos

$$\Phi_n(z) = \phi_n(z)/\kappa_n.$$

En efecto, se tiene

$$
\begin{aligned}
\Phi_{n+1}(z) &= z\Phi_n(z) - \overline{\alpha_n} \Phi_n^*(z), & (3.21) \\
\Phi_{n+1}^*(z) &= \Phi_n^*(z) - \alpha_n z \Phi_n(z), & (3.22)
\end{aligned}
$$

donde

$$\alpha_n = -\overline{\Phi_{n+1}(0)} = -\overline{-\phi_{n+1(0)}}\kappa_{n+1}.$$

Los coeficientes $\{\alpha_n\}$ son llamados los *coeficientes de recursión*.

Si se elimina $\phi_n^*(z)$ entre (3.19) y (3.20), se obtiene la relación de recurrencia de tres términos

$$\begin{aligned} & \kappa_n\phi_n(0)\phi_{n+1}(z) + \kappa_{n-1}\phi_{n+1}(0)z\phi_{n-1}(z) \\ = \quad & [\kappa_n\phi_{n+1}(0) + \kappa_{n+1}\phi_n(0)z]\phi_n(z). \end{aligned}$$

Note que los coeficientes de recursión en (3.19), (3.20) y (3.23) se pueden escribir en términos de los determinantes de momentos pos medio de (3.5).

□

## 3.2. Relación con polinomios ortogonales en un intervalo real

Se llega al final de éste trabajo; todo lo visto hasta el momento es la base que soporta el entendimiento del teorema que abarca esta sección; este teorema fue tomado de [7] y la demostración fue hecha con base en [9], y sera sorprendente evidenciar como, los polinomios de Chebyshev (vistos en la primera sección del capítulo 2) intervienen en la relación entre los polinomios en $\mathbb{R}$ y en el círculo unitario.

**Teorema 3.35.** *Sea $w(x)$ una función peso en un intervaLo real $-1 \leq x \leq 1$, y sean $\phi_n$ los polinopmios ortonormales con respecto a $w(\cos\theta)$ en el círculo unitario. Se asume además que $\{t_n(x)\}$ y $\{u_n(x)\}$ son sucesiones ortonormales de polinomios cuyas medidas de ortogonalidad son $w(x)$ y $c_2(1-x^2)w(x)$ en $-1 \leq x \leq 1$ respectivamente. Con $x = (z+z^{-1})/2$, se tiene*

$$\begin{aligned} t_n(x) &= [1+\phi_{2n}(0)/\kappa_{2n}]^{\frac{-1}{2}}[z^-n\phi_{2n}(z) + z^n\phi_{2n}(1/z)] && (3.23) \\ &= [1-\phi_{2n}(0)/\kappa_{2n}]^{\frac{-1}{2}}[z^{-n+1}\phi_{2n-1}(z) + z^{n-1}\phi_{2n-1}(1/z)], && (3.24) \\ u_n(x) &= \frac{z^{-n-1}\phi_{2n+2}(z) + z^{n+1}\phi_{2n+2}(1/z)}{\sqrt{1-\phi_{2n+2}(0)/\kappa_{2n+2}}/(z-1/z)} && (3.25) \\ &= \frac{z^{-n}\phi_{2n+1}(z) + z^n\phi_{2n+}(1/z)}{\sqrt{1+\phi_{2n+2}(0)/\kappa_{2n+2}}/(z-1/z)} && (3.26) \end{aligned}$$

*Demostración.* Para probar (3.24) con $x = \cos\ \theta$ y $z = e^{i\theta}$; se demuestra primero que

$$\int_0^\pi [z^{-n}\phi_{2n}(z) + z^n\phi_{2n}(1/z)]T_k(\cos\theta)\ d\mu(\cos\theta) = 0,$$

con $0 \leq n$. Así, de la ortogonalidad de $\phi_n(z)$

$$\begin{aligned}\int_0^\pi & [\quad z^{-n}\phi_{2n}(z) + z^n\phi_{2n}(1/z)]T_k(\cos\theta)\ d\mu(\cos\theta) \\ &= \int_{-\pi}^\pi [z^{-n}\phi_{2n}(z) + z^n\phi_{2n}(1/z)][z^k + z^{-k}]\ d\mu(\cos\theta) \\ &= \int_{-\pi}^\pi \phi_{2n}(z)(z^{n-k} + z^{-(n+k)})\ d\mu(\cos\theta) = 0.\end{aligned}$$

Cómo

$$z^{-n}\phi_{2n}(z) + z^n\phi_{2n}(1/z) = [\kappa_{2n} + \phi_{2n}(0)][z^n + z^{-n} + \ldots]$$

con $z = e^{i\theta}$, $x = \cos\theta$, se ve que

$$\begin{aligned}\int_{-1}^1 & |\quad z^{-n}\phi_{2n}(z) + z^n\phi_{2n}(1/z)|^2\ d\mu(x) \\ &= [\kappa_{2n} + \phi_{2n}(0)] \int_{-1}^1 [z^{-n}\phi_{2n}(z) + z^n\phi_{2n}(1/z)][2T_n(x)]\ d\mu(x) \\ &= [\kappa_{2n} + \phi_{2n}(0)] \int_{-\pi}^\pi [\phi_{2n}(z) + z^{2n}\overline{\phi_{2n}(z)}]\ d\mu(\cos\theta) \\ &= [\kappa_{2n} + \phi_{2n}(0)]/\kappa_{2n},\end{aligned}$$

cuando $n > 0$. Ésto establece la primera línea de (3.24). Similarmente,

$$\int_{-\pi}^\pi [z^{1-n}\phi_{2n-1}(z) + z^{n-1}\phi_{2n-1}(1/z)]T_k(\cos\theta)\ d\mu(\cos\theta) = 0$$

para $0 \leq k < n$. Por o tanto, la segunda línea de (3.26) es una constante múltiple de la primera y la constante múltiple puede ser determinada igualando coeficientes de $x^n$. la prueba de (3.24) es similar.

$\square$

Para ver detalles de esta demostracion ver [9] pp. 229.

CAPÍTULO 4

# Conclusiones

- Los estudios que se realizaron con respecto a álgebra lineal, análisis funcional y teoría de la medida, permitieron entender cómo estaban formados los polinomios, el espacio $L_2(a,b)$, los polinomios ortogonales y la medida con la que se hace su producto interno; la definición de éste y la aproximación por polinomios.
- Al comprender las propiedades que los polinomios ortogonales en $\mathbb{R}$, el estudio de lo polinomios ortogonales en el círculo unitario se simplificó, ya que éstos heredan las propiedades de los primeros, con simples modificaciones.
- Los polinomios ortogonales en $\mathbb{R}$ y en el círculo unitario, estan relacionados por medio de dos sucesiones de polinomios ortonormales, cada una con cierta función peso. La construcción de dichos polinomios, se hace utilizando los polinomios ortonormales mónicos en el círculo unitario y polinomios de Chebyshev, los estudiados en el primer capítulo.

# Bibliografía

[1] K. HOFFMAN y R. KUNZE, *Álgebra Lineal* $1^a$ ed., México D.F, 1993.

[2] E. KREYSZIG, *Introductory Functional Analysis with Applications* $1^a$ ed., Estados Unidos de América, 1978.

[3] T.M. APOSTOL, *Calculus Vol II* $2^a$ ed., España, 2002.

[4] R.G. BARTLE, *The Elements of Integration and Lebesgue Measure* $1^a$ ed., Estados Unidos de América, 1995.

[5] T.S. CHIHARA, *An Introduction to Orthogonal Polynomials* $1^a$ ed., New York, Estados Unidos de América, 1978.

[6] ULF. GREEN y G. SZEGÖ, *Toeplitz forms and their applications* $2^a$ ed., Estados Unidos de América, 1984.

[7] G. SZEGÖ, *Orthogonal Polynomials* $4^a$ ed., Estados Unidos de América, 1939.

[8] P. GONZÁLEZ, L. DARUIS, *Ortogonalidad y cuadratura sobre la circunferencia unidad* $1^a$ ed., Venezuela, 2005.

[9] M. E. H. ISMAIL, *Classical and Quantum Orthogonal Polynomials in One Variable* $4^a$ ed., Reino Unido, 2005.