

Universidad Surcolombiana

Programa de Licenciatura en
Matemáticas

Facultad de Educación
Tópicos sobre Espacios de Hilbert, Teoría
Wavelets y Método Wavelet Haar en la
Solución Numérica de Ecuaciones
Diferenciales Ordinarias de Orden
Superior

ALEXANDER LOSADA ALMONACID

EDINSON OSWALDO DELGADO RIVAS

Neiva (Huila)

Enero de 2.012

Universidad Surcolombiana

Programa de Licenciatura en
Matemáticas

Facultad de Educación

Tópicos sobre Espacios de Hilbert, Teoría
Wavelets y Método Wavelet Haar en la
Solución Numérica de Ecuaciones
Diferenciales Ordinarias de Orden
Superior

Trabajo de grado

para optar al título de

LICENCIADO EN MATEMÁTICAS

ALEXANDER LOSADA ALMONACID

Código: 2006136390

EDINSON OSWALDO DELGADO RIVAS

Código: 2006263712

Asesor

MSc. LUIS ARTURO POLANÍA QUIZA

Neiva (Huila)

Enero de 2.012

Índice general

INTRODUCCIÓN	I
JUSTIFICACIÓN	III
OBJETIVOS	IV
1. Preliminares	1
1.1. Espacios Funcionales	5
1.2. Espacios Topológicos	6
1.3. Espacios Métricos	7
1.4. Axiomas de separabilidad	8
1.5. Espacio Compacto	9
1.6. Espacio Vectorial	10
1.7. Espacios Normados	11
1.8. Espacios de Banach	13
1.9. Algunos ejemplos de Espacios de Banach	13
2. Teoría de La Medida e Integral de Lebesgue	16
2.1. Aspectos Históricos	17
2.2. Construcción de la Integral de Lebesgue	21
2.3. Teoría de la Medida	22
2.4. Sigma álgebra, (Medida)	23
2.5. La Función Medida	24
2.6. Propiedades de las Medidas	25
2.7. Medidas, Sigma-Finitas	26
2.8. Completitud	27
2.9. La Medida del conteo	28

2.10. La Medida de Lebesgue	28
2.11. Integración	30
2.12. Una Interpretación Intuitiva	32
2.13. Teoremas Integrales Básicos	33
3. Producto Interno y Espacios de Hilbert	35
3.1. Introducción Histórica	35
3.2. Espacios con Producto Interno	38
3.3. Norma en un Espacio con Producto Interno	39
3.4. Ejemplos de Espacios con Producto Interno	39
3.5. Caracterización de Espacios Normados	39
3.6. Espacios de Hilbert	40
3.7. Algunos Ejemplos de Espacios de Hilbert	40
4. Sistemas Wavelet	42
4.1. Sistemas wavelet de primera generación	42
4.2. Características de sistemas wavelet	42
4.3. Función escala	43
4.4. Función Wavelet	47
5. Wavelet Haar	49
5.1. Función Haar Escala	49
5.1.1. Relación ortogonalidad y normalidad	55
5.2. Función Wavelet Haar	57
5.2.1. Relación de ortogonalidad y normalidad	62
6. Matrices Haar	63
7. Solución Numérica EDO'S Usando Matrices Wavelet Haar	70
7.1. Método Wavelet Haar para Ecuaciones Diferenciales Ordinarias .	70
8. Conclusiones	81



INTRODUCCIÓN

En el curso de ecuaciones diferenciales que tomamos durante nuestro pregrado, estudiamos, algunos métodos analíticos para la solución de Ecuaciones Diferenciales Ordinarias (EDO's) y problemas de valor inicial, de segundo orden, como máximo, que modelizan tanto sistemas mecánicos como sistemas electromagnéticos.

Dentro de los métodos conocidos para solucionar problemas de valor inicial, como por ejemplo: el método de las series de potencias, la transformada de Laplace y la transformada de Fourier, permiten encontrar una solución **exacta** para dichas ecuaciones diferenciales.

En la Teoría de Wavelets (Ondículas u Ondoletas) dentro de sus muchas aplicaciones, se encuentra una dedicada a la solución **numérica** de EDO'S de orden superior basada en la wavelet Haar propuesta en 1.909 por Alfred Haar conocida como la más antigua y sencilla de las ondículas.

Las wavelets Haar están formadas por pares de funciones constantes a trozos y son matemáticamente las más sencillas de las wavelets ortonormales con soporte compacto, esto es, están definidas para cierto intervalo, y fuera de este su valor es cero. Debido a la sencillez de su definición matemática, el método wavelet Haar ha resultado ser una herramienta eficaz para la solución de muchos problemas, entre ellos, la compresión de imágenes, el filtrado de señales, la detección de variaciones de una señal en el plano tiempo-frecuencia, como también **en la solución numérica de ecuaciones diferenciales ordinarias de orden superior**, entre otros.

Se sabe, las ondas Haar no son continuas; y no se pueden diferenciar en los puntos de discontinuidad. Por tal razón, no es posible aplicar las wavelets Haar directamente para resolver ecuaciones diferenciales ordinarias. Una posibilidad para solucionar este inconveniente, fue propuesta por Chen y Hsiao en 1.997. Ellos recomiendan expandir en la serie de Haar no la función como tal, pero

sí la mayor derivada que aparece en la ecuación diferencial. Las otras derivadas (y la función) se obtienen a través de integraciones.

El sistema es discretizado por Galerkin o método de colocación, el cual fue desarrollado en el libro *Un Enfoque Cualitativo a Las EDO'S y Teoría Wavelets*, escrito por el magister en matemáticas Luis Arturo Polanía Quiza, profesor de tiempo completo adscrito a la Facultad de Ciencias Exactas y Naturales de la Universidad Surcolombiana.

En este trabajo se solucionan algunas EDO'S de segundo orden, pero los beneficios del método de Haar, se muestra especialmente en el caso de EDO'S de **orden superior**.

El presente trabajo esta constituido de la siguiente manera:

Unos preliminares en los que se dan algunas definiciones como espacio vectorial, espacio de funciones, espacios normados, espacio producto interno, espacio de Hilbert, espacio de Banach, $l^2(\mathbb{N})$, $L^2(\mathbb{R})$, bases en un espacio de funciones, bases wavelets, base Haar, teoría de la medida, luego, se definen los sistemas wavelet, la wavelet Haar, las matrices Haar y por último el método numérico como tal, con el cual se solucionan algunas ecuaciones diferenciales de segundo y cuarto orden, para evidenciar la eficiencia y conveniencia del método numérico.



JUSTIFICACIÓN

Como requisito parcial para optar al título de Licenciados en Matemática, presentamos este trabajo, resultado de uno de los temas desarrollados en el semillero DINUSCO.

La motivación por este tema inició con la participación en el cursillo: Introducción a la Teoría de Onditas-Wavelets- y sus aplicaciones, desarrollado en el marco del XVII Congreso Colombiano de Matemáticas Cali 3 al 6 de Agosto de 2.009, al cual asistió el magister en matemáticas Luis Arturo Polanía Quiza director del semillero y otros dos estudiantes de ingeniería electrónica, con los cuales nos iniciamos en el estudio de esta rica e innovadora teoría.

Como fruto de este trabajo se realizó una participación, como ponentes en el ICAMI (Conferencia Internacional de Matemática Aplicada e Informática), del 28 de Noviembre al 3 de Diciembre de 2.010, en San Andres islas.

Con el fin de mejorar lo hecho en el ICAMI, nos propusimos estudiar más a fondo las bases teóricas que fundamentan este método numérico para la solución de Ecuaciones Diferenciales Ordinarias y problemas de valor inicial, de orden superior, dejando un inicio para seguir con este método aplicado a ecuaciones diferenciales ordinarias no lineales y ecuaciones diferenciales parciales.

OBJETIVOS

Objetivo General:

- Implementar el método Wavelet Haar en la solución numérica de ecuaciones diferenciales ordinarias de orden superior.

Objetivos Específicos:

- a. Conocer algunos espacios de funciones entre ellos los de Hilbert, base para la teoría wavelet.
- b. Comprender algunos conceptos elementales propios del análisis funcional.
- c. Mostrar algunos tópicos acerca de la teoría wavelet.
- d. Presentar la familia de wavelets Haar y una de sus aplicaciones en el campo de los métodos numéricos.

¿Cómo surgieron los espacios de funciones particularmente los Espacios de Hilbert, y que aplicaciones puede tener la Teoría Wavelets en la resolución de Ecuaciones Diferenciales Ordinarias lineales de orden superior?

Este trabajo pretende sumergirnos en la cultura de algunos Espacios de Funciones que utilizamos frecuentemente, y otros también de gran interés y utilidad; la directriz propuesta es describir medianamente la evolución de los espacios clásicos de funciones comenzando por el espacio básico, **El Espacio de Banach** sobre el cual estudiaremos algunos ejemplos.

También estudiaremos otros espacios de funciones tanto en el escenario teórico, como en el escenario de las aplicaciones particularmente **El Espacio de Hilbert** $L^2(\mathbb{R})$ que es un espacio necesario para construir y desarrollar la **Teoría Wavelets**.

Se quiere comprender de la mejor manera, cómo los Espacios de Funciones surgieron en el escenario de la matemática. Posiblemente los espacios más familiares para un estudiante de matemática sean los espacios de Banach y los de Hilbert, ya que estos espacios fueron profundamente estudiados en el siglo pasado y permitieron el desarrollo de nuevas teorías en variados sectores de la matemática aplicada como La Teoría Wavelet y la Teoría de Marcos, entre otras enriqueciendo de esta manera el campo de las aplicaciones.

Historia de Vida de Algunos Matemáticos que Contribuyeron al Desarrollo del Análisis Funcional, la Teoría Wavelets y sus Aplicaciones.

Jean-Baptiste-Joseph Fourier, (Auxerre, 21 de marzo de 1768 - París 16 de mayo de 1830) Famoso matemático que en 1827, escribió un tratado en el que



Joseph Fourier

indicaba que la acumulación de gases en la atmósfera provocaría una subida de temperatura en la Tierra, efecto que, años más tarde, pasaría a denominarse “Efecto Invernadero”. Introdujo también el concepto del “Calor Negro”, que se basa en un balance entre la radiación infrarroja ganada y perdida de los planetas mediante cambios en la temperatura.



También formuló el teorema de descomposición de funciones periódicas en series trigonométricas conocido como Teorema de las Series de Fourier. Participó en la Revolución Francesa y viajó a Egipto con Napoleón, llegando a ser parte importante de su gobierno.

Fue alumno de Lagrange y Laplace, y en su labor investigadora destacan sus trabajos en el campo de la termodinámica.

Murió en París en 1830, siendo secretario perpetuo de las secciones de Matemáticas y Física de la Real Academia de las Ciencias Francesa y dejando a medio resolver una teoría sobre ecuaciones que fue editada y publicada por Navier, alumno suyo.

Johann Carl Friedrich Gauss, (Brunswick, 30 de abril de 1777 - Göttingen, 23 de febrero de 1855), matemático, astrónomo y físico alemán que contribuyó significativamente en muchos campos, incluida la teoría de números, el análisis matemático, la geometría diferencial, la geodesia, el magnetismo y la óptica. Considerado el príncipe de las matemáticas y el matemático más grande desde la antigüedad, Gauss ha tenido una influencia notable en muchos campos de la matemática y de la ciencia, y es considerado uno de los matemáticos que más influencia ha tenido en la historia. Fue de los primeros en extender el concepto de divisibilidad a otros conjuntos.



Gauss fue un niño prodigio de quien existen muchas anécdotas acerca de su asombrosa precocidad, e hizo sus primeros grandes descubrimientos mientras era apenas un adolescente. Completó su magnum opus, “Disquisitiones Arithmeticae” a los veintiún años, aunque no sería publicado hasta 1801, trabajo que fue fundamental para que la teoría de los números se consolidara y ha moldeado esta área hasta los días presentes.

Georg Friedrich Bernhard Riemann, (Breselenz, Alemania, 17 de septiembre de 1826 - Verbania, Italia, 20 de julio de 1866) matemático alemán que realizó contribuciones muy importantes en análisis y geometría diferencial, algunas de ellas fundamentales para el desarrollo de la relatividad general. Su nombre está conectado con la función Zeta, la integral de Riemann, el lema de Riemann, las variedades de Riemann, las superficies de Riemann y la geometría de Riemann.

Stefan Banach (Cracovia, Imperio austrohúngaro 30 de marzo de 1892 - Leópolis, Polonia 31 de agosto de 1945, actual Ucrania), fue un matemático



polaco, uno de los destacados de la Escuela de Matemáticas de Lwow (Lwowska Szkoła Matematyki) en la Polonia previa a la guerra. Fue un autodidacto en matemáticas; su talento fue descubierto accidentalmente por Juliusz Mien y posteriormente por Hugo Steinhaus.

Cuando la Segunda Guerra Mundial comenzó, Banach era el presidente de la Sociedad Matemática Polaca y profesor en la Universidad de Leópolis (Uniwersytet Lwowski). Era un miembro de la Academia de las Ciencias de la República Socialista Soviética de Ucrania, y por otra parte mantenía una buena relación con los matemáticos soviéticos, y se le permitió permanecer en su cargo a pesar de la ocupación soviética, desde 1939, de la ciudad. Banach sobrevivió la posterior ocupación alemana desde julio de 1941 hasta febrero de 1944, ganándose la vida alimentando un piojo con su sangre para el Instituto de Investigación sobre el Tifus del profesor Rudolf Weigl. Su salud empeoró durante la ocupación, y desarrolló un cáncer de pulmón. Tras la guerra, Leópolis se incorporó a la Unión Soviética, y Banach murió allí antes de que pudiera ser repatriado a Cracovia, Polonia.



Stefan Banach

David Hilbert, (23 de enero de 1862, Königsberg, Prusia Oriental - 14 de febrero de 1943, Göttingen, Alemania) fue un matemático alemán, reconocido como uno de los más influyentes del siglo XIX y principios del XX. Estableció su reputación como gran matemático y científico inventando o desarrollando un gran abanico de ideas, como la teoría de invariantes, la axiomatización de la geometría y la noción de espacio de Hilbert, uno de los fundamentos del análisis funcional.

Hilbert y sus estudiantes proporcionaron partes significativas de la infraestructura matemática necesaria para la mecánica cuántica y la relatividad general. Fue uno de los fundadores de la teoría de la demostración, la lógica matemática y la distinción entre matemática y metamatemática.

Adoptó y defendió vivamente la teoría de conjuntos y los números transfinitos de Cantor y en 1900 presentó un conjunto de problemas que establecieron el curso de gran parte de la investigación matemática del siglo XX.

En la pugna por demostrar correctamente algunos de los errores cometidos por Einstein, en la teoría general de la relatividad, David Hilbert se adelantó a las correcciones de Einstein, sin embargo nunca quiso otorgarse el mérito.

Ingrid Daubechies, (Bélgica, 17 de agosto de 1954) física y matemática, actualmente es profesora en los departamentos de matemática aplicada en la Universidad de Princeton. En enero de 2011 se movió a la Universidad de Duke como profesora de matemáticas. Ella es la primera mujer Presidente de la Unión Matemática Internacional (2011-2014), conocida por su trabajo con wavelets en la compresión de imágenes digitales.



David Hilbert





Ingrid Daubechies

Stéphane Mallat G, (París , Francia) hizo algunas contribuciones fundamentales al desarrollo de la Teoría Wavelet a finales de 1980 y principios de 1990. También ha realizado trabajos en las matemáticas aplicadas, procesamiento de señales, compresión de música e imágenes digitales.

Colaboró con Yves Meyer para desarrollar el Análisis Multiresolución (MRA), apoyó la construcción de ondas de forma compacta, lo que hizo la aplicación de las wavelets práctica de ingeniería de aplicaciones mediante la demostración de la equivalencia de la cabrilla bases y conjugado de espejo de filtros utilizados en imágenes discretas, multifrecuencia de bancos de filtros en procesamiento de señales. También desarrolló con Sifen Zhong la Transformada Wavelet Módulo Máxima, Método para la caracterización de la imagen, que utiliza el máximo local de los coeficientes wavelets a varias escalas para reconstruir las imágenes.

Mallat autor del texto “Un Paseo Wavelet de Procesamiento de Señales” común en algunos cursos de matemáticas aplicadas y de ingeniería. Ha sido profesor en la New York University, Massachusetts Institute of Technology, Universidad de Tel Aviv y en la actualidad en la École Polytechnique.

Para un mayor dominio, comprensión y ubicación contextual de cada uno de los temas tratados en éste trabajo, de manera general diremos que:

1. **El análisis real** es la rama de la matemática que se ocupa de los números reales y sus funciones. Se puede ver como una extensión rigurosa del cálculo, que estudia más profundamente las sucesiones y sus límites, continuidad, derivación, integración, y las sucesiones de funciones. Además empieza un proceso de abstracción cuyo sendero pasa por la topología.
2. **El espacio topológico** es la noción de base de la topología elemental, dominio que sólo depende de la teoría de conjuntos (no está construido a partir de otra cosa), y que tiene aplicaciones importantes en el campo del análisis porque permite definir rigurosamente la continuidad y los límites.
3. **Un espacio métrico** es un tipo particular de espacio topológico, donde está definida una distancia entre puntos. Corresponde al caso muy común en que se dispone de una noción de distancia sobre el espacio.
4. **El álgebra abstracta** es el campo de la matemática que estudia las estructuras algebraicas como las de grupo, anillo, cuerpo o espacio vectorial. Muchas de estas estructuras fueron definidas formalmente en el siglo XIX, y, de hecho, el estudio del álgebra abstracta fue motivado por la necesidad de más exactitud en las definiciones matemáticas. El estudio del álgebra abstracta ha permitido observar con claridad lo intrínseco de las afirmaciones lógicas en las que se basa la matemática y las ciencias naturales. Además, a lo largo de la historia, los algebristas descubrieron que estructuras lógicas aparentemente diferentes muy a menudo pueden caracterizarse de la misma forma con un pequeño conjunto de axiomas.



Stéphane Mallat



El término álgebra abstracta se usa para distinguir este campo del álgebra elemental o del álgebra de la escuela secundaria que muestra las reglas correctas para manipular fórmulas y expresiones algebraicas que conciernen a los números reales y números complejos. El álgebra abstracta fue conocida durante la primera mitad del siglo XX como **álgebra moderna**.

5. **El análisis funcional** es la rama de las matemáticas, mas específicamente del análisis, que trata del estudio de espacios de funciones. Tienen sus raíces históricas en el estudio de transformaciones tales como transformación de Fourier y en el estudio de las ecuaciones diferenciales y ecuaciones integrales. La palabra funcional se remonta al cálculo de variaciones, implicando una función cuyo argumento es una función. Su uso en general se ha atribuido a Volterra.

En la visión moderna inicial, se consideró el análisis funcional como el estudio de los espacios vectoriales normados completos sobre los reales o los complejos. Tales espacios se llaman Espacios de Banach. Un ejemplo importante es el espacio de Hilbert, donde la norma surge de un producto escalar. Estos espacios son de importancia fundamental en la formulación matemática de la mecánica cuántica.

Más general y modernamente, el análisis funcional incluye el estudio de los espacios de Fréchet y otros espacios vectoriales localmente convexos y aún topológicos. Un objeto importante de estudio en análisis funcional son los operadores lineales continuos definidos en los espacios de Banach y de Hilbert.

1.1. Espacios Funcionales

En matemáticas, un espacio funcional es un conjunto de funciones de un conjunto X a un conjunto Y de una clase dada. Se llama un espacio porque en la mayoría de las aplicaciones, es un espacio topológico o un espacio vectorial. Los espacios funcionales aparecen en varias áreas de las matemáticas:

1. **En la teoría de conjuntos**, el conjunto de partes de un conjunto X se puede identificar con el conjunto de todas las funciones de X en $\{0, 1\}$ (funciones características);
2. **En el álgebra lineal** el conjunto de toda las transformaciones lineales del espacio vectorial de V en otro, W , sobre el mismo cuerpo, es en sí mismo un espacio vectorial;
3. **En el análisis funcional** se ve lo mismo para las transformaciones lineales continuas, incluyendo topologías en los espacios vectoriales subyacentes, y muchos de los ejemplos principales son espacios funcionales con topología;



4. **En la topología**, uno puede procurar poner una topología en las funciones continuas del espacio topológico X a otro E , cuya utilidad depende de la naturaleza de los espacios; en la topología algebraica, el estudio de la teoría de la homotopía es esencialmente el de invariantes discretos de espacios funcionales; en la teoría del proceso estocástico, el problema técnico básico es cómo construir una medida de probabilidad en un espacio funcional de trayectorias del proceso (funciones del tiempo);
5. **En la teoría de categorías** el espacio funcional aparece como bifunctor canónico de representación pero como funtor simple de tipo $[X, -]$ como funtor adjunto, a un funtor del tipo $(Xx-)$ en objetos;
6. **en el cálculo lambda y la programación funcional**, tipos de espacio funcional se utilizan para expresar la idea de función de orden superior.
7. **En la teoría de dominios**, la idea básica es encontrar construcciones de un orden parcial que pueda modelar cálculo lambda, creando una buena categoría cartesiano cerrada.
8. Otra idea relacionada desde **la física** es el espacio de configuración. Esto no tiene un significado único, pero para N partículas moviéndose en una variedad M puede ser el espacio de posiciones M^N o el subespacio donde no hay dos posiciones iguales.

Las configuraciones de una curva serían un espacio funcional de alguna clase. En la mecánica cuántica una formulación acentúa las historias como configuraciones. En breve, un espacio de configuración es típicamente *la mitad* (distribución lagrangiana) del espacio de fase que se construye desde un espacio funcional. Además se relacionan con *la teoría de trenzas*, también, puesto que la condición en una cuerda de no pasar por sí misma es formulada cortando diagonales de los espacios funcionales.

1.2. Espacios Topológicos

Los conceptos de conjunto abierto y vecindad pueden definirse mediante una métrica en el espacio considerado. Otro camino es, sin definir métrica alguna en un conjunto dado, definir directamente, mediante axiomas, el sistema de conjuntos abiertos. Este camino conduce a los espacios topológicos; respecto de éstos, los espacios métricos representan un caso especial.

Definición 1.2.1 *Sea X un conjunto cualquiera. Se llama topología en X a todo sistema τ de subconjuntos G de X que verifica las siguientes condiciones:*

1. *El propio conjunto X y el conjunto vacío pertenecen a τ .*
2. *La unión $\cup_{\alpha} G_{\alpha}$ de un número cualquiera (finito o infinito) y la intersección $\cap_{k=1}^n G_{\alpha}$ de un número finito de conjuntos de τ , pertenecen a τ*



El par (X, τ) se denomina espacio topológico. Los conjuntos pertenecientes al sistema τ se denominan abiertos.

Ejemplo 1.2.1 Sea T un conjunto arbitrario. Consideramos como abiertos a todos sus subconjuntos. Es obvio entonces que se cumplen los axiomas de la definición y obtenemos entonces un espacio topológico. En él todos los conjuntos son a la vez abiertos y cerrados y por lo tanto, cada uno coincide con su clausura.

Ejemplo 1.2.2 En el otro extremo se puede considerar en un conjunto abierto arbitrario X la topología compuesta por sólo dos conjuntos: el propio conjunto X y el conjunto vacío. Aquí la clausura de todo conjunto no vacío coincide con todo X . Este espacio topológico, suele llamarse “espacio de puntos pegados”.

Definición 1.2.2 Sean X e Y dos espacios topológicos, y f una función de X en Y . Se dice que f es un homeomorfismo de X en Y si y sólo si se cumple lo siguiente:

1. f es biyectiva (uno a uno y sobreyectiva).
2. f es continua.
3. La inversa f^{-1} es también continua.

Si $f : X \rightarrow Y$ es un homeomorfismo, Y se dice homeomorfo a X . Si dos espacios son homeomorfos entonces tienen exactamente las mismas propiedades topológicas.

1.3. Espacios Métricos

En matemática, un espacio métrico es un tipo particular de espacio topológico, donde está definida una distancia entre puntos, y corresponde al caso muy común en que se dispone de una noción de distancia sobre el espacio.

Definición 1.3.1 (Espacio métrico) Un espacio métrico es un par (X, d) donde X es un conjunto arbitrario no vacío y $d : X \times X \rightarrow \mathbb{R}$ es una aplicación, llamada distancia o métrica, tal que, para cualesquiera $x, y, z \in X$, se verifica:

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \Leftrightarrow x = y$
3. $d(x, y) = d(y, x)$ (simetría)
4. $d(x, y) \leq d(x, z) + d(z, y)$ (desigualdad triangular).

Observación: Los conjuntos abiertos de cualquier espacio métrico verifican los axiomas de la definición de un espacio topológico, por lo tanto, todo espacio métrico es un espacio topológico.



Otras definiciones

Definición 1.3.2 Se llama **bola abierta** centrada en $a \in M$ y de radio $r > 0$, al conjunto $\{x \in M \mid d(x, a) < r\} \subset M$. Se denota usualmente como $B(a, r)$ o como $B_r(a)$.

Definición 1.3.3 Se llama **bola cerrada** centrada en $a \in M$ y de radio $r > 0$, al conjunto $\{x \in M \mid d(x, a) \leq r\} \subset M$. Se denota usualmente como $\overline{B}c(a, r)$ o por $\overline{B}(a, r)$, donde B se lee como la clausura de B .

En análisis matemático, una *sucesión de Cauchy* es una sucesión tal que la distancia entre dos elementos se va reduciendo a medida que se avanza. Se llama así en honor al matemático francés **Augustin Louis Cauchy**. Su interés radica en que se puede verificar que una sucesión es de Cauchy sin conocer el punto de convergencia.

Definición 1.3.4 (Sucesión de Cauchy) En un espacio métrico (X, d) , una sucesión $\{x_k\}$ se dice de Cauchy si para todo $\varepsilon > 0$ existe un N_ε en los naturales, tal que para todo $n, m > N_\varepsilon$ se verifica que la distancia entre dos elementos $d(x_n, x_m)$ es inferior a ε . Simbólicamente,

$$(\forall \varepsilon > 0) (\exists N_\varepsilon \in \mathbb{N}) (d(x_n, x_m) < \varepsilon \text{ si } n, m > N_\varepsilon)$$

Ejemplo 1.3.1 En $(0, 2)$, la sucesión $\frac{1}{k}$ es de Cauchy, sin embargo no es convergente en dicho espacio, pues su límite es cero y éste no está en el conjunto donde definimos la sucesión.

1.4. Axiomas de separabilidad

Aunque muchos conceptos principales de la teoría de espacios métricos (conjuntos abiertos, entornos, bases, etc.) se extienden fácilmente a cualquiera espacio topológico, sin embargo, un espacio topológico arbitrario representa un ente demasiado general desde el punto de vista de los problemas del Análisis. En estos espacios se producen, a veces, situaciones que difieren de modo sustancial de lo que puede ocurrir en un espacio métrico. Así, por ejemplo, en un espacio topológico un conjunto finito de puntos puede no ser cerrado.

Entre los espacios topológicos se distinguen espacios que por sus propiedades se aproximan a los espacios métricos. Para ello es necesario agregar a los axiomas ya establecidos en la definición de espacios topológicos algunas condiciones adicionales, por ejemplo condiciones de numerabilidad que permiten estudiar la topología del espacio a través del concepto de convergencia. Otro tipo de condiciones adicionales y de naturaleza distinta son los *axiomas de separabilidad*.

Axioma 1.4.1 Primer axioma de separabilidad T_1 : Para dos puntos diferentes cualesquiera x e y del espacio T existe una vecindad O_x del punto x , que



no contiene al punto y , y una vecindad O_y del punto y que no contiene al punto x .

Los espacios que verifican este axioma se denominan T_1 – *espacios*. En un T_1 – *espacio*, todo punto es un conjunto cerrado. Por lo tanto se puede demostrar que en un T_1 – *espacio* todo conjunto formado por un número finito de puntos es cerrado. Además, puede demostrarse que el axioma T_1 es equivalente a pedir que todos estos conjuntos sean cerrados. El segundo axioma es una acentuación del primero.

Axioma 1.4.2 Segundo axioma de separabilidad T_2 o axioma de Hausdorff: Para dos puntos cualesquiera x e y del espacio topológico T existen vecindades O_x y O_y de intersección vacía.

Los espacios topológicos que verifican este axioma se denominan T_2 – *espacio* o espacios de Hausdorff. Todo espacio de Hausdorff es un T_1 – *espacio*, pero el recíproco no es cierto.

1.5. Espacio Compacto

En el análisis matemático desempeña un papel fundamental el siguiente hecho, conocido como **Lema de Heine-Borel**:

Lemma 1 De cualquier cubrimiento del segmento $[a, b]$ de la recta numérica por medio de intervalos se puede extraer un subcubrimiento finito.

Esta afirmación continúa siendo válida cuando en lugar de intervalos se consideran conjuntos abiertos cualesquiera: de todo cubrimiento abierto del segmento $[a, b]$ se puede extraer un subcubrimiento finito. Partiendo de esta propiedad del segmento de la recta numérica, introducimos el siguiente concepto importante:

Definición 1.5.1 (Espacio compacto) Un espacio topológico se denomina compacto, cuando cualquier cubrimiento abierto suyo contiene un subcubrimiento finito. Un espacio topológico compacto que verifica el axioma de separabilidad de Hausdorff se llama un compacto.

La propiedad de compacidad la tienen. Además de los segmentos, todos los subconjuntos cerrados acotados de un espacio euclídeo de cualquier dimensión. Por el contrario, la recta, el plano y el espacio \mathbb{R}^3 son ejemplos elementales de espacios no compactos.

Definición 1.5.2 Un sistema de conjuntos $\{A_i\}_{i \in I}$ del conjunto T se denomina centrado, cuando cualquier intersección finita $\bigcap_{i=1}^n A_i$ de elementos del sistema es no vacía.

Teorema 1.5.1 Para que un espacio topológico T sea compacto, es necesario y suficiente que verifique la condición: Todo sistema centrado de subconjuntos cerrados de este espacio tiene intersección no vacía.



Algunas propiedades

Teorema 1.5.2 *La imagen continua de un espacio compacto es compacto.*

Teorema 1.5.3 *Todo subconjunto cerrado de un espacio compacto es un compacto.*

Teorema 1.5.4 *Todo compacto resulta cerrado en cualquier espacio de Hausdorff que lo contiene.*

Los dos últimos teoremas indican que en los espacios de Hausdorff la compacidad es una propiedad interna del espacio, es decir que todo compacto continúa siendo compacto, aunque sea sumergido en espacios de Hausdorff cada vez mas amplios.

1.6. Espacio Vectorial

Definición 1.1 *Espacio vectorial.* Un conjunto $X \neq \Phi$ se llama **espacio vectorial** (y sus elementos se llamarán vectores) respecto al cuerpo E (a cuyos elementos llamaremos escalares), en X se definen dos operaciones suma y multiplicación por escalar, con las propiedades algebraicas siguientes:

i $(\forall x, y \in X)(x + y \in X)$.

ii $(\forall x, y, z \in X)((x + y) + z = x + (y + z))$.

iii Existe un único 0 tal que $x + 0 = 0 + x, \forall x \in X$ (no confundirlo con el elemento neutro de E).

iv $(\forall x \in X)(\exists(-x) \in X)$ tal que $x + (-x) = 0$. Al elemento $-x \in X$ se le llama el opuesto de x .

v $(\forall x, y \in X)(x + y = y + x)$.

vi Si $x \in X, \alpha \in E$, entonces $\alpha \cdot x \in X$.

vii $\forall x, y \in X, \alpha \in E, \alpha(x + y) = \alpha x + \alpha y$.

viii $\forall x \in X, \alpha, \beta \in E, (\alpha + \beta)x = \alpha x + \beta x$.

ix $\forall x \in X, \alpha, \beta \in E, \alpha(\beta x) = (\alpha\beta)x$.

x $(\forall x \in X)(1 \cdot x = x)$. 1 es la identidad o elemento identidad de E .



1.7. Espacios Normados

En algebra lineal, análisis funcional y otras áreas de la matemática, **la norma es una función** que asigna un valor real positivo a cada uno de los vectores del espacio vectorial, que no sea el vector nulo. **Una seminorma (o pseudo-norma)** en cambio, puede asignar longitudes nulas a vectores que no sean el vector nulo. Ejemplos simples son el espacio euclídeo bidimensional \mathbb{R}^2 con la norma Euclídea.

Definición 1.7.1 (Norma) Sea $(V, +, \cdot)$ un espacio vectorial sobre un subcuerpo F de los números complejos (F puede ser el cuerpo de los números complejos o los números racionales, o los números reales). Se denomina **seminorma** en V a una función $\rho : V \rightarrow \mathbb{R}$ tal que a cada $x \in V$ le asigna un número real $\rho(x)$, tal que satisface las siguientes propiedades:

1. $\rho(v) \geq 0, \forall v \in V$. (Positividad)
2. $\rho(av) = |a| \rho(v), \forall a \in F$ y $\forall v \in V$. (Homogeneidad positiva o escalabilidad positiva)
3. $\rho(u + v) \leq \rho(u) + \rho(v)$ (Desigualdad triangular o subaditividad)

La positividad es una consecuencia de los dos últimos axiomas de la definición. Una norma es una seminorma con la condición adicional que: (Definida positiva)

$$\rho(v) = 0 \text{ sí y sólo si } v = 0.$$

Una norma habitualmente se indica como $\|v\|$ en lugar de $\rho(v)$.

Definición 1.7.2 (Espacio normado) Se denomina espacio vectorial seminormado al par (V, ρ) , donde V es un espacio vectorial y ρ una seminorma en V .

Se denomina espacio vectorial normado (o espacio normado) al par $(V, \|\cdot\|)$, donde V es un espacio vectorial y $\|\cdot\|$ una norma en V .

Notación: Suele omitirse ρ o $\|\cdot\|$ y referirse sólo al “espacio (semi)normado V ”, quedando claro del contexto cuál es la (semi)norma considerada.

Aún cuando todo espacio vectorial es seminormado, con la seminorma trivial ($\rho(x) = 0 \forall x \in V$), pero no tiene por que ser normado. Todo espacio vectorial V con seminorma ρ puede convertirse en normado construyendo el espacio cociente V/W , donde W es el subespacio de V consistente en todos los vectores $v \in V$ tales que $\rho(v) = 0$. La norma inducida en V/W está dada por $\|W + v\| = \rho(v)$ y se puede verificar que está bien definida. Si ρ es una seminorma, vale que,

$$\rho(u \pm v) \geq |\rho(u) - \rho(v)|, \quad (1.1)$$

para todo $u, v \in V$. Además, toda norma es una seminorma.



Ejemplos

Ejemplo 1.7.1 *La seminorma trivial: $\rho(v) = 0$ para todo $v \in V$.*

Ejemplo 1.7.2 *El valor absoluto es una norma en \mathbb{R} .*

Ejemplo 1.7.3 *Toda forma lineal $f : V \rightarrow \mathbb{R}$ en un espacio vectorial V define una seminorma en V haciendo $\rho(x) = |f(x)|$.*

Ejemplo 1.7.4 *Las normas Euclideas en \mathbb{R}^n :*

$$\|x\| = \left\{ \sum_{i=1}^n x_i^2 \right\}^{\frac{1}{2}}, \quad \text{para } x \in \mathbb{R}^n \quad (1.2)$$

y en \mathbb{C}^n :

$$\|z\| = \left\{ \sum_{i=1}^n |z_i|^2 \right\}^{\frac{1}{2}}, \quad \text{para } z \in \mathbb{C}^n \quad (1.3)$$

Ejemplo 1.7.5 *En un espacio con producto interno, la norma inducida por el producto interno,*

$$\|x\|^2 = \langle x, x \rangle$$

.

Ejemplo 1.7.6 *La norma-1:*

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad (1.4)$$

también denominada **norma Manhattan**.

Ejemplo 1.7.7 *La norma-p, para $p \geq 1$, en \mathbb{R} :*

$$\|x\|_p = \left\{ \sum_{i=1}^n x_i^p \right\}^{\frac{1}{p}}, \quad (1.5)$$

Note que la norma Euclidea y la norma 1 son casos particulares de la norma-p. Ver también los espacios L^p en sección siguiente.

Ejemplo 1.7.8 *La norma infinito o norma máximo:*

$$\|x\|_\infty = \max_{i=1, \dots, n} |x_i|. \quad (1.6)$$



1.8. Espacios de Banach

Los espacios de Banach son uno de los objetos centrales de estudio del análisis funcional y son así denominados en honor al matemático polaco **Stefan Banach** (1892-1945) quien las estudió. Muchos de los espacios funcionales de dimensión infinita estudiados en este área de la matemática son ejemplos de espacios de Banach.

Definición 1.8.1 (Espacio vectorial normable) *Un espacio vectorial topológico se dice normable (seminormable) si la topología del espacio puede ser inducida por una norma (seminorma).*

Definición 1.8.2 (Espacios de Banach) *Los espacios de Banach son espacios vectoriales normados completos. Sea $(V, +, \cdot)$ un espacio vectorial sobre el campo de los números reales o de los números complejos, con una norma $\|\cdot\|$ y la métrica inducida por ella en $V : d(x, y) = \|x - y\|$. Se dice que $V(+, \cdot, \|\cdot\|)$ es un espacio de Banach si toda sucesión de Cauchy es convergente en V (y lo hace a un elemento de V).*

1.9. Algunos ejemplos de Espacios de Banach

Sea K el cuerpo de los números reales o el de los números complejos.

Ejemplo 1.9.1 *Los espacios euclídeos de dimensión n , K_n , donde la norma de $x = (x_1, \dots, x_n)$ es la norma cuadrática usual,*

$$\|x\| = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}, \quad (1.7)$$

son espacios de Banach.

Ejemplo 1.9.2 *El espacio de todas las funciones continuas $f : [a, b] \rightarrow K$ definidas en un intervalo real cerrado $[a, b]$ es un espacio de Banach bajo la norma del supremo, definida por $\|f\| = \sup\{|f(x)| : x \in [a, b]\}$. Esta es efectivamente una norma, dado que las funciones continuas definidas en intervalos cerrados son acotadas. El espacio es completo bajo esta norma y el espacio de Banach resultante se indica $C([a, b])$.*

*Este ejemplo puede generalizarse a $C(X)$ de todas las funciones continuas de X en K , donde X es un espacio compacto, o en el espacio de todas las funciones acotadas de X en K , donde X es cualquier espacio topológico, o en el espacio $B(X)$ de todas las funciones acotadas de X en K , cuando X es cualquier conjunto. En cualquiera de estos casos, podemos multiplicar funciones y el producto está en el mismo espacio (propiedad de clausura), haciendo que todos estos sean ejemplos también de lo que se denominan **álgebras de Banach**.*



Ejemplo 1.9.3 Si $p \geq 1$ es un número real, consideramos el espacio de todas las sucesiones $x = (x_1, x_2, x_3, \dots)$ de elementos de K tales que la serie,

$$\sum_{i=1}^{\infty} |x_i|^p \quad (1.8)$$

sea convergente (sea finita). Se define la **norma-p** de la sucesión x como,

$$\|x\|_p = \left[\sum_{i=1}^{\infty} |x_i|^p \right]^{\frac{1}{p}}, \quad (1.9)$$

El espacio de todas las sucesiones $x = (x_1, x_2, x_3, \dots)$ de elementos de K tales que la serie (1.8) sea convergente, dotado de la norma-p es un espacio de Banach y se indica L^p .

Ejemplo 1.9.4 El espacio de todas las sucesiones acotadas de elementos en K es el espacio de Banach L^∞ con la norma del supremo: $\|x\|_{sup}$, es decir, el supremo de los valores absolutos de los elementos de la sucesión, para $n \in \mathbb{N}$.

Ejemplo 1.9.5 Si $p \geq 1$ es un número real, se puede considerar el conjunto de todas las funciones $f : [a, b] \rightarrow K$ tales que $|f|^p$ es **integrable Lebesgue**. Se define la norma-p de f como:

$$\|f\|_p = \left[\int_{[a,b]} |f(x)|^p d\mu(x) \right]^{\frac{1}{p}}, \quad (1.10)$$

Este espacio no es en si mismo un espacio de Banach porque existen funciones no idénticamente nulas cuya norma-p es cero $\|f\|_p = 0$. Sin embargo, si se define una **clase de equivalencia** entre aquellas funciones f y g cuya diferencia tienen norma-p igual a cero, $\|f - g\|_p = 0$, entonces el conjunto de las clases de equivalencia constituye un espacio de Banach que se indica con $L^p[a, b]$ y se habla del espacio L^p .

Observación: Aquí es importante observar que la integral utilizada es la integral de Lebesgue, dado que con la de Riemann no se obtiene un espacio completo.

Ejemplo 1.9.6 Si X y Y son dos espacios de Banach, su suma directa $X \oplus Y$ también es un espacio de Banach.

Ejemplo 1.9.7 Todo producto interno da lugar a una norma asociada. Un espacio vectorial con producto interno se llama un espacio de Hilbert si su norma asociada es completa. Por lo tanto, todo espacio de Hilbert es, por definición, un espacio de Banach. El recíproco no siempre es cierto.



Ejemplo 1.9.8 *El espacio vectorial normado: $(C([0, 1]), \mathbb{R}, \|\cdot\|_2)$ no es completo. Pues, considere la sucesión de funciones continuas*

$$g_n(x) = \begin{cases} 0 & ; \text{si } x \in [0, \frac{1}{2}] \\ (n+1)(x - \frac{1}{2}) & ; \text{si } x \in (\frac{1}{2}, \frac{1}{2} + \frac{1}{n+1}) \\ 1 & ; \text{si } x \in (\frac{1}{2} + \frac{1}{n+1}, 1) \end{cases}$$

esta sucesión tiene como límite la función no continua

$$g(x) = \begin{cases} 0 & ; \text{si } x \in [0, \frac{1}{2}] \\ 1 & ; \text{si } x \in (\frac{1}{2}, 1] \end{cases}$$



Teoría de La Medida e Integral de Lebesgue

En los textos avanzados y artículos científicos de uso corriente en ingeniería y matemática aplicada es frecuente encontrar referencia a funciones medibles, a la integral de Lebesgue y a funciones de L^1 , L^2 o L^p . Aclaramos que no tratamos de realizar un desarrollo profundo, sino que, siguiendo diversos textos presentamos conceptos básicos y explicaremos las ideas centrales de modo tal que sean fácilmente comprendidas para poder abordar con posterioridad el tema central de este trabajo.

En matemática, la integral de una función no negativa puede mirarse, en el caso más simple, como el área comprendida entre el gráfico de dicha función y el eje x . La integración de Lebesgue es una construcción matemática que extiende la noción de integral a una clase de funciones más amplia; también extiende el dominio en el cual dichas funciones pueden estar definidas y que usa la medida y el concepto de “en casi todas partes” o de “para casi todo punto”.

Se entiende que para funciones no negativas, con un gráfico suficientemente suave, tales como las funciones continuas en intervalos acotados cerrados, el área bajo la curva puede definirse como la integral de Riemann en dicho intervalo y calculada mediante la técnica aproximación de la región mediante polígonos. De todos modos, cuando surge la necesidad de considerar funciones más generales (por ejemplo, como resultado del proceso de límite del análisis matemático y en la teoría matemática de probabilidad) queda claro que es necesario disponer de técnicas de aproximación más cuidadosas para poder definir adecuadamente la integral.

La integral de Lebesgue juega un rol importante en la rama de la matemática denominada análisis real y en muchos otros campos de las ciencias matemáticas.

La integral de Lebesgue se denomina de este modo en honor al matemático

francés Henri León Lebesgue (1875-1941).

2.1. Aspectos Históricos

Como dijéramos, la integración de manera usual es una operación matemática que se corresponde informalmente con la idea de encontrar el área bajo el gráfico de una función en un intervalo dado. La primera teoría de integración fue desarrollada por Arquímedes en el siglo III a.C, con el denominado método de cuadraturas, pero pudo aplicarse sólo a un número limitado de casos con un alto grado de simetrías geométricas. En el siglo XVII, Isaac Newton y Gottfried Wilhelm Leibniz descubrieron, independientemente, que la integración era por así decirlo, la operación inversa de la derivación y sentaron las bases de lo que hoy es el cálculo diferencial e integral. La contribución principal de ambos fue el Teorema Fundamental del Cálculo. Proporcionaron los elementos simbólicos básicos para los posteriores desarrollos de la matemática y física moderna. Sin embargo, a diferencia del método de Arquímedes, que se basaba en la geometría Euclideana, el método propuesto por Newton y Leibnitz no tenía bases rigurosas.

En el siglo XIX, Augustín Cauchy desarrolló finalmente una teoría rigurosa de límite, y Bernhard Riemann (1826-1866) la continuó formalizando, dando lugar a lo que hoy es conocido como la integral de Riemann. Para definir esta integral, se cubre el área bajo el gráfico de la función con rectángulos de tamaño cada vez menor y se toma el límite de las sumas de las áreas de los rectángulos en cada etapa. Sin embargo, para algunas funciones, el área total de dichos rectángulos no tiene límite (no se aproxima a un número), y por lo tanto no poseen integral de Riemann.

Dos ejemplos simples de funciones que no son integrables Riemann son las funciones $f(x) = \frac{1}{x}$ en el intervalo real $[0, b]$ y la $g(x) = \frac{1}{x^2}$ en cualquier intervalo que contenga al cero. Son intrínsecamente no integrables, dado que el área que la integral de Riemann debería representar es infinita. En otros ejemplos, el integrando posee muchas discontinuidades. Un ejemplo extremo de esta situación es la función característica de los números racionales.

Si consideramos el área bajo la curva definida por $f(x) = \frac{1}{x}$ en un intervalo entre $-a$ y b para a y b positivos, el área será infinita a ambos lados de 0. Sin embargo, es posible definir el área en este caso y obtener el área neta por límite, haciendo

$$\int_{-a}^{-d} \frac{1}{x} dx \cdot \int_d^b \frac{1}{x} dx, \quad \text{para } d > 0 \quad (2.1)$$

y tomando el límite para $d \rightarrow 0$. En este caso hemos calculado el valor principal de la integral. Esto es posible para funciones como $f(x) = \frac{1}{x}$ cuyas áreas infini-



tas tienen igual magnitud y signos opuestos en parte del intervalo de integración y pueden cancelarse mutuamente.

En 1907 Lebesgue propuso un nuevo método de integración para resolver este tipo de problemas. En lugar de utilizar las áreas de rectángulos, que pone la atención en el dominio de la función, Lebesgue buscaba en el codominio de la función su unidad fundamental de área. La idea de Lebesgue era construir primero la integral para lo que denominaba *funciones simples o funciones medibles* que sólo tomaran un número finito de valores. Luego la definió para funciones más complicadas, como la menor cota superior de todas las integrales de funciones simples que fueran menores que la función en cuestión.

La integral de Lebesgue tiene la propiedad de que toda función que posea integral de Riemann también es integrable según Lebesgue y para dichas funciones, ambas integrales coinciden. Pero existen muchas funciones que son integrables según Lebesgue, pero no poseen integral Riemann.

Una manera didáctica de entender la diferencia entre la integral de Riemann y la de Lebesgue, fue propuesta por el mismo Lebesgue. Supongamos que tenemos una canasta con billetes de \$1.000, \$2.000, \$5.000, \$10.000, \$20.000 y \$50.000 y deseamos saber cuánto dinero tenemos. Podemos contar el dinero de dos formas diferentes:

1. Sacamos los billetes de a uno y vamos sumando sus valores.
2. Agrupamos los billetes por su valor y contamos cuantos tenemos en cada grupo, multiplicando luego dicha cantidad por el valor correspondiente y sumamos todo.

Aunque ambas formas de contar nos va a dar el mismo resultado, es claro que la segunda es más eficiente que la primera. La primera forma se corresponde con la idea de cómo se calcula la integral de Riemann y la segunda con la de Lebesgue. Sin embargo, para funciones que no sean escalonadas (es decir, que no sean de las denominadas funciones simples), no es evidente que el método de Lebesgue de el mismo resultado que el de Riemann.

Como parte de los desarrollos integrales de Lebesgue, el matemático inventó el concepto de medida de Lebesgue, que extiende la idea de longitud de los intervalos a una clase mucho más grande de conjuntos, denominados conjuntos medibles (por lo tanto, las funciones simples son funciones que toman sólo un número finito de valores y cada valor se toma en un conjunto medible). La técnica de Lebesgue para ajustar una medida en una integral se generaliza fácilmente a muchas otras situaciones, dando lugar al campo de la matemática conocido como Teoría de la medida.



Limitaciones de la Integral de Riemann.

La integral de Riemann sólo está definida en intervalos acotados y su extensión a intervalos infinitos sólo se puede realizar como una integral impropia, haciendo:

$$\int_{-\infty}^{\infty} f(t)dt = \lim_{x \rightarrow \infty} \int_{-x}^x f(t)dt. \quad (2.2)$$

Pero esta extensión no funciona de manera apropiada. Una propiedad básica, la invariancia por traslaciones, no es satisfecha. ¿Qué quiere decir esto?. La integral Riemann de una función no debería cambiar si se traslada la función hacia la derecha o hacia la izquierda. Por ejemplo, sea $f(x) = 1$ para $x > 0$, $f(0) = 0$, y $f(x) = -1$ para $x < 0$. Entonces,

$$\int_{-x}^x f(t)dt = \int_{-x}^0 f(t)dt + \int_0^x f(t)dt = -x + x = 0 \quad (2.3)$$

para todo x . Pero, si trasladamos $f(x)$ hacia la derecha una unidad para tener $f(x-1)$, resulta

$$\int_{-x}^x f(t-1)dt = \int_{-x}^1 f(t-1)dt + \int_1^x f(t-1)dt = -(x+1) + (x-1) = -2 \quad (2.4)$$

Para todo $x > 1$.

Como esto no es aceptable, podemos intentar otra alternativa:

$$\int_{-\infty}^{\infty} f(t)dt = \lim_{a \rightarrow -\infty} \lim_{b \rightarrow +\infty} \int_a^b f(t)dt \quad (2.5)$$

Sin embargo, si tratamos de integrar la función anterior de esta forma, obtenemos $+\infty$, dado que tomamos primero el límite para b tendiendo a infinito. Si cambiamos el orden de los límites, tenemos $-\infty$.

Observamos que para que exista la integral, deberíamos obtener el mismo valor, independientemente del orden.

Otro problema es que la integral de Riemann no conmuta con límites uniformes. Por ejemplo, si consideramos $f_n(x) = \frac{1}{n}$ en el $[0, n]$ y 0 para cualquier otro x , tendremos que para todo n , $\int_{\mathbb{R}} f_n dx = 1$. Por otra parte f_n converge uniformemente a cero, entonces la integral de $\lim f_n$ es cero. En consecuencia,

$$\int f(x) dx \neq \lim \int f_n dx \quad (2.6)$$



Esto muestra que el importante criterio de intercambio de los signos de límite y de integral (propia) es falso cuando trabajamos con integrales impropias, haciendo a la integral Riemann no apropiada para muchas aplicaciones.

De otro lado, las series de Fourier, dieron origen a diferentes problemas relacionados con las integrales, cuya solución satisfactoria requería intercambiar los signos de sumatorias infinitas de funciones y los de integración. Aparecieron dificultades relacionadas con las condiciones bajo las cuales las integrales, eran iguales en el contexto Riemann.

$$\sum_k \int f_k(x) dx, \quad y \quad \int \sum_k f_k(x) dx \quad (2.7)$$

Existen también otras dificultades técnicas con la integral de Riemann, vinculadas con el límite.

1. Falla en la convergencia monótona

Como se mostró antes, la función característica de los racionales X_Q no es integrable Riemann. En particular falla el teorema de la convergencia monótona. Para verlo, consideremos a_k una sucesión de números racionales en $[0, 1]$.

$$g_k(x) = \begin{cases} 1 & \text{si } x = a_k \\ 0 & \text{en otro caso} \end{cases}$$

Entonces hagamos,

$$f_k = g_1 + g_2 + \cdots + g_k.$$

La función f_k es cero en todo punto, excepto en un conjunto finito de puntos; entonces su integral Riemann es cero. La sucesión f_k también es claramente no negativa y monotonamente decreciente a X_Q , que no es integrable Riemann.

2. Inadecuación de los intervalos no acotados:

La integral de Riemann sólo puede integrar funciones en intervalos acotados. La extensión mas simple consiste en definir,

$$\int_{-\infty}^{+\infty} f(x) dx = \lim_{a \rightarrow +\infty} \int_{-a}^a f(x) dx \quad (2.8)$$

cuando los límites existen. De todos modos, esto destruye la deseable propiedad de invarianza ante traslaciones. Con esta definición de la integral impropia (denominada el valor principal de Cauchy de la integral



impropia entorno al cero), las funciones $f(x) = (1 \text{ si } x > 0, -1 \text{ para otro valor})$ y $g(x) = (1 \text{ si } x > 1, -1 \text{ para otro valor})$ son traslaciones una de la otra y sin embargo sus integrales impropias son diferentes:

$$\int f(x) dx = 0, \quad \int g(x) dx = -2 \quad (2.9)$$

2.2. Construcción de la Integral de Lebesgue

La integral de una función f entre los límites a y b puede interpretarse como el área bajo del gráfico de f . Esto es fácil de interpretar en funciones polinómicas, pero no es tan intuitivo para funciones más raras. En general, ¿para qué tipo de funciones es que tiene sentido pensar en el “área bajo la curva”? La respuesta a esta pregunta posee importancia teórica y práctica.

Recordemos que, conceptualmente, la integral de Riemann (conocida como la integral definida de f en el intervalo $[a, b]$) lo que hace es construir una sucesión de integrales de fácil cálculo que convergen a la integral de la función dada. Esta definición es exitosa en el sentido que proporciona respuesta a muchos problemas ya resueltos y da resultados útiles para muchos otros problemas.

Sin embargo, la integración de Riemann, no interactúa adecuadamente cuando, al tomar los límites de las sucesiones de funciones, el proceso se torna complicado de analizar. De principal importancia son, por ejemplo, en el estudio de las series de Fourier, en las transformadas de Fourier y otros temas. La integral de Lebesgue describe mejor cómo y cuándo es posible tomar límites bajo un signo de integral. La definición de Lebesgue considera una clase diferente de integrales, de fácil cálculo y permite calcular la integral de un espectro más amplio de funciones que la integral de Riemann. Así por ejemplo, la función de Dirichlet (la función característica de los números racionales \mathbb{Q}), que vale 1 cuando el argumento es racional y 0 para cualquier otro valor, posee integral de Lebesgue, pero no posee integral de Riemann.

Recordemos que, la integral de Riemann de una función $f : [a, b] \rightarrow \mathbb{R}$ se construye particionando el intervalo $[a, b]$ en sub-intervalos $I_{k=1}^{\infty}$ y aproximando el valor de la función f en cada intervalo por un valor constante $f(\xi_k^*)$, donde $\xi_k^* \in I_k$. Entonces, la integral $\int_a^b f(x) dx$ se obtiene por aproximación, como el límite de la suma de las áreas de los rectángulos cuyas bases se corresponden con los intervalos I_k cuyas alturas son las correspondientes constantes, por exceso y por defecto, $f(\xi_k^*)$.

Esta idea de ir cortando en rebanadas el área bajo el gráfico de una función f es natural y puede realizarse de dos maneras posibles: una es como se hace para calcular la integral de Riemann, es decir con rebanadas paralelas al eje de las abscisas; la otra es cortando en forma paralela al eje de las ordenadas. Esta



última forma fue la adoptada por Lebesgue. Es así que conforme a Lebesgue, la integral

$$\int_{-\infty}^{\infty} f(x) dx, \quad (2.10)$$

puede pensarse como,

$$\int_{-\infty}^{\infty} f(x) dx \approx \sum_{y \in \mathbb{R}} y. \quad (2.11)$$

Observemos que, la colección de los “valores de x para los cuales $f(x) = y$ ” no es otra cosa que:

$$f^{-1}(\{y\}) = \{x \in \mathbb{R} : f(x) = y\}. \quad (2.12)$$

Es por ésto que en la teoría de la integral de Lebesgue es de principal importancia la medida de conjuntos del tipo $f^{-1}(A)$, siendo $A \subset \mathbb{R}$ un subconjunto en el eje y .

Observemos que para calcular la integral de Riemann de una función es necesario que esta tenga variaciones suaves. Esto no es necesario para la integral de Lebesgue.

2.3. Teoría de la Medida

En esta sección presentaremos una introducción a los principales conceptos y propiedades de la integral de Lebesgue, distinguiendo dos partes:

1. Teoría de conjuntos medibles y medidas de dichos conjuntos.
2. Teoría de funciones medibles e integración de dichas funciones.

La teoría de la medida fue creada originalmente para permitir un análisis detallado de la noción de longitud de subconjuntos de la recta real y más generalmente, de los conceptos de área y volúmen de subconjuntos en espacios Euclídeos. Proporciona una forma sistemática de responder a la pregunta sobre qué subconjuntos de \mathbb{R} tienen una longitud. El desarrollo de la teoría de conjuntos logró mostrar que es imposible asignar una longitud a todos los subconjuntos de \mathbb{R} , de modo tal que se preserven las propiedades naturales de aditividad e invariancia por traslaciones. Esto sugiere que es un requisito esencial el tomar una clase adecuada de subconjuntos medibles.

La integral de Riemann implícitamente utiliza la noción de longitud. En efecto, el elemento de cálculo de la integral de Riemann es el rectángulo $[a, b] \times [c, d]$, cuyo área se calcula como $(b-a)(d-c)$. La cantidad $b-a$ es la longitud de la base del rectángulo y $d-c$ es la altura del rectángulo. Riemann podía sólo utilizar rectángulos planos para aproximar el área bajo una curva dado que no existía



una teoría adecuada para medir conjuntos mas generales. En el desarrollo de la teoría en textos modernos (a partir de 1950), el enfoque de medida e integración es axiomático. Esto quiere decir que una medida es cualquier función μ definida en ciertos subconjuntos E de un conjunto X que satisfaga ciertas propiedades.

Un ejemplo claro es un espacio de probabilidad, que describe nuestra incertidumbre sobre cierto experimento y consiste en el espacio muestral de las posibles salidas y una medida de probabilidad que cuantifica cuantas oportunidades chances tiene de salir cada uno de ellos. Un evento es un conjunto de salidas del experimento. La medida de probabilidad obedece tres axiomas: es no negativa, su suma (o integral) es la unidad (normalización) y es aditiva para eventos disjuntos.

2.4. Sigma álgebra, (Medida)

Una σ -álgebra (se lee “sigma-álgebra”) o σ -campo sobre un conjunto X es una colección Σ de subconjuntos de X que es cerrada bajo las operaciones de complementación y uniones numerables de sus elementos. Es el análogo numerable de un álgebra de Boole y toda σ -álgebra representa un álgebra Booleana. El uso principal de las σ -álgebra es en la definición de medidas en X , constituyendo un concepto importante en análisis matemático y en teoría de probabilidad.

Definición 2.4.1 (σ -álgebra) Dado un conjunto X , sea $P(X)$ el conjunto de todos sus subconjuntos. Se dice que un subconjunto no vacío Σ de $P(X)$ es un σ -álgebra si y sólo si satisface las siguientes propiedades:

1. si $A \in \Sigma$, entonces su complemento en X , el conjunto $X \setminus A$ también pertenece a Σ ,
2. si $\{A_n\}_{n \in \mathbb{N}}$ es una colección numerable de elementos de Σ , su unión $\cup_{n \in \mathbb{N}} A_n$ también estará en Σ .

Esto quiere decir que un subconjunto Σ deberá satisfacer las siguientes tres condiciones para calificar como un σ -álgebra:

1. Ser cerrado bajo complementación
2. Ser cerrado bajo uniones numerables
3. Contener al conjunto vacío

De estos axiomas sigue que X y el conjunto vacío; están en Σ (dado que Σ no es vacío) y que la σ -álgebra es también cerrada bajo intersecciones numerables, por las leyes de Morgan:

$$\begin{aligned}(A \cap B)^C &= A^C \cup B^C \\ (A \cup B)^C &= A^C \cap B^C\end{aligned}$$



2.5. La Función Medida

Una medida en X es una función que asigna un número real mayor o igual a cero a cada subconjunto de X , permitiendo pensar de manera mas precisa en la noción de tamaño o de volumen de un conjunto. Uno podría desear asignar tal tamaño a todo subconjunto de X , sin embargo, el axioma de elección (de la teoría de conjuntos) implica que cuando el tamaño en consideración es la longitud estandar de subconjuntos de la recta real, entonces, existen conjuntos (conocidos como **conjuntos de Vitali**) para los cuales no existe dicho tamaño. Por esta razón se hace necesario considerar en cambio un colección mas pequeña de subconjuntos de X cuya medida esté definida y dichos conjuntos constituyen la σ -álgebra.

Definición 2.5.1 (Medida) Una *medida* μ es una función definida en una σ -álgebra sobre un conjunto X y que toma valores en el intervalo extendido $[0, +\infty)$, tal que verifica las siguientes propiedades:

1. El conjunto vacío tiene medida cero:

$$\mu(\Phi) = 0;$$

2. Es numerablemente aditiva o σ -aditiva: Si E_1, E_2, E_3, \dots es una sucesión numerable de conjuntos mutuamente disjuntos en Σ , es decir si son tales que:

$$E_i \in \Sigma \quad \forall i,$$

y además

$$E_i \cap E_j = \Phi \text{ si y solo si } i \neq j$$

entonces se verifica que la medida de la unión de todos los E_i es igual a la suma de las medidas de cada uno de los E_i :

$$\mu(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$$

Algunos ejemplos

Ejemplo 2.5.1 Si X es un conjunto cualquiera, entonces, la familia compuesta por sólo el conjunto vacío y X es una σ -álgebra trivial sobre X .

Ejemplo 2.5.2 Otra σ -álgebra sobre X está dada por todo el conjunto de partes de X , $\wp(X)$.

Ejemplo 2.5.3 Una colección de subconjuntos de X que sea numerable y cuyos complementos sean numerables es una σ -álgebra, diferente de la de partes de X si y sólo si X no es numerable.



Ejemplo 2.5.4 Si Σ_a es una familia de σ -álgebras sobre X , entonces la intersección de todas las Σ_a también es una σ -álgebra sobre X .

Ejemplo 2.5.5 Un ejemplo importante es el álgebra de Borel sobre un espacio topológico cualquiera: la σ -álgebra generada por los conjuntos abiertos (o, de modo análogo, por los conjuntos cerrados), es decir generada por la misma topología de X . En teoría de probabilidades, el álgebra de Borel constituye un ejemplo particular. Dada una variable aleatoria, definida en un espacio de probabilidad, su distribución de probabilidad es por definición, también una medida en el álgebra de Borel. Este álgebra es la menor σ -álgebra en \mathbb{R} que contiene todos los intervalos. Observemos que esta σ -álgebra no es en general el conjunto de las partes completo. Un ejemplo no trivial es el llamado conjunto de Vitali.

Ejemplo 2.5.6 En los espacios euclídeos \mathbb{R}^n , una σ -álgebra de importancia es la de los conjuntos medibles Lebesgue. Dicha σ -álgebra contiene más conjuntos que la σ -álgebra de Borel en \mathbb{R}^n y es preferida en teoría de integración, dado que provee un espacio de medida completo.

Definición 2.5.2 (Espacio de medida y conjuntos medibles) El triplete (X, Σ, μ) se denomina **espacio de medida**, y los elementos de Σ se denominan **conjuntos medibles**.

Definición 2.5.3 (Espacio medible y funciones medibles) El par ordenado (X, Σ) , donde X es un conjunto y Σ es un σ -álgebra sobre X , se dice es un **espacio medible**.

Una función entre dos espacios medibles se dice es una **función medible** si la preimagen de todo conjunto medible es medible.

Las medidas son definidas como cierto tipo de funciones de un σ -álgebra en $[0, +\infty)$. Para indicar las σ -álgebras se suele usar letras mayúsculas en lugar de Σ , con el objeto de evitar la confusión con el operador sumatoria, en este caso tendremos por ejemplo (X, A) , en lugar de (X, Σ) .

2.6. Propiedades de las Medidas

Algunas de las propiedades más importantes de las medidas son:

1. Monotonía:

Una medida μ es monótona:

Si E_1 y E_2 son conjuntos medibles, tales que $E_1 \subset E_2$, entonces

$$\mu(E_1) \leq \mu(E_2) \quad (2.13)$$

2. Medidas de uniones infinitas de conjuntos de A :

Si E_1, E_2, E_3, \dots es una sucesión numerable de conjuntos en A , no necesariamente disjuntos, entonces:



$$\mu(\cup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} \mu(E_i) \quad (2.14)$$

3. Medidas de uniones infinitas de conjuntos medibles:

Si E_1, E_2, E_3, \dots son conjuntos medibles encajados de modo tal que E_n es un subconjunto de E_{n+1} para todo n , es decir si:

$$E_n \subset E_{n+1}, \quad \forall n, \quad (2.15)$$

entonces la unión de los conjuntos E_n también es medible y

$$\mu(\cup_{i=1}^{\infty} E_i) = \lim_{i \rightarrow \infty} \mu(E_i) \quad (2.16)$$

4. Medida de intersecciones infinitas de conjuntos medibles:

Si E_1, E_2, E_3, \dots son conjuntos medibles encajados de modo tal que E_{n+1} es un subconjunto de E_n para todo n , es decir si:

$$E_{n+1} \subset E_n, \quad \forall n, \quad (2.17)$$

entonces la intersección de los conjuntos E_n también es medible. Más aún, si al menos uno de los E_n tiene medida finita, entonces:

$$\mu(\cap_{i=1}^{\infty} E_i) = \lim_{i \rightarrow \infty} \mu(E_i) \quad (2.18)$$

Observación: Esta última propiedad es falsa sin la hipótesis de que al menos uno de los conjuntos tenga medida finita. En efecto, sea por ejemplo la sucesión $\{E_n\}$, tal que para cada $n \in \mathbb{N}$,

$$E_n = [n, \infty) \subseteq \mathbb{R},$$

que satisfacen (2.17) y poseen todos medida infinita. Pero, dado que su intersección es vacía, resulta $\mu(\cap_{i=1}^{\infty} E_i) = 0$.

La propiedad (2.15) y (2.16) será de fundamental importancia en las propiedades del análisis multiresolución de la Transformada Wavelets.

2.7. Medidas, Sigma-Finitas

Definición 2.7.1 (Espacio de medida finito) Sea (X, Σ, μ) un espacio de medida.

1. (X, Σ, μ) se dice un **espacio de medida finito** si $\mu(X)$ es un número real finito.



2. (X, Σ, μ) se dice **σ -finito** si X puede descomponerse en una unión enumerable de conjuntos medibles de medida finita.
3. Un conjunto en un espacio de medida posee **medida σ -finita** si es una unión de conjuntos de medida finita.

Ejemplo 2.7.1 Los números reales con la medida de Lebesgue estándar son σ -finitos pero no finitos. En efecto, considere los intervalos cerrados $[k, k+1]$ para todos los enteros k . Se tiene una cantidad numerable de tales intervalos, cada uno posee medida 1, y la unión de todos ellos es la recta real.

Ejemplo 2.7.2 Consideremos ahora los números reales con la medida de conteo, que asigna a cada conjunto finito de números reales, el número de puntos en el conjunto. Este espacio de medida no es σ -finito, porque todo conjunto con medida finita contiene a lo sumo un número finito de puntos, y sería necesario un cantidad no numerable de tales conjuntos para cubrir toda la recta real.

2.8. Completitud

Definición 2.8.1 Un conjunto medible X se dice un conjunto nulo si $\mu(X) = 0$. Un subconjunto de un conjunto nulo se llama despreciable.

Un conjunto despreciable no necesita ser medible, pero todo conjunto medible despreciable es automáticamente un conjunto nulo.

Definición 2.8.2 Una medida μ se dice completa si todo conjunto despreciable es medible.

Una medida puede extenderse a completa considerando la σ -álgebra de los subconjuntos Y que difieren en un conjunto despreciable de un conjunto medible X , esto es, si la diferencia simétrica de X e Y está contenida en un conjunto nulo. Entonces, por definición se dice que $\mu(Y)$ es igual a $\mu(X)$.

Ejemplos

Ejemplo 2.8.1 La medida de conteo, definida por $\mu(S)$ igual al número de elementos en S .

Ejemplo 2.8.2 Medida del ángulo circular, es invariante ante rotaciones.

Ejemplo 2.8.3 La medida de Haar, para grupos topológicos localmente compactos, es una generalización de la medida de Lebesgue y tiene similares propiedades de unicidad.

Ejemplo 2.8.4 La medida de Hausdorff que es un refinamiento de la medida de Lebesgue para algunos conjuntos fractales.



Ejemplo 2.8.5 *Todo espacio de probabilidad da lugar a una medida que toma el valor 1 en todo el espacio (y en consecuencia toma todos sus valores en el intervalo $[0, 1]$). Una medida de este tipo se denomina una medida de probabilidad.*

Ejemplo 2.8.6 *La medida de Lebesgue, que es una medida completa invariante por traslaciones en una σ -álgebra que contenga intervalos de \mathbb{R} , tal que $\mu([0, 1]) = 1$.*

Ejemplo 2.8.7 *La medida de Dirac μ_α está dada por $\mu_\alpha(S) = X_s(a)$, donde X_s es la función característica de S .*

Otras medidas son: medida de Borel, medida de Jordan, medida Ergódica, medida de Euler, medida de Gauss, medida de Baire, medida de Radon.

2.9. La Medida del conteo

La medida de conteo es una forma intuitiva de dar medida a cualquier conjunto: el “tamaño” de un subconjunto se toma como el número de elementos del subconjunto si éste es finito, y en caso de no serlo se dice que es de tamaño infinito.

Formalmente, sea Ω un conjunto y considerese la σ -álgebra X en Ω consistente en todos los subconjuntos de Ω . Se define una medida μ en dicho σ -álgebra haciendo $\mu(A) = |A|$ ($|A|$ indica el cardinal de A , es decir, el número de elementos del conjunto) si A es un subconjunto finito de Ω y $\mu(A) = \infty$ si A es un subconjunto infinito de Ω . Entonces, (Ω, X, μ) es un espacio de medida.

La medida de conteo permite traducir muchas afirmaciones sobre los espacios L^p en estructuras más familiares. Si $\omega = \{1, 2, \dots, n\}$ y S es el espacio de medida con medida de conteo en, entonces $L^p(S)$ coincide con \mathbb{R}^n (o con \mathbb{C}^n), con *norma* definida por,

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \quad (2.19)$$

para $x = (x_1, x_2, x_3, \dots, x_n)$.

De manera similar, si se toma Ω como el conjunto de números naturales y S es el espacio de medida con la medida de conteo en Ω , entonces $L^p(S)$ consiste en todas las sucesiones $x = (x_n)$ para las cuales (2.19) es finito. Este espacio se escribe generalmente como L^p .

2.10. La Medida de Lebesgue

La medida de Lebesgue es la manera estandar de asignar un longitud, área o volúmen a subconjuntos de un espacio euclídeo. Se utiliza, en particular, para



definir la integración en el sentido de Lebesgue. Aquellos conjuntos a los cuales se les puede asignar un volúmen, se denominan medibles Lebesgue. Indicamos con $\lambda(A)$ el volúmen o medida de Lebesgue de un conjunto medible A .

Ejemplos

Ejemplo 2.10.1 Si $A = [a, b] \subset \mathbb{R}$, entonces su medida de Lebesgue es $b - a$. El intervalo abierto (a, b) tiene la misma medida, dado que la diferencia entre ambos conjuntos tienen medida cero (un punto tiene medida nula).

Ejemplo 2.10.2 Si A es el producto cartesiano de los intervalos $[a, b]$ y $[c, d]$, entonces es un rectángulo y su medida de Lebesgue es el área $(b - a)(d - c)$.

Ejemplo 2.10.3 *El conjunto de Cantor*, es un ejemplo de un conjunto no numerable que tiene medida de Lebesgue cero.

Algunas propiedades de las medidas de Lebesgue

Las medidas de Lebesgue en \mathbb{R}^n poseen las siguientes propiedades:

1. Si A es un producto cartesiano de intervalos $I_1 \times I_2 \times \cdots \times I_n$, entonces A es medible Lebesgue y $\lambda(A) = |I_1| \cdot |I_2| \cdots |I_n|$. Donde $|I|$ indica la longitud del intervalo I .
2. Si A es una unión disjunta de una cantidad finita o numerable de conjuntos disjuntos medibles Lebesgue, entonces A es también medible Lebesgue y $\lambda(A)$ es igual a la suma (o serie infinita) de las medidas de los conjuntos medibles involucrados.
3. Si A es un conjunto medible Lebesgue, también lo es su complemento.
4. $\lambda(A) \geq 0$ para todo conjunto medible Lebesgue A .
5. Si A y B son conjuntos medibles Lebesgue tales que $A \subset B$, entonces $\lambda(A) \leq \lambda(B)$. (Consecuencia de 2, 3 y 4).
6. Las uniones e intersecciones enumerables de conjuntos Lebesgue son Lebesgue-medibles. (Consecuencia de 2 y 3)
7. Si A es un subconjunto abierto o cerrado de \mathbb{R}^n , entonces A es medible Lebesgue.
8. Si A es un conjunto medible Lebesgue, entonces es “aproximadamente abierto” y “aproximadamente cerrado” en el sentido de Lebesgue (para esto debe verse el teorema de **regularidad de las medidas**).
9. La medida de Lebesgue es al mismo tiempo localmente finita e internamente regular, por lo cual es una medida de Radon.



10. La medida de Lebesgue es estrictamente positiva, y por lo tanto su soporte es todo \mathbb{R}_n .
11. Si A es un conjunto medible Lebesgue, con $\lambda(A) = 0$ (el conjunto nulo), entonces, todo subconjunto de A también es un conjunto nulo.
12. Si A es medible Lebesgue y x es un elemento de \mathbb{R}_n , entonces la traslación de A por x definida por $A + x = \{a + x : a \in A\}$, también es medible Lebesgue y tiene la misma medida que A : $\lambda(A + x) = \lambda(A)$.

El listado anterior se puede resumir en el siguiente resultado:

La medida de Lebesgue de conjuntos medibles constituye una σ -álgebra que contiene todos los intervalos producto y λ es la única medida completa, invariante por traslaciones en dicho σ -álgebra tal que $\lambda([0, 1] \times [0, 1] \times \cdots \times [0, 1]) = 1$.

2.11. Integración

Trabajaremos en el siguiente contexto abstracto:

1. Σ es una σ -álgebra de subconjuntos de X .
2. μ es una medida no negativa en X .

Por ejemplo, X es un espacio Euclídeo \mathbb{R}^n o cierto subconjunto medible Lebesgue, Σ es la σ -álgebra de todos los subconjuntos medibles Lebesgue de Σ , y μ es una medida de Lebesgue. En teoría de probabilidad, μ sería una medida de probabilidad del espacio de probabilidades X . En la teoría de Lebesgue, las integrales están limitadas a la clase de funciones llamadas funciones medibles.



Lebesgue

Se puede demostrar que la definición de función medible es equivalente a requerir que la pre-imagen de cualquier subconjunto Borel de \mathbb{R} esté en Σ . El conjunto de funciones medibles es cerrado bajo operaciones algebraicas. El límite puntual de sucesiones:

$$\liminf_{k \in N} f_k, \quad \limsup_{k \in N} f_k$$

Es medible si la sucesión original $\{f_k\}$, consiste en funciones medibles.

Construimos una integral,

$$\int_X f d\mu$$

para funciones medibles a valores reales f definidas en X en cuatro pasos:



1. Función Característica: Para asignar un valor a la integral de una función característica de un conjunto medible S consistente con una medida dada μ , se elige:

$$\int X S d\mu = \mu S$$

2. Funciones Simples: Extendemos por linealidad la definición anterior:

$$\int \left(\sum_k a_k X S_k \right) d\mu = \sum_k a_k \int X S_k d\mu$$

donde la suma es finita y los coeficientes a_k son números reales. Aún cuando una función simple puede escribirse de diferentes maneras como combinación lineal de funciones características, puede demostrarse que la integral simple será la misma.

3. Funciones no Negativas: Sea f una función medible no negativa en X que puede tomar el valor $[0, \infty)$, es decir, que toma valores en el semieje real positivo extendido. Definimos,

$$\int_X f d\mu := \sup \left\{ \int_X S d\mu : S \leq f \right\}, \mathbf{S} \text{ simple}$$

Esta integral coincide con la precedente para el caso particular en que f sea una función simple. Además, coincide con la integral de Riemann si f es integrable Riemann. Para algunas funciones $\int_E f d\mu$, podrá ser infinito.

4. Funciones con Signo: Para trabajar con funciones con signo necesitamos mas definiciones.

Si f es una función definida en un conjunto medible X , entonces se puede escribir:

$$f = f^+ - f^-$$

donde,

$$f^+(x) = \begin{cases} f(x) & \text{si } f(x) > 0 \\ 0 & \text{en otro caso} \end{cases}$$

$$f^-(x) = \begin{cases} -f(x) & \text{si } f(x) < 0 \\ 0 & \text{en otro caso} \end{cases}$$



De otro lado podemos observar que f^+ y f^- son funciones no negativas, además, $|f| = f^+ + f^-$, siempre que $\int |f| d\mu < \infty$, entonces se dice que f es integrable Lebesgue. En este caso, ambas integrales satisfacen,

$$\int f^+ d\mu < \infty, \quad \int f^- d\mu < \infty \quad (2.20)$$

y tiene sentido definir:

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu \quad (2.21)$$

Es claro que las definiciones anteriores dan a la integral de Lebesgue todas las propiedades deseables. En el caso de funciones a valores complejos, basta considerar la parte real y la parte imaginaria, de manera separada.

2.12. Una Interpretación Intuitiva

Para lograr cierta intuición sobre los diferentes aspectos de la integración, supongamos que se desea encontrar el volumen de una montaña sobre el nivel del mar y que las fronteras de la montaña están claramente marcados (serán la región de integración).

El enfoque de Riemann (o Riemann-Darboux): cortar la montaña en torres verticales, cada una con base cuadrada al nivel de mar. Tomar un par de puntos dentro de ese cuadrado, uno donde la altura sea máxima y otro donde sea mínima. Asociar a esas alturas los volúmenes por exceso y por defecto obtenidos al multiplicar las alturas por el área del cuadrado de la base. La suma Riemann por exceso es la suma de los volúmenes de las torres por exceso y de modo similar se tiene la suma Riemann por defecto. Si las sumas por exceso y por defecto convergen a un mismo valor, a medida que el lado de los cuadrados decaiga a cero, entonces se dice que existe la integral de Riemann.

El enfoque Lebesgue: Dibuje un mapa de nivel de la montaña con tantas líneas de nivel como sea posible. Para cada contorno (o conjunto de nivel) con mínima altura, encuentre el área total encerrada (dentro de ese mapa) por ese conjunto de niveles, es decir encuentre la medida de ese conjunto de niveles. Multiplique dicha medida por la altura representada por el conjunto de niveles: el producto será un sumando de la suma de Lebesgue.

Luego encuentre el nivel o conjunto de niveles que están a la altura siguiente (menor altura de los niveles restantes). Calcule la medida del área encerrada por ellos. Multiplique la medida por la diferencia en altura con respecto al nivel anterior. Dicho producto será otro sumando de la suma de Lebesgue.



Repita este procedimiento para sucesivos niveles de contorno, cada vez más altos, hasta que haya procesado el nivel de contorno más elevado. La suma resultante será el generado lineal: cada contorno corresponde a una función característica (del correspondiente conjunto de nivel). La suma puede definirse sumando los contornos inmediatos: dividiendo por la mitad las diferencias entre alturas sucesivas y recalculando la suma. La integral de Lebesgue es el límite de dicho proceso.

Ejemplos 2.12.1 *La función característica de los números racionales X_Q no es continua y no es derivable.*

1. X_Q no es integrable Riemann en $[0, 1]$: Cualquier partición de $[0, 1]$ en subintervalos contiene al menos un número racional y al menos un número irracional, dado que la unión del conjunto de números racionales con el de irracionales da lugar a la densidad del conjunto de los números reales. Entonces las sumas superiores serán todas iguales a 1 y las por defecto serán todas iguales a cero. Por lo tanto, al no coincidir las integrales superiores (por exceso) con las inferiores (por defecto), no existe la integral en el sentido Riemann de la función X_Q .

2. X_Q es integrable Lebesgue en $[0, 1]$: En efecto, dado que los números racionales son numerables en el $[0, 1]$, se tiene que,

$$\int_{[0,1]} X_Q d\mu = \mu(Q \cap [0, 1]) = 0 \quad (2.22)$$

2.13. Teoremas Integrales Básicos

La integral de Lebesgue no distingue entre funciones que sólo difieren en un conjunto de μ -medida nula.

Definición 2.1 *Sean f y g dos funciones definidas sobre un subconjunto de un espacio medible E . Se dice que f y g son iguales para casi todo punto y se indica $f = g$, si y sólo si $\mu(\{x \in E : f(x) \neq g(x)\}) = 0$.*

Teorema 2.13.1 *Si f y g son funciones no negativas (que pueden tomar valores $\pm \infty$) tales que $f = g$, entonces $\int f \mu d\mu = \int g \mu d\mu$*

Teorema 2.13.2 *Si f y g son tales que $f = g$, entonces f es integrable si y sólo si g es integrable y $\int f \mu d\mu = \int g \mu d\mu$*

Propiedades

La integral de Lebesgue posee las siguientes propiedades:



1. Linealidad: Si f y g son funciones integrables y a y b son números reales, entonces $af + bg$ es integrable y

$$\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu \quad (2.23)$$

2. Monotonía: Si $f \leq g$, entonces

$$\int f \mu d\mu \leq \int g \mu d\mu \quad (2.24)$$

Teorema 2.13.3 (*Teorema de la Convergencia monótona de Lebesgue*)

Sea $\{f_k\}_{k \in \mathbb{N}}$ una sucesión de funciones medibles no negativas tales que:

$$f_k(x) \leq f_{k+1}(x) \quad \forall x \in E, \quad \forall k \in \mathbb{N} \quad (2.25)$$

Entonces,

$$\lim_k \int f_k \mu d\mu = \int \sup_k f_k \mu d\mu \quad (2.26)$$

Observación: El valor de cualquiera de las integrales puede ser infinito.

Lema 2.1 (*Lema de Fatou*): Sea $\{f_k\}_{k \in \mathbb{N}}$ una sucesión de funciones medibles no negativas. Si $f = \liminf f_k$, entonces

$$\int f \mu d\mu \leq \liminf_k \int f_k d\mu \quad (2.27)$$

Nuevamente, el valor de cualquiera de las integrales puede ser infinito.

Teorema 2.13.4 (*Teorema de la convergencia Dominada de Lebesgue*):

Sea $\{f_k\}_{k \in \mathbb{N}}$ una sucesión de funciones medibles con límite puntual f . Si existe una función integrable g tal que $|f_k| \leq g$ para todo k , entonces f es integrable y

$$\lim_k \int f_k d\mu = \int f d\mu \quad (2.28)$$

Con estos teoremas, definiciones, propiedades y ejemplos tenemos las herramientas necesarias para comprender con mayor claridad los capítulos siguientes.



Producto Interno y Espacios de Hilbert

3.1. Introducción Histórica

Las siguientes ecuaciones en derivadas parciales han sido fundamentales en el desarrollo de los espacios de Hilbert:

1. Ecuación del calor: $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ (su solución corresponde a la temperatura $u(x, t)$ de un alambre en el instante t y en el punto x).
2. Ecuación de ondas: $\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$ (la solución $u(x, t)$ es ahora la altura de una onda en propagación para el instante t en un punto cuya proyección sobre el eje es x).
3. Ecuación de Laplace: $\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = \rho$ (ahora u es la energía potencial de un cuerpo y ρ es la densidad de masas).



David Hilbert

Nos limitaremos a estudiar la ecuación del calor. Para buscar sus soluciones, suponemos que u es de variables separadas, es decir de la forma,

$$u(t, x) = \varphi(t)\xi(x), \quad (3.1)$$

entonces,

$$u_x = \varphi'(t)\xi(x) \quad (3.2)$$

y,

$$u_{xx} = \varphi(t)\xi''(x), \quad (3.3)$$

luego,

$$\varphi'(t)\xi(x) = \varphi(t)\xi''(x), \text{ o bien } \frac{\varphi'(t)}{\varphi(t)} = \frac{\xi''(x)}{\xi(x)} \quad (3.4)$$

Como el primer miembro no depende de x y el segundo miembro no depende de t , deben ser constantes. Debemos resolver entonces las siguientes ecuaciones:

$$\frac{\varphi'(t)}{\varphi(t)} = k, \quad \frac{\xi''(x)}{\xi(x)} = p \quad (3.5)$$

con k y p constantes.

La solución general de la primera ecuación es $\varphi(x) = ce^{kt}$ y de la segunda $\xi(x) = c_1 \cos(\delta x) + c_2 \operatorname{sen}(\delta x)$ donde $p = -\delta^2$ (si $p > 0$, δ es complejo).

Supongamos por simplificar que el alambre es circular y tiene longitud 2π . Esto hace que ξ sea periódica de período 2π , y que δ deba ser entero; resulta,

$$\xi(x) = c_1 \cos(nx) + c_2 \operatorname{sen}(nx)$$

con $n \in \mathbb{Z}$.

Como las soluciones de la ecuación del calor se obtienen haciendo $K = n^2$ tenemos la siguiente familia de soluciones:

$$u_n(t, x) = e^{-n^2 t} \cos(nx), \quad v_n(t, x) = e^{-n^2 t} \operatorname{sen}(nx) \quad (3.6)$$

Además, cualquier combinación lineal finita de ellas,

$$u(t, x) = \sum_{n=n_0}^{n_1} [a_n e^{-n^2 t} \cos(nx) + b_n e^{-n^2 t} \operatorname{sen}(nx)] \quad (3.7)$$

es solución.

Fourier pensaba que toda solución puede escribirse como suma infinita de senos y cosenos (lo que lleva involucrado un problema de convergencia).

Por otra parte, cuando $t = 0$, obtenemos una función de una variable

$$u(0, x) = f(x) = \sum_{n=n_0}^{n_1} [a_n e^{-n^2 t} \cos(nx) + b_n e^{-n^2 t} \operatorname{sen}(nx)] \quad (3.8)$$

De acuerdo con esto, Fourier suponía también que toda función periódica se puede escribir como suma infinita de senos y cosenos, lo cual no es cierto puntualmente.

Otra forma de escribir las series anteriores se obtiene con las fórmulas

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}), \quad \operatorname{sen} x = \frac{1}{2i}(e^{ix} - e^{-ix}) \quad (3.9)$$

Así,



Fourier



$$u(t, x) = \sum_{n \in \mathbb{N}} e^{-n^2} \left[\left(\frac{a_n}{2} + \frac{b_n}{2i} \right) e^{inx} + \left(\frac{a_n}{2} - \frac{b_n}{2i} \right) e^{-inx} \right] = \sum_{n \in \mathbb{N}} c_n e^{-n^2} e^{inx}$$

donde, para $t = 0$, tenemos $f(x) = \sum_{n \in \mathbb{Z}} c_n e^{inx}$

Aquí se plantean las siguientes preguntas:

1. ¿Las exponenciales trigonométricas serán base de algún espacio?
Esto permitiría obtener la función f cuando se conocen sus coeficientes respecto a dicha base.
2. ¿Cómo obtener los coeficientes c_n cuando la función f es conocida? Para responder a esta pregunta, aplicamos las siguientes fórmulas:

$$\frac{1}{2\pi} \int_0^{2\pi} e^{inx} dx = \frac{1}{2\pi in} (e^{in2\pi} - 1) = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n \neq 0, \end{cases}$$

$$\frac{1}{2\pi} \int_0^{2\pi} e^{imx} e^{-inx} dx = \frac{1}{2\pi in} (e^{i(m-n)2\pi} - 1) = \begin{cases} 1 & \text{si } m = n \\ 0 & \text{si } m \neq n, \end{cases}$$

Con lo anterior tenemos,

$$\int_0^{2\pi} f(x) dx = \int_0^{2\pi} \sum_{n \in \mathbb{Z}} c_n e^{inx} dx = \sum_{n \in \mathbb{Z}} c_n \int_0^{2\pi} e^{inx} dx = 2\pi c_0 \quad (3.10)$$

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx$$

$$\int_0^{2\pi} f(x) e^{-imx} dx = \int_0^{2\pi} \sum_{n \in \mathbb{Z}} c_n e^{inx} e^{-imx} dx = \sum_{n \in \mathbb{Z}} c_n \int_0^{2\pi} e^{inx} e^{-imx} dx = 2\pi c_m \quad (3.11)$$

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-imx} dx$$

Con esto, podemos calcular los coeficientes de la función f .



3.2. Espacios con Producto Interno

Una clase importante de espacios normados permite generalizar muchas propiedades de la geometría del espacio euclídeano que dependen de la noción de ángulo, especialmente las que se refieren a la perpendicularidad. Recordamos que, si $x, y \in \mathbb{R}^n$ y α es el ángulo que forman se define el producto escalar por,

$$\langle x, y \rangle = \|x\| \cdot \|y\| \cdot \cos \alpha \quad (3.12)$$

Queremos obtener una noción abstracta de producto escalar en espacios normados generales que permita extender la fórmula anterior. Como es lógico pensar, los espacios con producto escalar son históricamente anteriores a los espacios normados generales y mantienen todavía muchas propiedades de los espacios euclídeos.

Definición 3.2.1 *Espacios con Producto Interno o Pre-Hilbertiano.* Un espacio con producto interno también llamado pre-Hilbertiano es un espacio vectorial X en el que se define una aplicación, $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{C}$ con las siguientes propiedades:

1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$; (Propiedad aditiva)
2. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$; (Propiedad homogénea)
3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$; (Propiedad Hermítica)

Observación: Sea $x + iy \in \mathbb{C}$, con $y \neq 0$, entonces

$$\begin{aligned} x + iy &= \overline{x + iy} \\ x + iy &= x - iy \\ iy &= -iy \\ i &= -i \end{aligned}$$

Contradicción. Por lo tanto $y = 0$, entonces $x + iy = x \in \mathbb{R}$.

En conclusión: Si $x \in X$ entonces, $\langle x, x \rangle \in \mathbb{C}$.

4. $\langle x, x \rangle = \overline{\langle x, x \rangle}$ entonces $\langle x, x \rangle \in \mathbb{R}$, o, $\langle x, x \rangle \geq 0$
5. $\langle x, y \rangle \geq 0$, $y, \langle x, x \rangle = 0 \iff x = 0$ (Propiedad de la no negatividad)

De los axiomas se deducen inmediatamente las siguientes propiedades:

1. $\langle x, 0 \rangle = 0, \forall x \in X$
2. $\langle x, y \rangle = 0, \forall x \in X \iff y = 0$
3. $\langle x, \alpha y_1 + \beta y_2 \rangle = \overline{\alpha} \langle x, y_1 \rangle + \overline{\beta} \langle x, y_2 \rangle$



3.3. Norma en un Espacio con Producto Interno

Un Espacio con producto interno es un espacio normado, donde la norma asociada se define como,

$$\|x\| = \sqrt{\langle x, x \rangle}, \quad (3.13)$$

además se tiene que,

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

y por tanto es también un espacio métrico, con la distancia,

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}, \quad (3.14)$$

Teorema 3.3.1 (Desigualdad de Cauchy-Schwarz-Buniakowski). Para cualesquiera $x, y \in X$, (X un espacio con producto Interno) se tiene,

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad (3.15)$$

la igualdad $|\langle x, y \rangle| = \|x\| \|y\|$ se cumple si y sólo si x, y son linealmente dependientes.

3.4. Ejemplos de Espacios con Producto Interno

Ejemplo 3.4.1 El espacio de los Números Complejos. Este ejemplo es simple pero muy importante puesto que uno de los productos internos está definido por,

$$\langle x, y \rangle = x \bar{y}$$

Ejemplo 3.4.2 El espacio \mathbb{C}^n de las n -tuplas ordenadas $x = (x_1, \dots, x_n)$ de los Números Complejos. Uno de los productos internos puede estar definido por,

$$\langle x, y \rangle = \sum_{k=1}^n x_k \bar{y}_k, \quad x = (x_1, \dots, x_n) \quad y = (y_1, \dots, y_n)$$

3.5. Caracterización de Espacios Normados

Proposición 3.1 (Identidad del Paralelogramo). Si $(X, \langle \cdot, \cdot \rangle)$ es un espacio pre-Hilbertiano, la norma asociada verifica,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (3.16)$$

El siguiente resultado muestra cómo la norma asociada permite a su vez definir el producto interior.

Proposición 3.2 (Identidad de Polarización). Sea $(X, \langle \cdot, \cdot \rangle)$ un espacio pre-Hilbertiano arbitrario,



1. Si X es real, $\langle x, y \rangle = \frac{1}{4}[\|x + y\|^2 + \|x - y\|^2]$

2. Si X es complejo,

$$\operatorname{Re}\langle x, y \rangle = \frac{1}{4}[\|x + y\|^2 + \|x - y\|^2]$$

$$\operatorname{Im}\langle x, y \rangle = \frac{1}{4}[\|x + iy\|^2 + \|x - iy\|^2] \quad (3.17)$$

Esta propiedad sugiere una forma de definir un producto interior a partir de una norma. Sin embargo, será necesaria la identidad del paralelogramo.

3.6. Espacios de Hilbert

Definición 3.6.1 *Espacios de Hilbert.* Un espacio de Hilbert es un espacio pre-Hilbert completo (respecto a la métrica asociada). Esto significa además que todo espacio de Hilbert es un espacio de Banach en el que se ha definido un producto interno.

3.7. Algunos Ejemplos de Espacios de Hilbert

Ejemplo 3.7.1 \mathbb{R}^n es un espacio de Hilbert si definimos el producto interno, $\langle x, y \rangle = x_1y_1 + x_2y_2 + \dots + x_ny_n$. Como la métrica asociada,

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

es la euclídea, es ya sabido que el espacio es completo.

Ejemplo 3.7.2 \mathbb{C}^n es un espacio de Hilbert si definimos el producto interno, $\langle x, y \rangle = x_1\overline{y_1} + x_2\overline{y_2} + \dots + x_n\overline{y_n}$, donde

$$\|x\| = \sqrt{|x_1|^2 + \dots + |x_n|^2}$$

Ejemplo 3.7.3 El espacio $l^p(\mathbb{N})$, o espacio normado discreto.

$$l^p(\mathbb{N}) := \left\{ x = (x_1, x_2, \dots, x_n, \dots) / \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}, 1 \leq p < \infty \quad l^p(\mathbb{N}) = l^p$$

con la norma

$$\|x\|_{l^p} = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} \quad y \quad \|x\|_{l^\infty} = \sup_i |x_i| < \infty$$

Estos espacios **discretos** son entre otras cosas útiles en el estudio de las ondículas.



Para $p = 2$ tenemos el espacio normado

$$l^2(\mathbb{N}) \equiv l^2 = \left\{ x = (x_1, x_2, \dots, x_n, \dots) / \sum_{i=1}^{\infty} |x_i|^2 < \infty \right\}$$

Si $x, y \in l^2$ consideramos

$$\langle x, y \rangle = \sum_{n \in \mathbb{Z}} x(n) \overline{y(n)}$$

Si $x \in l^2$

$$\|x\| = \left(\sum_{n \in \mathbb{N}} |x(n)|^2 \right)^{\frac{1}{2}}$$

Ejemplo 3.7.4 Los espacios $L^p(A)$. Sea A un subconjunto medible (según Lebesgue) de \mathbb{R} y f una función medible sobre A . Para $1 \leq p \leq \infty$ se tiene que

$$L^p(A) = \left\{ f : A \rightarrow \mathbb{C} / \int_A |f(x)|^p dx < \infty \right\}$$

con la norma

$$\|f\|_p = \left(\int_A |f(x)|^p dx \right)^{\frac{1}{p}}$$

Ejemplo 3.7.5 El espacio $L^2([0, 2\pi])$ donde:

$$L^2([0, 2\pi]) = \left\{ f : [0, 2\pi] \rightarrow \mathbb{C} / \int_0^{2\pi} |f(\theta)|^2 d\theta < \infty \right\}$$

$L^2([0, 2\pi])$ y

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) \overline{g(\theta)} d\theta, \quad f, g \in L^2([0, 2\pi])$$

con la norma

$$\|f\| = \left(\frac{1}{2\pi} \int_0^{2\pi} |f(\theta)|^2 d\theta \right)^{\frac{1}{2}}$$

Ejemplo 3.7.6 El espacio $L^2(\mathbb{R})$: con f medible según Lebesgue.

$$L^2(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C} / \int_{\mathbb{R}} |f(t)|^2 dt < \infty \right\}$$

y

$$\langle f, g \rangle = \int_{\mathbb{R}} f(t) \overline{g(t)} dt$$

con la norma

$$\|f\| = \left(\int_{\mathbb{R}} |f(t)|^2 dt \right)^{\frac{1}{2}}$$



En este capítulo se estudian los sistemas wavelets de primera generación. Se definen la función escala y la función wavelet, y se demuestran algunas de sus propiedades.

4.1. Sistemas wavelet de primera generación

Los sistemas wavelet de primera generación son todos aquellos generados sólo por traslaciones enteras y escalamientos de una única función wavelet ψ , llamada wavelet madre o wavelet generadora, que da origen a una familia de funciones de la forma:

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k) \quad j, k \in \mathbb{Z} \quad (4.1)$$

donde el factor $2^{\frac{j}{2}}$ mantiene una norma constante independiente de la escala j . Esta familia de funciones es llamada el conjunto de expansión wavelet.

La wavelet madre ψ , trae siempre asociada consigo una función escala ϕ . Con estas dos funciones podremos aproximar cualquier función o señal $f \in L^2(\mathbb{R})$, mediante una de las funciones o mediante ambas, de la forma

$$f(t) = \sum_k \sum_j c_{j,k} \phi(t) + \sum_k \sum_j d_{j,k} \psi(t) \quad (4.2)$$

4.2. Características de sistemas wavelet

El conjunto de expansión wavelet no es único, existen muchos y muy diferentes sistemas wavelet, pero todos tienen las siguientes características:

1. Un sistema wavelet puede describirse de una manera “amigable”, como un conjunto de ladrillos (que para cada sistema pueden tener diferente forma) que sirven para reconstruir o representar una señal o función. Este

conjunto es una expansión bi-dimensional, que suele ser una base para alguna clase de señal de una o más dimensiones. En otras palabras, si el set de expansión esta dado por $\{\psi_{j,k}(t) \mid j, k \in \mathbb{Z}\}$, una expansión lineal puede ser

$$f(t) = \sum_k \sum_j d_{j,k} \psi_{j,k}(t) \quad (4.3)$$

para algún conjunto de coeficientes $d_{j,k}$.

2. La expansión wavelet entrega una localización tiempo-frecuencia instantánea de la señal, representación que puede explicarse como un pentagrama musical, donde la localización y forma de la figura musical nos dice cuando ocurre el tono y cual es su frecuencia. Esto quiere decir que la mayor parte de la energía de la señal es bien representada por unos pocos coeficientes $d_{j,k}$. Mientras un coeficiente de Fourier representa un componente que dura todo el tiempo en que se extiende la señal. Un coeficiente de expansión wavelet representa un componente bien definido en un intervalo de tiempo, esto es, un coeficiente wavelet es en sí bien localizado en el tiempo.
3. Los sistemas wavelet satisfacen las condiciones de multi-resolución. Esto significa que si un conjunto de señales puede ser representado por una suma de $\phi(t - k) \mid k \in \mathbb{Z}$, un conjunto más amplio de señales (que incluye el conjunto original) puede ser representado por una suma $\phi(2t - k) \mid k \in \mathbb{Z}$.
4. Los coeficientes de más baja resolución pueden ser calculados a partir de los coeficientes de más alta resolución, mediante un algoritmo en forma de árbol, llamado banco de filtros. Esto permite un eficiente cálculo de los coeficientes de expansión (también conocida como la Transformada Discreta Wavelet).
5. El tamaño de los coeficientes de expansión wavelet disminuye rápidamente con j y k .¹
6. Las wavelets son ajustables y adaptables. Debido a que existen muchos wavelet, estos pueden ser diseñados para adaptarse a una aplicación particular.
7. La generación de wavelets y el cálculo de la Transformada Discreta Wavelet es bien realizada por una computadora, pues estos cálculos se remiten sólo a multiplicaciones y sumas.

4.3. Función escala

Sea $\phi \in L^2(\mathbb{R})$, una función escala, que trasladada y escalada genera una familia de funciones $\{\phi_{j,k}(t) \mid j, k \in \mathbb{Z}\}$ definida como:

¹Debido a esto wavelet es una efectiva herramienta en compresión y denoising (limpieza) de señales.



$$\phi_{j,k}(t) = 2^{\frac{j}{2}} \phi(2^j t - k) \quad j, k \in \mathbb{Z} \quad (4.4)$$

$\forall j \in \mathbb{Z}$ la función escala define un subespacio $V_j \subset L^2(\mathbb{R})$ como

$$V_j = \overline{\text{Span}_{k \in \mathbb{Z}} \{ \phi_{j,k}(t) \}}$$

Entonces una función f estará en V_j si puede escribirse como

$$f(t) = \sum_{k \in \mathbb{Z}} c_{j,k} \phi_{j,k}(t) \quad (4.5)$$

con

$$c_{j,k} = \langle f(t), \phi_{j,k}(t) \rangle = \int_{-\infty}^{\infty} |f(t) \phi_{j,k}(t)| dt$$

esta última ecuación se demuestra haciendo el producto interno de 4.5 con $\phi_{j,k}(t)$, lo que nos entrega:

$$\begin{aligned} \langle f(t), \phi_{j,k}(t) \rangle &= c_{j,-\infty} \langle \phi_{j,-\infty}(t), \phi_{j,k}(t) \rangle + \cdots + c_{j,k} \langle \phi_{j,k}(t), \phi_{j,k}(t) \rangle \\ &+ \cdots + c_{j,\infty} \langle \phi_{j,\infty}(t), \phi_{j,k}(t) \rangle \end{aligned}$$

pero como veremos luego, la función escala debe cumplir con ciertos requisitos, y uno de ellos es ser ortonormal, por lo que sólo uno de los productos puntos de la ecuación anterior es distinto de cero $\langle \phi_{j,k}(t), \phi_{j,k}(t) \rangle = 1$ lo que nos entrega

$$c_{j,k} = \langle f(t), \phi_{j,k}(t) \rangle$$

Otra propiedad importante que resulta a partir de la definición de la función escala es que

$$f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}$$

lo que se demuestra como sigue

$$\begin{aligned} f(t) \in V_j &\Leftrightarrow f(t) = \sum_{k \in \mathbb{Z}} c_{j,k} \phi_{j,k}(t) \\ &\Leftrightarrow f(2t) = \sum_{k \in \mathbb{Z}} c_{j,k} \phi_{j,k}(2t) \\ &\Leftrightarrow f(2t) = \sum_{k \in \mathbb{Z}} c_{j,k} 2^{\frac{j}{2}} \phi(2^{j+1}t - k) \\ &\Leftrightarrow f(2t) = \sum_{k \in \mathbb{Z}} \frac{c_{j,k}}{\sqrt{2}} 2^{\frac{j+1}{2}} \phi(2^{j+1}t - k) \\ &\Leftrightarrow f(2t) = \sum_{k \in \mathbb{Z}} \frac{c_{j,k}}{\sqrt{2}} 2^{\frac{j+1}{2}} \phi_{j+1,k}(t) \\ &\Leftrightarrow f(2t) \in V_{j+1} \end{aligned}$$

Características de una función escala

Una función $\phi \in L^2(\mathbb{R})$ es considerada como una “buena” función de escala si cumple con las siguientes condiciones:



1. Para cada $j \in \mathbb{Z}$, $\{\phi_{j,k}/k \in \mathbb{Z}\}$ forma una base ortonormal para el subespacio $V_j \subset L^2(\mathbb{R})$.
2. Los subespacios V_j están anidados, o sea

$$\forall j \in \mathbb{Z}, \{V_j \subset V_{j+1}\}$$

3. La función ϕ tiene soporte compacto. Esto es, existe un subconjunto del dominio de ϕ donde esta no es cero, $\text{sup}\{\phi\} = \{x \in \mathbb{Z}/\phi(x) \neq 0\}$.

La normalidad exigida en la condición 1 se demuestra como:

$$\begin{aligned} \|\phi_{j,k}(t)\| &= \sqrt{\langle \phi_{j,k}(t), \phi_{j,k}(t) \rangle} \\ &= \sqrt{\int_{-\infty}^{\infty} \phi_{j,k}^2(t) dt} \\ &= \sqrt{\int_{-\infty}^{\infty} \left(2^{\frac{j}{2}}\right)^2 \phi^2(2^j t - k) dt} \\ &= \sqrt{\int_{-\infty}^{\infty} 2^j \phi^2(u) \frac{1}{2^j} du} \\ &= \sqrt{\int_{-\infty}^{\infty} \phi^2(u) du} \\ &= \|\phi(t)\| \end{aligned}$$

con

$$u = 2^j t - k \quad du = 2^j dt$$

La condición de ortonormalidad exige que para cada $j \in \mathbb{Z}$ tengamos:

$$\langle \phi_{j,k}(t), \phi_{j,l}(t) \rangle = \delta(k-l) = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases}$$

La condición 2 nos dice que los subespacios V_j incluyen más funciones de $L^2(\mathbb{R})$ a medida que j crece.

Proposición 4.1 *El anidamiento de los espacios se puede expresar como:*

$$\forall j \in \mathbb{Z}, \{V_j \subset V_{j+1}\} \Leftrightarrow \phi(t) \in V_j \quad (4.6)$$

Demostración: Demostramos ambas implicaciones por separado, comenzando por la más simple.

Por demostrar que si $\forall j \in \mathbb{Z}, \{V_j \subset V_{j+1}\} \Rightarrow \phi(t) \in V_j$ (primera implicación de 4.6).



Como $\phi(t) = \phi_{0,0}(t)$, tenemos que $\phi(t) \in V_0$ y como por hipótesis $V_0 \subset V_1$, debería darse que $\phi(t) \in V_1$.

Por demostrar $\phi(t) \in V_j \Rightarrow \forall j \in \mathbb{Z}, \{V_j \subset V_{j+1}\}$ (segunda implicación de 4.6).

Se tiene por hipótesis $\phi(t) \in V_j$, esto es

$$f(t) \in V_j \Leftrightarrow \sum_{k=-\infty}^{\infty} c_{j,k} 2^{\frac{j}{2}} \phi(2^j t - k)$$

además existen coeficientes $h(n)$ tales que

$$\phi(m') = \sum_{n \in \mathbb{Z}} h(n) \sqrt{2} \phi(2m' - n)$$

con $m' = 2^j t - k$

$$f(t) = \sum_{k=-\infty}^{\infty} c_{j,k} 2^{\frac{j}{2}} \sum_{n=-\infty}^{\infty} h(n) \sqrt{2} \phi(2(2^j t - k) - n)$$

$$f(t) = \sum_{k=-\infty}^{\infty} c_{j,k} \sum_{n=-\infty}^{\infty} 2^{\frac{j+1}{2}} h(n) \phi(2^{j+1} t - 2k - n)$$

sea $m = 2k + n$

$$f(t) = \sum_{k=-\infty}^{\infty} c_{j,k} \sum_{m=-\infty}^{\infty} 2^{\frac{j+1}{2}} h(m - 2k) \phi(2^{j+1} t - m)$$

$$f(t) = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} 2^{\frac{j+1}{2}} c_{j,k} h(m - 2k) \phi(2^{j+1} t - m)$$

como se mostrará más adelante, sólo un número finito de términos en la sumatoria son distintos de cero, intercambiando sumatorias en la última expresión se tiene

$$f(t) = \sum_{m=-\infty}^{\infty} 2^{\frac{j+1}{2}} \phi(2^{j+1} t - m) \sum_{k=-\infty}^{\infty} c_{j,k} h(m - 2k)$$

$$f(t) = \sum_{m=-\infty}^{\infty} \phi_{j+1,m}(t) \sum_{k=-\infty}^{\infty} c_{j,k} h(m - 2k)$$

sea $c_{j+1,m} = \sum_{k=-\infty}^{\infty} c_{j,k} h(m - 2k)$ entonces $f(t)$ puede escribirse como

$$f(t) = \sum_{m=-\infty}^{\infty} c_{j+1,m} \phi_{j+1,m}(t) \Rightarrow f(t) \in V_{j+1}$$



lo que implica que existen coeficientes $\{h(n)\}_{n \in \mathbb{Z}}$, tal que $\phi(t)$ puede ser expresada en términos de una suma de $\phi(2t)$ ajustada y trasladada como

$$\phi(t) = \sum_{k \in \mathbb{Z}} h(n) \sqrt{2} \phi(2t - k) \quad (4.7)$$

llamada ecuación básica de recursión o ecuación de escala, donde $\phi(t)$ tiene soporte compacto sobre $0 \leq t \leq p - 1$, los coeficientes $h(n)$ pueden ser una secuencia de números reales o complejos, llamados coeficientes de función escala (o filtro de escalamiento o vector de escalamiento) y la $\sqrt{2}$ mantiene la normalidad de la función escala en V_1 .

Proposición 4.2 *Si ϕ es una función escala como se define en 4.7, que tiene soporte compacto sobre $0 \leq t \leq p - 1$ y $\{\phi(2t - k)/k \in \mathbb{Z}\}$ es base ortonormal para V_1 , entonces $h(n)$ también tiene soporte compacto sobre $0 \leq t \leq p - 1$, por lo que la máxima longitud de la secuencia $h(n)$ es p .*

Demostración: Se debe mostrar que

$$\sqrt{2}h(n) = \int_{-\infty}^{\infty} \phi(t)\phi(2t - n)dt$$

por hipótesis la integral de la derecha es cero para $n < 0$ y $n > p - 1$, entonces queda demostrado que $h(n)$ tiene soporte en $0 \leq t \leq p - 1$.

Se tiene entonces por 4.7 que:

$$\int_{-\infty}^{\infty} \phi(t)\phi(2t - n)dt = \int_{-\infty}^{\infty} \sum_{k \in \mathbb{Z}} h(k) \sqrt{2} \phi(2t - k) \phi(2t - n) dt$$

como por hipótesis $\phi(t)$ tiene soporte compacto, sólo finitos coeficientes son no nulos, por lo que se puede intercambiar la sumatoria con la integral

$$\begin{aligned} \int_{-\infty}^{\infty} \sum_{k \in \mathbb{Z}} h(k) \sqrt{2} \phi(2t - k) \phi(2t - n) dt &= \sum_{k \in \mathbb{Z}} h(k) \sqrt{2} \int_{-\infty}^{\infty} \phi(2t - k) \phi(2t - n) dt \\ \Rightarrow \sum_{k \in \mathbb{Z}} h(k) \sqrt{2} \int_{-\infty}^{\infty} \phi(2t - k) \phi(2t - n) dt &= \sum_{k \in \mathbb{Z}} h(k) \sqrt{2} \delta(k - n) = h(n) \sqrt{2} \end{aligned}$$

luego, queda demostrada la proposición 4.2.

4.4. Función Wavelet

Se define

$$W_j = \overline{\text{Span}_{k \in \mathbb{Z}} \{\psi_{j,k}(t)\}}$$



como el complemento ortogonal de V_j en V_{j+1} , esto significa que todos los miembros de V_j son ortogonales a todos los miembros de W_j . Entonces se requiere que

$$\langle \phi_{j,k}(t), \psi_{j,k}(t) \rangle = \int_{-\infty}^{\infty} \phi_{j,k}(t) \psi_{j,k}(t) dt = 0 \quad j, k \in \mathbb{Z}$$

con

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k) \quad j, k \in \mathbb{Z}$$

y además

$$V_1 = V_0 \oplus W_0$$

donde cualquier función $f(t) \in W_j$ puede ser representada como

$$f(t) = \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t)$$

la función básica wavelet $\psi(t) = \psi_{0,0}(t) \in W_0$, y como $V_1 = V_0 \oplus W_0$, entonces $\psi(t) \in V_1$.

Por lo tanto ψ se conoce como *wavelet madre*, la cual se representa como

$$\psi(t) = \sum_{k \in \mathbb{Z}} \sqrt{2} h(n) \phi(2t - k) \quad (4.8)$$



Se define entonces un conjunto de funciones de escalamiento en términos de traslaciones enteras de la función básica de escalamiento $\phi(t)$:

$$\phi_k(t) = \phi(t - k) = \begin{cases} 1 & \text{si } t_1 = k \leq t < k + 1 = t_2 \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.2)$$

donde

$$k \in \mathbb{Z} \quad \wedge \quad \phi \in L^2(\mathbb{R})$$

el subespacio de $L^2(\mathbb{R})$ generado por esta función es definido como

$$V_0 = \overline{\text{Span}_k\{\phi_k(t)\}} \quad (5.3)$$

donde $\text{Span}_k\{\phi_k(t)\}$ significa la expansión de la familia de funciones $\phi_k(t)$, para todos los enteros k desde $-\infty$ a ∞ . El superrayado denota la clausura, esto significa que

$$f(t) = \sum_{k \in \mathbb{Z}} c_k \phi_k(t) \quad (5.4)$$

para cualquier $f \in V_0 \in L^2(\mathbb{R})$. O sea, cualquier función f que esté en V_0 puede ser representada por una combinación lineal del conjunto de funciones ϕ_k con sus respectivos coeficientes c_k . Para un rápido cálculo de estos coeficientes es necesario que ϕ_k sea ortonormal. Para probar esta propiedad de la familia de funciones en cuestión, se escribe

$$\phi_m(t) = \phi(t - m) = \begin{cases} 1 & \text{si } t_3 = m \leq t < m + 1 = t_4 \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.5)$$

entonces el producto punto entre 5.2 y 5.5 debe ser cero para cumplir con la condición de ortogonalidad, es decir

$$\langle \phi_k(t), \phi_m(t) \rangle = \int \phi_k(t) \phi_m(t) dt = 0 \quad \text{con } k \neq m \quad (5.6)$$

esto puede probarse, ya que 5.2 y 5.5 están definidas en distintos intervalos de tiempo, entonces se tiene que

$$\int \phi_k(t) \phi_m(t) dt = \int_{t_1}^{t_2} \phi_k(t) \phi_m(t) dt + \int_{t_3}^{t_4} \phi_k(t) \phi_m(t) dt \quad (5.7)$$

$$\int_{t_1}^{t_2} 1 * 0 dt + \int_{t_3}^{t_4} 0 * 1 dt = 0 \quad (5.8)$$

lo cual demuestra la ortogonalidad de $\phi_k(t)$ con $k \in \mathbb{Z}$, con lo que sólo nos resta demostrar la normalidad de este conjunto de funciones.

Sea

$$\langle \phi_k(t), \phi_k(t) \rangle = \int_{t_1}^{t_2} \phi_k(t) \phi_k(t) dt \quad (5.9)$$



por 5.2

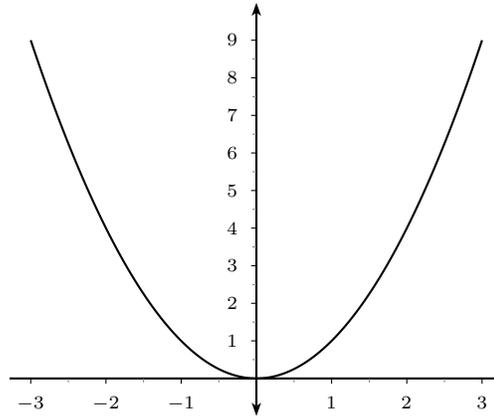
$$\int_{t_1}^{t_2} \phi_k(t)\phi_k(t) dt = t|_{t_1}^{t_2} = (k+1) - k = 1 \quad (5.10)$$

se demuestra así la ortonormalidad de $\phi_k(t)$.

Ejemplo 1: Representar la función

$$\begin{aligned} f : [-3, 3] &\longrightarrow \mathbb{R} \\ t &\longrightarrow t^2 \end{aligned}$$

mediante la función de escalamiento $\phi_k(t)$ con $t \in [-3, 2]$.



$f(t) = t^2$ definida sobre el intervalo $[-3, 3]$.

Solución:

Representaremos f , como una combinación lineal de la forma

$$f(t) = c_{-3}\phi(t+3) + c_{-2}\phi(t+2) + c_{-1}\phi(t+1) + c_0\phi(t) + c_1\phi(t-1) + c_2\phi(t-2) \quad (5.11)$$

Ahora calcularemos el coeficiente c_{-3} para lo que haremos la multiplicación vectorial de $\phi(t+3)$ por 5.11

$$\langle t^2, \phi(t+3) \rangle = \langle c_{-3}\phi(t+3), \phi(t+3) \rangle + \cdots + \langle c_2\phi(t-2), \phi(t+3) \rangle \quad (5.12)$$

como se mostró anteriormente, $\phi_k(t)$ es ortogonal, por lo que todos los productos puntos de la parte derecha de 5.12 serán cero, excepto $\langle c_{-3}\phi(t+3), \phi(t+3) \rangle$

$$\langle t^2, \phi(t+3) \rangle = \langle c_{-3}\phi(t+3), \phi(t+3) \rangle$$

$$\langle t^2, \phi(t+3) \rangle = \int_{-3}^{-3+1} c_{-3}\phi(t+3)\phi(t+3) dt$$



donde los límites de integración fueron calculados a partir de 5.2

$$\langle t^2, \phi(t+3) \rangle = c_{-3} t \Big|_{-3}^{-2} = c_{-3}[-2+3] = c_{-3} \quad (5.13)$$

o bien,

$$c_{-3} = \int_{-3}^3 t^2 \phi(t+3) dt \quad (5.14)$$

$$c_{-3} = \int_{-3}^{-2} t^2 \phi(t+3) dt + \int_{-2}^3 t^2 \phi(t+3) dt$$

como $\phi(t+3)$ está definida sólo en $[-3,-2]$

$$c_{-3} = \int_{-3}^{-2} t^2 dt + \int_{-2}^3 t^2 * 0 dt \quad (5.15)$$

$$\Rightarrow c_{-3} = \frac{t^3}{3} \Big|_{-3}^{-2}$$

$$c_{-3} = 6.\bar{3} \quad (5.16)$$

Ahora calculamos c_{-2}

$$\langle t^2, \phi(t+2) \rangle = \langle c_{-3}\phi(t+3), \phi(t+2) \rangle + \cdots + \langle c_2\phi(t-2), \phi(t+2) \rangle \quad (5.17)$$

como se mostró anteriormente, $\phi_k(t)$ es ortogonal, por lo que todos los productos puntos de la parte derecha de 5.17 serán cero, excepto $\langle c_{-2}\phi(t+2), \phi(t+2) \rangle$

$$\langle t^2, \phi(t+2) \rangle = \langle c_{-2}\phi(t+2), \phi(t+2) \rangle \quad (5.18)$$

$$\langle t^2, \phi(t+2) \rangle = \int_{-2}^{-2+1} c_{-2}\phi(t+2)\phi(t+2) dt \quad (5.19)$$

donde los límites de integración fueron calculados a partir de 5.2

$$\langle t^2, \phi(t+2) \rangle = c_{-2} t \Big|_{-2}^{-1} = c_{-2}[-1+2] = c_{-2} \quad (5.20)$$

$$c_{-2} = \int_{-3}^3 t^2 \phi(t+2) dt \quad (5.21)$$

$$c_{-2} = \int_{-3}^{-2} t^2 \phi(t+2) dt + \int_{-2}^{-1} t^2 \phi(t+2) dt + \int_{-1}^3 t^2 \phi(t+2) dt$$

como $\phi(t+2)$ está definida sólo en $[-2,-1]$

$$c_{-2} = \int_{-3}^{-2} t^2 * 0 dt + \int_{-2}^{-1} t^2 dt + \int_{-1}^3 t^2 * 0 dt \quad (5.22)$$



$$c_{-2} = \frac{t^3}{3} \Big|_{-2}^{-1}$$

$$c_{-2} = 2.\bar{3} \quad (5.23)$$

del cálculo de estos dos coeficientes queda claro que cualquier coeficiente c_k puede ser calculado como

$$c_k = \langle f(t), \phi_k(t) \rangle = \int_{t_1}^{t_2} f(t) dt \quad (5.24)$$

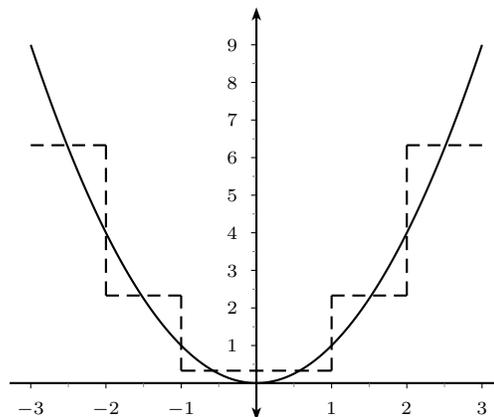
siempre que $f(t) \in L^2(\mathbb{R})$ y $\phi_k(t)$ sea ortonormal. Se calculan entonces los coeficientes restantes utilizando 5.24

$$c_{-1} = \int_{-1}^{-1+1} t^2 \phi(t+1) dt = \frac{t^3}{3} \Big|_{-1}^0 = 0.\bar{3}$$

En la siguiente tabla se encuentran todos los coeficientes calculados

Coficiente	Valor obtenido
c_{-3}	6.33
c_{-2}	2.33
c_{-1}	0.33
c_0	0.33
c_1	2.33
c_2	6.33

Ahora que se tienen los seis coeficientes buscados, se puede reconstruir $f(t) = t^2$ mediante 5.11, y se obtiene la gráfica mostrada en la siguiente figura



En esta figura apreciamos que la aproximación realizada por la función de escalamiento es bastante tosca, debido al espacio en que trabajamos.

Para una mejor representación de cualquier señal perteneciente a $L^2(\mathbb{R})$ por medio de la función de escalada, debemos ser capaces de crear una nueva familia de funciones, que sea posible trasladar, y también de ser escaladas para lograr así una mejor resolución. Definimos entonces un nuevo conjunto de funciones de escalamiento

$$\phi_{j,k}(t) = \phi(2^j t - k) = \begin{cases} 1 & \text{si } t_1 = \frac{k}{2^j} \leq t < \frac{k+1}{2^j} = t_2 \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.25)$$

como $\phi_{j,k}(t) \in V_j$, que se define como

$$V_j = \overline{\text{Span}_k\{\phi_k(2^j t)\}} = \overline{\text{Span}_k\{\phi_{j,k}(t)\}} \quad (5.26)$$

donde $\phi_{j,k}(t)$ es la función básica de escalamiento trasladada y escalada, y V_j es un subespacio de $L^2(\mathbb{R})$ generado por 5.25.

Ejemplo 2:

Representar la función $f(t) = t^2$ con $t \in [-3, 3]$, mediante la familia de funciones $\phi_{j,k}(t)$.

Solución:

Para representar esta función trabajaremos en el espacio V_1 , o lo que es lo mismo, haciendo $j = 1$. Luego,

$$t^2 = \sum_{k=-6}^5 c_k \phi(2t - k)$$

siguiendo un procedimiento análogo al del ejemplo 1 (revisar desde la ecuación 5.12 a 5.15 tenemos que

$$\langle t^2, \phi(2t + 6) \rangle = \langle c_{-6} \phi(2t + 6), \phi(2t + 6) \rangle = \int_{-3}^{-2.5} c_{-6} dt$$

donde los límites de integración están dados por 5.25. Así que

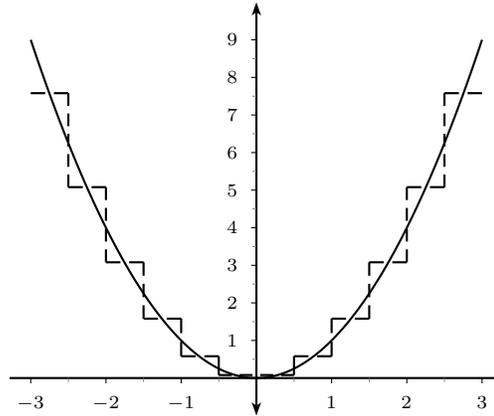
$$\langle c_{-6} \phi(2t + 6), \phi(2t + 6) \rangle = \frac{c_{-6}}{2} = \langle t^2, \phi(2t + 6) \rangle \langle$$

$$c_{-6} = 2 \int_{-3}^{-2.5} t^2 dt = \frac{t^3}{3} \Big|_{-3}^{-2.5} = 7,58\bar{3}$$

los demás coeficientes serán calculados de la misma forma, estos son mostrados en la siguiente tabla y en la figura se muestra la reconstrucción de $f(t) = t^2$ usando $\phi_{j,k}(t)$ en V_1



Coefficientes	Valor obtenido
c_{-6}	7,58
c_{-5}	5,08
c_{-4}	3,08
c_{-3}	1,58
c_{-2}	0,58
c_{-1}	0,08
c_0	0,08
c_1	0,58
c_2	1,58
c_3	3,08
c_4	5,08
c_5	7,58



En esta figura se presenta una clara mejoría en la resolución de la representación de la función cuadrática, debido que $\phi_{j,k}(t)$ en V_1 esta definida en un intervalo de tiempo más pequeño que en V_0 .

5.1.1. Relación ortogonalidad y normalidad

Del ejemplo anterior se puede observar que para V_1 los coeficientes quedaban divididos por 2, lo que indica que las funciones de escalamiento de este subespacio no son ortonormales. Demostraremos ahora que esto no ocurre solo en V_1 , si no que en todo V_j . En efecto,

$$\langle \phi_{j,k}(t), \phi_{j,k}(t) \rangle = \int_{t_1}^{t_2} \phi_{j,k}(t) \phi_{j,k}(t) dt \quad \text{con } t_1 = \frac{k}{2^j} \quad t_2 = \frac{k+1}{2^j} \quad (5.27)$$

$$\langle \phi_{j,k}(t), \phi_{j,k}(t) \rangle = t \Big|_{t_1}^{t_2} = \frac{1}{2^j}$$



lo cual demuestra la no normalidad de esta familia de funciones. Por lo tanto debemos encontrar una constante que nos permita hacer estas funciones ortonormales, entonces, para V_j multipliquemos las funciones base por un número cualquiera r_j

$$\langle r_j \phi_{j,k}(t), r_j \phi_{j,k}(t) \rangle = 1$$

$$\langle r_j \phi_{j,k}(t), r_j \phi_{j,k}(t) \rangle = \int_{t_1}^{t_2} r_j \phi_{j,k}(t) r_j \phi_{j,k}(t) dt = \frac{r_j^2}{2^j} = 1$$

asi que

$$r_j = 2^{\frac{j}{2}}$$

con lo que obtenemos lo que denominaremos la constante de normalización.

Con esta mejora de $\phi_{j,k}(t)$ definimos una nueva familia de funciones

$$\phi_{j,k}(t) = 2^{\frac{j}{2}} \phi(2^j t - k) = \begin{cases} 2^{\frac{j}{2}} & \text{si } t_1 = \frac{k}{2^j} \leq t < \frac{k+1}{2^j} = t_2 \\ 0 & \text{caso contrario} \end{cases} \quad (5.28)$$

entonces de ahora en adelante cuando hagamos referencia a $\phi_{j,k}(t)$, estaremos hablando de 5.28.

Una importante propiedad de la función de escalamiento $\phi_{j,k}(t)$, es que puede ser representada por una combinación lineal de ella misma, trasladada y escalada. Si elegimos $j = 1$, y de acuerdo con 4.8 obtenemos que

$$\Rightarrow \phi(t) = \sqrt{2}h(1)\phi(2t) + \sqrt{2}h(2)\phi(2t - 1) \quad (5.29)$$

se escogen $\phi_{1,0}(t)$ y $\phi_{1,1}(t)$ ya que necesitamos cumplir con el intervalo de tiempo en que está definida $\phi(t)$, esto es, $\phi_{1,0}(t)$ esta definida en $t^2 \in [0, \frac{1}{2})$ y $\phi_{1,1}(t)$ esta definida $t^2 \in [\frac{1}{2}, 1)$, como se muestra en 5.25.

Calculemos $h(1)$ y $h(2)$

$$\begin{aligned} h(1) &= \sqrt{2} \int \phi(t) \cdot \phi(2t) dt \\ &= \sqrt{2} \left[\int_0^{0,5} 1 \cdot 1 dt + \int_{0,5}^1 1 \cdot 0 dt \right] \\ &= \sqrt{2} t \Big|_0^{0,5} = \frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}} \end{aligned}$$

ahora,

$$\begin{aligned} h(2) &= \sqrt{2} \int \phi(t) \cdot \phi(2t - 1) dt \\ &= \sqrt{2} \left[\int_{0,5}^1 0 \cdot 1 dt + \int_{0,5}^1 1 \cdot 1 dt \right] \\ &= \sqrt{2} t \Big|_{0,5}^1 = \frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}} \end{aligned}$$



entonces tenemos que la función básica de escalamiento puede ser representada por sí misma, escalada y trasladada, como sigue

$$\phi(t) = \phi(2t) + \phi(2t - 1) \quad (5.30)$$

5.2. Función Wavelet Haar

Como se observó en la sección anterior, se obtiene una mejor aproximación de la señal utilizando las funciones de escalamiento que ocupan el espacio V_1 que utilizando la función escalamiento que ocupa el espacio V_0 . Sin embargo, las características de una señal pueden ser mejor descritas, no incrementado el tamaño del espacio de las funciones escalamiento, sino definiendo un espacio W_j de funciones levemente diferentes a las funciones escalamiento, que representen la diferencia que existe entre un espacio V_j y un espacio V_{j+1} , tal que

$$V_{j+1} = V_j \oplus W_j \quad (5.31)$$

por lo que ya estamos en condiciones de decir que el espacio W_0 corresponde al complemento del espacio V_0 en el espacio V_1 .

Ahora bien, la función que expande el espacio W_0 se conoce como función wavelet, y se define de la forma

$$\psi(t) = \begin{cases} 1 & \text{si } 0 \leq t < \frac{1}{2} \\ -1 & \text{si } \frac{1}{2} \leq t \leq 1 \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.32)$$

que al igual que la función de escalamiento que expande el espacio V_0 , puede ser representada sobre el intervalo $[0,1]$ como una combinación lineal de las funciones de escalamiento que expanden el espacio V_1 de la siguiente manera

$$\psi(t) = h_1(1)\sqrt{2}\phi(2t) + h_1(2)\sqrt{2}\phi(2t - 1) \quad (5.33)$$

de tal manera que al realizar el producto interno de 5.33 por $\sqrt{2}\phi(2t)$ y luego por $\sqrt{2}\phi(2t - 1)$, obtenemos respectivamente que

$$h_1(1) = \sqrt{2} \int_0^{\frac{1}{2}} \psi(t) dt \quad (5.34)$$

y

$$h_1(2) = \sqrt{2} \int_{\frac{1}{2}}^1 \psi(t) dt \quad (5.35)$$

respectivamente, y calculando los coeficientes h_1 y h_2 obtenemos como resultado que

$$h_1(1) = -h_1(2) = \frac{1}{\sqrt{2}} \quad (5.36)$$



observando que estos coeficientes nos permiten mantener la normalidad de la función $\psi(t)$.

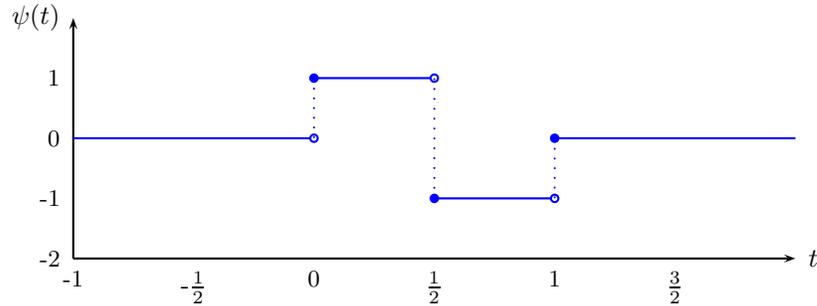
Como los espacios V_0 y V_1 son ortogonales entonces cualquier espacio V_j con $j = 0, \pm 1, \pm 2, \dots$ también lo es, entonces el espacio W_0 al ser el complemento de V_0 en V_1 es ortogonal. Por lo tanto, al igual que con la función escalamiento, es posible obtener una representación de la diferencia que existe entre aproximar una señal con un nivel de resolución j y aproximar la misma señal con un nivel de resolución $j + 1$, mediante el producto interno de esta señal con un conjunto de funciones que expandan el espacio W_j donde j será elegido de acuerdo al grado de aproximación que se desee. La propiedad de ortogonalidad de W_0 y por ende de la función wavelet $\psi(t)$, puede ser demostrada en forma análoga a como se demostró con la función de escalamiento.

Si definimos la función wavelet como

$$\psi_k(t) = \begin{cases} 1 & \text{si } k \leq t < k + \frac{1}{2} \\ -1 & \text{si } k + \frac{1}{2} \leq t \leq k + 1 \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.37)$$

donde

$$k \in \mathbb{Z} \quad \text{y} \quad \psi \in L^2(\mathbb{R})$$



Función wavelet en W_0 como combinación lineal de las funciones escalamiento que expanden V_1 y V_0 .

que corresponde a la misma función pero desplazada en el tiempo por una constante k , y definimos de la misma forma otra función $\psi_m(t)$ con $m > k$, entonces realizando el producto interno entre ellas obtenemos

$$\langle \psi_k(t), \psi_m(t) \rangle = \int_k^{m+1} \psi_k(t) \psi_m(t) dt \quad (5.38)$$

obtenemos que

$$\langle \psi_k(t), \psi_m(t) \rangle = \int_k^{k+1} \psi_k(t) * 0 dt + \int_m^{m+1} 0 * \psi_m(t) dt = 0 \quad m, k \in \mathbb{Z} \quad (5.39)$$



con lo que se da por finalizada la demostración.

Ejemplo 3:

Aproximar la función $f(t) = t^2$ sobre el intervalo $[-3,3]$ en forma análoga a como se hizo con la función escalamiento.

Solución:

La primera aproximación la haremos en W_0 de tal manera que las funciones wavelet de este espacio serán de la forma $\psi_k(t) = \psi(t - k)$ con $k \in [-3, 2)$, y por lo tanto

$$f(t) = d_{-3}\psi(t+3) + d_{-2}\psi(t+2) + d_{-1}\psi(t+1) + d_0\psi(t) + d_1\psi(t-1) + d_2\psi(t-2) \quad (5.40)$$

ahora, para calcular el primer coeficiente d_{-3} haremos el producto punto de $\psi(t+3)$ con 5.40

$$\langle t^2, \psi(t+3) \rangle = \langle d_{-3}\psi(t+3), \psi(t+3) \rangle + \cdots + \langle d_2\psi(t-2), \psi(t+3) \rangle \quad (5.41)$$

entonces, como demostramos anteriormente que $\psi_k(t)$ es ortogonal, todos los productos puntos de la parte derecha de 5.41 serán cero, excepto $\langle d_{-3}\psi(t+3), \psi(t+3) \rangle$, por lo tanto

$$\langle t^2, \psi(t+3) \rangle = \int_{-3}^{-3+1} d_{-3}\psi(t+3)\psi(t+3) dt$$

$$\int_{-3}^{-3+1} t^2, \psi(t+3) dt = \int_{-3}^{-3+1} d_{-3}\psi(t+3)\psi(t+3) dt$$

que para $\psi(t+3) = 1$ nos queda

$$\int_{-3}^{-3+\frac{1}{2}} t^2 dt = \frac{d_{-3}^+}{2} \quad (5.42)$$

para $\psi(t+3) = -1$

$$\int_{-3+\frac{1}{2}}^{-2} -t^2 dt = \frac{d_{-3}^-}{2} \quad (5.43)$$

y para $\psi(t+3) = 0$ la solución es trivial.

La resolución de las integrales de 5.42 y 5.43 nos entregan dos coeficientes, los cuales son

$$d_{-3}^+ = \frac{91}{12} = 7,58 \quad d_{-3}^- = -\frac{61}{12} = -5,08$$

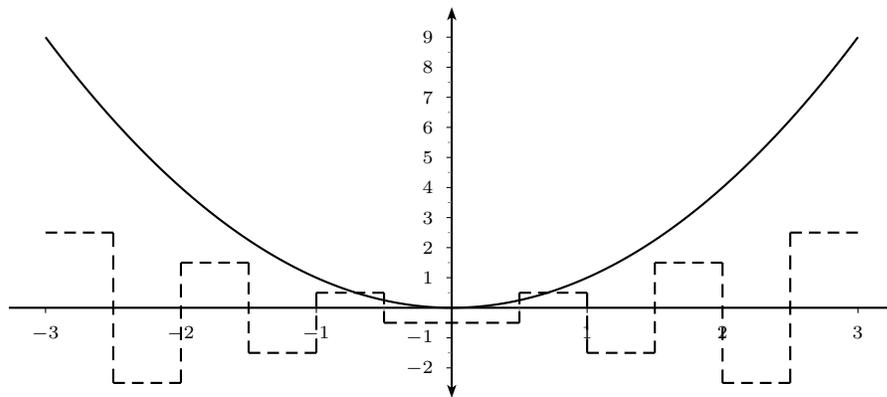
donde d^+ y d^- representan los coeficientes obtenidos debido a la parte positiva de la función wavelet Haar y la parte negativa de la función wavelet Haar respectivamente.

Realizando el cálculo de los otros coeficientes en forma análoga, observamos sus resultados en la siguiente tabla y en la figura 6.6.



Coefficientes	Valores obtenidos
d_{-3}	2.5
d_{-2}	1.5
d_{-1}	0.5
d_0	-0.5
d_1	-1.5
d_2	-2.5

Coefficientes obtenidos para la representación de la señal t^2 con la Haar Wavelet en el espacio W_0 .



Aproximación de la señal t^2 mediante las función wavelet del espacio W_0 .

Ejemplo 4:

Aproximar la función $f(t) = t^2$ con $t \in [-3, 3]$ mediante las funciones $\psi_{j,k}(t)$ con un valor de $j = 1$.

Solución:

Representar esta función t^2 para $j = 1$, significa que obtendremos una aproximación de ella en el subespacio V_1 , por lo tanto la representación de $f(t) = t^2$ mediante la combinación lineal de $\psi_{1,k}(t)$, queda

$$t^2 = d_{-6}\psi(2t+6) + d_{-5}\psi(2t+5) + d_{-4}\psi(2t+4) + \dots + d_5\psi(2t-5) \quad (5.44)$$

el cálculo de los coeficientes se realiza de forma análoga al ejemplo anterior, de tal manera que nuestro primer coeficiente lo podemos expresar como

$$\int_{-3}^{-3+\frac{1}{2}} t^2 \psi(2t+6) dt = \int_{-3}^{-3+\frac{1}{2}} \psi(2t+6) \psi(2t+6) dt \quad (5.45)$$

que para $\psi(2t+6) = 1$ nos queda

$$\int_{-3}^{-3+\frac{1}{4}} t^2 dt = \frac{d_{-6}^+}{4}$$



y para $\psi(2t + 6) = -1$

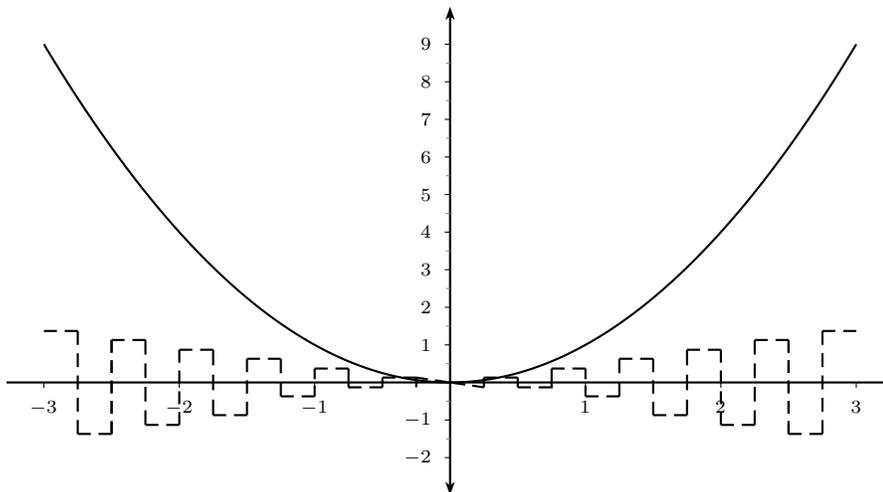
$$\int_{-3+\frac{1}{4}}^{-3+\frac{1}{2}} -t^2 dt = \frac{d_{-6}^-}{4}$$

Mediante la resolución de las anteriores integrales obtenemos los respectivos valores de los coeficientes d_{-6}^+ y d_{-6}^-

$$d_{-6}^+ = 8,27 \quad d_{-6}^- = -6,9 \quad (5.46)$$

El resto de los coeficientes se calcula en forma análoga, obteniéndose al final una cantidad de 12 coeficientes (mostrados en la siguiente tabla), que corresponden al doble de coeficientes obtenidos en el ejemplo anterior. La representación de la función original a través de estos coeficientes puede observarse en la siguiente figura

Coeficientes	Valores obtenidos
d_{-6}	1.37
d_{-5}	1.13
d_{-4}	0.87
d_{-3}	0.63
d_{-2}	0.37
d_{-1}	0.13
d_0	-0.13
d_1	-0.37
d_2	-0.63
d_3	-0.87
d_4	-1.13
d_5	-1.37



Representación de la función $f(t) = t^2$ en el subespacio W_1 .



5.2.1. Relación de ortogonalidad y normalidad

De los ejemplos anteriores hay que observar que para W_0 los coeficientes quedaban divididos por 2, y que para W_1 los coeficientes quedaban divididos por 4, lo que indica que las funciones wavelets de estos subespacios no son ortonormales. Por lo tanto debemos encontrar una constante que nos permita hacer estas funciones ortonormales, entonces, para W_0 multipliquemos las funciones base por un número cualquiera r_0 y de forma análoga llevémoslo a cabo para W_1 con r_1 , de tal manera que

$$\begin{aligned}\langle r_0\psi(t), r_0\psi(t) \rangle &= \pm \frac{1}{2} \\ \langle r_1\psi(2t), r_1\psi(2t) \rangle &= \pm \frac{1}{4}\end{aligned}\tag{5.47}$$

entonces como sabemos que 6.3 debe dar 1 para que $\psi(t)$ y $\psi(2t)$ sean ortonormales, el valor de r_0 debe ser $\sqrt{2}$ y el valor de r_1 debe ser 2. Además, podemos observar que es posible escribir r_1 en término de r_0 de la forma $r_1 = \sqrt{2} * \sqrt{2}$. Por lo tanto para un subespacio W_j podemos establecer que

$$r_j = \sqrt{2} * \sqrt{2} * \sqrt{2} * \dots \quad \text{hasta la } j - \text{ésima multiplicación por } \sqrt{2}$$

es decir

$$r_j = 2^{\frac{j}{2}}$$

con lo que obtenemos lo que denominaremos la constante de normalización. De esta manera, al igual que las funciones escalamiento, ya estamos en condiciones de definir un set de funciones bases ortonormales mediante escalamiento y traslación de la forma

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k) = \begin{cases} 2^{\frac{j}{2}} & \text{si } \frac{k}{2^j} \leq t < \frac{k}{2^j} + \frac{1}{2^{j+1}} \\ -2^{\frac{j}{2}} & \text{si } \frac{k}{2^j} + \frac{1}{2^{j+1}} \leq t < \frac{k+1}{2^j} \\ 0 & \text{caso contrario} \end{cases}\tag{5.48}$$

con $\psi_{j,k}(t) \in W_j$ que se define como

$$W_j = \overline{\text{Span}_k\{\psi_k(2^j t)\}} = \overline{\text{Span}_k\{\psi_{j,k}(t)\}}\tag{5.49}$$

donde $\psi_{j,k}(t)$ es la función wavelet base trasladada y escalada, y W_j es un subespacio de $L^2(\mathbb{R})$ generado por 6.5.



6

Matrices Haar

Tomado del libro *Un Enfoque Cualitativo a Las EDO'S y Teoría Wavelets* del magister Luis Arturo Polanía Quiza, capítulo 6, sección 6.6, [14].

En el análisis numérico, los algoritmos basados en wavelets se hacen una herramienta importante a causa de las propiedades de localización. Una de las familias más conocidas de wavelet es la Wavelet Haar. Debido a lo sencillo, estas se hacen una herramienta efectiva para resolver muchos problemas, entre los que están las Ecuaciones Diferenciales Ordinarias de orden superior. En estos trabajos, el operador o representación matricial es expandida en una base wavelet.

Haciendo un recorrido a través de la historia de la solución de ecuaciones diferenciales ordinarias usando el método relacionado con la Wavelet Haar, Chen C. F. y Hsiao C. H. han obtenido una matriz operacional de integración basada en la Wavelet Haar. También, Ülo Lepik ha usado el método Wavelet Haar para solucionar ecuaciones diferenciales de alto orden; así como también, EDO'S no lineales. Con los aportes realizados por Chen y Hsiao y Lepik; junto a los aportes de Phang Chang, Phang Piasu se muestra una presentación del procedimiento construido por dichas personas, para solucionar ecuaciones diferenciales ordinarias las cuales hacen uso de las potencias de wavelets.

La función Wavelet Haar es la primera wavelet conocida y fue propuesta en 1.909 por Alfred Haar. Ésta una función escalón $\psi(t)$, definida por

$$\psi(t) = \begin{cases} 1, & \text{si } 0 \leq t \leq \frac{1}{2} \\ -1, & \text{si } \frac{1}{2} \leq t \leq 1 \\ 0, & \text{en los otros casos.} \end{cases}$$

como se mostró en el capítulo 2 y cuya gráfica aparece en la sección 2.2. Tam-

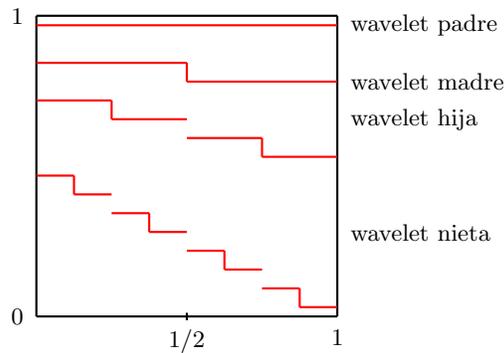
bién es llamada wavelet madre. Con el fin de ejecutar transformada wavelet, la wavelet Haar usa traslaciones y dilataciones de la función, es decir, la transformada hace uso de la siguiente función

$$\psi(t) = \psi(2^j t - k)$$

$$\text{Traslación } \psi(t) = \psi(t - k)$$

$$\text{Escalamiento (dilatación) } \psi(t) = \psi(2^j t)$$

Mediante los procesos de traslación y dilatación dados por las dos últimas ecuaciones, se pueden obtener la wavelet padre, wavelet hija, y wavelet nieta como aparece en la siguiente figura:

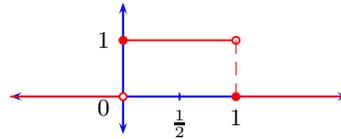


wavelet Haar
(bajo 2 niveles de resolución)

En lo que sigue se presenta el proceso que permite obtener la figura anterior:

Wavelet padre

$$\phi(t) = \begin{cases} 1, & \text{si } 0 \leq t < 1 \\ 0, & \text{si } t \notin [0, 1] \end{cases}$$

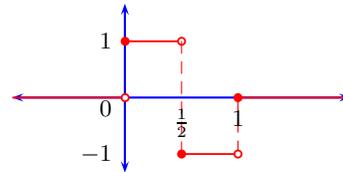


Función escala ó función padre ϕ

Wavelet madre o función madre

$$\psi(t) = \phi(2t) - \phi(2t - 1)$$

$$\psi(t) = \begin{cases} 1, & \text{si } 0 \leq t < 1/2 \\ -1, & \text{si } 1/2 \leq t < 1 \\ 0, & \text{en los otros casos} \end{cases}$$



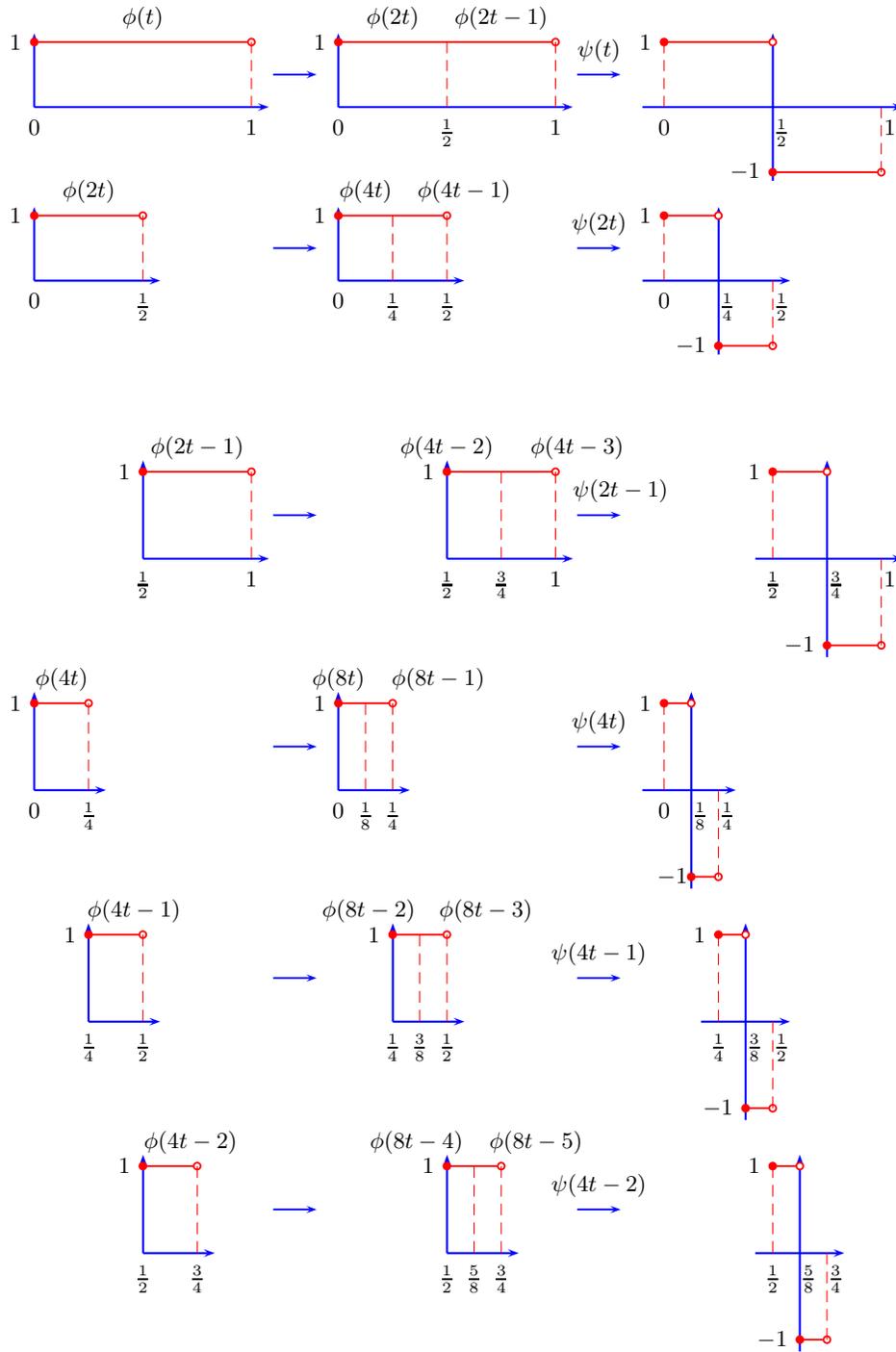
La Wavelet de Haar ψ

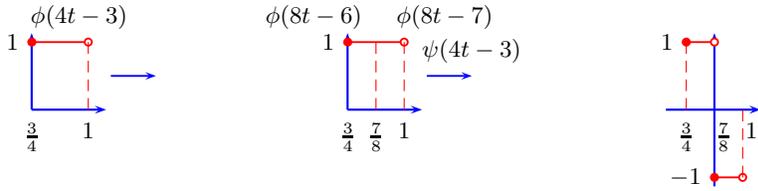
De las definiciones de ϕ y ψ se obtiene las siguientes relaciones de escala:

$$\phi(t) = \phi(2t) + \phi(2t - 1) \quad \text{ecuación de dilatación}$$

$$\psi(t) = \phi(2t) - \phi(2t - 1)$$







Según Ü. Lepik en su artículo titulado “Numerical solution of differential equations using Haar wavelets”, en la página 129, define la familia wavelet Haar para $t \in [0, 1]$ como sigue:

$$h_i(t) = \begin{cases} 1, & \text{para } t \in \left[\frac{k}{m}, \frac{k + \frac{1}{2}}{m} \right] \\ -1, & \text{para } t \in \left[\frac{k + \frac{1}{2}}{m}, \frac{k + 1}{m} \right] \\ 0, & \text{en las otras partes de la recta.} \end{cases}$$

donde el entero $m = 2^j$, con $j = 0, 1, \dots, J$ indica el nivel de la wavelet; el entero $k = 0, 1, \dots, m - 1$ es el parámetro de traslación. El nivel maximal de resolución es J . El índice i de $h_i(t)$ se calcula conforme a la fórmula $i = m + k + 1$; en el caso del nivel minimal de resolución $j = 0$, se tiene que $m = 2^0 = 1$, y, $k = 1 - 1 = 0$, $i = 2$. Luego, $h_2(t) = \psi(t)$, es decir, es la función Haar (wavelet madre). El valor máximo de i es $i = 2^J + 2^J - 1 + 1 = 2 \cdot 2^J$.

Se definen los puntos de colocación $t_l = \frac{l - \frac{1}{2}}{2 \cdot 2^J}$, $l = 1, 2, \dots, 2 \cdot 2^J$; y se discretiza la función Haar $h_i(t)$; para de esta manera obtener la matriz de entrada $(h_i(t_l))$ de tamaño $2 \cdot 2^J \times 2 \cdot 2^J$.

Para resolución bajo 2 niveles, en la forma matricial, la **matriz Haar** es de tamaño $2 \cdot 2^2 \times 2 \cdot 2^2$, y está dada por

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}_{8 \times 8} \begin{matrix} \text{wavelet} \\ \text{padreHaar} \\ \text{madreHaar} \\ \text{hija} \\ \text{hija} \\ \text{nieta} \\ \text{nieta} \\ \text{nieta} \\ \text{nieta} \end{matrix}$$

Nótese que la familia wavelet Haar son funciones constantes a trozos en primer lugar; en segundo lugar, es la más simple wavelets ortonormal (recuerdese que no todas las wavelets son ortonormales). Tercero, tiene soporte compacto en el intervalo cerrado $[0, 1]$, pues es un conjunto cerrado y acotado de la recta real (con la métrica usual de la recta). Las características primera y tercera



hacen que la wavelet Haar no pueda ser aplicada directamente para solucionar ecuaciones diferenciales ordinarias pues las funciones constantes a trozos no son continuas, y así no pueden ser diferenciables en los puntos de discontinuidad.

Para superar este inconveniente los Ph.D. Chuan F. Chen de la universidad de Arizona, y C.H. Hsiao, en uno de sus artículos titulado “Haar wavelet method for solving Lumped and distributed-parameter systems”, proponen expandir el orden mayor de la derivada de la ecuación diferencial ordinaria dada, en series Haar. Las otras derivadas son obtenidas directamente por integración. Todo el sistema es discretizado por el método de colocación. Estos son procesos de segmentación, es decir, la ecuación diferencial ordinaria

$$\frac{d^2x}{dt^2} = F\left(t, x, \frac{dx}{dt}\right), \quad t \in [0, T] \quad (6.1)$$

es equivalente al sistema de ecuaciones diferenciales de primer orden

$$\begin{aligned} \frac{dx}{dt} &= y \\ \frac{dy}{dt} &= F(t, x, y) \end{aligned} \quad (6.2)$$

que puede ser solucionado.

Se divide el intervalo temporal en N segmentos; la longitud del n -ésimo segmento es denotada por $d^{(n)}$. Si $t^{(n)}$ es la coordenada del punto de la n -ésima división, entonces

$$\begin{aligned} t^{(n+1)} &= t^{(n)} + d^{(n)}, \quad n = 1, 2, \dots, N. \\ \text{con } t^{(1)} &= 0, \quad t^{(N+1)} = 1. \end{aligned}$$

Se considera el n -ésimo segmento, y se introduce el tiempo local

$$\tau = \frac{t - t^{(n)}}{d^{(n)}} \quad (6.3)$$

y se escoge $2M$ puntos de colocación

$$\tau_j = \frac{1}{2M} \left(j - \frac{1}{2} \right), \quad j = 1, 2, \dots, 2M. \quad (6.4)$$

donde $M = 2^J$, J nivel maximal de resolución.

El sistema 6.2 en tiempo local toma la forma

$$\begin{aligned} \frac{dx}{dt} &= d^{(n)}g \\ \frac{dy}{dt} &= d^{(n)}F(t^{(n)} + d^{(n)}\tau, x, y) \end{aligned} \quad (6.5)$$



En lo siguiente es conveniente considerar las variables x, y como vectores fila con las componentes $x_j = x(\tau_j)$, $y_j = y(\tau_j)$, $j = 1, 2, \dots, 2^{J+1}$.

De acuerdo con C.F. Chen, C.H. Hsiao en sus artículos “*Haar wavelet method for solving Lumped and distributed-parameter systems*”, y, “*Wavelet approach to optimising dynamic systems*” de Teoría de Control, la solución es solicitada en la forma

$$\begin{aligned} \dot{x} &= aH, & x &= aPH + x_n E \\ \dot{y} &= bH, & y &= bPH + y_n E \end{aligned} \quad (6.6)$$

Aquí a y b son vectores fila 2^{J+1} dimensionales, E es un vector unitario de la misma dimensión; los símbolos x_n , y_n representan los valores de x e y en la frontera $t = t_n$.

Sustituyendo la ecuación 6.6 en 6.5, se obtiene:

$$\begin{aligned} aH &= d^{(n)}(bPH + y_n E) \\ bH &= d^{(n)}F(t^{(n)} + d^{(n)}\tau, \quad aPH + x_n E, \quad bPH + y_n E) \end{aligned} \quad (6.7)$$

De estas ecuaciones matriciales los vectores a y b son calculados; después que x e y pueden ser evaluados de la ecuación 6.6.

Para proceder para el próximo segmento los valores x_{n+1} , y_{n+1} deben ser calculados de acuerdo a 6.6

$$\begin{aligned} x_{n+1} &= a(PH)_{\tau=1} + x_n E, \\ y_{n+1} &= b(PH)_{\tau=1} + y_n E. \end{aligned}$$

De la ecuación $(PH)_{il} = \int_0^{t_i} h_i(t)dt$ que define la matriz operacional de integración, que es una matriz cuadrada de tamaño $2^{J+1} \times 2^{J+1}$, se sigue que $(PH)_{\tau=1} = Y$, donde Y denota el vector de dimensión 2^{J+1} , $Y = [1, 0, 0, \dots, 0]$.

Así se tiene:

$$\begin{aligned} x_{n+1} &= a_1 + x_n E, \\ y_{n+1} &= b_1 + y_n E \end{aligned} \quad (6.8)$$

donde a_1 , y , b_1 son las primeras componentes de los vectores a y b .

Por ejemplo, considere la ecuación diferencial ordinaria lineal de segundo orden

$$\frac{d^2x}{dt^2} + \alpha \frac{dx}{dt} + \beta x = \varepsilon f(t) \quad (6.9)$$

donde α , β , ε son constantes y $f(t)$ es una función.

Llevando a cabo las mismas operaciones las cuales se necesitaron para derivación



de 6.5, la ecuación 6.9 adquiere la forma

$$\begin{aligned}\frac{dx}{dt} &= d^{(n)}y, \\ \frac{dy}{dt} &= -py - qx + s\varphi(t),\end{aligned}\tag{6.10}$$

donde $p = \alpha d^{(n)}$, $q = \beta d^{(n)}$, $s = \varepsilon d^{(n)}$, $\varphi(t) = f(t^{(n)} + d^{(n)}\tau)$.

Usando las ecuaciones 6.6 y sustituyendo estas ecuaciones en 6.10, y posteriormente multiplicando todos los términos con la matriz H^{-1} y tomando en cuenta la condición $Y = EH^{-1}$, la ecuación 6.10 puede ser formulada en la forma:

$$\begin{aligned}a - d^{(n)}bP &= d^{(n)}y_n Y \\ qaP + b(I + pP) &= -(py_n + qx_n)Y + s\varphi(\tau)H^{-1}\end{aligned}$$

donde I es la matriz idéntica de tamaño $2M \times 2M$, y, $p = \alpha d^{(n)}$.



Solución Numérica de Ecuaciones Diferenciales Ordinarias Usando Matrices Wavelet Haar

7.1. Método Wavelet Haar para Ecuaciones Diferenciales Ordinarias

En este capítulo, se aplicará el método Wavelet Haar para resolver algunas Ecuaciones Diferenciales Ordinarias de orden 2 y 4 con condiciones iniciales o de frontera. Se usarán para realizar los cálculos representaciones matriciales cuyo procedimiento consiste en tomar las sumas Haar para las diferenciales de alto orden de la derivada e integrar luego estas sumas.

Para solucionar una ecuación diferencial ordinaria lineal de n -ésimo orden, digamos

$$A_1 \frac{d^n x}{dt^n} + A_2 \frac{d^{n-1} x}{dt^{n-1}} + \cdots + A_n \frac{dx}{dt} + A_{n+1} x(t) = f(t),$$

donde $t \in [A, B]$ y con condiciones iniciales

$$x^{(n-1)}(A), x^{(n-2)}(A), \dots, x(A) \quad \text{conocidas,}$$

se sigue el trabajo de Ülo Lepik, en donde se intenta desarrollar hasta el nivel J de resolución. Para este nivel de resolución la matriz Haar es de tamaño $2 \cdot 2^J \times 2 \cdot 2^J$, resultado obtenido en una página anterior. Luego, se coloca $m = 2 \cdot 2^J$, y el intervalo $[A, B]$ será dividido en m subintervalos, cada uno de los cuales de longitud $\Delta x = \frac{B-A}{m}$. El procedimiento de solución básicamente consiste de 5 pasos como sigue:

(i) Se hace

$$\frac{d^n x}{dt^n} = \sum_{i=1}^m a_i h_i(t) \quad (7.1)$$

donde h es la matriz Haar y a_i son los coeficientes wavelet.

(ii) Se obtiene un v orden apropiado de $x(t)$ usando

$$\frac{d^v x}{dt^v} = \sum_{i=1}^m a_i P_{n-v,i}(t) + \sum_{\sigma=0}^{n-v-1} \frac{1}{\sigma!} (t-A)^\sigma x_0^{(v+\sigma)} \quad (7.2)$$

(iii) Se reemplaza $\frac{d^n x}{dt^n}$ y todo valor de $\frac{d^v x}{dt^v}$ en el problema.

(iv) Se calculan los coeficientes wavelet, a_i .

(v) Se obtiene la solución numérica para $x(t)$.

Nota. Obsérvese que 7.2 se obtiene integrando 7.1 $n - v$ veces. Luego, es necesario calcular las integrales:

$$P_{\alpha,i}(x) = \underbrace{\int_A^x \int_A^x \cdots \int_A^x}_{\alpha \text{-veces}} h_i(t) dt^\alpha = \frac{1}{(\alpha-1)!} \int_A^x (x-t)^{\alpha-1} h_i(t) dt$$

$$\alpha = n - v = 1, 2, \dots, n. \quad i = 1, 2, \dots, m.$$

Tales integrales se calculan en lo que sigue usando la definición de la i -ésima wavelet Haar. Esto es,

$$h_i(t) = \begin{cases} 1 & \text{para } t \in [\beta_1(i), \beta_2(i)], \\ -1 & \text{para } t \in [\beta_2(i), \beta_3(i)], \\ 0 & \text{en los otros casos.} \end{cases}$$

donde

$$\begin{aligned} \beta_1(i) &= A + 2k\mu\Delta x, & \beta_2(i) &= A + (2k+1)\mu\Delta x, \\ \beta_3(i) &= A + (2k+2)\mu\Delta x, & \mu &= \frac{2^J}{2^j} \end{aligned}$$

Los parámetros introducidos son $j = 0, 1, \dots, J$, y, $k = 0, 1, 2, \dots, (m-1)$.

En el caso $i = 1$, se tiene $\beta_1 = A$, $\beta_2 = \beta_3 = B$; y este caso corresponde a la función escala

$$h_1(t) = \begin{cases} 1 & \text{para } t \in [A, B], \\ 0 & \text{en los otros casos,} \end{cases}$$

por lo tanto,

$$\begin{aligned} P_{\alpha,1}(x) &= \frac{1}{(\alpha-1)!} \int_A^x (x-t)^{\alpha-1} \cdot 1 dt = \frac{1}{(\alpha-1)!} \frac{(-1)(x-t)^\alpha}{\alpha} \Big|_A^x \\ &= \frac{(x-A)^\alpha}{\alpha!} \end{aligned}$$



Para $i > 1$, se presentan cuatro casos:

$x < \beta_1(i)$, luego $A \leq t < x < \beta_1(i)$, $h_i(t) = 0$. En consecuencia,

$$P_{\alpha,i}(x) = \frac{1}{(\alpha-1)!} \int_A^x (x-t)^{\alpha-1} \cdot 0 dt = 0$$

$x \in [\beta_1(i), \beta_2(i)]$, luego $A \leq t \leq x \leq \beta_2(i)$, $h_i(t) = 1$. En consecuencia,

$$P_{\alpha,i}(x) = \frac{1}{(\alpha-1)!} \int_A^x (x-t)^{\alpha-1} \cdot 1 dt = \frac{1}{(\alpha-1)!} \int_A^x (x-t)^{\alpha-1} dt = \frac{(x-A)^\alpha}{\alpha!}$$

Ahora, $x \in [\beta_2(i), \beta_3(i)]$, luego $\beta_1(i) = A \leq t < x$; pero $\beta_2(i) \leq x \leq \beta_3(i)$ y así,

$$\int_A^x (x-t)^{\alpha-1} h_i(t) dt = \int_{\beta_1(i)}^{\beta_2(i)} (x-t)^{\alpha-1} \cdot 1 dt + \int_{\beta_2(i)}^x (x-t)^{\alpha-1} \cdot (-1) dt$$

En consecuencia,

$$\begin{aligned} P_{\alpha,i}(x) &= \frac{1}{(\alpha-1)!} \left[\int_{\beta_1(i)}^{\beta_2(i)} (x-t)^{\alpha-1} \cdot 1 dt + \int_{\beta_2(i)}^x (x-t)^{\alpha-1} \cdot (-1) dt \right] \\ &= \frac{1}{(\alpha-1)!} \left[-\frac{(x-t)^\alpha}{\alpha} \Big|_{\beta_1(i)}^{\beta_2(i)} + (-1)(-1) \frac{(x-t)^\alpha}{\alpha} \Big|_{\beta_2(i)}^x \right] \\ &= \frac{1}{(\alpha-1)!} \left(\frac{1}{\alpha} \right) [(-1)[(x-\beta_2(i))^\alpha - (x-\beta_1(i))^\alpha] + 0 - (x-\beta_2(i))^\alpha] \\ &= \frac{1}{\alpha!} [(x-\beta_1(i))^\alpha - (x-\beta_2(i))^\alpha - (x-\beta_2(i))^\alpha] \\ &= \frac{1}{\alpha!} [(x-\beta_1(i))^\alpha - 2(x-\beta_2(i))^\alpha] \end{aligned}$$

Por último, $x > \beta_3(i)$, luego $A = \beta_1(i) \leq t < x$; pero $x > \beta_3(i)$ por lo tanto,

$$\int_A^x (x-t)^{\alpha-1} h_i(t) dt = \int_{\beta_1(i)}^{\beta_2(i)} \boxed{} dt + \int_{\beta_2(i)}^{\beta_3(i)} \boxed{} dt + \int_{\beta_3(i)}^x \boxed{} dt$$

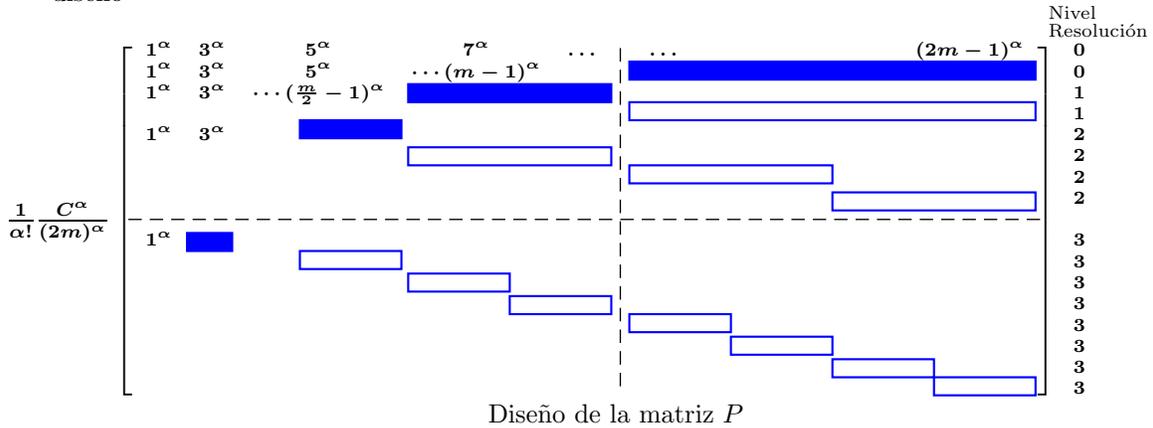
donde $\boxed{} \equiv (x-t)^{\alpha-1} h_i(t)$.



, es decir,

$$\begin{aligned}
 P_{\alpha,i}(x) &= \frac{1}{(\alpha-1)!} \int_A^x (x-t)^{\alpha-1} h_i(t) dt \\
 &= \frac{1}{(\alpha-1)!} \left[\int_{\beta_1(i)}^{\beta_2(i)} (x-t)^{\alpha-1} \cdot 1 dt + \int_{\beta_2(i)}^{\beta_3(i)} (x-t)^{\alpha-1} \cdot (-1) dt + \int_{\beta_3(i)}^x (x-t)^{\alpha-1} \cdot 0 dt \right] \\
 &= \frac{1}{(\alpha-1)!} \left[-\frac{(x-t)^\alpha}{\alpha} \Big|_{\beta_1(i)}^{\beta_2(i)} + (-1)(-1) \frac{(x-t)^\alpha}{\alpha} \Big|_{\beta_2(i)}^{\beta_3(i)} + 0 \right] \\
 &= \frac{1}{(\alpha-1)!} \cdot \frac{1}{\alpha} \{ -[(x-\beta_2(i))^\alpha - (x-\beta_1(i))^\alpha] + [(x-\beta_3(i))^\alpha - (x-\beta_2(i))^\alpha] \} \\
 &= \frac{1}{\alpha!} [(x-\beta_1(i))^\alpha - 2(x-\beta_2(i))^\alpha + (x-\beta_3(i))^\alpha].
 \end{aligned}$$

El segundo paso (ii) es la clave del procedimiento donde la matriz $P_{n-v,i}(t)$ será calculada. Si se intenta hacer el cálculo hasta el nivel J de resolución, se obtendrá la matriz $P_{n-v,i}(t)$ como en el modelo que se muestra en el siguiente diseño



Diseño de la matriz P

donde $\alpha = n - v$, $C = B - A$, $m = 2 \cdot 2^j$, j nivel de resolución

■ indica los elementos que son calculados mediante la expresión

$$\left(\frac{m}{2^L} + 2l - 1 \right)^\alpha - 2(2l - 1)^\alpha,$$

Siendo $L = 0, 1, 2, \dots, j$; $l = 1, 2, 3, \dots, \frac{m}{2(2^L)}$.

indica los elementos de la matriz P con el mismo valor en el mismo nivel.

El escalar $\frac{C^\alpha}{\alpha!(2m)^\alpha}$ es el factor común de todos los elementos de la matriz P que son calculados mediante el algoritmo:

$$\frac{C^\alpha}{\alpha!(2m)^\alpha} \left[\left(\frac{m}{2^L} + 2l - 1 \right)^\alpha - 2(2l - 1)^\alpha \right]$$



Algoritmo sugerido por Phang Chang, y, Phang Piau para calcular los elementos requeridos en la matriz P , en su artículo “Haar Wavelet Matrices Designation in Numerical Solution of Ordinary Differential Equations”. Nótese que el modelo del diseño de las matrices P es semejante al gráfico de la familia wavelet Haar mostrado o construido anteriormente (bajo dos niveles de resolución).

Observación. En los siguientes ejemplos de ilustración, todos los cálculos son realizados usando 3 niveles de la wavelet Haar. Esto es, $L = 0$, $L = 1$, $L = 2$, y, $L = 3$; puesto que $j = 3$.

Ejemplo 7.1.1 Solucionar la ecuación $y''(t) + y(t) = \text{sen}(t) + t \cos(t)$,

$$t \in [0, 1] \quad \text{con} \quad y(0) = y'(0) = 1$$

Solución Exacta: Usando Transformada de Laplace

$$\mathcal{L}[y''(t) + y(t)] = \mathcal{L}[\text{sen}(t) + t \cos(t)]$$

En virtud de la propiedad de linealidad, se obtiene

$$\begin{aligned} \mathcal{L}[y''(t)] + \mathcal{L}[y(t)] &= \mathcal{L}[\text{sen}(t)] + \mathcal{L}[t \cos(t)], \iff, \\ s^2 \mathcal{L}[y(t)] - sy(0) - y'(0) + \mathcal{L}[y(t)] &= \frac{1}{s^2 + 1} + \frac{s^2 - 1}{(s^2 + 1)^2}, \iff, \\ (s^2 + 1)\mathcal{L}[y(t)] &= \frac{1}{s^2 + 1} + \frac{s^2 - 1}{(s^2 + 1)^2} + s + 1, \iff, \\ \mathcal{L}[y(t)] &= \frac{1}{(s^2 + 1)^2} + \frac{s^2 - 1}{(s^2 + 1)^3} + \frac{s}{s^2 + 1} + \frac{1}{s^2 + 1} \\ \mathcal{L}[y(t)] &= \frac{s}{s^2 + 1} + \frac{s^2 + 1 + s^2 - 1 + (s^2 + 1)^2}{(s^2 + 1)^3} \\ \mathcal{L}[y(t)] &= \frac{s}{s^2 + 1} + \frac{s^2 + s^2 + 2s^2 + s^4 + 1}{(s^2 + 1)^3} \\ \mathcal{L}[y(t)] &= \frac{s}{s^2 + 1} + \frac{\frac{6}{4}s^2 + \frac{5}{4}s^4 - \frac{1}{4}s^4 + \frac{10}{4}s^2 + \frac{5}{4} + \frac{1}{4} - \frac{2}{4}}{(s^2 + 1)^3} \\ \mathcal{L}[y(t)] &= \frac{s}{s^2 + 1} + \frac{(\frac{6}{4}s^2 - \frac{2}{4}) + (\frac{5}{4}s^4 + \frac{10}{4}s^2 + \frac{5}{4}) + (-\frac{1}{4}s^4 + \frac{1}{4})}{(s^2 + 1)^3} \\ \mathcal{L}[y(t)] &= \frac{s}{s^2 + 1} + \frac{\frac{1}{4}(6s^2 - 2) + \frac{5}{4}(s^2 + 1)^2 - \frac{1}{4}(s^4 - 1)}{(s^2 + 1)^3} \\ \mathcal{L}[y(t)] &= \frac{s}{s^2 + 1} + \frac{1}{4} \left[\frac{6s^2 - 2}{(s^2 + 1)^3} \right] + \frac{5}{4} \left(\frac{1}{s^2 + 1} \right) - \frac{1}{4} \left[\frac{s^2 - 1}{(s^2 + 1)^2} \right] \end{aligned}$$

Ahora, aplicando transformada inversa de Laplace, se obtiene:

$$y(t) = \cos(t) + \frac{1}{4}t^2 \text{sen}(t) + \frac{5}{4} \text{sen}(t) - \frac{1}{4}t \cos(t).$$



en la matriz P ; y corresponden a los niveles de la wavelet Haar $L = 0$, $L = 1$, $L = 2$, y $L = 3$. Esto es, para $L = 0$, se obtiene $l = 1, 2, 3, 4, 5, 6, 7, \frac{16}{2(2^0)}$; y en consecuencia 8 elementos resaltados en la segunda fila y calculados con la expresión $(\frac{m}{2^L} + 2l - 1)^\alpha - 2(2l - 1)^\alpha$. Por ejemplo,

$$\text{para } l = 1, \text{ se obtiene } \left(\frac{16}{2^0} + 2 \cdot 1 - 1\right)^2 - 2(2 \cdot 1 - 1)^2 = 287$$

$$\text{para } l = 2, \text{ se obtiene } \left(\frac{16}{2^0} + 2 \cdot 2 - 1\right)^2 - 2(2 \cdot 2 - 1)^2 = 343$$

$$\text{para } l = 3, \text{ se obtiene } \left(\frac{16}{2^0} + 2 \cdot 3 - 1\right)^2 - 2(2 \cdot 3 - 1)^2 = 391$$

$$\text{para } l = 4, \text{ se obtiene } \left(\frac{16}{2^0} + 2 \cdot 4 - 1\right)^2 - 2(2 \cdot 4 - 1)^2 = 431$$

$$\text{para } l = 5, \text{ se obtiene } \left(\frac{16}{2^0} + 2 \cdot 5 - 1\right)^2 - 2(2 \cdot 5 - 1)^2 = 463$$

y así sucesivamente hasta llegar a $l = 8$.

Para $L = 1$, se obtiene $l = 1, 2, 3, 4$; donde el máximo de l se obtiene de $\frac{m}{2^{L+1}} = \frac{16}{2^2} = 4$. En consecuencia, aparece 4 elementos resaltados en la tercera fila y calculados con la expresión $(\frac{m}{2^L} + 2l - 1)^\alpha - 2(2l - 1)^\alpha$. Estos son:

$$\text{para } l = 1, \left(\frac{16}{2} + 2 - 1\right)^2 - 2(2 - 1)^2 = 79$$

$$\text{para } l = 2, \left(\frac{16}{2} + 3\right)^2 - 2 \cdot 3^2 = 103$$

$$\text{para } l = 3, \left(\frac{16}{2} + 5\right)^2 - 2 \cdot 5^2 = 119$$

$$\text{para } l = 4, \left(\frac{16}{2} + 7\right)^2 - 2 \cdot 7^2 = 127$$

Para $L = 2$, se obtiene $l = 1, 2$. En consecuencia, 2 elementos aparecerán resaltados en la quinta fila y calculados con la misma expresión. Esto es,

$$\text{para } l = 1, \left(\frac{16}{4} + 1\right)^2 - 2(2 - 1)^2 = 23$$

$$\text{para } l = 2, \left(\frac{16}{4} + 3\right)^2 - 2(4 - 1)^2 = 31$$

Para $L = 3$, se obtiene $l = 1$. En consecuencia, 1 elemento aparecerá resaltado en la novena fila. Esto es, $(\frac{16}{8} + 1)^2 - 2(2 - 1)^2 = 7$.

(v) Se obtiene la solución numérica para $y(t)$, dada mediante la siguiente tabla:



t (/32)	Solución Numérica	Solución Exacta	Error Absoluto
1	1.030777	1.030767	0.000010
3	1.089527	1.089496	0.000031
5	1.144880	1.144700	0.000180
7	1.196844	1.196643	0.000201
9	1.247001	1.245594	0.001407
11	1.293240	1.291819	0.001421
13	1.337136	1.335577	0.001559
15	1.378678	1.377118	0.001560
17	1.427640	1.416676	0.010964
19	1.465381	1.454467	0.010914
21	1.501598	1.490681	0.010917
23	1.536265	1.525485	0.010780
25	1.570336	1.559012	0.011324
27	1.602462	1.591364	0.011098
29	1.633465	1.622605	0.010860
31	1.663308	1.652763	0.010545

Ejemplo 7.1.2 Solucionar la ecuación $y^{(4)}(t) + ty(t) = 16 \cos(2t) + t \cos(2t)$, $t \in [0, 1]$ con $y(0) = 1$, $y'(0) = 0$, $y''(0) = -4$, $y'''(0) = 0$.

Solución Exacta: Consiste de la suma de la solución general de la homogénea $y^{(4)}(t) + ty(t) = 0$; que es la función $y(t) = 0$ para todo $t \in [0, 1]$, y una solución particular de la no -homogénea, $y_p(t) = \cos(2t)$ para todo $t \in [0, 1]$. En consecuencia, la solución exacta es $Y(t) = \cos(2t)$.

Solución Numérica: Usando el Método Wavelet Haar

$$(i) \quad y^4(t) = \sum_{i=1}^m a_i h_i(t)$$

$$\begin{aligned}
 (ii) \quad y(t) &= \sum_{i=1}^m a_i P_{4,i}(t) + \sum_{\sigma=0}^{4-1} \frac{1}{\sigma!} (t-0)^\sigma y_0^{(0+\sigma)} \\
 &= \sum_{i=1}^m a_i P_{4,i}(t) + 1 \cdot 1 \cdot 1 + t y_0' + \frac{1}{2} t^2 y_0'' + \frac{1}{6} t^3 y_0''' \\
 &= \sum_{i=1}^m a_i P_{4,i}(t) + 1 + t \cdot 0 + \frac{1}{2} t^2 \cdot (-4) + \frac{1}{6} t^3 \cdot 0 \\
 &= \sum_{i=1}^m a_i P_{4,i}(t) + 1 - 2t^2
 \end{aligned}$$

puesto que $y_0 = y(0) = 1$, $y_0' = y'(0) = 0$, $y_0'' = y''(0) = -4$, $y_0''' = y'''(0) = 0$.

(iii) Sustituyendo (i) y (ii) en la ecuación $y^{(4)}(t) + ty(t) = 16 \cos(2t) +$



$t \cos(2t)$, se obtiene:

$$\sum_{i=1}^m a_i h_i(t) + t \sum_{i=1}^m a_i P_{4,i}(t) + 1 - 2t^2 = 16 \cos(2t) + t \cos(2t),$$
$$\sum_{i=1}^m a_i [h_i(t) + t P_{4,i}(t)] + 1 - 2t^2 = 16 \cos(2t) + t \cos(2t)$$

(iv) Para calcular los coeficientes wavelet a_i es necesario la matriz P . Para conocerla, se sabe que $\alpha = n - v = 4 - 0 = 4$; $C = B - A = 1 - 0 = 1$, $m = 2 \cdot 2^3 = 16$. Luego la matriz P está dada por las entradas:





$$\frac{1}{25.165.824}$$

1	3 ⁴	5 ⁴	7 ⁴	9 ⁴	11 ⁴	13 ⁴	15 ⁴	17 ⁴	19 ⁴	21 ⁴	23 ⁴	25 ⁴	27 ⁴	29 ⁴	31 ⁴
1	3 ⁴	5 ⁴	7 ⁴	9 ⁴	11 ⁴	13 ⁴	15 ⁴	83.519	130.159	193.231	275.039	377.503	502.159	650.159	822.271
1	3 ⁴	5 ⁴	7 ⁴	6.559	14.479	27.311	45.823								
								1	3 ⁴	5 ⁴	7 ⁴	6.559	14.479	27.311	45.823
1	3 ⁴	623	2.239												
				1	3 ⁴	623	2.239								
								1	3 ⁴	623	2.239				
												1	3 ⁴	623	2.239
1	79														
		1	79												
				1	79										
						1	79								
								1	79						
										1	79				
												1	79		
														1	79

7.1. MÉTODO WAVELET HAAR PARA ECUACIONES DIFERENCIALES ORDINARIAS⁷⁹

donde,

el factor común $\frac{1}{25.165.824}$ se calculó con la expresión $\frac{C^\alpha}{\alpha!(2m)^\alpha}$. Como $j = 3$, los elementos calculados son los resaltados en la matriz P ; que corresponden a los niveles de la Wavelet Haar $L = 0$, $L = 1$, $L = 2$, y, $L = 3$. Esto es,

Para $L = 0$, l toma los valores, $l = 1, 2, 3, 4, 5, 6, 7, 8$. En consecuencia, en la segunda fila de la matriz P aparece 8 elementos resaltados y calculados mediante la expresión $(\frac{m}{2^L} + 2l - 1)^\alpha - 2(2l - 1)^\alpha$.

Para $L = 1$, l toma los valores $l = 1, 2, 3, 4$. En consecuencia, aparece 4 elementos resaltados en la tercera fila, y calculados con la misma expresión.

Para $L = 2$, l toma los valores $l = 1, 2$. En consecuencia, aparece 2 elementos resaltados en la quinta fila, calculados con la expresión ya conocida.

Para $L = 3$, $L = 1$. Aparece el elemento 79 resaltado en la novena fila.

(v) Se obtiene la solución numérica para $x(t)$, mediante la siguiente tabla:

t (/32)	Solución Numérica	Solución Exacta	Error Absoluto
1	0.062459	0.062459	0.000000
3	0.186405	0.186403	0.000002
5	0.307463	0.307439	0.000024
7	0.423743	0.423676	0.000067
9	0.533859	0.533303	0.000556
11	0.635637	0.634607	0.001030
13	0.727687	0.726009	0.001678
15	0.808581	0.806081	0.002500
17	0.887019	0.873575	0.013444
19	0.946546	0.927437	0.019109
21	0.992728	0.966827	0.025901
23	1.024951	0.991129	0.033822
25	1.042978	0.999966	0.043012
27	1.046518	0.993198	0.053320
29	1.035738	0.970932	0.064806
31	1.010991	0.933514	0.077477



8

Conclusiones

1. Las funciones bases de wavelet son versiones dilatadas y trasladadas de una función de soporte compacto (o finito), llamada wavelet madre, denotada comúnmente por ψ .
2. Un conjunto de versiones dilatadas y trasladadas, tanto de una función escala ϕ como de una función wavelet ψ , ambas pertenecientes al espacio L^2 , son capaces de aproximar cualquier solución de una ecuación diferencial ordinaria de orden superior.
3. La función escala ϕ es la encargada de analizar el comportamiento general de la solución, mientras que la función wavelet ψ se encarga de analizar el comportamiento del detalle de la solución ecuación diferencial ordinaria de orden superior.
4. La ortonormalidad de las bases empleadas, es una propiedad esencial, debido a que convierte el cálculo de los coeficientes en una tarea rápida y sencilla.
5. La obtención de los coeficientes escala como de los coeficientes wavelet se realiza mediante el producto interno entre las versiones dilatadas y trasladadas de la función escala y wavelet con la solución a calcular. De esta forma el coeficiente obtenido representa el grado de correlación que existe entre la función escala y wavelet con la solución en un intervalo finito en el espacio del tiempo.

Bibliografía

- [1] C.H. Hsiao. *Haar Wavelet Approach to Linear Stiff Systems*. Department of Computer and Information Sciences, Aletheia University, 32 Chen-Li Street, Tamsui, Taipei 25103, Taiwan, ROC., Mathematics and Computers in Simulation 64 (2004) 561-567.
- [2] Chen, C.F. and Hsiao C.H., *Haar Wavelet Method for Solving Lumped and Distributed -parameter systems*. IEE Proc. Control Theory Appl. Vol. 144, 1997, pp. 87-94.
- [3] Christian Gargour, Marcel Gabrea, Venkatanarayana Ramachandran and Jean Marc Lina, *A Short Introduction to Wavelets and Their Applications*. University Montreal, Quebec, Canadá.
- [4] David F. Walnut, *An Introduction to Wavelet Analysis*. Advisory Board, 2002 Birkhäuser Boston.
- [5] Doris Hinestroza G., *Introduction to Wavelets and Applications*. Departamento de Matemáticas, Facultad Ciencias Exactas y Naturales, Universidad del Valle, 2009.
- [6] Ingrid Daubechies, *10 lecturas por: Ingrid Daubechies*. Rutgers University.
- [7] Ingrid Daubechies, Member IEEE, *The Wavelet Transform, Time-Frequency Localization and Signal Analysis*. Vol. 36, No 5, September 1990.
- [8] Ingrid Daubechies, Member IEEE, *Where Do Wavelets Come From?-A personal point of view* .Department of Mathematics and Program in Applied and Computational Mathematics, Princeton University, princeton usa, Volúmen 84 proceedings of the IEEE, April 1996.

-
- [9] Michael W. Frazier, *An Introduction to Wavelets Through Linear Algebra*. Springer-Verlag, New York, Inc, 1999.
- [10] Noelia Rillo Pastor, *Introducción a la Teoría de Wavelets*. Departamento de Matemática Aplicada y Análisis. Universidad de Barcelona, 2006.
- [11] Oliver Rioul and Martin Vetterli, *Wavelets and Signal Processing*. IEEE SP Magazine, October 1991.
- [12] Paulo Cupertino de Lima, *Wavelets: Uma Introdução*. Matemática Universitaria, 2002.
- [13] Phang Chang, Phang Piau, *Haar Wavelet Matrices Designation in Numerical Solution of Ordinary Differential Equations*. International Journal of Applied Mathematics, 2008.
- [14] Polania Quiza, Luis Arturo. *Un Enfoque Cualitativo a Las EDO'S y Teoría Wavelets*. Facultad Ciencias Exactas y Naturales, Universidad Surcolombiana 2.011.
- [15] Ü. Lepik. *Numerical Solution of Differential Equations Using Haar Wavelets*. Institute of Applied Mathematics, University of Tartu, Lüvi2, 50409 Tartu Estonia, *Mathematics and Computers in Simulation* 68 (2005) 127-143.

