UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

CALCULO OPERACIONAL EN ESPACIOS DE HILBERT Y ALGUNAS DE SUS APLICACIONES

T E S I S

QUE PARA OBTENER EL TITULO DE

M A T E M A T I C O

P R E S E N T A

HUMBERTO MADRID DE LA VEGA

MEXICO, D. F. 1 9 7 3

A MI MADRE

A MI HERMANO

I N D I C E

PROLOGO

CAPITULO 0. INTRODUCCION 1

I.1 Motivación 1

I.2 La formulación del problema 5

CAPITULO 1. CALCULO OPERACIONAL EN ESPACIOS VECTORIALES DE DIMENSION FINITA 6

1.0 Introducción 6

1.1 Un caso particular del ejemplo (D) 7

1.2 Definición de Q(A) cuando Q(x) es racional 8

1.3 El caso en que f(z) es analítica 14

1.4 Solución del ejemplo A 17

CAPITULO 2. OPERADORES EN ESPACIOS DE HILBERT 18

2.0 Introducción 18

2.1 Motivación 18

2.1.1 Un núcleo separable determina un operador cuyo rango es de dimensión finita 18

2.1.2 Los espacios L_2 19

2.1.3 Todo elemento de $L_2(\Omega)$ determina un operador integral de L_2 (0, 1) en sí mismo 22

2.1.4 Todo operador integral en $L_2(0, 1)$ puede ser aproximado por operadores con núcleo separable 24

2.2 Espacios de Hilbert 25

2.3 Algunas convenciones sobre operadores lineales 28

2.4 Operadores acotados 29

2.5 Operadores de dimensión finita y operadores compactos 31

2.5.1 Un operador de dimensión finita no acotado 31

2.5.2 Definiciones de F(A) y de su cerradura K(H) 32

2.5.3 Una caracterización de K(H) 33

2.5.4 Otra caracterización de K(H) 35

2.5.5 Operadores compactos 35

2.6 Relaciones entre operadores acotados y operadores compactos **35**

2.7 Un operador no acotado **36**

2.8 Operadores cerrados **38**

2.9 Operadores autoadjuntos **40**

2.9.1 El adjunto de un operador integral 40

2.9.2 El adjunto de un operador 40

2.9.3 Operadores autoadjuntos 42

CAPITULO 3. CALCULO OPERACIONAL PARA FUNCIONES RACIONALES 46

3.0 Introducción **46**

3.1 Definiciones y convenciones **47**

3.2 Definición de Q(A) cuando A es un operador de un espacio de dimensión finita **47**

3.3 Definición de Q(A) cuando A es un operador compacto **48**

3.4 Definición de Q(A) cuando A es un operador acotado **50**

3.4.1 Un operador acotado con espectro no numerable 50

3.4.2 El espectro de un operador acotado 51

3.5 Una observación importante **54**

3.6 El espectro de un operador autoadjunto **54**

3.7 El espectro de un operador cerrado **55**

3.7.1 Un operador cuyo espectro no es acotado 56

3.7.2 El espectro de un operador cerrado es cerrado 56

CAPITULO 4. CALCULO OPERACIONAL PARA OPERADORES ACOTADOS. CASO GENERAL 57

4.0 Introducción **57**

4.1 Funciones analíticas con valores en L(H) **57**

4.1.1 Introducción 57

4.1.2 Definiciones y propiedades básicas 59

4.2 Definición de f(A)	61
4.2.1 Una definición restringida	61
4.2.2 Hacia una definición general	62
4.2.3 Definición de f(A)	63
4.3 Cálculo operacional	65
4.4 Proyecciones	70
4.5 Descomposición espectral para operadores en espacios de dimensión finita	73
4.5.1 El caso general	73
4.5.2 El caso en que A tiene n valores propios distintos	73
4.5.3 Representación espectral para operadores autoadjuntos	75
4.6 Descomposición espectral para operadores compactos	77
4.6.1 El caso general	77
4.6.2 Descomposición espectral para operadores compactos autoadjuntos	78
CAPITULO 5. TEOREMA ESPECTRAL PARA OPERADORES AUTOADJUNTOS ACOTADOS	81
5.0 Introducción	81
5.1 Algunos resultados previos	84
5.1.1 Proyecciones ortogonales	74
5.1.2 Una relación de orden parcial para operadores autoadjuntos	86
5.2 Teorema espectral para operadores acotados autoadjuntos	88
5.2.1 Introducción	88
5.2.2 Construcción de la familia E(λ)	90
5.2.3 El teorema espectral	97
5.2.4 Consecuencias del teorema espectral	102
CAPITULO 6. CALCULO OPERACIONAL PARA OPERADORES CERRADOS Y TEOREMA ESPECTRAL PARA OPERADORES AUTOADJUNTOS NO ACOTADOS	105
6.0 Introducción	105
6.1 Cálculo operacional para operadores cerrados	105

6.2 Teorema espectral para operadores autoadjuntos no acotados 110

CAPITULO 7. APLICACIONES 113

7.0 Introducción 113

7.1 Operadores con espectro discreto 114

7.1.1 Definición y ejemplos 114

7.1.2 Los operadores con espectro discreto están determinados por sus valores propios y vectores propios 115

7.1.3 El operador $A = -\ d^2/dx^2$ es positivo definido 117

7.2 Aplicación a los ejemplos del capítulo 0 118

7.2.1 Una observación importante 118

7.2.2 La ecuación de calor 118

7.2.3 La cuerda vibrante 119

APENDICE A. INTEGRAL DE RIEMANN–STIELTJES PARA FUNCIONES CON VALORES DE L(H) 121

A.0 Introducción 121

A.1 La integral cuando el integrando toma valores en L(H) 122

A.2 La integral cuando la función integradora toma valores en L(H) 125

APENDICE B. FUNCIONES ANALITICAS CON DOMINIO EN C Y VALORES EN L(H) 127

BIBLIOGRAFIA 133

PROLOGO

La formación del matemático es frecuentemente tan abstracta que le impide el contacto con la gente que trabaja en otras ramas de la ciencia y que necesita la colaboración de los matemáticos. Un reflejo de este hecho es que las tesis profesionales, generalmente parten de un nivel bastante abstracto y tienen como finalidad demostrar uno o más hechos aún más abstractos.

Creemos que este aislamiento debe romperse. Es necesario tratar de ligarse con otros terrenos de la ciencia. Al mismo tiempo, hay que tratar de partir de problemas concretos y a partir de estos, desarrollar la teoría de forma tal, que podamos aplicarla a dichos problemas concretos. Al respecto, estamos de acuerdo con Richard Courant cuando afirma que "Las teorías generales surgen de consideraciones sobre lo concreto y ellas carecen de sentido si no sirven para aclarar y ordenar precisamente lo concreto. La interacción entre generalidad y particularidad, deducción y construcción, lógica e imaginación es la esencia profunda de las matemáticas vivas". Creemos también que lo concreto debe estar en la medida de lo posible, relacionando con la problemática local, con los problemas que se plantean en la práctica en este país.

Este trabajo intenta recorrer este camino. Sin embargo, la misma formación del autor ha limitado el mismo trabajo e impedido lograr los objetivos deseados.

De acuerdo con esta concepción, en este trabajo, se pueden distinguir dos estructuras. La primera de ellas es la estructura general del trabajo. Se ha intentado partir de problemas concretos y llegar a la teoría de una manera natural, desarrollar esta y finalmente regresar a los problemas concretos. En el capítulo 0 se plantean, en términos de transformaciones lineales, ciertos problemas relacionados con ecuaciones diferenciales e integrales, en particular con ecuaciones diferenciales en la Física. Estos problemas motivan el estudio del Cálculo Operacional. En los siguientes capítulos se desarrolla la teoría para finalmente volver en el último capítulo a los problemas del capítulo 0. Ciertamente los problemas planteados en dicho capítulo no son tan concretos como se hubiera deseado, pero al menos están relacionados con problemas concretos.

La segunda estructura es la del desarrollo de la teoría. Aquí también se ha procurado que los resultados teóricos aparezcan de una manera natural y que los temas tratados estén relacionados entre sí. Por ejemplo, el Cálculo Operacional es tratado primero en el caso finito y después en el caso general. Se ha tratado de ir de lo particular a lo general para luego regresar a casos particulares.

En el último capítulo se menciona que las soluciones de las ecuaciones diferenciales obtenidas por medio del Cálculo Operacional coinciden con las que se obtienen por medio del método de separación de variables. Esta coincidencia no es accidental. De hecho el método de separación de variables puede ser justificado y generalizado con base en las técnicas desarrolladas en el presente trabajo. Lamentablemente, debido a una serie de limitaciones de distinta índole, no fue posible desarrollar dicho tema en este trabajo. Sin embargo, podemos citar que B. Friedman [4] indica el camino a seguir para su desarrollo.

Queremos agradecer al Dr. Pablo Barrera su valiosa y decisiva ayuda prestada a todos los niveles para la elaboración de este trabajo.

Finalmente, se agradece al CIMASS* la beca otorgada al autor para el desarrollo del presente trabajo.

* Centro de Investigación en Matemáticas Aplicadas, Sistemas y Servicios

CAPITULO 0. INTRODUCCION

I.1 Motivación

Para motivar el presente trabajo veremos algunos ejemplos que plantean un problema común.

Ejemplo A. Una ecuación diferencial matricial

El primer problema es encontrar una función $\overline{x}(t) : [a, b] \to R^n$, tal que

$$\dot{\overline{x}} = A\overline{x} \qquad (I.1)$$

con $\overline{x}(o) = \overline{x}_o$ y donde A es una matriz con elementos constantes.

Para el caso escalar, se tendrá la ecuación

$$\dot{x} = a\,x$$

donde $x(o) = x_o$ y a es una constante real. En este caso la solución viene dada por

$$x(t) = x_o \; e^{at} .$$

Lo anterior sugiere la posibilidad de que la solución del caso general pueda expresarse de forma parecida. Esto es, de la forma

$$\bar{x}(t) = e^{tA} \; \bar{x}_o \qquad (I.2)$$

donde e^{tA} debe ser un operador. Aquí el problema es ver si es posible darle un sentido a este operador de forma tal que (I.2) nos dé efectivamente la solución de la ec (I.1).

Ejemplo B. La ecuación de calor

Consideremos una barra cilíndrica de material homogéneo cuyos extremos están colocados en $x = 0$ y $x = 1$ (fig 1). Supongamos que el diámetro de la barra es lo suficientemente pequeño para que podamos considerar que la temperatura en cada sección transversal dependa sólo de x; además supongamos que la temperatura inicial está dada por una función $f(x)$, que la barra está colocada en un medio ambiente de temperatura $\theta = 0$ y que a través de los extremos no hay flujo de calor.

La temperatura de la barra $u(x, t)$ satisface la ecuación

$$\frac{\partial u}{\partial t} = k \; \frac{\partial^2 u}{\partial x^2} , \qquad (I.3)$$

con condiciones a la fontera

$$\frac{\partial u}{\partial x}(o, t) = \frac{\partial u}{\partial x}(1, t) = 0 \;\; y \;\; u(x, o) = f(x).$$

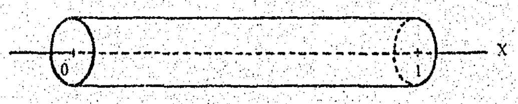


Fig 1.

En este ejemplo tenemos dos operadores: $\frac{\partial}{\partial t}$ y $k \frac{\partial^2}{\partial x^2}$. Aquí el problema es ver si es posible considerar a uno de estos operadores como constante, resolver la ecuación para el otro operador, e interpretando adecuadamente el resultado, obtener la solución del problema original. Por ejemplo, si consideramos al operador $A = k \frac{\partial^2}{\partial x^2}$ como una constante, resolviendo la ecuación

$$u' = Au$$

con $u(x, o) = f(x)$, obtenemos

$$u(x, t) = e^{tA} f(x) \qquad (I.4)$$

El problema es ver si se le puede dar un sentido a e^{tA} de forma tal que (I.4) sea la solución de (I.3).

Ejemplo C. La cuerda vibrante

La ecuación

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u,$$

donde Δ es el Laplaciano, representa en general la ecuación de propagación de ondas.

Como caso particular, consideremos una cuerda fija por sus extremos a los puntos $x = 0$ y $x = 1$. La cuerda es llevada a una posición dada por una función $f(x)$, y soltada para que empieze a vibrar (fig 2). La ecuación del movimiento de la cuerda viene dada por

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \qquad (I.5)$$

con $u(0, t) = u(l, t) = 0$, $u(x, 0) = f(x)$

y $\frac{\partial u}{\partial t}(x, 0) = 0$

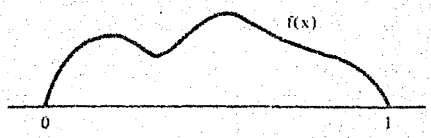


Fig 2

Igual que en el ejemplo anterior, tenemos dos operadores; de nuevo nos preguntamos si será posible encontrar la solución considerando un operador, digamos $A = c^2 \frac{\partial^2}{\partial x^2}$, como constante y entonces resolver la ecuación

$$u'' = Au$$

con $u(x, 0) = f(x)$. La solución a esta ecuación vendría dada por

$$u(x, t) = \cos(t\sqrt{-A})\, f(x). \qquad (I.6)$$

El problema es entonces ver si se le puede dar un sentido al operador $\cos(t\sqrt{-A})$ de forma tal que (I.6) sea la solución de (I.5).

Ejemplo D. Una ecuación integral

Dada una función $f(x)$, consideremos el problema de encontrar $u(x)$ satisfaciendo la ecuación integral

$$u(x) = f(x) + \lambda \int_0^1 K(x, t)\, u(t)\, dt. \qquad (I.7)$$

Si denotamos por A al operador $Af = g$ donde

$$g(x) = \int_0^1 K(x, t)\, f(t)\, dt,$$

la ecuación integral (I.7) puede escribirse como

$$u = f + \lambda Au.$$

o bien como

$$(I - \lambda A)\, u = f.$$

Esto sugiere que la solución de (I.7) venga dada por

$$u = (I - \lambda A)^{-1} f \qquad (I.8)$$

En este ejemplo, el problema es ver si se le puede dar sentido apropiado a $(I - \lambda A)^{-1}$ para que (I.8) sea la solución de (I.7).

I.2 La formulación del problema

Los ejemplos anteriores plantean el problema de definir f(A) para alguna función f(x), donde A es una transformación lineal.

En el ejemplo (A), la función es $f(x) = e^{tx}$ y A es una matriz. En (B) la función es la misma, pero A es un operador diferencial. En los ejemplos (C) y (D), las funciones son $f(x) = \cos t\sqrt{-x}$ y $f(x) = (1 - \lambda x)^{-1}$, respectivamente.

Es importante hace notar que las funciones e^{tx} y $\cos t\sqrt{-x}$, para t fija, son analíticas y que la función $f(x) = (1 - \lambda x)^{-1}$ es una función racional.

El objetivo de este trabajo es estudiar algunas clases de operadores y de funciones analíticas para las cuales se puede definir f(A). Esto es llamado Cálculo Operacional.

CAPITULO 1. CALCULO OPERACIONAL EN ESPACIOS VECTORIALES DE DIMENSION FINITA

1.0 Introducción

En este capítulo veremos como, dada una matriz cuadrada A, podemos encontrar una clase de funciones analíticas para las cuales se puede definir f(A). Así mismo veremos las propiedades de esta expresión.

Esto hará posible darle un sentido a e^{tA} de tal manera que efectivamente, en el ejemplo (A) de la introducción, la expresión (I.2) sea solución de la ecuación (I.1).

También es posible aplicar los resultados que se obtendrán en este capítulo a un caso particular del ejemplo (D) de la introducción, como haremos ver a continuación.

1.1 Un caso particular del ejemplo (D)

En la ecuación $(I - \lambda A)\, u = f$, el operador $Af = g$, donde

$$g(x) = \int_0^1 K(x, t)\, f(t)\, dt,$$

está determinado por la función $K(x,t)$ llamada el núcleo de A. Supongamos que el núcleo $K(x, t)$ es separable (o degenerado), es decir, que $K(x, t)$ es de la forma

$$K(x, t) = \sum_{i=1}^{n} a_i(x)\, b_i(t)\,.$$

Entonces

$$u(x) = f(x) + \lambda \sum_{i=1}^{n} a_i(x) \int_0^1 b_i(t)\, u(t)\, dt.$$

Para determinar $u(x)$, sólo nos hace falta conocer las

$$c_i = \int_0^1 b_i(t)\, u(t)\, dt, \qquad i = 1, \ldots, n.$$

Sustituyendo en el integrando de c_i el valor de $u(x)$ se obtiene

$$c_i = f_i + \lambda \sum_{j=1}^{n} K_{ij}\, C_j\,, \quad i = 1, \ldots, n,$$

siendo

$$f_i = \int_0^1 b_i(t)\, f(t)\, dt\ ; \quad K_{ij} = \int_0^1 b_i(t)\, a_j(t)\, dt$$

En notación matricial esto se puede escribir como

$$\overline{C} = \overline{f} + \lambda\, K\, \overline{C}$$

o bien como

$$(I - \lambda K)\, C = f$$

donde los vectores C y f tienen como componentes a las C_i y f_i, respectivamente, y $K = (K_{ij})$.

Entonces en este caso el problema se reduce a encontrar $(I - \lambda K)^{-1}$. Es decir, a encontrar f(A), donde $f(x) = (1 - \lambda x)^{-1}$ es una función racional.

A continuación veremos cuándo es posible definir Q(A) cuando Q(x) es una función racional.

1.2 Definición de Q(A) cuando Q(x) es racional

En el caso particular de un polinomio

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n ,$$

donde $a_o, a_1, \ldots, a_n$ son números reales, y A es una matriz con elementos reales, podemos definir un dificultad p(A) como

$$p(A) = a_o I + a_1 A + \ldots + a_n A^n .$$

Si $Q(x) = p(x)/q(x)$, con p(x) y q(x) polinomios, es natural definir $Q(A) = p(A)\, q(A)^{-1}$, siempre y cuando $q(A)^{-1}$ exista. Antes de continuar, cabe hacer una aclaración. También puede parecer natural definir Q(A) como $Q(A) = q(A)^{-1} p(A)$, veremos que es lo mismo que $p(A)\, q(A)^{-1}$. En efecto como $p(A)q(A) = q(A)\, p(A)$ – ya que p(x) y q(x) son polinomios – tenemos que $p(A) = q(A)\, p(A)\, q(A)^{-1}$ y de aquí que $q(A)^{-1}\, p(A) = p(A)\, q(A)^{-1}$.

Vamos a considerar a A como una matriz n x n con elementos reales. Además consideraremos a A indistintamente como matriz y como transformación lineal de R^n en sí mismo.

Ahora consideremos el caso particular en que $Q(x) = 1/x$. Podemos definir Q(A) cuando A es invertible. Para el desarrollo posterior de la teoría, es conveniente caracterizar cuándo A es invertible desde un cierto punto de vista.

Sea pues, $A : R^n \to R^n$ lineal. Puesto que las componentes de $A\bar{x}$ dependen linealmente de las componentes de $\bar{x}$, entonces A es continua en R^n. El conjunto $S = \left\{ \bar{x} \in R^n \mid |\bar{x}| = 1 \right\}$ es cerrado y acotado en R^n y por consiguiente la función $|A\bar{x}|$ alcanza su máximo y su mínimo en S.

Definamos la norma de A como

$$\|A\| = \max_{\bar{x} \in S} |A\bar{x}| .$$

Por consiguiente $|A\bar{x}| \leqslant \|A\| \, |\bar{x}|$ para toda $\bar{x} \in R^n$.

Por otro lado sea $m = \min_{\bar{x} \in S} |A\bar{x}|$, entonces $m \geqslant 0$ y $|A\bar{x}| \geqslant m \, |\bar{x}|$ para cualquier $x \in R^n$.

Si $m > 0$, entonces A es inyectiva y por lo tanto invertible y A^{-1} es lineal. Recíprocamente si A^{-1} existe, entonces $\|A^{-1}\| > 0$ y $|A^{-1}\bar{y}| \leqslant \|A^{-1}\| \, |\bar{y}|$ para toda $\bar{y} \in R^n$. Pero $\bar{y} = A\bar{x}$ y por lo tanto

$$|\bar{x}| \leqslant \|A^{-1}\| \, |A\bar{x}|$$

Entonces $m \geqslant 1/\|A^{-1}\| > 0$. Esto lo podemos resumir en el siguiente teorema.

Teorema 1.1. A es invertible si y solo si

$$m = \min |A\bar{x}| > 0, \quad \bar{x} \in S .$$

Hemos así encontrado una condición necesaria y suficiente para definir Q(A) cuando $Q(x) = 1/x$.

A continuación estudiaremos el caso en el que $Q(x) = 1/p(x)$, $p(x)$ un polinomio de primer grado que por comodidad lo supondremos de la forma $p(x) = a_0 + x$. Entonces $p(A) = a_0 I + A$.

Por el teorema 1.1 sabemos que p(A) es invertible si y solo si $m = \min |p(A)\,\bar{x}| > 0$, donde $\bar{x} \in S$. Pero esto es equivalente a que $a_0 \bar{x} + A\bar{x} \neq 0$ para toda $\bar{x} \in S$.

Es de esperarse que existan números a_0 satisfaciendo la propiedad anterior. El preguntarse la existencia de estos números es equivalente a preguntarse si existen números λ tales que $A\bar{x} = \lambda\bar{x}$ para alguna $\bar{x} \in S$.

Sean $\bar{x}_1, \bar{x}_2 \in S$ tales que $A\bar{x}_1 = \lambda_1\bar{x}_1$ y $A\bar{x}_2 = \lambda_2\bar{x}_2$. Si $\lambda_1 \neq \lambda_2$ entonces $\bar{x}_1$ y $\bar{x}_2$ son linealmente independientes, ya que si

$$\alpha\bar{x}_1 + \beta\bar{x}_2 = 0,$$

aplicando A a esta ecuación obtenemos

$$\alpha\lambda_1\bar{x}_1 + \beta\lambda_2\bar{x}_2 = 0.$$

De estas ecuaciones se sigue que $\alpha = \beta = 0$.

Como la dimensión de R^n es precisamente n, no puede haber más de n reales λ tales que $A x = \lambda x$ para alguna $x \in S$ y por lo tanto habrá una infinidad de números a_0 para los cuales $A x \neq -a_0 x$ para toda $x \in S$.

Definamos el espectro de A como $\sigma_r(A) = \{ \lambda \in R : A x = \lambda x$ para alguna $\bar{x} \in S \}$. A los elementos de $\sigma_r(A)$ los llamaremos valores propios de A y a los vectores asociados vectores propios de A. El subíndice r indica que $\sigma_r(A)$ es un subconjunto de los números reales.

La discusión anterior la podemos resumir en el siguiente teorema.

Teorema 1.2. $p(A) = aI + A$ **es invertible si y solo si** $-a \notin \sigma_r(A)$

Corolario 1. $p(A) = aI + A$ **es invertible si y solo si** $p(x) \neq 0$ **para toda** $x \in \sigma_r(A)$

Corolario 2. $\lambda \in \sigma_r(A)$ **si y solo si** $\lambda I - A$ **no es invertible.**

Consideremos ahora polinomios de la forma $p(x) = (\lambda_1 - x)(\lambda_2 - x)$, entonces $p(A) = (\lambda_1 I - A)(\lambda_2 I - A) = (\lambda_2 I - A)(\lambda_1 I - A)$. Si $\lambda_1 I - A$ y $\lambda_2 I - A$ son invertible, es decir, si $\lambda_1, \lambda_2 \notin \sigma_r(A)$, entonces p(A) es invertible. Recíprocamente, si p(A) es invertible, también lo son $\lambda_1 I - A$ y $\lambda_2 I - A$, ya que si, por ejemplo, $(\lambda_1 I - A)\bar{x} = 0$ para alguna $\bar{x} \neq 0$, entonces $p(A)\bar{x} = (\lambda_2 I - A)(\lambda_1 I - A)\bar{x} = 0$, lo que contradice la hipótesis de que p(A) es invertible.

Resulta, pues, que en este caso p(A) es invertible si y sólo si $\lambda_1, \lambda_2 \notin \sigma_r(A)$; o equivalentemente, si $p(x) \neq 0$ para toda $x \in \sigma_r(A)$.

De los razonamientos anteriores, resulta fácil ver que si $p(x) = \prod_{i=1}^{n} (\lambda_i - x)^{a_i}$, entonces una condición necesaria y suficiente para que p(A) sea invertible, es que $p(x) \neq 0$ para toda $x \in \sigma_r(A)$.

Por lo que hemos visto, todo parece indicar que para cualquier polinomio $p(x)$ se tendrá que una condición necesaria y suficiente para que $p(A)$ sea invertible, es que $p(x) \neq 0$ para toda $x \in \sigma_r(A)$. Sin embargo este resultado es falso. Para ver esto, consideremos el polinomio $p(x) = x^2 + 1$. Evidentemente $p(x) \neq 0$ para todo real x, en particular para los elementos del espectro de cualquier transformación lineal. Sea A dada por la matriz, que también denotaremos por A,

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

entonces

$$A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$$

por lo tanto $p(A) = A^2 + I = 0$ y obviamente $p(A)$ no es invertible.

Trataremos de ver a qué se debe esta situación. Investiguemos primeramente el espectro de A. Supongamos que $A\bar{x} = \lambda\bar{x}$ para alguna $\bar{x} \in S$, entonces $A^2\ \bar{x} = A(A\bar{x}) = \lambda^2\ \bar{x}$, pero por otro lado $A^2\ \bar{x} = -\bar{x}$. De aquí que $\lambda^2 = -1$, o sea que $\lambda = \pm i$. En consecuencia $\sigma_r(A) = \phi$. Sin embargo, nótese que $\lambda = \pm i$ son las raíces de $p(\lambda)$, y si aceptamos la posibilidad de que A tenga valores propios complejos, tendremos que $p(x)$ se anula en el espectro de A. Esto nos da lugar a pensar que si aceptamos que A tenga valores propios complejos, entonces $p(A)^{-1}$ existe si y solo si $p(x) \neq 0$ para toda x en el espectro de A. Pero para considerar que A tenga valores propios complejos debe tener sentido $\lambda\bar{x}$ con $\lambda \in C$, ya que $\lambda\bar{x}$ es un vector con coordenadas complejas, lo que nos saca automáticamente de R^n.

Necesitamos entonces trabajar en un espacio vectorial sobre los complejos que contenga a R^n, por ejemplo C^n, o cualquier espacio vectorial sobre los complejos de dimensión n. Trabajaremos nosotros con C^n.

Sea pues $A : C^n \to C^n$ lineal, y redefinamos el espectro de A:

$$\sigma(A) = \{ \lambda \in C \mid A\bar{x} = \lambda \bar{x} \text{ para alguna } \bar{x} \in S \},$$

donde

$$S = \left\{ \bar{x} \in C^n \mid \| \bar{x} \| = 1 \right\}.$$

Los resultados obtenidos para R^n siguen valiendo en este caso.

Ahora sea p(x) un polinomio con coeficientes complejos. Si $A\bar{x} = \lambda\bar{x}$ para alguna $\bar{x} \in S$, es evidente que $p(A)\bar{x} = p(\lambda)\bar{x}$ y de aquí que

$$\sigma(p(A)) \supset p(\sigma(A)) = \left\{ p(\lambda) \mid \lambda \in \sigma(A) \right\}.$$

La otra contención es también válida, esto es,

$$\sigma(p(A) = P(\sigma(A)).$$

La demostración la veremos después en casos más generales.

De lo anterior vemos que si $p(\lambda)$ se anula en algún punto de $\sigma(A)$, entonces p(A) tendrá al cero como valor propio, y en consecuencia no es invertible. Recíprocamente si p(A) no es invertible, entonces p(A) tiene al cero como valor propio, esto es, $p(\lambda) = 0$ para alguna $\lambda \in \sigma(A)$. Esto demuestra el siguiente teorema.

Teorema 1.3 **$p(A)$ es invertible si y sólo si $p(x) \neq 0$ para toda $x \in \sigma(A)$.**

Hemos logrado encontrar bajo qué condiciones $p(A)^{-1}$ existe. Con esta caracterización podemos definir $Q(A)$ donde $Q(x) = p(x)/q(x)$

1.3 El caso en que f(z) es analítica

La función racional $f(z) = 1/1 - z$ es una función analítica para $|z| < 1$. En efecto, si $|z| < 1$, se tiene que $f(z) = 1 + z + \ldots + z^n + \ldots$ Esto hace pensar que podamos definir, bajo algunas condiciones, $f(A)$ siendo $f(z)$ analítica. Por ejemplo, en el caso anterior parece que si $\| A \| < 1$ podemos definir $f(A) = I + A + \ldots + A^n + \ldots$ Pero esto lleva implícita la noción de convergencia de operadores la cual no hemos definido pero que ahora precisaremos.

Sea

$$L = \left\{ T : C^n \to C^n \mid T \text{ es lineal} \right\}.$$

Este conjunto es un espacio vectorial normado con la norma dada por

$$\| T \| = \sup_{\|x\|=1} \| T x \| .$$

Además L es completo con la métrica inducida por esta norma.

Sea $f(z) = \sum_{k=0}^{\infty} a_k z^k$ para $|z| < R$. Necesitamos ver cuando $f_n(A) = \sum_{k=0}^{\infty} a_k A^k$ converge en L. Sabemos que si $m > n$

$$\| f_n(A) - f_m(A) \| \leqslant \sum_{k=n+1}^{m} | a_k | \| A \|^k .$$

Si f(z) converge absolutamente para alguna z tal que $|z| = \|A\|$, entonces por la desigualdad anterior, fn(A) convergerá en L. Pero es sabido que f(z) converge absolutamente para $|z| \leq \alpha R$, donde $0 < \alpha < 1$, o sea que si $\|A\| < R$, la serie $\sum_{n=0}^{\infty} a_n A^n$ converge en L. Podemos entonces definir

$$f(A) = \sum_{k=0}^{\infty} a_n A^k \quad \text{si } \|A\| < R$$

Por otro lado, sea $\lambda \epsilon \sigma(A)$ y $\bar{x} \epsilon S$ tal que $A\bar{x} = \lambda \bar{x}$, entonces

$$|\lambda| = |A\bar{x}| \leq \|A\|$$

esto es, $\sigma(A)$ está contenido en el círculo con centro en el origen y radio $\|A\|$. En consecuencia, si f es analítica en un círculo de radio R y $R > \|A\|$, podemos definir

$$f(A) = \sum_{k=0}^{\infty} a_k A^k.$$

En este caso, el círculo de convergencia de f(z) contiene al espectro de A.

Si f(z) es de tal clase y si $A\bar{x} = \lambda \bar{x}$ para alguna $\bar{x} \epsilon S$ obtenemos

$$f_n(A)\,\bar{x} = f_n(\lambda)\,\bar{x},$$

y tomando límites,

$$f(A)\,\bar{x} = f(\lambda)\,\bar{x}$$

Es decir, $\sigma(f(A)) \supset f(\sigma(A))$. De nuevo, la otra contención es válida, la demostración se hará en casos más generales. Como consecuencia de esto se tiene que

Teorema 1.4. f(A) es invertible si y sólo si $f(z) \neq 0$ para toda $z \in \sigma(A)$

Como un ejemplo, consideremos la función $f(z) = (\lambda - z)^{-1}$, esta función es analítica para $|z| < |\lambda|$ y

$$f(z) = \lambda^{-1} + \lambda^{-2}z + \ldots + \lambda^{-n}z^{n-1} + \ldots$$

por lo tanto

$$f(A) = \sum_{n=1}^{\infty} \lambda^{-n} A^{n-1} \quad \text{si} \quad |\lambda| > \|A\|$$

Supongamos que f(z) es analítica en un círculo de radio $R > \|A\|$, y sea Γ una circunferencia con centro en el origen y radio r tal que $\|A\| < r < R$. Entonces para cualquier punto z en el interior de Γ, se tiene

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma} f(\xi)\ (\xi - z)^{-1}\, d\xi$$

Como $\|A\| < |\xi|$ para cada $\xi \in \Gamma$, $(\xi - A)^{-1}$ existe, y esto sugiere que se pueda definir f(A) por medio de la fórmula anterior. Es decir, como

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(\xi)\ (\xi - A)^{-1}\, d\xi.$$

obteniendo así otra representación para f(A). Por supuesto, hay que precisar el significado de una integral de este tipo, cosa que haremos posteriormente. Este tipo de representación es particularmente útil, ya que no necesitamos conocer el desarrollo de f(z) para calcular f(A).

En capítulos posteriores haremos una reconstrucción del cálculo operacional en casos más generales.

1.4 Solución del ejemplo A

Para cada t

$$e^{tz} = \sum_{n=0}^{\infty} \frac{t^n z^n}{n!}$$

y el radio de convergencia es infinito. De aquí que podamos definir e^{tA}, siendo A una matriz cuadrada, y

$$e^{tA} = I + tA + \frac{t^2}{2!} A^2 + \ldots + \frac{t^n}{n!} A^n + \ldots$$

para toda t. De aquí que

$$e^{tA} \bar{x}_o = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n \bar{x}_o$$

La cual es, en efecto, la solución de la ecuación diferencial dada por I.1.

Esto nos hace ver que vamos por buen camino. En los siguientes capítulos se desarrollarán técnicas que nos permitirán resolver, vía Cálculo Operacional, los otros ejemplos.

CAPITULO 2. OPERADORES EN ESPACIOS DE HILBERT

2.0 Introducción

El objetivo de este capítulo es estudiar algunas propiedades de operadores en espacios de Hilbert.

Primeramente haremos ver la necesidad de trabajar con dicho tipo de operadores.

2.1 Motivación

2.1.1 Un núcleo separable determina un operador cuyo rango es de dimensión finita

Tomamos de nuevo el ejemplo (D) y en particular el caso en que el operador A está determinado por un núcleo separable, esto es, cuando

$$K(x, t) = \sum_{k=1}^{n} a_k(x)\, b_k(t)$$

Si g = Af entonces

$$g(x) = \int_0^1 \sum_{k=1}^{n} a_k(x)\, b_k(t)\, f(t)\, dt = \sum_{k=1}^{n} \alpha_k\ a_k(x) \qquad (2.1)$$

donde

$$\alpha_k = \int_0^1 b_k(t)\, f(t)\, dt.$$

La ecuación (1.1) demuestra que el rango de A es de dimensión finita. Hemos visto que el problema de encontrar u tal que $(I - \lambda A)\, u = f$ se reduce, en esta situación, al problema matricial $(I - \lambda K)\, \bar{c} = \bar{f}$. Sabemos también cuándo $(I - \lambda K)^{-1}$ existe.

Existe una amplia clase de funciones K(x, t) que pueden aproximarse por medio de núcleos separables. A continuación trataremos de aclarar qué funciones tienen dichas propiedades.

2.1.2 Los espacios L_2

Consideremos los siguientes espacios vectoriales

$$L_2\ (0, 1) = \left\{ f\colon [0, 1] \to C \mid \int_0^1 |f(x)|^2\, dx < \infty \right\},$$

$$L_2\ (\Omega) = \left\{ K\colon [0,1] \times [0,1] \to C \mid \int_0^1 \int_0^1 |K(x,y)\, dx\, dy < \infty \right\},$$

donde $\Omega = [0,1] \times [0,1]$ y la integración es en el sentido de Lebesgue.

Veremos primero algunas propiedades de $L_2\ (0, 1)$ y después las generalizaremos a $L_2\ (\Omega)$.

Haciendo la convención de que $f = g$ si

$$\int_0^1 | f(x) - g(x) |^2 \, dx = 0$$

podemos definir un producto escalar en $L_2(0, 1)$ dado por

$$< f, g > = \int_0^1 f(x) \overline{g(x)} \, dx.$$

Esto induce una norma en $L_2(0, 1)$, a saber

$$\| f \| = < f, f >^{1/2} = (\int_0^1 | f(x) |^2 \, dx)^{1/2}$$

Se puede demostrar que con esta norma $L_2(0, 1)$ es completo.

Las propiedades anteriores y las que se mencionana a continuación aparecen normalmente en los libros de Análisis Matemático.

Existe una sucesión $u_1, u_2, \ldots, u_n, \ldots$ de elementos de $L_2(0, 1)$ tal que

1) $$<u_i, u_j> = \int_0^1 u_i(x) \overline{u_j(x)} \, dx = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j . \end{cases}$$

2) Cualquier $f \in L_2(0, 1)$ puede escribirse como

$$f(x) = \sum_{k=1}^{\infty} a_k \, u_k(x),$$

siendo $a_k = \langle f, u_k \rangle$. Esto significa que

$$\| f - \sum_{k=1}^{n} a_k u_k \| \to 0 \quad \text{si } n \to \infty$$

Esto es,

$$\int_0^1 | f(x) - \sum_{k=1}^{n} a_k u_k(x) |^2 \, dx \to 0 \quad \text{si } n \to \infty$$

Un ejemplo de un conjunto tal es $\sqrt{2}$ sen πx, $\sqrt{2}$ sen $2\pi x, \ldots,$ $\sqrt{2}$ sen $n\pi x, \ldots$. También el conjunto $1, \sqrt{2} \cos \pi x, \sqrt{2} \cos 2\pi x, \ldots, \sqrt{2} \cos n\pi x, \ldots$ tiene las mismas propiedades.

Un conjunto satisfaciendo (1) y (2) es llamado un sistema ortonormal completo.

En general, un espacio vectorial con un producto escalar y completo con la norma inducida por éste, es llamado un espacio de Hilbert. Si además el espacio posee un conjunto ortonormal completo se dice que es un espacio de Hilbert separable.

Entonces $L_2(0, 1)$ es un espacio de Hilbert separable.

Más generalmente, $L_2(a, b)$, con $-\infty \leqslant a < b \leqslant \infty$, sigue siendo un espacio de Hilbert.

En $L_2(\Omega)$ también puede definirse un producto escalar, a saber,

$$\langle K_1, K_2 \rangle = \int_0^1 \int_0^1 K_1(x, t) \, \overline{K_2(x, t)} \, dx \, dt,$$

el cual induce una norma

$$\| K \| = \langle K, K \rangle^{1/2} = \left(\int_0^1 \int_0^1 |K(x, t)|^2 \, dx \, dt \right)^{1/2}$$

$L_2(\Omega)$ es también un espacio de Hilbert separable. Cuando Ω es cualquier rectángulo en R^2, aun si $\Omega = R^2$, $L_2(\Omega)$ sigue siendo un espacio de Hilbert separable

2.1.3 Todo elemento de $L_2(\Omega)$ determina un operador integral de $L_2(0, 1)$ en sí mismo

Sean $\Omega = [0, 1] \times [0, 1]$, $K(x, t)$ en $L_2(\Omega)$, $f(t)$ en $L_2(0, 1)$ y

$$g(x) = \int_0^1 K(x, t)\, f(t)\, dt. \qquad (2.2)$$

Para cada x fija, $K(x, t)$ se puede considerar en $L_2(0, 1)$ y entonces

$$g = \langle K(x, .), f \rangle$$

De aquí se sigue que

$$|g(x)|^2 = |\langle K(x,.), f \rangle|^2$$

y utilizando la desigualdad de Schwarz

$$|g(x)|^2 \leqslant \int_0^1 |K(x, t)|^2 \, dt \int_0^1 |f(t)|^2 \, dt$$

Entonces

$$\int_0^1 |g(x)|^2\, dx \leq \int_0^1 \int_0^1 |K(x, t)|^2\, dx\, dt \int_0^1 |f(t)|^2\, dt < \infty$$

En consecuencia $g \epsilon L_2\ (0, 1)$. Esto significa que

$$A : L_2\ (0, 1) \rightarrow L_2\ (0, 1)$$

donde $Af = g$, siendo

$$g(x) = \int_0^1 K(x, t)\ f(t)\ dt$$

En resumen: A es un operador lineal en un espacio de Hilbert y además, de (2.3),

$$\| Af \| \leq \| K \|\ \ \| f \| \qquad (2.3)$$

Es importante hacer notar que si $K_1(x, t)$ y $K_2(x, t)$ están en $L_2(\Omega)$, y A_1, A_2 son los respectivos operadores integrales, entonces $K_1 + K_2$ determina un operador integral dado por $A_1 + A_2$. Así mismo, si $a \epsilon$ C, K(x, t) está en $L_2(\Omega)$ y A es el operador integral determinado por K(x, t), entonces aA es el operador integral determinado por (aK) (x, t).

Ahora sea

$$L = \left\{ A : L_2 (0, 1) \rightarrow L_2 (0, 1) \mid Af = g,\ g \text{ dado por } (2.2) \right\}$$

Este conjunto constituye un espacio vectorial y en él se puede definir una norma de la siguiente manera:

$$\| A \| = \inf \left\{ M \mid \| Af \| \leq M\ \| f \| \quad \text{para toda} \quad f \epsilon L_2\ (0, 1) \right\}$$

En particular, de (2.4), se sigue que

$$\| A \| \leqslant \| K \| \qquad (2.5)$$

2.1.4 Todo operador integral en L_2 (0, 1) puede ser aproximado por operadores con núcleo separable

Un hecho del que haremos uso, aunque no lo demostraremos es el siguiente [9]: Toda función $K(x, t)$ en $L_2(\Omega)$ se puede escribir como

$$K(x,t) = \sum_{n=1}^{\infty} a_n(x)\ b_n(t)$$

con $a_n, b_n \in L_2(0, 1)$, $n = 1, 2, \ldots$

Si

$$K_n(x, t) = \sum_{k=1}^{n} a_k(x)\ b_k(t)$$

la ecuación anterior nos dice que $K_n \to K$, o lo que es lo mismo, que $\| K_n - K \| \to 0$. Si A y A_n son los operadores integrales determinados por K y K_n, respectivamente, entonces de (2.5) se sigue que

$$\| A - A_n \| \to 0, \quad n \to \infty .$$

Esto significa que A puede ser aproximado por operadores de rango de dimensión finita, los cuales ya conocemos.

Resumiendo: Todo $A \in L$ puede ser aproximado por operadores cuyo rango es de dimensión finita. Problemas en los que interviene un operador integral pueden ser aproximados por operadores con núcleo degenerado, y problemas en los que intervienen este último tipo de operadores pueden ser reducidos a problemas matriciales.

Hemos visto que el ejemplo (D) puede ser formulado en términos de operadores en cierto espacio de Hilbert. En este caso el problema puede ser atacado, a final de cuentas, vía el caso matricial. Sin embargo existen operadores en espacios de Hilbert que no tienen las propiedades de los operadores estudiados en esta sección. Esto lo haremos ver en las siguientes secciones.

2.2 Espacios de Hilbert

Hemos visto en el capítulo 1 la necesidad de trabajar con espacios vectoriales sobre el campo de los complejos.

Sea H un espacio vectorial complejo. Un producto escalar en H es una función

$$< , > : H \times H \to C$$

tal que

1) $< x, y > = \overline{< y, x >}$.
2) $< a_1 x_1 + a_2 x_2, y > = a_1 <x_1, y> + a_2 < x_2, y >$.
3) $< x, x > \geqslant 0$ y $< x, x > = 0$ si y solo si $x = 0$.

Este producto escalar induce una norma en H dada por

$$\| x \| = < x, x >^{1/2}$$

Definición. Un espacio vectorial complejo H con un producto escalar, el cual es completo con la norma inducida por dicho producto, es llamado un espacio de Hilbert complejo.

En forma análoga se puede definir el concepto de espacio de Hilbert real.

Ejemplo 1. C^n

Ejemplo 2. L_2 (a, b), $-\infty \leqslant a < b \leqslant \infty$. También $L_2(\Omega)$ siendo Ω un rectángulo en R^2.

Ejemplo 3. $l_2 = \left\{ \{x_n\} \subset C \mid \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}$. Aquí $<x, y> = \sum_{n=1}^{\infty} x_n \overline{y_n}$ donde $x = \{x_n\}$, $y = \{y_n\}$.

Los siguientes teoremas pueden ser demostrados sin dificultad siguiendo técnicas rutinarias.

Teorema 2.1. (Desigualdad de Schwarz). Para toda x, y ϵH,

$$|< x, y >| \leqslant \|x\| \ \|y\|.$$

La igualdad vale si y solo si x, y son linealmente dependientes.

Teorema 2.2. Si $\{x_n\} \to x$ y $\{y_n\} \to y$, entonces $< x_n, y_n > \to < x, y >$.

Definición. Un conjunto $\{x_1, x_2, \ldots\}$ se dice que es linealmente independiente si $\{x_1, \ldots, x_n\}$ es linealmente independiente para toda n.

El conjunto se dice que es ortogonal si $< x_i, x_j > = 0$ para $i \neq j$. Si además $< x_i, x_i > = 1$ para toda i, el conjunto se llama ortonormal.

Teorema 2.3. Un conjunto ortonormal $\{ x_1, x_2, \ldots \}$ es linealmente independiente. Además

$$\| \sum_{k=1}^{n} a_k \, x_k \|^2 = \sum_{k=1}^{n} | a_k |^2 .$$

Teorema 2.4. Si $< x, y > = 0$, entonces $\| x + y \|^2 = \| x \|^2 + \| y \|^2$.

Teorema 2.5. Sea $\{ x_1, x_2, \ldots \}$ un conjunto ortonormal. La serie $\sum_{n=1}^{\infty} a_n x_n$ converge si y solo si $\sum_{n=1}^{\infty} | a_n |^2$ converge. En este caso

$$\| \sum_{n=1}^{\infty} a_n \, x_n \|^2 = \sum_{n=1}^{\infty} | a_n |^2 .$$

Teorema 2.6. Si $\{ x_1, x_2, \ldots \}$ es un conjunto ortonormal y $x \in H$, entonces

$$\sum_{n=1}^{\infty} |< x, x_n >|^2 \leq \| x \|^2 .$$

La igualdad se satisface si y solo si $x = \sum_{n=1}^{\infty} < x, x_n > x_n$.

Definición. Un conjunto ortonormal $\{ x_1, x_2, \ldots \}$ tal que

$$x = \sum_{n=1}^{\infty} < x, x_n > x_n$$

para toda $x \in H$ es llamado un conjunto ortonormal completo.

Definición. Un espacio de Hilbert que contiene a un conjunto numerable $\{ x_1, x_2, \ldots \}$, el cual es ortonormal completo, es llamado un espacio de Hilbert separable.

En el presente trabajo sólo usaremos espacios de Hilbert separables. Un ejemplo típico de espacio separable es L_2 (0,1). Un conjunto ortonormal completo es el formado por las funciones $u_n(x) = \sqrt{2}$ sen $n\pi x$.

Definición. Un subconjunto $D \subset H$, el cual es un espacio vectorial, es llamado un subespacio de H. Si además D es cerrado se dice que es un subespacio cerrado de H.

Definición. Una función $f : H \to C$ tal que

i) $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ para toda $x, y \in H$ y toda $\alpha, \beta \in C$

ii) $|f(x)| \leqslant M \; \|x\|$ para algún real M y para toda $x \in H$,

es llamada una funcional lineal acotada en H. El conjunto de funcionales lineales acotadas en H es denotado por H'.

Teorema 2.7. $f \in H'$ si y solo si existe un único vector $y \in H$ tal que $f(x) = \langle x, y \rangle$ para toda $x \in H$.

2.3 Algunas convenciones sobre operadores lineales

Definición. Una transformación $A : D \subset H \to H$, donde D es un subespacio de H, se dice que es un operador lineal si $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$ para toda $x, y \in H$ y toda $\alpha, \beta \in C$. D es llamado el dominio de A.

Definición. Al conjunto de valores Ax, con $x \in D$, lo llamaremos el rango de A y será denotado por R(A). Si S es un subconjunto de D, al conjunto de puntos Ax con $x \in S$ lo denotaremos por A(S).

Definición. Si existe un operador lineal $B : R(A) \rightarrow H$ tal que $AB = I_D$ y $BA = I_{R(A)}$ - donde I_D e $I_{R(A)}$ denotan al operador identidad en D y R(A), respectivamente entonces se dice que A es invertible y su inverso es $B = A^{-1}$.

Convención. Puesto que únicamente trabajaremos con operadores lineales, a estas los llamaremos simplemente operadores.

2.4 Operadores acotados

En la sección 2.1 vimos que el operador integral

$$A : L_2\ (0,\ 1) \rightarrow L_2\ (0.\ 1)$$

dado por $Af = g$, donde

$$g(x) = \int_0^1 K(x,\ t)\ f(t)\ dt,$$

tiene las siguientes propiedades:

i) Existe M tal que

$$\| Af \| \leqslant M\ \| f \| \qquad (2.6)$$

ii) A puede ser aproximado en norma por operadores cuyo rango es de dimensión finita.

Es inmediato que si un operador $A : H \rightarrow H$ (i.e., $D = H$) tiene la propiedad (2.6), entonces tranforma conjuntos acotados en conjuntos acotados.

Definición. Un operador $A : H \to H$ se dice que es acotado si existe M tal que

$$\| Ax \| \leqslant M \| x \|$$

para toda $x \in H$. Si A es acotado definimos la norma de A como

$$\| A \| = \inf \left\{ M \mid \| Ax \| \leqslant M \| x \| \text{ para toda } x \in H \right\}.$$

Teorema 2.8. Si A es acotado, entonces

$$\| A \| = \sup_{\|x\|=1} \| Ax \| .$$

La demostración es inmediata.

Notación. $L(H)$ denotará al conunto de operadores acotados en H.

Las propiedades más importantes de $L(H)$ vienen dadas por los siguientes resultados.

Teorema 2.9. $L(H)$ es un espacio vectorial normado y completo. Además

i) $\| A + B \| \leqslant \| A \| + \| B \|$, $A, B \in L(H)$

ii) $\| \alpha A \| = |\alpha| \, \| A \|$, $\alpha \in C$, $A \in L(H)$

iii) Si $A, B \in L(H)$, entonces $AB \in L(H)$ y

$$\| AB \| \leqslant \| A \| \, \| B \|$$

Para la demostración de la primera afirmación del teorema nos remitiremos a [20]. Las afirmaciones (i), (ii) y (iii) son consecuencias inmediata de la definición de la norma.

Observación. Este teorema nos dice. en lenguaje algebráico que L(H) es una álgebra normada.

De la definición de operador acotado se sigue inmediatamente que todo operador acotado es continuo. El recíproco es también cierto: todo operador es continuo es acotado. Esto constituye un resultado clásico del Análisis Funcional.

Teorema 2.10. Un operador $A : H \to H$ es acotado si y solo si es continuo.

2.5 Operadores de dimensión finita y operadores compactos

Definición. A un operador $A : H \to H$ cuyo rango es de dimensión finita lo llamaremos operador de dimensión finita.

Si recordamos que toda $A : H \to H$ es acotada cuando H es de dimensión finita, podemos pensar que cuando A es de dimensión finita, A es acotado. Sin embargo esto no sucede como veremos a continuación.

2.5.1 Un operador de dimensión finita no acotado.

Sea $A : L_2 (0, 1) \to C$ dado por

$$A(f) = f(1)$$

El rango de A tiene dimensión 1. A es lineal, pero no es continuo como veremos a continuación.

Sea

$$fn(x) = x^n ,$$

entonces

$$\| fn \|^2 = \int_0^1 x^{2n} \, dx = \frac{1}{2n + 1},$$

de donde

$$\| fn \| \to 0 \quad \text{si} \quad n \to \infty ,$$

pero

$$A(fn) = fn(1) = 1$$

para toda n. Es decir $A(fn) \not\to 0$. Esto demuestra que A no es continuo.

2.5.2 Definiciones de F(H) y de su cerradura K(H)

Definición. Definimos F(H) como el conjunto de operadores acotados de dimensión finita en H.

Es fácil ver que F(H) es un subespacio de L(H). En general F(H) no es cerrado, y el que realmente es importante es su cerradura.

Definición. Definimos K(H) como la cerradura de F(H).

Los elementos de K(H) son, pues, límite (en norma) de operadores acotados de dimensión finita.

A continuación veremos una caracterización de K(H)

2.5.3 Una caracterización de K(H)

Notemos que si $A \in F(H)$, entonces dada una sucesión acotada $\{x_n\}$ en H, la sucesión $\{Ax_n\}$ tiene una subsucesión convergente ya que $\{Ax_n\}$ es una sucesión acotada en un subespacio de dimensión finita.

Veremos cómo esta propiedad caracteriza a K(H). Para ello haremos uso del siguiente lema [20].

Lema. Si $D \subset H$ es un subespacio cerrado, D distinto de H, y si a es un número real tal que $0 < a < 1$, entonces existe un vector $x_o \in H$ tal que $\| x_o \| = 1$ y $\| x - x_o \| \geq a$ para toda $x \in D$.

Teorema 2.11. $A \in K(H)$ si y solo si dada una sucesión acotada $\{x_n\} \subset H$ es posible extraer una subsucesión convergente de $\{Ax_n\}$.

Demostración.$\Rightarrow$Supongamos que $\| A - A_n \| \to 0$ con $A \in K(H)$ y $A_n \in F(H)$. Sea $\{x_n\}$ tal que $\| x_n \| \leqslant C$ para toda n. Para cada k, existe $\{x_{n_k}\}$ tal que $\{A_k x_{n_k}\}$ converge. Si $z_n = x_{n_n}$ entonces $\{A_k z_n\}$ converge para cada k. Ahora

$$\| Az_i - Az_j \| \leqslant \| Az_i - A_k z_i \| + \| A_k z_i - A_k z_j \| + \| A_k z_j - Az_j$$

$$\leqslant 2C \, \| A - A_k \| + \| A_k z_i - A_k z_j \| .$$

De aquí que $\{Az_i\}$ converge.

$\Leftarrow$ Necesitamos demostrar que existe una sucesión $\{A_n\}$ en F(H) tal que $\| A - A_n \| \to 0$ si $n \to \infty$. Sea $\{u_1, u_2, \ldots\}$ un conjunto ortonormal completo en H. Como $Ax = \sum_{k=1}^{\infty} < Ax, u_k > u_k$ podemos definir $A_n \in F(H)$ como $A_n x = \sum_{k=1}^{n} < Ax, u_k > u_k$. Obviamente $\| Ax - A_n x \| \to 0$ cuando $n \to \infty$, para toda $x \in H$. Además para cada n, A_n es de dimensión finita. Mostraremos que $\| A - A_n \| \to 0$ cuando $n \to \infty$.

Para esto, supongamos que lo anterior no es cierto, entonces existe $\epsilon > 0$ tal que $\| A - A_n \| \geqslant \epsilon$ para una infinidad de naturales n. Sin perder generalidad podemos suponer que $\| A - A_n \| \geqslant \epsilon$ para toda n. Puesto que $\| A - A_n \| = \sup \| Ax - A_n x \|$ con $\| x \| = 1$, entonces para cada n existe $x_n \in H$ tal que $\| x_n \| = 1$ y $\| (A - A_n) x_n \| \geqslant \epsilon/2$. Como $\{x_n\}$ es acotada, existe $\{x_{n_k}\}$ tal que $\{Ax_{n_k}\}$ converge. Denotemos al límite de esta sucesión por ω. Entonces $\omega = \sum_{k=1}^{\infty} <\omega, u_k> u_k$.

Por otro lado

$$\| (A - A_n) x_n \| \leqslant \| (A - A_n) x_n - (\omega - \omega_n) \| + \| \omega - \omega_n \|$$

donde $\omega_n = \sum_{k=1}^{n} < \omega, u_k > u_k$. De la definición de A_n y de ω_n obtenemos

$$\| (A - A_n) x_n - (\omega - \omega_n) \|^2 = \sum_{k=n+1}^{\infty} |< Ax_n - \omega, u_k >|^2 \leqslant \| A x_n - \omega \|^2$$

y por lo tanto

$$\| (A - A_n) x_n \| \leqslant \| Ax_n - \omega \| + \| \omega - \omega_n \|$$

Pero entonces es posible escoger una n_k tal que $\| (A - A_n) x_{n_k} \| < \epsilon/2$ lo cual es una contradicción. De aquí que $\| A - A_n \| \to 0$ si $n \to \infty$. Esto demuestra el teorema.

2.5.4 Otra caracterización de K(H)

Consideremos ahora un conjunto S acotado en H y sea $A \in K(H)$. Toda sucesión en A(S) tiene entonces una subsucesión convergente (cuyo límite no está necesariamente en A(S)). Un conjunto con tal propiedad es llamado relativamente compacto (ya que su cerradura es un conjunto compacto). Entonces A transforma conjuntos acotados en conjuntos relativamente compactos. Recíprocamente, si A transforma conjuntos acotados en conjuntos relativamente compactos entonces $A \in K(H)$. Esto nos da la siguiente caracterización de K(H).

Teorema 2.12. $A \in K(H)$ si y sólo si A transforma conjuntos acotados en conjuntos relativamente compactos.

2.5.5 Operadores compactos

Definición. A los elementos de K(H) los llamaremos operadores compactos.

El ejemplo más importante de operadores compactos son los operadores integrales ya examinados al principio de este capítulo.

Es importante hacer notar que no todo operador acotado es compacto. Como ejemplo tenemos al operador identidad. En efecto, tomemos un conjunto ortonormal numerable $\{ x_1, x_2, \ldots \}$. este constituye una sucesión acotada, pero no tiene subsucesiones convergentes ya que $\| x_i - x_j \| = \sqrt{2}$ para $i \neq j$.

2.6 Relaciones entre operadores acotados y operadores compactos

Veremos ahora una importante relación entre L(H) y K(H)

Lema. Si $A \in K(H)$ y $B \in L(H)$, entonces AB y BA son compactos.

Demostración. Sea S un conjunto acotado en H, entonces $B(S)$ es acotado y por lo tanto $AB(S)$ es relativamente compacto. Entonces AB es compacto. Por otro lado $A(S)$ es relativamente compacto, y por la continuidad de B, $BA(S)$ es también relativamente compacto. De aquí que $BA \in K(H)$ y el lema queda demostrado.

Observación. Este resultado nos dice que $K(H)$ es un ideal de $L(H)$. (Véase la observación al teorema 2.9).

Teorema 2.13. Si A es compacto y A^{-1} existe, éste no puede ser acotado.

Demostración. Supongamos que A^{-1} es acotado, por el lema anterior $A\,A^{-1} = I$ es compacto, lo cual es falso.

Este teorema es de suma importancia ya que, por ejemplo, el inverso de un operador diferencial es un operador integral, éste es compacto y por lo tanto el operador diferencial no puede ser acotado.

2.7 Un operador no acotado

Como un ejemplo de lo anterior consideremos el operador $T : D \subset L_2\,(0,\,1) \to L_2\,(0,\,1)$, dado por $Tx = x'$, siendo D el conjunto de funciones que satisfacen $x(0) = x(1) = 0$ y tales que $x' \in L(0,1)$. Para ver que T no es acotado consideremos la sucesión $x_n(t) = \sqrt{2}$ sen $n\,\pi\,t$; entonces

$$\| Tx_n \|^2 = \int_0^1 2n^2 \cos^2 n\pi t \, dt = n^2$$

Como $\| x_n \| = 1$ para toda n, se sigue que T no es acotado.

El operador $Ax = y$, donde

$$y(t) = \int_0^t x(S)\, dS$$

es compacto y su inverso es precisamente T.

Dada la importancia, para nuestro trabajo, de los operadores diferenciales, es necesario que estudiemos una clase de operadores no acotados que contenga al ejemplo anterior. Esta clase es la de los operadores cerrados. Para dar una definición de estos operadores seguiremos analizando el operador que acabamos de ver, y veremos cómo este operador aunque no es continuo, conserva aún una propiedad común a los continuos.

Demostraremos que si $x_n \to x$ y $Tx_n \to y$, entonces $x \epsilon D$ y $Tx = y$.

Para esto sea

$$z(s) = \int_0^s y(t)\, dt$$

y recordemos que

$$x_n(s) = \int_0^s x_n'(t)\, dt.$$

Entonces

$$|x_n(s) - z(s)| \leq \int_0^s |x_n'(t) - y(t)|\, dt \leq \int_0^1 |x_n'(t) - y(t)|\, dt \qquad (2.7)$$

$$\leq \left(\int_0^1 |x_n'(t) - y(t)|^2\, dt\right)^{1/2} \left(\int_0^1 1^2\, dt\right)^{1/2} = \| x_n' - y \|,$$

la desigualdad es válida para toda $s \epsilon [0, 1]$. De aquí que

$$\| x_n - z \|^2 = \int_0^1 |x_n(S) - z(S)|^2 \, ds \leqslant \| x'_n - y \|^2 ,$$

y por lo tanto $x_n \to z$. En consecuencia $x = z$ (en el sentido de igualdad en $L_2(0,1)$). Además –como consecuencia de (2.7)– $z(0) = 0$, $z(1) = 0$ y $z' \epsilon L_2(0,1)$. Entonces $x \epsilon D$ y $Tx = x' = y$.

En este ejemplo, aun cuando $x_n \to x$ no podemos asegurar que $\{ Tx_n \}$ converja, sí sabemos que cuando $x_n \to x$ y $Tx_n \to y$ necesariamente $x \epsilon D$ y $Tx = y$.

Estudiaremos a continuación operadores con esta última propiedad.

2.8 Operadores cerrados

Definición. Un operador $A : D \subset H \to H$ se dice que es cerrado si dada una sucesión $\{ x_n \}$ en D tal que

$$x_n \to x, \quad Tx_n \to y ,$$

entonces necesariamente

$$x \epsilon D \qquad y \qquad Tx = y.$$

Una consecuencia inmediata de esta definición es que un operador acotado (en cuyo caso $D = H$) es cerrado, ya que es continuo. Tenemos, pues el siguiente resultado.

Teorema 2.14. Todo operador acotado es cerrado.

A continuación veremos que si A es cerrado e invertible, A^{-1} también es cerrado.

Teorema 2.15. Si $A : D \subset H \to H$ es cerrado y si A^{-1} existe, entonces A^{-1} es también cerrado.

Demostración. Sea $A^{-1} : R(A) \subset H \to H$ y supongamos que $\{ y_n \} \subset R(A), y_n \to y$, además $A^{-1} y_n \to x$. Sea $x_n = A^{-1} y_n$; entonces $\{ x_n \} \subset D$, $x_n \to x$ y $Ax_n \to y$. Por ser A cerrado, $x \in D$ y $Ax = y$. De aquí que $y \in R(A)$ y $y = A^{-1} x$. Así queda demostado el teorema.

Un interesante problema es el de determinar cuándo A^{-1} es acotado, siendo A cerrado. Para que A^{-1} sea acotado basta que $R(A) = H$ como se verá en los siguientes resultados.

Teorema 2.16. Si $A : H \to H$ (i.e., $D = H$) es cerrado, entonces A es acotado.

Para la demostración de este teorema nos remitiremos a [20].

Los dos siguientes teoremas son consecuencia inmediata del teorema anterior.

Teorema 2.17. Si $A : D \subset H \to H$ es cerrado, $R(A) = H$ y A^{-1} existe, entonces A^{-1} es acotado.

Teorema 2.18. Si $A : H \to H$ es acotado, $R(A) = H$ y A^{-1} existe, entonces A^{-1} es acotado.

2.9 Operadores autoadjuntos

2.9.1 El adjunto de un operador integral

Consideremos de nuevo el oprador $Af = g$ donde

$$g(x) = \int_0^1 K(x, t) f(t) dt$$

y supongamos que $K(x, t) = x^2 t$. Entonces

$$<Au, v> = \int_0^1 \int_0^1 x^2 \, t \, u(t)dt \quad \overline{v(x)} \, dx = \int_0^1 u(t) \int_0^1 x^2 t \, \overline{v(x)} \, dx \, dt$$

Si definimos $Bf = g$ como

$$g(x) = \int_0^1 x \, t^2 \; f(t) \; dt$$

obtendremos

$$<Au, v> = \int_0^1 u(t) \; \overline{B v(t)} \; dt = <u, Bv>$$

Es decir, existe un operador B tal que

$$<Au, v> = <u, Bv> \qquad (2.8)$$

para toda $u, v \epsilon L_2 \; (0, 1)$

B es llamado el adjunto de A.

2.9.2 El adjunto de un operador

Nos preguntamos ahora si en general, dado un operador $A : D \subset H \rightarrow H$, éste satisfacerá una ecuación como (2.8). Veremos que la respuesta es afirmativa si D es denso en H. Primero veremos que existen y's en H con la propiedad de que

$$<Ax, y> = <x, y^*> \qquad (2.9)$$

para alguna $y^* \in D$. En efecto, $y = 0$ satisface (2.9) tomando $y^* = 0$. Si y satisface (2.4) entonces y^* es única. Porque supongamos que $y_1{}^*$ también satiface (2.9) para la misma y: entonces $< x, y^*> = < x, y_1{}^* >$, esto es, $<x, y^* - y_1{}^*> = 0$ para toda $x \in D$. Como D es denso en H y el producto escalar es una función continua, se sigue que $< x, y^* - y_1{}^*> = 0$ para toda $x \in H$, lo cual implica que $y^* = y_1^*$.

Sea $D^* = \left\{ y \in H \mid y \text{ satisface } (2.9) \right\}$. Se puede probar, directamente de (2.9), que D^* es un subespacio de H y que $A^* : D^* \subset H \rightarrow H$, definida por $A^* y = y^*$ es lineal. Hemos construido, pues, un operador A^* –que llamaremos el adjunto de A– satisfaciendo

$$< Ax, y> = < x, A^* y>$$

para toda $x \in D$ y toda $y \in D^*$. Siempre y cuando D sea denso en H.

Teorema 2.19. A^* es cerrado

Demostración. Sea $\{ x_n \} \subset D^*$ tal que $x_n \rightarrow x_o$ y $A^*x_n \rightarrow y_o$. Entonces $< Ax, x_n> = < x, A^* x_n>$. Tomando límite en ambos lados obtenemos $< Ax, x_o> = < x, y_o >$ para toda $x \in D$. Por definición $x_o \in D$ y $A^*x_o = y_o$, y esto prueba que A^* es cerrado.

Teorema 2.20. Si $A : D \subset H \rightarrow H$ es un operador, siendo D denso en H; si $A^{-1} : R(A) \subset H \rightarrow H$ existe y además R(A) es denso en H, entonces $(A^{-1})^* = (A^*)^{-1}$.

Demostración. Puesto que R(A) es el dominio de A^{-1}, $R(A)^*$ denotará el dominio de $(A^{-1})^*$. Sean $x \in D$ y $y \in R(A)^*$. Entonces

$$< x, y > \;=\; < A^{-1} Ax, y > \;=\; < Ax, (A^{-1})^* y > .$$

esto significa que $A^* (A^{-1})^* y = y$, es decir $A^*(A^{-1})^* = I_{R(A)^*}$. Por otro lado, supongamos que $x \in R(A)$ y que $y \in D^*$. En este caso tenemos

$$< x, y > \;= < A A^{-1} x, y > = < A^{-1} x, A^* y > ,$$

lo cual implica que $A^* y \in R(A)^*$ y que $(A^{-1})^* A^* y = y$. Entonces $(A^*)^{-1}$ existe. La igualdad anterior muestra que la restricción de $(A^{-1})^*$ a $R(A^*)$ es precisamente $(A^*)^{-1}$, es decir $R(A^*) \subset R(A)^*$. Por otro lado, si $y \in R(A)^*$, sabemos que $A^*(A^{-1})^* y = y$, esto es, $y \in R(A^*)$. Esto junto con la contención anterior muestra que $R(A^*) = R(A)^*$. En consecuencia $(A^*)^{-1} = (A^*)^{-1}$.

2.9.3 Operadores autoadjuntos

Definición. Un operador es llamado autoadjunto si $A = A^*$.

Nótese que si A es autoadjunto, necesariamente su dominio es denso en H y $D = D^*$. Del teorema 2.19 se sigue que todo operador autoadjunto es cerrado.

A continuación mencionaremos dos ejemplos de operadores autoadjuntos.

Ejemplo 1. $A : L_2 (0, 1) \to L_2 (0, 1)$, $Af = g$ con

$$g(x) = \int_0^1 K(x, t)\, f(t)\, dt,$$

donde $K(x, t) = K(t, x)$.

Ejemplo 2. $A : D \subset L_2(0, 1) \to L_2(0, 1)$, siendo $A = - d^2/dx^2$ y $D = \left\{ u \in L_2(0, 1) \mid u'' \in L_2(0, 1) \text{ y } u(0) = u(1) = 0 \right\}$. Que A es autoadjunto, es consecuencia del método de integración por partes.

Teorema 2.21. Si A es autoadjunto y A^{-1} existe, entonces $R(A)$ es denso en H y A^{-1} es autoadjunto.

Demostración. Supongamos que $R(A)$ no es denso en H, entonces existe $z \neq 0$ tal que $z \in R(A)^{\perp}$. Por lo tanto $< Ax, z > = 0$ para toda $x \in D$; pero $<Ax, y>=<x, 0>$ y esto implica que $z \in D^* = D$ y $A^*z = Az = 0$, lo cual contradice la hipótesis de que A^{-1} existe. Por el teorema 2.20 sabemos que

$$A^{-1} = (A^*)^{-1} = (A^{-1})^*$$

lo que prueba que A^{-1} es autoadjunto.

Ejemplo. El inverso del operador diferencial $-d^2/dx^2$, $u(0) = u(1) = 0$, es el operador integral con núcleo

$$K(x, t) = \begin{cases} x(1 - t), & x \leq t \\ (1 - x)t & x \geq t \end{cases}$$

el cual es autoadjunto.

Nótese que si A es autoadjunto, $< Ax, x >$ es real para toda $x \in D$ ya que

$$< Ax, x > = < x, Ax > = \overline{< Ax, x >}$$

Definición. Si A es autoadjunto y si

$$m = \inf < Ax, x > , \quad \| x \| = 1, x \in D,$$

es positivo, entonces se dice que A es positivo definido.

Teorema 2.22. Si A es autoadjunto y positivo definido, entonces A^{-1} existe, $R(A) = H$ y A^{-1} es acotado.

Demostración. Como A es positivo definido

$$< Ax, x > \geqslant m < x, x >$$

para toda $x \in D$, con $m > 0$. Aplicando la desigualdad de Schwarz obtenemos que $m \| x \| \leqslant \| Ax \|$ para toda $x \in D$. De aquí se sigue que A^{-1} existe y además

$$\| A^{-1} y \| \leqslant m^{-1} \| y \|$$

para toda $y \in R(A)$. Por el teorema 2.21 sabemos que R(A) es denso en H. Ahora veremos que $R(A) = H$. Sea $y \in H$, entonces existe una sucesión $\{ y_n \} \subset R(A)$ tal que $y_n \to y$. Por otro lado,

$$\| A^{-1} y_n - A^{-1} y_m \| \leqslant m^{-1} \| y_n - y_m \|$$

lo cual implica que $\{ A^{-1} y_n \}$ es de Cauchy y entonces converge a un límite x. Esto es, $y_n \to y$ y $A^{-1} y_n \to x$. Como A^{-1} es cerrado, por ser autoadjunto, se sigue que $y \in R(A)$ y $A^{-1} y = x$. Esto muestra que $R(A) = H$. Como A^{-1} es cerrado, se sigue del teorema 2.16, que A^{-1} es acotado. Así queda demostrado el teorema.

Si además de ser autoadjunto, A es también acotado, tenemos

$$| < Ax, x > | \leqslant \| A \| < x, x >$$

y tomando x tal que $\| x \| = 1$, obtenemos

$$| < Ax, x > | \leqslant \| A \| .$$

y de aquí que

$$\sup_{\|x\|=1} |< Ax, x >| \leqslant \|A\|$$

De hecho se da la igualdad.

Teorema 2.23. Si A es acotado y autoadjunto, entonces

$$\|A\| = \sup_{\|x\|=1} |< Ax, x >| .$$

Para la demostración de este teorema nos remitimos a [11].

Corolario. Si A es acotado y autoadjunto, entonces

$$\|A\| = \text{máx} \left\{ |m|, \ |M| \right\}$$

donde $m = \inf_{\|x\|=1} < Ax, x >$ y $M = \sup_{\|x\|=1} < Ax, x >$

Desarrollada la herramienta necesaria, pasaremos al estudio de Cálculo Operacional para operadores en espacios de Hilbert.

CAPITULO 3. CALCULO OPERACIONAL PARA FUNCIONES RACIONALES

3.0 Introducción

Estamos interesados en definir f(A), siendo $f : \Delta \subset C \to C$. En este capítulo trataremos de definir f(A) para polinomios y funciones racionales. Veremos cómo esto nos plantea la necesidad de conocer el espectro de A y cómo este conjunto juega un papel esencial para poder definir f(A).

Si f(z) es un polinomio de grado n no tenemos dificultad de definir f(A), siempre y cuando el rango de A^k esté contenido en el dominio de A para $K < n$. Si $f(z) = a_0 + a_1 z + \ldots + a_n z^n$ definimos

$$f(A) = a_o I + a_1 A + \ldots + a_n A^n$$

Si $Q(z) = p(z)/q(z)$, donde p(k) y q(x) son polinomios, es natural definir Q(A) como $Q(A) = p(A)\ Q(A)^{-1}$. Las dificultades aparecen de inmediato ya que $q(A)^{-1}$ puede no existir. Necesitamos, pues, determinar cuándo es posible definir Q(A) donde $Q(z) = 1/p(z)$.

En los casos que analizaremos en este capítulo, estudiaremos primero el caso $Q(z) = 1/p(z)$ con $p(z)$ de primer grado, y a continuación el caso general.

3.1 Definiciones y convenciones

Por comodidad supondremos que $p(z) = \lambda - z$, cuando $p(z)$ sea de primer grado.

Pondremos una fuerte restricción en este problema. Pediremos que $(\lambda I - A)^{-1}$ sea acotado, y para esto es necesario que $R(\lambda I - A) = H$.

Definición. Al conjunto de puntos en C tales que $(\lambda I - A)^{-1}$ existe y es acotado lo llamaremos el conjunto resolvente de A y lo denotaremos por $\rho(A)$.

Definición. Definimos el espectro de A, $\sigma(A)$, como el espectro de A.

Definición. Si $\lambda \epsilon \rho(A)$, a $(\lambda I - A)^{-1}$ le llamaremos la resolvente de A.

Definición. Si $\lambda \epsilon \sigma(A)$ y $Ax = \lambda x$ para alguna $x \neq 0$ diremos que λ es un valor propio de A y x su correspondiente vector propio.

3.2 Definición de Q(A) cuando A es un operador en un espacio de dimensión finita

En el caso matricial, los resultados, ya obtenidos, que nos interesan los enunciaremos a continuación.

Teorema 3.1. Si H es de dimensión finita y $A : H \rightarrow H$ es lineal, entonces $(\lambda I - A)^{-1}$ existe y es acotado para toda $\lambda \epsilon C$ excepto para un número finito de puntos. Esto es, $\sigma(A)$ es un conjunto finito.

Este resultado nos resuelve el problema de definir $Q(z) = 1/p(z)$ cuando $p(z)$ es de primer grado.

Teorema 3.2. Si $p(z)$ es cualquier polinomio, entonces

$$P(\sigma(A)) = \sigma(p(A)).$$

Corolario. Si $p(z)$ es cualquier polinomio, entonces $p(A)^{-1}$ existe si y sólo si $p(z) \neq 0$ para toda $z \in \sigma(A)$.

Este corolario resuelve el problema de definir $Q(A)$ cuando $Q(z) = 1/p(z)$, $p(z)$ cualquier polinomio, y en consecuencia también para cuando $Q(z) = p(z)/q(z)$.

3.3 Definición de Q(A) cuando A es un operador compacto

Sea ahora A compacto. Recordando que A se puede aproximar por operadores de dimensión finita es de esperarse que pase algo parecido a lo que pasa en el caso matricial. Pero por ese mismo hecho no podemos esperar que $\sigma(A)$ siga siendo finito, pero sí a lo más numerable como lo muestran los siguientes teoremas que enunciamos sin demostración (véase [20]).

Teorema 3.3. Sea $A : H \to H$ compacto. Si $\lambda \neq 0$, entonces λ es un valor propio de A o λ es un elemento de $\rho(A)$.

Teorema 3.4. El conjunto de valores propios de un operador compacto es a lo más numerable, y cero es el único posible punto de acumulación de este conjunto.

Corolario. El espectro de un operador compacto consiste de la unión del cero y el conjunto de valores propios del operador. En consecuencia $\sigma(A)$ es a lo más numerable.

Esto nos resuelve el problema de definir $Q(A)$ cuando $Q(z) = (\lambda - z)^{-1}$.

Consideremos ahora un polinomio $p(z)$ de grado n, y sea λ un valor propio de A, puesto que $Ax = \lambda x$ para alguna $x \neq 0$, entonces $p(A)x = p(\lambda)x$; esto es, $p(\lambda) \epsilon\, \sigma(p(A))$. Además, como $p(A)$ es compacto, $o\ \epsilon\ \sigma(p(A))$.

Lo anterior significa que

$$p(\sigma(A)) \subset \sigma\ (p(A))$$

Veremos que la contención en el otro sentido también es válida. De paso quedará demostrada la misma afirmación para el caso matricial.

Teorema 3.5. Si A es compacto y p(z) es un polinomio, entonces

$$p(\sigma(A)) = \sigma(p(A))$$

Demostración. Sólo nos resta demostrar una contención. El polinomio $\lambda - p(z)$ puede descomponerse como

$$\lambda - p(z) = C(z - \lambda_1) \ldots (z - \lambda_n),$$

donde $C \neq 0$ y las λ_i son las raíces de $\lambda - p(z)$. Entonces

$$\lambda I - p(A) = C(A - \lambda_1 O) \ldots (A - \lambda_n I).$$

Si $\lambda_1, \lambda_2, \ldots, \lambda_n \,\epsilon\, \rho(A)$, entonces $\lambda\, \epsilon\, \rho(p(A))$. De aquí que si $\lambda\, \epsilon\, \sigma(p(A))$ entonces para alguna k, $\lambda_k \in \sigma(A)$, y en consecuencia $\lambda = p(\lambda_k)$. Esto demuestra que $\sigma(p(A)) \subset p(\sigma(A))$.

Como consecuencia inmedi.ta tenemos:

Teorema 3.6. Si A es compacto y $p(z)$ es cualquier polinomio entonces $p(A)^{-1}$ existe si y solo si $p(z) \neq 0$ para toda $z \in \sigma(A)$.

Esto nos resuelve el problema de definir $Q(A)$, siendo A compacto y $Q(z)$ cualquier función racional.

3.4 Definición de Q(A) cuando A es un operador acotado

Sea A acotado. De la cardinalidad del espectro de A no podemos asegurar nada. Ya vimos, en las secciones anteriores, que éste puede ser finito o numerable, pero en general puede no ser numerable como lo muestra el siguiente ejemplo.

3.4.1 Un operador acotado con espectro no numerable

Sea $E : L_2(a, b) \rightarrow L_2(a, b)$ dado por $(Ef)(x) = xf(x)$. Es fácil ver que E es acotado $-\alpha = \text{máx}\{|a|, |b|\}$ es una cota para E $-$ y que E es autoadjunto.

Veremos que todo punto de $[a, b]$ es un elemento de $\sigma(E)$. La idea es la siguiente: Si $\lambda \in [a, b]$, veremos que $(\lambda I - E)$ es inyectiva, entonces $(\lambda I-E)^{-1}$ existe en $R(\lambda I - E)$, pero $(\lambda I - E)^{-1}$ no es acotado en $R(\lambda I - E)$. Sea, pues, $\lambda \in [a, b]$ y f tal que $(\lambda I - E) f = 0$, entonces

$$\|(\lambda I - E)f\|^2 = \int_a^b |\lambda - x|\,|f(x)|^2\, dx = 0$$

Como $|\lambda - x| > 0$ para toda $x \in [a, b]$ excepto para $x = \lambda$, se concluye que $f(x)=0$ para casi toda $x \in [a, b]$. De aquí que $(\lambda I - E)^{-1}$ existe en $R(\lambda I - E)$. Veremos que este operador no es acotado. Para esto construiremos una sucesión $\{g_n\}$ en $R(\lambda I - E)$ tal que $\| g_n \| = 1$ y $\|(\lambda I - E)^{-1} g_n\| \longrightarrow \infty$. Primero construiremos $\{f_n\}$ tal que $\|(\lambda I - E)f_n\| \longrightarrow 0$. Sea

$$C_n(x) = \begin{cases} 1 \text{ si } x \in [\lambda - 1/n, \lambda + 1/n] \quad [a, b] \\ 0 \text{ si } x \notin [\lambda - 1/n, \lambda + 1/n] \quad [a, b] \end{cases}$$

Entonces $\|C_n\| > 0$ y definimos $f_n = C_n / \|C_n\|$. En consecuencia

$$\|(\lambda I - E) f_n\|^2 = \int_a^b |\lambda - x|^2 \, |f_n(x)|^2 \, dx \leqslant \frac{1}{n^2} \int_a^b |f_n(x)|^2 \, dx = \frac{1}{n^2}$$

Sea ahora $g_n = (\lambda I - A) f_n / \|(\lambda I - A) f_n\|$, $g_n \in R(\lambda I - E)$ $\|g_n\| = 1$ y

$$\|(\lambda I - E)^{-1} g_n\| = \frac{1}{\|(\lambda I - E) f_n\|} > n.$$

Entonces $(\lambda I - E)^{-1}$ no es acotado y por lo tanto $\lambda \in \sigma(E)$ para toda $\lambda \in [a, b]$. De hecho $\sigma(E) = [a, b]$, pero no veremos la otra contención ([7]).

3.4.2 El espectro de un operador acotado

El espectro de A en los casos anteriores tiene una propiedad común: la de ser un conjunto compacto. En el caso matricial el espectro es un conjunto finito y por lo tanto compacto. Cuando A es compacto, el espectro es finito o es infinito formando, en este último caso, una sucesión convergente a cero, y cero es un punto del espectro; luego entonces el espectro es compacto.

Este resultado sigue siendo válido en el caso presente. Para demostrar esto nos basaremos en los siguientes lemas. (Véanse [11] y [20], respectivamente).

Lema 1. Si A es acotado y $\|A\| < 1$, entonces $(I - A)^{-1} \in L(H)$ y

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k .$$

Lema 2. Si A es acotado, entonces $\rho(A)$ es abierto y $\rho(A) \neq C$.

Teorema 3.7. El espectro de un operador acotado es un conjunto compacto.

Demostración. Probaremos que $\sigma(A)$ es cerrado y acotado. Que es cerrado se sigue del Lema 2. Para ver que es acotado nótese que si λ es un valor propio y x su correspondiente vector propio normalizado, entonces $|\lambda| = |\lambda x| = |Ax| \leqslant \|A\|$. Esto es, el conjunto de valores propios de A está acotado por $\|A\|$. Veremos que $\|A\|$ constituye una cota para $\sigma(A)$. Para esto es suficiente mostrar que si $|\lambda| > \|A\|$, $\lambda \epsilon \rho(A)$. Si $|\lambda| > \|A\|$ entonces $(\lambda I - A) = \lambda(I - \lambda^{-1}A)$ y $\|\lambda^{-1}A\| < 1$, de aquí que $(I - \lambda^{-1} A)^{-1} \epsilon\ L(H)$ y en consecuencia $(\lambda I - A)^{-1} \epsilon\ L(H)$; esto es, $\lambda \epsilon \rho(A)$. Esto demuestra el teorema.

Con estos resultados, ya sabemos cuándo podemos definir $Q(A)$ si $Q(z) = (\lambda - z)^{-1}$.

Teorema 3.8. Si A es acotado y $p(z)$ es un polinomio, entonces

$$p(\sigma(A)) = \sigma(p(A)).$$

Demostración. La demostración de que $\sigma(p(A)) \subset p(\sigma(A))$ es idéntica al caso en que A es compacto. (Teorema 3.5). Veamos la otra contención. El polinomio $\lambda - p(z)$ puede ser expresado como

$$\lambda - p(z) = C\ (z - \lambda_1) \ldots (z - \lambda_n),$$

donde $C \neq 0$ y las λ_i son las raíces de $\lambda - p(z)$, entonces

$$\lambda I - p(A) = C(A - \lambda_1 I) \ldots (A - \lambda_n I). \qquad (3.1)$$

Supongamos que $\lambda \in p(\sigma(A))$, esto significa que $\lambda = p(z_0)$ (o bién $\lambda - p(z_0) = 0$) para alguna $z_0 \in \sigma(A)$, pero como las λ_i son las raices de $\lambda - p(z)$, entonces $z_0 = \lambda_k$ para alguna k. Supongamos que $k = 1$, esto es, $\lambda = p(\lambda_1)$ y $\lambda_1 \in \sigma(A)$. Esto último implica que $A - \lambda_1 I$ no tiene inverso acotado, y se presentan, para $A - \lambda_1 I$, dos posibilidades:

i) $A - \lambda_1 I$ no tiene inverso. En este caso existe $x \neq 0$ tal que $(A-\lambda_1 I)x = 0$. Si intercambiamos, en (3.1), $A - \lambda_1 I$ y $A - \lambda_n I$, obtendremos que

$$(\lambda I - p(A))\, x = 0$$

para una $x \neq 0$. Esto implica que $\lambda \in \sigma(p(A))$.

ii) Puede suceder que $A - \lambda_1 I$ tenga inverso, pero en este caso su rango no es todo H. De esto y de la ec (3.1), se sigue que el rango de $\lambda I - p(A)$ no es todo H. Entonces $\lambda \in \sigma(p(A))$. Así queda demostrado el teorema.

Como consecuencia inmediata de esto tenemos el siguiente resultado.

Teorema 3.9. Si A es acotado y p(z) es cualquier polinomio, entonces $p(A)^{-1}$ existe si y sólo si $p(z) \neq 0$ para toda $z \in \sigma(A)$.

Esto resuelve, nuevamente, el problema de definir Q(A) cuando Q(z) es una función racional.

3.5 Una observación importante

Hemos visto que el poder definir Q(A) depende del espectro de A. En los casos anteriores se pudo definir Q(A) para cuando el denominador de Q(z) era distinto de cero en un conjunto compacto, a saber $\sigma(A)$.

En los siguientes casos a estudiar, es más difícil determinar $\sigma(A)$ por lo que sólo veremos algunas de las propiedades más importantes del espectro.

3.6 El espectro de un operador autoadjunto

Cuando H es de dimensión finita es conocido que A posee valores propios, esto es, el espectro de cualquier operador autoadjunto es no vacío. Además el espectro es real ya que

$$\lambda <x, x> = <\lambda x, x> = <Ax, x> = <x, Ax> = <x, \lambda x> = \bar{\lambda} <x, x> \qquad (3.2)$$

Si A es compacto autoadjunto su espectro es no vacío ([7]).

Teorema 3.10. Sea A un operador compacto autoadjunto. Entonces existe un valor propio λ_o de A, donde $|\lambda_o| = \|A\|$.

Corolario. Si A es compacto autoadjunto

$$|\lambda_o| = \|A\| \geqslant |\lambda|$$

para toda $\lambda \in \sigma(A)$.

Teorema 3.11. Si A es compacto y autoadjunto, entonces su espectro, $\sigma(A)$, es real.

Demostración. Por el corolario del teorema 3.3, sabemos que $\sigma(A)$ consta del cero y de sus valores propios. Entonces sólo resta demostrar que los valores propios de A son reales. Esto se sigue de la ec (3.2).

En este par de casos que acabamos de ver, al demostrar que los valores propios son reales – y por consiguiente el espectro– se usó únicamente el hecho de que A es autoadjunto. De aquí que los valores propios de cualquier operador autoadjunto sean reales. Esto hace plausible que el espectro de cualquier operador autoajunto sea real. La demostración de este hecho es, por supuesto, más complicada y no la haremos aquí.

Para la demostración de los siguientes resultados nos remitiremos a [1].

Teorema 3.12. Si A es acotado y autoadjunto, entonces $\sigma(A)$ es real y además

i) $\sigma(A) \subset [m, M]$ donde

$$m = \inf_{\|x\|=1} \langle Ax, x\rangle \ , \quad M = \sup_{\|x\|=1} \langle Ax, x\rangle$$

ii) $m, M \in \sigma(A)$

Teorema 3.13. Si A es autoadjunto, entonces $\sigma(A)$ es un conjunto cerrado.

3.7 El espectro de un operador cerrado

Una de las dificultades que presenta el determinar $\sigma(A)$, en este caso general, es que $\sigma(A)$ puede no ser acotado como veremos en el siguiente ejemplo.

3.7.1 Un operador cuyo espectro no es acotado

Consideremos el siguiente operador:

$$Au = -\frac{d^2u}{dx^2}$$

con $u(0) = u(1) = 0$. Veremos que el espectro de este operador no es acotado. Para esto basta fijarse en que sus valores propios forman la sucesión $\{ n^2 \pi^2 \}$, siendo sus correspondientes vectores propios las funciones sen $n\pi x$.

3.7.2 El espectro de un operador cerrado es cerrado

Aunque en general es espectro de un operador cerrado no es acotado, sigue conservando su cualidad de ser cerrado. Esto nos lo asegura el siguiente teorema que no demostraremos [20]

Teorema 3.14. Si A es cerrado, entonces $\sigma(A)$ es un conjunto cerrado.

CAPITULO 4. CALCULO OPERACIONAL PARA OPERADORES ACOTADOS. CASO GENERAL

4.0 Introducción

El propósito de este capítulo es, primeramente, definir f(A) cuando A es acotado y $f : \Delta \subset C \rightarrow C$ es una función analítica. Después veremos, como una aplicación de lo anterior, una clase importante de operadores -proyecciones- que nos llevarán a final de cuentas a obtener una representación para A en términos de dichas proyecciones en el caso en A sea autoadjunto.

4.1 Funciones analíticas con valores en L(H)

4.1.1 Introducción

En el capítulo anterior veíamos que la definición de Q(A), cuando Q(z) es racional, dependía de la posibilidad de que pudiésemos definir la resolvente de A

$$R_\lambda = (\lambda I - A)^{-1}$$

Por otro lado en el capítulo 1 veíamos cómo se sugería definir f(A) –siendo A una matriz y f(z) analítica– utilizando la fórmula de Cauchy; esto es, definir f(A) como

$$f(A) = \frac{1}{2\pi i}\int_{\Gamma} f(z)\ (zI - A)^{-1}\ dz$$

Expresión en la cual aparece la resolvente de A.

Es este camino que seguiremos en el presente capítulo para definir f(A).

Para abordar el problema haremos primero una reflexión sobre la resolvente R_λ.

Si λ es un complejo tal que $|\lambda| \geqslant \|A\|$, sabemos por el teorema 3.7 que $\lambda \in \rho(A)$ y además, por el lema anterior a dicho teorema, tenemos que

$$R_\lambda = (\lambda I - A)^{-1} = \sum_{k=1}^{\infty} \lambda^{-n} A^{n-1}.$$

O sea que en el exterior del círculo con centro en el origen y radio $r = \|A\|$, R_λ se puede desarrollar en serie de potencias en λ con coeficiente en L(H).

Lo anterior sugiere la posibilidad de considerar a R_λ como una función analítica, cuyo dominio es un subconjunto del plano complejo, con valores en L(H).

Esto nos lleva a tratar de precisar el concepto de función analítica cuando

$$F : \Delta \subset C \to L(H).$$

4.1.2 Definiciones y propiedades básicas

Definición. Decimos que $F : \Delta \subset C \rightarrow L(H)$, Δ abierto, es diferenciable en un punto $\lambda_o \in \Delta$ si existe $A \in L(H)$ tal que

$$\left\| \frac{F(\lambda) - F(\lambda_o)}{\lambda - \lambda_o} - A \right\| \rightarrow 0 \quad \text{si } \lambda \rightarrow \lambda_o$$

y en este caso definimos $F'(\lambda_o) = A$.

Definición. Sea $F : \Delta \subset C \rightarrow L(H)$, Δ abierto. Decimos que $F(\lambda)$ es analítica en Δ si $F(\lambda)$ es diferenciable para cada $\lambda \in \Delta$.

A continuación enunciaremos una serie de propiedades de este tipo de funciones. La demostración de estas propiedades se hará en los Apéndices A y B.

Teorema 4.1. Sea $F : \Delta \subset C \rightarrow L(H)$, Δ abierto. Una condición necesaria y suficiente para que $F(\lambda)$ sea analítica en Δ es que $<F(\lambda)\, x,\ y>$ sea analítica para toda $x, y \in H$.

Este resultado permite que las principales propiedades de la teoría de funciones analíticas se sigan conservando. En particular podemos reconstruir la integral de contorno y como consecuencia de esto, el teorema de Cauchy.

En el Apéndice A veremos que es posible definir la integral

$$\int_{\Gamma} F(z)\, dz \tag{4.1}$$

donde Γ es una curva rectificable, y se demostrará el siguiente resultado.

Teorema 4.2. La integral dada por (4.1) es un elemento de L(H) y

$$< \quad F(z)\ dz\ x,\ y> = \int_{\Gamma} <F(z)\ x,\ y>\ dz.$$

La existencia de la integral del lado derecho garantiza la existencia de (4.1).

Teorema 4.3. Si $F : \Delta \subset C \rightarrow L(H)$ es analítica, entonces

$$\int_{\Gamma} F(z)\ dz = 0$$

para cualquier curva Γ simple cerrada en Δ, cuyo interior esté contenido en A.

Corolario. Sea $F : \Delta \subset C \longrightarrow L(H)$ analítica y sean Γ_1, Γ_2 dos curvas simples cerradas, que no se interesectan entre sí, tales que la región acotada por ellas esté contenida en Δ. Entonces

$$\int_{\Gamma_1} F(z)\ dz = \int_{\Gamma_2} F(z)\ dz.$$

El teorema de Cauchy sigue valiendo en sus casos más generales.

También se pueden reconstruir resultados clásicos de la teoría de funciones analíticas tales como el teorema de Liouville, los desarrollos en series de Taylor y de Laurent, el intercambio de sumatorias con la integral etc.

Teorema 4.4. La resolvente $R_z : \rho(A) \subset C \rightarrow L(H)$ es una función analítica en $\rho(A)$.

4.2 Definición de f(A)

4.2.1 Una definición restringida

Sean $r > \|A\|$, Δ es disco abierto con centro en el origen y radio r, y $f: \Delta \subset C \rightarrow C$ analítica. Entonces

$$f(z) = \sum_{n=0}^{\infty} a_n \, z^n, \qquad |z| < r. \tag{4.2}$$

Esto sugiere que definamos f(A) como

$$f(A) = \sum_{n=0}^{\infty} a_n \, A^n ,$$

pero debemos asegurarnos de que esta serie sea convergente en L(H).

Como $\|A\| < r$, entonces

$$\sum_{n=0}^{\infty} a_n \, z^n$$

converge absolutamente para $|z| = \|A\|$, esto es,

$$\sum_{n=0}^{\infty} |a_n| \, \|A\|^n$$

es convergente. Ahora bien

$$\| \sum_{k=n+1}^{p} a_k \, A^k \| \leq \sum_{k=n+1}^{p} |a_k| \, \|A\|_k$$

y entonces

$$\sum_{n=0}^{\infty} a_n \, A^n$$

converge en L(H).

Definición. Siendo f(z) dada por (4.2), definimos

$$f(A)= \sum_{n=0}^{\infty} a_n \; A^n . \qquad (4.3)$$

4.2.2 Hacia una definición general

Sea Γ un círculo con centro en el origen y radio mayor que $\| A \|$ – esto es, Γ está contenido en $\rho(A)$ – entonces

$$(zI - A)^{-1} = \sum_{n=1}^{\infty} z^{-n} \; A^{n-1}$$

si $z \epsilon \Gamma$. De aquí que

$$z\,(zI - A)^{-1} = \sum_{n=0}^{\infty} z^{-n} \; A^n$$

En vista de que Γ es un conjunto compacto, podemos intercambiar esta sumatoria con la integral y obtenemos

$${}_{\Gamma}\; z\,(zI - A)^{-1} dz = \sum_{n=0}^{\infty} \int_{\Gamma} z^{-n} \; A^n \, dz = \sum_{n=0}^{\infty} \int_{\Gamma} z^{-n} \, dz \quad A^n = 2\,\pi_i \, A$$

Esto es

$$A = \frac{1}{2\,\pi i} \int_{\Gamma} z\,(zI - A)^{-1} dz$$

Por un razonamiento análogo obtenemos:

$$A^n = \frac{1}{2\,\pi i} \int_{\Gamma} z^n \; (zI - A)^{-1} \, dz$$

Las fórmulas anteriores siguen siendo válidas para cualquier curva simple cerrada que contenga en su interior a $\sigma(A)$, ya que $R_z = (zI - A)^{-1}$ es analítica en $\rho(A)$. Tenemos, pues, el siguiente resultado.

Teorema 4.5. Si $A : H \to H$ es acotado, entonces

$$A^n = \frac{1}{2\pi i} \int_\Gamma z^n \, (zI - A)^{-1} \, dz \qquad (4.4)$$

siendo Γ una curva simple cerrada conteniendo a $\sigma(A)$ en su interior.

Combinando (4.2) con (4.4) obtenemos

$$f(A) = \sum_{n=0}^{\infty} \frac{a_n}{2\pi i} \int_\Gamma z^n \, (zI - A)^{-1} dz = \frac{1}{2\pi i} \int_\Gamma \sum_{n=0}^{\infty} a_n \, z^n \, (zI - A)^{-1} \, dz$$

de donde

$$f(A) = \frac{1}{2\pi i} \int_\Gamma f(z) \, (zI - A)^{-1} \, dz, \qquad (4.5)$$

siendo Γ un círculo con centro en el origen, radio menor que r – r tomada como en la sección 4.2.1 – y que contiene a $\sigma(A)$ en su interior.

Notando que la integral en (4.5) no cambia si consideramos a Γ como una curva simple cerrada que contenga a $\sigma(A)$ en su interior y esté contenida en el círculo con centro en el origen y radio r, vemos que mediante la fórmula (4.5) es posible definir f(A) para una clase más amplia de funciones analíticas. A continuación veremos que esto es posible hacerlo para cualquier función cuyo dominio contenga a $\sigma(A)$.

4.2.3 Definición de f(A)

Si $f : \Delta \subset C \to C$ es analítica en Δ, Δ conteniendo a $\sigma(A)$ y si existe un círculo Γ contenido en Δ, con centro en el origen y conteniendo a $\sigma(A)$

en su interior (fig 2), podemos definir f(A) por medio de la fórmula (4.5).

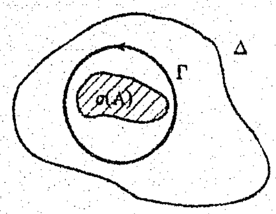


Fig 3

Aun más, si Δ es tal que podemos encontrar una curva simple cerrada Γ en Δ ∩ ρ(A) (fig 4), la integral del lado derecho de (4.5) esta bien definida y podemos definir f(A) por medio de dicha fórmula.

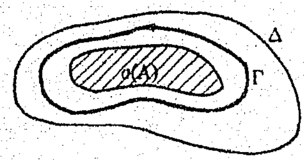


Fig 4

Pero se puede presentar una situación más general. Pudiese ser que Δ y σ(A) sean como los de la fig 5. Si $\Gamma = \Gamma_1 \cup \Gamma_2$, la integral de (4.5) sigue estando bien definida y podemos definir f(A) por medio de esa fórmula.

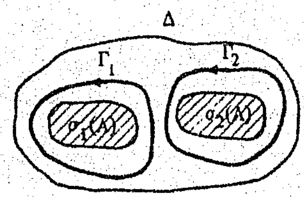


Fig 5

En cada uno de los casos anteriores Δ es un conjunto abierto que contiene a σ(A). Haremos ver que si f tiene como dominio a un conjunto abierto Δ conteniendo a σ(A), podemos definir f(A) por medio de la fórmula (4.5).

El problema es asegurarnos de que podemos encontrar una curva Γ —contenida en Δ ∩ ρ(A)— adecuada para que la integral en (4.5) tenga sentido. El hecho de que σ(A) sea compacto nos va a garantizar que podamos encontrar una curva tal. Esto nos lo asegura el siguiente lema que no demostraremos ([19]).

Lema. Dado $S \subset C$ compacto y $\Omega \supset S$ abierto, existe un conjunto abierto, existe un conjunto abierto ω tal que

$$S \subset \omega \subset \bar{\omega} \subset \Omega$$

y su frontera, $\partial\omega$, consiste de un número finito de curvas simples cerradas que no se intersectan entre sí.

Definición. Si $f : \Delta \subset C \to C$ es analítica, Δ abierto y $\Delta \supset \sigma(A)$ entonces definimos

$$f(A) = \frac{1}{2\pi i} \int_{\partial\omega} f(z)\ (zI - A)^{-1}\ dz, \qquad (4.6)$$

siendo $\partial\omega$ como en el lema anterior.

Ahora, para que f(A) quede bien definida, debemos asegurarnos que la integral en (4.6) no depende del conjunto ω escogido. Esto es una aplicación del teorema de Cauchy y la demostración puede encontrarse en [20].

4.3 Cálculo operacional

Lo primero que podemos observar de (4.6) es que $f(A) \in L(H)$.

De la definición de f(A) se sigue, sin mucha dificultad que

$$(f + g)\ (A) = f(A) + g(A)$$

y que para cualquier $\alpha \in C$

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

Un hecho que no es trivial –y que demostraremos a continuación– es que

$$(fg)\ (A) = f(A)\ g(A)$$

Esta propiedad implicará que si $f(A)^{-1}$ existe

$$f(A)^{-1} = g(A)$$

donde $g(z) = 1/f(z)$.

Teorema 4.6. Si $f : \Delta \subset C \to C$ es analítica y $\Delta \supset \sigma(A)$ es abierto, entonces

i) $(f + g)(A) = f(A) + g(A)$

ii) $(\alpha f)(A) = \alpha f(A),\ \alpha \in C$

iii) $(fg)(A) = f(A)\, g(A) = g(A)\, f(A)$

iv) Si $f(z) \neq 0$ para toda $z \in \sigma(A)$, entonces $f(A)$ es invertible y

$$f(A)^{-1} = \frac{1}{2\pi i}\int_{\partial\omega} \frac{1}{f(z)}(zI - A)^{-1}\, dz$$

donde $\partial\omega$ es la frontera de un conjunto abierto ω con la propiedad de que $\sigma(A) \subset \omega \subset \overline{\omega}$ y tal que $f(z)$ no se anula en $\overline{\omega}$.

Demostración.

iii) La demostración se basa en la igualdad

$$R_z - R_\lambda = (\lambda - z)\, R_z\, R_\lambda$$

que se demuestra en el Apéndice B.

Sean f con dominio Ω, g con dominio Δ. Escojamos conjuntos abiertos ω_1, ω_2 tales que

$$\sigma(A) \subset \omega_1\ , \quad \overline{\omega}_1 \subset \omega_2\ , \quad \overline{\omega}_2 \subset \Omega \cap \Delta$$

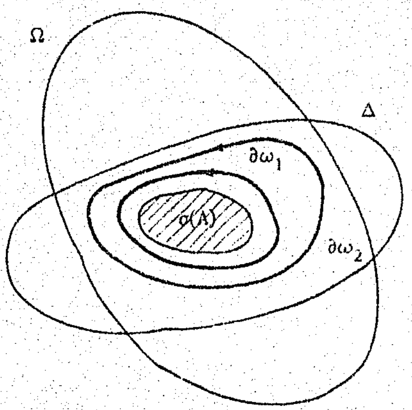


Fig 6

(ver fig (6). Entonces

$$f(A) = \frac{1}{2\pi i}\int_{\partial\omega_1} f(z)\, R_z \; dz$$

$$g(A) = \frac{1}{2\pi i}\int_{\partial\omega_2} g(\lambda)\, R_\lambda \; d\lambda$$

y en consecuencia

$$f(A)\; g(Z) = \frac{1}{2\pi i}\int_{\partial\omega_1} f(z)\; R_z \left\{ \frac{1}{2\pi i}\int_{\partial\omega_2} g(\lambda)\; R_\lambda \; d\lambda \right\} \; dz.$$

$$= \frac{1}{2\pi i}\int_{\partial\omega_1} f(z) \left\{ \frac{1}{2\pi i}\int_{\partial\omega_2} g(\lambda)\; R_z \; R_\lambda \; d\lambda \right\} \; dz$$

Sustituyendo $R_\lambda \; R_z$ por

$$\frac{R_z - R_\lambda}{\lambda - z} = \frac{R_z}{\lambda - z} + \frac{R_\lambda}{z - \lambda}$$

obtenemos

$$f(A)\; g(A) = \frac{1}{2\pi i}\int_{\partial\omega_1} f(z)\, R_z \left\{ \frac{1}{2\pi i}\int_{\partial\omega_2} \frac{g(\lambda)}{\lambda - z} \; d\lambda \right\} \; dz$$

$$+ \frac{1}{2\pi i}\int_{\partial\omega_1} f(z) \left\{ \frac{1}{2\pi i}\int_{\partial\omega_2} \frac{g(\lambda)\, R_\lambda}{z - \lambda} \; d\lambda \right\} dz$$

$$= \frac{1}{2\pi i}\int_{\partial\omega_1} f(z)\, R_z \left\{ \frac{1}{2\pi i}\int_{\partial\omega_2} \frac{g(\lambda)}{\lambda - z} \; d\lambda \right\} \; dz$$

$$+ \frac{1}{2\pi i}\int_{\partial\omega_2} g(\lambda)\, R_\lambda \left\{ \frac{1}{2\pi i}\int_{\partial\omega_1} \frac{f(z)}{z - \lambda} \; dz \right\} \; d\lambda$$

Pero como z está en el interior de $\partial\omega_2$ y λ en el exterior de $\partial\omega_1$,

tenemos que

$$\frac{1}{2\pi i}\int_{\partial\omega_2}\frac{g(\lambda)}{\lambda - z}\,d\lambda = g(z) : \frac{1}{2\pi i}\int_{\partial\omega_1}\frac{f(z)}{z-\lambda}\,dz = 0$$

Luego entonces

$$f(A)\ g(A) = \frac{1}{2\pi i}\int_{\partial\omega_1} f(z)\ g(z)\ R_z\ dz = (fg)\ (A)\ ,$$

y como

$$(fg)\ (A) = (gf)\ (A) = g(A)\ f(A)$$

tenmos demostrada la proposición.

vi) Sea $f : \Delta \subset C \to C$, $\Delta \supset \sigma(A)$ abierto. Si $f(z) \neq 0$ en $\sigma(A)$, podemos construir un conjunto abierto ω tal que

$$\sigma(A) \subset \omega \subset \bar{\omega} \subset \Delta$$

y $f(z) \neq 0$ en $\bar{\omega}$.

Veremos cómo construir ω. Por ser f continua, para cada $z \in \sigma(A)$ existe un abierto conteniendo a z y contenido en Δ —esto por ser Δ abierto— en el cual f no se anula. La unión Ω de estos conjuntos es un conjunto abierto contenido en Δ y conteniendo a $\sigma(A)$. Sea ω un conjunto abierto tal que

$$\sigma(A) \subset \omega \subset \bar{\omega} \subset \Omega \subset \Delta\ .$$

Este conjunto ω es el buscado.

De lo anterior se concluye que $g(z) = 1/f(z)$ es analítica en Ω. Por lo tanto g(A) está definida y

$$f(A)\,g(A) = g(A)\,f(A) = \frac{1}{2\pi i}\int_{\partial\omega} f(z)\,g(z)\,R_z\,dz =$$

$$= \frac{1}{2\pi i}\int_{\partial\omega} R_z\,dz = I$$

Luego entonces

$$g(A) = f(A)^{-1}$$

y

$$f(A)^{-1} = \frac{1}{2\pi i}\int_{\partial\omega} \frac{1}{f(z)}(zI - A)^{-1}\,dz.$$

con lo que queda demostrado el teorema.

Un resultado importante es el llamado teorema de transformación del espectro que demostraremos a continuación.

Teorema 4.7. **(Transformación del espectro).** Si $f : \Delta \subset C \to C$ es analítica, $\Delta \supset \sigma(A)$ es abierto, entonces

$$\sigma(f(A)) = f(\sigma(A)).$$

Es decir, $\mu \in \sigma(f(A))$ si y sólo si $\mu = f(\lambda)$ para alguna $\lambda \in \sigma(A)$.

Demostración. Mostraremos primero que $\mu \in \sigma(f(A))$ implica $\mu = f(\lambda)$ para alguna $\lambda \in \sigma(A)$. Supongamos que esto no es cierto, esto es, que $f(z) - \mu \neq 0$ para toda $z \in \sigma(A)$. Por el teorema 4.6. $f(A) - \mu I$ tiene un inverso acotado y por lo tanto $\mu \in \rho(f(A))$; lo cual es una contradicción.

Supongamos ahora que $\mu = f(\lambda)$ para alguna $\lambda \in \sigma(A)$. Definamos

$$g(z) = \begin{cases} \dfrac{f(z) - \mu}{z - \lambda} & \text{si } z \neq \lambda \\[2ex] f'(\lambda) & \text{si } z = \lambda\,. \end{cases}$$

Entonces $g(z)$ es analítica en Δ y además

$$g(z)\,(z - \lambda) = f(z) - \mu$$

para toda $z \in \Delta$. De aquí que

$$g(A)\,(A - \lambda I) = (A - \lambda I)\,g(A) = f(A) - \mu I.$$

Si μ fuera un punto de $\rho(f(A))$, se tendría que

$$h(A)\,(A - \lambda I) = (A - \lambda I)\,h(A) = I$$

donde

$$h(A) = g(A)\,[f(A) - \mu I]^{-1}$$

Pero esto significaría que λ fuese un punto de $\lambda(A)$, lo cual contradice la hipótesis. Entonces $\mu \in \sigma(f(A))$ y esto completa la demostración.

4.4 Proyecciones

Puede suceder que $\sigma(A)$ no sea conexo – como en el caso en que A es compacto – y entonces el dominio Δ de $f(z)$ no tiene por qué ser conexo. En este caso $f(z)$ puede estar definida de distintas formas en las diferentes componentes conexas de Δ.

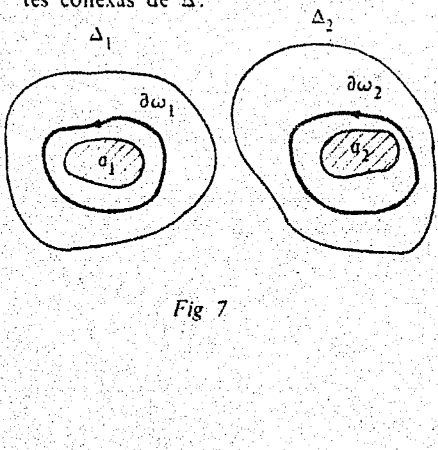


Fig 7

Supongamos, por ejemplo que $\sigma(A) = \sigma_1 \cup \sigma_2$, $\sigma_1 \cap \sigma_2 = \phi$, que existen abiertos Δ_1, Δ_2 tales que $\sigma_i \subset \Delta_i$, $i = 1, 2$, y $\Delta_1 \cap \Delta_2 = \phi$ (fig 7).

Definamos $f(z)$ como $f(z) \equiv 1$ en Δ_1 y $f(z) \equiv 0$ en Δ_2.

Entonces

$$P_1 = f(A) = \frac{1}{2\pi i}\int_{\partial\omega_1} (zI - A)^{-1}\, dz$$

donde ω_1 es un abierto tal que $\sigma_1 \subset \omega_1 \subset \bar{\omega}_1 \subset \Delta_1$ (Fig 7). Como $f(z)f(z)=f(z)$, tenemos que $P_1^2 = P_1$.

Definición. Un operador $P : H \to H$ se dice que es una proyección si $P^2 = P$.

Si ahora consideramos la función $g(z)$ definida como $g(z) \equiv 0$ en Δ_1 y $g(z) \equiv 1$ en Δ_2 obtenemos

$$P_2 = g(A) = \frac{1}{2\pi i}\int_{\partial\omega_2} (zI - A)^{-1}\, dz,$$

donde $\partial\omega_2$ es como en la fig 7. P_2 es también una proyección.

Como $f + g = 1$ en $\Delta = \Delta_1 \cup \Delta_2$, entonces

$$I = P_1 + P_2 . \qquad (4.6)$$

El rango de P_i, $R(P_i)$ consiste de aquellos puntos $x \in H$ tales que $x = P_i x$. En efecto, si $x = P_i z$ entonces $P_i x = P_i^2 z = x$.

Esto junto con (4.6) nos dice que

$$H = R(P_1) + R(P_2).$$

Aun más, si $x \in R(P_1) \cap R(P_2)$ entonces

$$x = P_1 x = P_2 x = (I - P_1)\, x = x - P_1 x$$

por lo tanto $P_1 x = 0$ y en consecuencia $x = 0$. Esto muestra el siguiente lema.

Lema. $H = R(P_1) \oplus R(P_2)$.

Así como H se puede descomponer en subespacios que dependen de proyecciones, el operador A también se puede descomponer en sumas de operadores determinados por las proyecciones. Para ver esto, supongamos que $x \in R(P_1)$, entonces $x = P_1 x$ y por lo tanto

$$P_1 A x = AP_1 x = A x$$

(el hecho de que $AP_1 = P_1 A$ proviene de que tanto A como P_1 son funciones del operador A y, por el teorema 4.6, conmutan). La igualdad anterior nos dice que $R(P_1)$ es invariante bajo A y que $A = AP_1$ en $R(P_1)$. Lo mismo sucede con $R(P_2)$.

Estas observaciones junto con el lema anterior nos dan el siguiente resultado.

Teorema 4.8. Bajo las hipótesis hechas a lo largo de esta sección, $R(P_1)$ y $R(P_2)$ son invariantes bajo A. En consecuencia, si $A_1 = AP_1$ y $A_2 = AP_2$, entonces A puede escribirse como

$$A = A_1 + A_2 .$$

Hemos descompuesto a A en la suma de operadores que en algunas ocasiones son más sencillos que A. Es importante observar que a final de cuentas los que determinan dichos operadores son el espectro de A y las proyecciones, que también dependen del espectro de A.

Si denotamos por A' y A" las restricciones de A a $R(P_1)$ y $R(P_2)$, respectivamente, es de esperarse que los espectros $\sigma(A')$ y $\sigma(A'')$ tengan relación con

σ_1 y σ_2. En efecto el siguiente resultado ([12]) nos establece dicha relación.

Teorema 4.9. $\sigma(A') = \sigma_1$; $\sigma(A'') = \sigma_2$.

Hemos recalcado que las transformaciones A_1 y A_2 en ocasiones son más sencillas que A. Con el fin de aclarar esto –que nos llevará a una representación de un operador autoadjunto en términos de proyecciones– analizaremos casos más particulares que el presente.

4.5 Descomposición espectral para operadores en espacios de dimensión finita

4.5.1 El caso general

El espectro de un operador en un espacio de dimensión finita consta únicamente de valores propios y constituye un conjunto finito. De aquí que sea posible encerrar a cada punto λ_i del espectro en un círculo Γ_i que no contenga a otros puntos del espectro (Fig 8). De aquí que A pueda representarse como

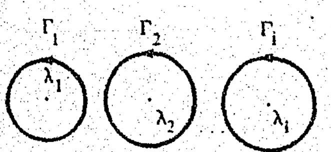


Fig 8

$$A = A_1 + \ldots + A_k$$

donde $A_i = AP_i$ y k es el número de elementos de $\sigma(A)$.

4.5.2 El caso en que A tiene n valores propios distintos

Supongamos que la dimensión de H es n y que A tiene precisamente n valores propios distintos. Puesto que

$$H = R(P_1) \oplus \ldots \oplus R(P_n)$$

se sigue que cada $R(P_i)$, $i = 1, \ldots, n$, es de dimensión 1

Utilizando el teorema 4.9 vemos que el espectro de la restricción de A a $R(P_i)$ es λ_i, y por ser $R(P_i)$ unidimensional, se sigue que dicha restricción es $\lambda_i I$. Esto es, para toda $x \in R(P_i)$

$$AP_i x = Ax = \lambda_i x = \lambda_i Px. \qquad (4.7)$$

Aun más, sea x un elemento cualquiera de H, entonces utilizando (4.7) se tiene que

$$AP_i x = A P_i^2 x = AP_i (P_i x) = A(P_i x) = \lambda_i P_i x$$

esto es

$$A_i = AP_i = \lambda_i P_i .$$

Luego entonces

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_n P_n . \qquad (4.8)$$

En este caso, A tiene una representación en términos de su espectro y de proyecciones asociadas a éste.

Una representación del tipo (4.8) es llamada una representación espectral de A.

En general un operador en un espacio de dimensión finita no tiene representación espectral, sin embargo cuakquier operador autoadjunto sí tiene siempre representación espectral.

4.5.3 Representación espectral para operadores autoadjuntos

Sea A es autoadjunto. Si λ_1 y λ_2 son valores propios de A; x_1, x_2 sus respectivos vectores propios y $\lambda_1 \neq \lambda_2$, entonces podemos asegurar que x_1 y x_2 son ortogonales. En efecto

$$\lambda_1 <x_1, x_1> = <\lambda_1 x_1, x_2> = <Ax_1, x_2> = <x_1, Ax_2> = \lambda_2 <x_1, x_2>$$

y de esta igualdad es evidente nuestra afirmación.

Lema. Si x_1, x_2 son vectores propios de A correspondientes a valores propios distintos, entonces x_1, x_2 son ortogonales.

Nota. Este resultado puede generalizarse fácilmente a un conjunto finito de vectores propios. Además, nótese que no estamos utilizando el hecho de que H es de dimensión finita, sino únicamente la propiedad de que A es autoadjunto, así que este resultado es válido aun en el caso en que H sea de dimensión infinita.

En el caso en que A tenga n valores propios distintos, la relación (4.8) sigue valiendo con algunas propiedades adicionales.

La primera de estas es que $R(P_i) \perp R(P_j)$ si $i \neq j$. Esto es evidente si notamos que

$$R(P_i) = \left\{ x \in H \mid Ax = \lambda_i x \right\}$$

y aplicamos el Lema anterior.

La otra propiedad es que los P_i son operadores autoadjuntos. En efecto, como los vectores propios de A, $x_1, x_2, \ldots, x_n$, forman una base ortogonal en H, si tomamos $x, y \in H$ tenemos

$$x = \sum_{k=1}^{n} \alpha_k \, x_k \quad , \qquad y = \sum_{k=1}^{n} \beta_k \, x_k$$

y de aquí que

$$<P_i x, y> = <\alpha_i x_i, \sum_{k=1}^{n} \beta_k \, x_k> = <\alpha_i x_i, \beta_i x_i> = <\sum_{k=1}^{n} \alpha_k \, x_k, \beta_i \, x_i> = <x, P_i \, y>$$

Los resultados anteriores siguen siendo esencialmente válidos para cualquier operador autoadjunto en un espacio de dimensión finita, aun cuando no todos sus valores propios sean distintos.

Se puede demostrar lo siguiente ([17]):

Teorema 4.10. Si $A : H \to H$ es autoadjunto, H de dimensión finita y si $\lambda_1, \ldots, \lambda_k$ son los distintos valores propios de A, entonces

i) $A = \sum_{i=1}^{k} \lambda_i \, P_i$

ii) $P_i^2 = P_i$: $P_i = P_i^*$; $P_i \, P_j = 0, \quad i \neq j.$

iii) $R(P_i) = \left\{ x \in H \mid Ax = \lambda_i x \right\}$.

iv) $H = R(P_1) \oplus \ldots \oplus R(P_k)$

v) $R(P_i) \perp R(P_j)$ si $i \neq j$.

Nota. Operadores teniendo la propiedad (ii) son llamadas proyecciones ortogonales.

En el caso en que H sea de dimensión infinita y A autoadjunto, seguirán valiendo resultados análogos a este teorema, pero con las debidas modificaciones debidas al caracter de H y al caracter del operador A.

4.6 Descomposición espectral para operadores compactos

4.6.1 El caso general

Sea $A : H \to H$ compacto. Si $\sigma(A)$ es un conjunto finito, entonces el análisis hecho en 4.5.1 se aplica al caso presente.

Supongamos, pues, que $\sigma(A)$ es infinito. Sin perder generalidad podemos suponer que $|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_n| \geqslant \ldots$ Escojamos un círculo Γ, con centro en el origen, tal que λ_i esté en el interior de Γ para $i > N$ y $\lambda_1, \ldots, \lambda_N$ estén en el exterior de Γ (supondremos que $\lambda_i \neq \lambda_j$ si $i \neq j$). Para cada λ_i $i \leqslant N$, podemos construir un círculo Γ_i con centro en λ_i y que no contenga en su interior a ningún otro punto de $\sigma(A)$ (Fig 9).

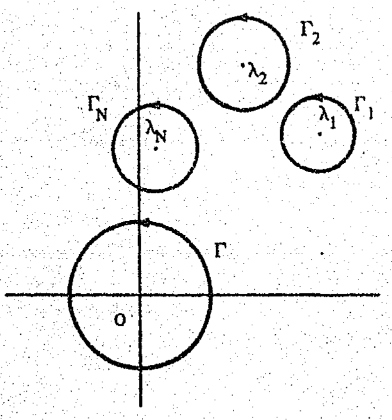


Fig 9

Tenemos entonces que

$$A = A_1 + \ldots + A_N + B$$

donde $A_k = AP_k$, $B = AP$ y

$$P_k = \frac{1}{2\pi i} \int_{\Gamma_k} (zI - A)^{-1} \, dz$$

$$P = \frac{1}{2\pi i} \int_{\Gamma} (zI - A)^{-1} \, dz$$

Nuevamente, si $R(P_i)$, $i = 1, \ldots, N$, es de dimensión uno, obtenemos

$$A = \lambda_1 P_1 + \ldots + \lambda_N P_N + B$$

Pero en general no tendremos esta situación. Sin embargo, en cualquier caso, el rango de P_i, $R(P_i)$ seguirá conservando su carácter finito.

Teorema 4.11. Si $A : H \to H$ es compacto; $\lambda_o \in \sigma(A)$, $\lambda_o \neq 0$; Γ_o es un círculo con centro en λ_o y que no contiene en su interior a ningun otro punto de $\sigma(A)$; y

$$P = \frac{1}{2\pi i} \int_{\Gamma_o} (zI - A)^{-1}\, dz$$

es la proyección asociada a λ_o. Entonces $R(P)$ es de dimensión finita.

Para una elegante demostración de este teorema véase [8].

4.6.2 Descomposición espectral para operadores compactos autoadjuntos

En el caso en que A es compacto autoadjunto seguimos teniendo

$$A = A_1 + \ldots + A_N + B$$

y

$$H = R(P_1) \oplus R(P_2) \oplus \ldots \oplus R(P_N) \oplus R(P)$$

y cada $R(P_i)$ es invariante bajo A.

Por el teorema 4.9, el espectro de la restricción de A a $R(P_i)$ es precisamente λ_i. Como la restricción de A a $R(P_i)$ es autoadjunto y $R(P_i)$ es de dimensión finita, se sigue que dicha restricción es $\lambda_i I$, esto es

$$Ax = \lambda_i x = \lambda_i P x$$

para toda $x \in R(P_i)$. Aun más sea $x \in H$, entonces

$$A_i x = AP_i x = \lambda_i P_i x.$$

Esto es

$$A_i = \lambda_i P_i$$

de donde

$$A = \lambda_1 P_1 + \ldots + \lambda_N P_N + B.$$

Podemos hacer tender a cero el radio del círculo Γ, equivalentemente, podemos hacer que N crezca. Parece más o menos natural que podamos tener

$$A = \sum_{i=1}^{\infty} \lambda_i P_i .$$

Veremos que esto es cierto. Tenemos

$$\| A - \sum_{i=1}^{N} \lambda_i P_i \| = \| B \|$$

Sólo resta mostrar que $\| B \| \to 0$ cuando $N \to \infty$.

Sabemos que el espectro de la restricción de A a R(P), llámemosle A', es $\{ \lambda_{N+1}, \ldots \} \cup \{ 0 \}$. De aquí que $\| A' \| = |\lambda_{N+1}|$ ya que A' es compacto autoadjunto. Ahora

$$\| Bx \| = \| APx \| = \| A'(Px) \| \leq |\lambda_{N+1}| \| Px \| \leq |\lambda_{N+1}| \| x \|$$

para toda $x \in H$. Entonces

$$\| B \| \leq |\lambda_{N+1}| .$$

Como el radio de Γ tiende a cero y λ_{N+1} está dentro de dicho círculo, entonces $|\lambda_{N+1}| \to 0$ y por lo tanto

$$A = \sum_{i=1}^{\infty} \lambda_i P_i .$$

Es natural esperar que

$$H = R(P_1) \oplus \ldots \oplus R(P_n) \oplus \ldots = \sum_{k=1}^{\infty} \oplus R(P_k)$$

en el sentido de que

$$x = \sum_{k=1}^{\infty} x_k$$

para toda $x \in H$ y donde $x_k \in R(P_k)$. Como $R(P_k)$ consiste de vectores propios de A correspondientes al valor propio λ_k y $R(P_k)$ es de dimensión finita, entonces en cada $R(P_k)$ podemos escoger una base ortonormal de vectores, y en vista de que $R(P_i) \perp R(P_j)$ si $i \neq j$, tenemos que juntando las bases ortonormales de cada $R(P_k)$, obtenemos una base completa ortonormal para H que consiste de vectores propios de A. En efecto, tenemos el siguiente teorema [7] .

Teorema 4.12. Si $A : H \to H$ es compacto autoadjunto, entonces

i) $A = \sum_{i=1}^{\infty} \lambda_i P_i$

ii) $H = \sum_{i=1}^{\infty} \oplus R(P_i)$

iii) Existe una base ortonormal completa e_k que consiste de vectores propios de A. Además

$$Ax = \sum_{i=1}^{\infty} \lambda_k < x, e_k > e_k$$

para toda $x \in H$ y donde e_k es un vector propio asociado a λ_k (aquí las λ_k pueden repetirse).

CAPITULO 5. TEOREMA ESPECTRAL PARA OPERADORES AUTOADJUNTOS ACOTADOS

5.0 Introducción

Sea $A : H \to H$ autoadjunto y H de dimensión finita. Si $\lambda_1, \lambda_2, \ldots, \lambda_k$ son los valores propios de A, consideremos el conjunto abierto

$$\Delta = \Delta_o \cup \Delta_1 \cup \ldots \cup \Delta_k$$

donde $\Delta_i \cap \Delta_j = \phi$, si $i \neq j$; Δ_i contiene a λ_i, $i = 1, \ldots, k$, y no contiene a ninguna otra λ_j; además supongamos que Δ_o no contiene a ningún punto del espectro de A, pero sí puntos del eje real.

En la sección 4.5.3 vimos que

$$A = \sum_{i=1}^{k} \lambda_i P_i \tag{5.1}$$

donde

$$P_i = \frac{1}{2\pi i} \int_{\Gamma_i} (zI - A)^{-1} \, dz.$$

Con la introducción de Δ_o podemos expresar (5.1) como

$$A = \sum_{i=0}^{k} \lambda_i P_i$$

ya que $P_o = 0$ porque $(zI - A)^{-1}$ es analítica en Δ_o. Cabe hacer la aclaración de que introdujimos Δ_o para poder escribir (5.1) de otra forma, pero también en términos de proyecciones. Esto último nos dará la posibilidad de generalizar nuestros resultados sobre descomposición espectral.

Definamos

$$E_i(z) = \begin{cases} 1 \text{ si } z \in \Delta_o \cup \ldots \cup \Delta_i \\ 0 \text{ si } z \in \Delta_{i+1} \cup \ldots \cup \Delta_k \end{cases}$$

para $i = 0, 1, \ldots, k$. Consideremos las proyecciones

$$E_i(A) \qquad i = 0, 1, \ldots, k.$$

Aquí $E_0(A) = 0$. No es difícil ver que

$$P_i = E_i(A) - E_{i-1}(A), \; i = 1, \ldots, k.$$

Por comodidad haremos un cambio de notación. Denotaremos a los $E_i(A)$ por

$$E_i(A) = E(\lambda_i) \qquad i = 1, \ldots, k.$$

También definimos

$$E_0(A) = E(\lambda_o)$$

para ser consistentes con el cambio de notación.

Aquí λ_0 es un punto cualquiera de Δ_0, pero sobre el eje real. $E(\lambda_0)$ está bien definido ya que $E_0(A) = 0$. Aquí recalcamos que $E_0(A)$ es un simple auxiliar para poder reescribir (5.1).

Con este cambio de notación tenemos que

$$P_i = E(\lambda_i) - E(\lambda_{i-1}) \qquad i = 1, \ldots, k.$$

De aquí que (5.1) se puede escribir como

$$A = \sum_{i=1}^{k} \lambda_i [E(\lambda_i) - E(\lambda_{i-1})] \tag{5.2}$$

Esto empieza a sugerir que A se pueda escribir como una integral de Riemann–Stieltjes.

Obsérvese que tanto los P_i como los $E(\lambda_i)$ están determinados por números reales, esto sugiere la posibilidad de atacar el problema de la representación espectral restringiéndonos a los números reales sin utilizar para nada los números complejos.

Si A es compacto autoadjunto puede hacerse lo mismo que hicimos en el caso de dimensión finita y obtenemos, en lugar de (5.2),

$$A = \sum_{i=1}^{\infty} \lambda_i [E(\lambda_i) - E(\lambda_{i-1})] \tag{5.3}$$

lo cual sugiere más la posibilidad de representar a A como una integral de Riemann–Stieltjes.

El problema para tener una representación de este tipo es que tenemos definida una función $E(\lambda)$ para un conjunto numerable de puntos y no la tenemos definida en todo un intervalo.

Sin embargo, si A es acotado autoadjunto y tomamos en cuenta que el espectro de A puede ser continuo, tenemos más posibilidades de definir una función

$$E(\lambda) : [a, b] \to L(H),$$

donde [a, b] contiene a $\sigma(A)$, de forma tal que A pueda ser representada como

$$A = \int_a^b \lambda \, d\, E(\lambda) \,. \tag{5.4}$$

Es natural esperar que $E(\lambda)$ sea una proyección ortogonal para cada λ.

Demostraremos que efectivamente A se puede representar como (5.4). Las condiciones para que dicha integral exista están tratadas en el Apéndice A. Para poder desarrollar la teoría necesitamos una serie de resultados previos que trataremos en la siguiente sección.

5.1 Algunos resultados previos

5.1.1 Proyecciones ortogonales

Definición. Una proyección tal que

$$N(P) = R(P)^{\perp}$$

es llamada una proyección ortogonal.

Teorema 5.1. Sea P una proyección. P es ortogonal si y sólo si P es autoadjunto.

Demostración. Supongamos primero que P es una proyección ortogonal. Sabemos que

$$H = R(P) \oplus N(P)$$

y

$$N(P) = R(P)^{\perp}.$$

Sean

$$z_1 = x_1 + y_1 \quad , \qquad z_2 = x_2 + y_2$$

con $x_1, x_2 \in R(P)$; $y_1, y_2 \in N(P)$. Entonces

$$<Pz_1, z_2> = <x_1, x_2 + y_2> = <x_1, x_2> ,$$

$$<z_1, Pz_2> = <x_1 + y_1, x_2> = <x_1, x_2>.$$

esto es,

$$<Pz_1, z_2> = <z_1, Pz_2>$$

para toda $z_1, z_2 \in H$. Esto muestra que P es autoadjunto.

Supongamos ahora que P es una proyección y que es autoadjunto. Sean $x \in R(P)$, $y \in N(P)$, entonces

$$<x, y> = <Px, y> = <x, Py> = <x, 0> = 0$$

de aquí se infiere que

$$N(A) = R(A)^{\perp}.$$

Lo que demuestra el teorema.

Además de ser autoadjunto, una proyección es acotado ([1]).

Teorema 5.2. Una proyección ortogonal P es un operador acotado. Si $P \neq 0$, $\|P\| = 1$ y

$$0 \leq \langle Px, x\rangle \leq 1$$

para toda $x \in H$ tal que $\|x\| = 1$.

5.1.2 Una relación de orden parcial para operadores autoadjuntos

De cierta manera los operadores autoadjuntos juegan dentro del conjunto de operadores lineales en H, el papel que los números reales juegan en los números complejos. Podemos, por ejemplo, definir una relación de orden parcial entre operadores autoadjuntos acotados.

Definición. Decimos que un operador acotado autoadjunto es positivo si

$$\langle Ax, x\rangle \geq 0$$

para toda $x \in H$. Esto lo denotamos por $A \geq 0$.

Definición. Decimos que $A \leq B$ si $B - A \geq 0$. Es decir si

$$\langle Ax, x\rangle \leq \langle Bx, x\rangle$$

para toda $x \in H$.

Ejemplo. Si P es una proyección ortogonal, el teorema 5.2 muestra que

$$0 \leq P \leq I \ .$$

Sabemos que en los números reales toda sucesión monótona y acotada es convergente. Esto también sucede en el espacio de operadores acotados autoadjuntos.

Teorema 5.3. Sea $\{A_n\}$ una sucesión de operadores acotados autoadjuntos. Sea $B \in L(H)$ tal que

$$A_n B = B A_n$$

para toda n. Supongamos también que

$$A_1 \leqslant A_2 \leqslant \ldots \leqslant A_n \leqslant \ldots \leqslant B$$

Entonces existe $A \in L(H)$ autoadjunto tal que

$$\|A_n - A\| \to 0 \quad \text{y} \quad A \leqslant B.$$

La correspondiente afirmación para sucesiones decrecientes también es válida. Para una demostración de tal hecho nos remitiremos a [1]. Ahí mismo se encuentra la demostración del siguiente teorema.

Teorema 5.4. Si A es acotado autoadjunto y $p(\lambda)$ es un polinomio con coeficientes reales tal que

$$p(\lambda) \geqslant 0$$

para toda $\lambda \in [m, M]$ – siendo $m = \min \sigma(A)$ y $M = \max \sigma(A)$ – entonces

$$p(A) \geqslant 0.$$

5.2 Teorema espectral para operadores acotados autoadjuntos

5.2.1 Introducción

El teorema espectral nos afirma que si A es acotado autoadjunto, entonces A puede ser expresado como

$$A = \int_a^b \lambda \, dE\,(\lambda)$$

donde $\sigma(A) \subset [a, b]$ y $E(\lambda)$ es una proyección ortogonal para cada $\lambda \epsilon [a, b]$. A continuación daremos las ideas en las que descansa la demostración. En la siguiente sección se verán los pasos principales de la demostración del teorema.

La idea es poder definir $E(\lambda) = E_\lambda(A)$ donde

$$E_\lambda(x) = \begin{cases} 1 & \text{si} \quad x \leqslant \lambda \\ 0 & \text{si} \quad x > \lambda \end{cases}$$

$E_\lambda(x)$ es mostrada en la fig 10. la dificultad que esto presenta es que $E_\lambda(x)$ no es continua. Sin embargo $E_\lambda(x)$ es una función semicontinua superiormente – concepto que se definirá más tarde - Pero lo más importante es que existe una sucesión de polinomios $\{ p_n(x) \}$ tales que

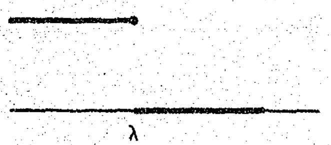


Fig 10

i) $Pm(x) \geqslant Pn(x) \qquad n < m \qquad$ para toda $x \epsilon [a, b]$

ii) $\lim\limits_{n \to \infty} Pn(x) = E_\lambda(x) \qquad$ para toda $x \epsilon [a, b]$.

Por el teorema 5.4 $Pn(A) \geqslant 0$ y $Pm(A) \geqslant Pn(A)$ si $n > m$.

Tenemos pues la siguiente sucesión de operadores acotados autoadjuntos

$$P_1(A) \geq P_2(A) \geq \ldots \geq P_n(A) \geq \ldots \geq 0.$$

Por el teorema 5.3 las $P_n(A)$ convergen en L(H) a un operador autoadjunto que definimos como

$$E_\lambda(A) = E(\lambda)$$

Como $E_\lambda(x)\, E_\lambda(x) = E_\lambda(x)$, resulta que $E(\lambda)$ es una proyección ortogonal.

Ahora, la función $I(x) = x$ puede ser aproximada por funciones escalonadas –que están dadas por las funciones $E_\lambda(x)$ – y esto es lo que permitirá llegar a la representación integral de A.

Para ver esto, sean $\lambda_o = a$, $\lambda_1 = b$, y definamos

$$\Phi_1(x) = \lambda_1\, [E_{\lambda_1}(x) - E_{\lambda_o}(x)]$$

Ahora sean, $\lambda_o = a$, $\lambda_1 = (a + b)/2$, $\lambda_2 = b$ y definamos

$$\Phi_2(x) = \lambda_1[E_{\lambda_1}(x) - E_{\lambda_o}(x)] + \lambda_2\, [E_{\lambda_2}(x) - E_{\lambda_1}(x)]$$

Por inducción definimos

$$\Phi_n(x) = \sum_{i=1}^{2^{n-1}} \lambda_i\, [E_{\lambda_i}(x) - E_{\lambda_{i-1}}(x)]$$

En la fig 11, $\Phi_2(x)$ es la función escalonada.

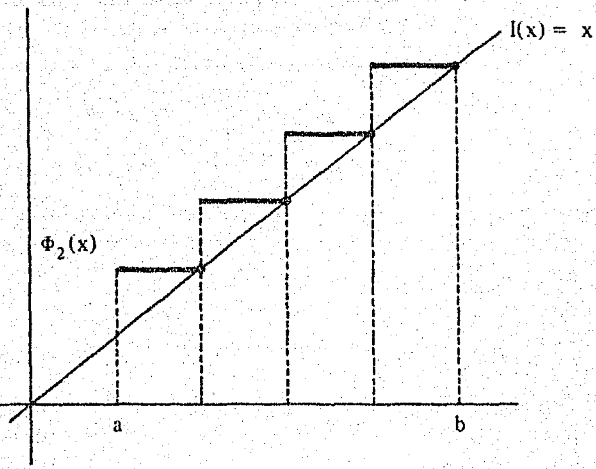


Fig 11

Es claro que $\Phi_m(x) \geqslant \Phi_n(x)$ si $n > m$ y que $\lim \Phi_n(x) = x$ para cada $x \epsilon [a, b]$. De aquí que

$$\Phi_1(A) \geqslant \Phi_2(A) \geqslant \ldots \geqslant \Phi_n(A) \geqslant A$$

Como $\Phi_n(A) A = A \Phi_n(A)$ para toda n, entonces $\Phi_n(A) \to A$ en L(H). Esto es

$$A = \int_a^b \lambda \; d \; E(\lambda).$$

La demostración que se dará del teorema espectral difiere ligeramente de la idea manejada en esta última parte, pero en esencia la idea es la misma.

5.2.2 Construcción de la familia E(λ)

Como ya dijimos, las funciones $E_{\lambda_i}(x)$ no son continuas, pero sí semicontinuas superiormente. Nuestro objetivo es entonces definir f(A) para f semicontinua superiormente y A acotado autoadjunto. Esto se hará en varios pasos.

Primer paso. Definición de p(A) cuando p(x) es un polinomio

Ya sabemos que si p(x) es un polinomio con coeficientes reales, p(A) es un elemento autoadjunto de L(H) cuando A es acotado autoadjunto. Tenemos pues, una función

$$\varphi_1 : P \to L(H)$$

donde P es el espacio de polinomios con coeficientes reales y

$$\varphi_1(p) = p(A) .$$

φ_1 tiene las siguientes propiedades:

1) $\varphi_1 (p_1 + p_2) = \varphi_1(p_1) + \varphi_1(p_2)$

2) $\varphi_1 (p_1 \cdot p_2) = \varphi_1(p_1)\, \varphi_1(p_2)$

3) $\varphi_1 (\alpha p) = \alpha\, \varphi_1 (p) \quad \alpha \in R$

4) Si $p(x) \geqslant 0$ en $[m, M]$, entonces $p(A) \geqslant 0$. Siendo $m = \inf \langle Ax, x\rangle$; $M = \sup \langle Ax, x\rangle$, $\| x \| = 1$.

Las tres primeras propiedades son consecuencia inmediata de la definición de φ_1 y la cuarta del teorema 5.4.

A continuación vamos a definir f(A) para una clase más amplia de funciones.

Segundo paso. Definición de f(A) cuando f(x) es semicontinua superiormente y no negativa.

Definición. Una función $f: [a, b] \to R$ se dice que es semicontinua superiormente en $x_0 \epsilon [a, b]$, si dada $\epsilon > 0$ existe $\delta > 0$ tal que

$$f(x) < f(x_0) + \epsilon$$

para toda $x \epsilon [a, b]$ tal que $|x - x_0| < \delta$.

Es claro, de la definición, que toda función continua es semicontinua superiormente. En la fig 12 se muestran algunos ejemplos de funciones de este tipo.

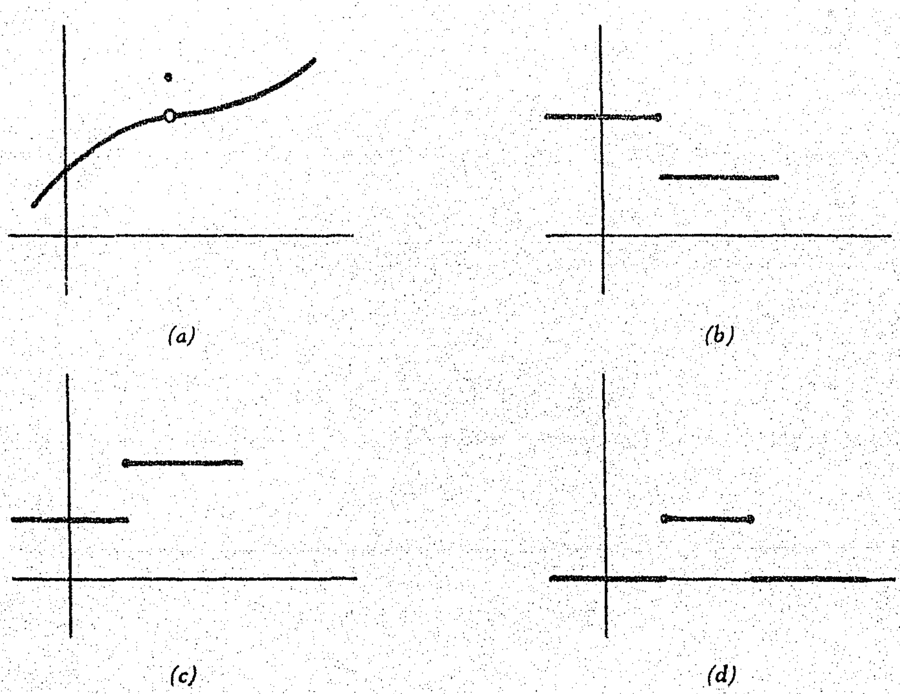


Fig 12

Como se puede ver, las funciones $E_{\lambda_i}(t)$ son funciones semicontinuas superiormente y no negativas.

La idea, en este paso, es definir f(A) para funciones semicontinuas superiormente no negativas aprovechando φ_1, es decir, tratar de aproximar este tipo de funciones por medio de polinomios. Esto es posible gracias al siguiente resultado.

Teorema 5.5. Si $f : [a, b] \to R$ es semicontínua superiormente y no negativa, entonces existe una sucesión $\{P_n(x)\}$ de polinomios con coeficientes reales tal que

i) $P_m(x) \geqslant P_n(x)$ si $n > m$

ii) $\lim\limits_{n \to \infty} P_n(x) = f(x)$ para toda $x \in [a, b]$

Esto lo denotaremos por

$$P_n(x) \downarrow f(x)$$

Demostración. Una demostración general puede encontrarse en [1]. Aquí nos limitaremos a ver la demostración para el caso particular en que $f(x) = E_\lambda(x)$.

La sucesión $\{f_n(x)\}$ de funciones continuas, donde $f_n(x)$ está descrita en la fig 13, es tal que

$$f_n(x) \downarrow E_\lambda(x)$$

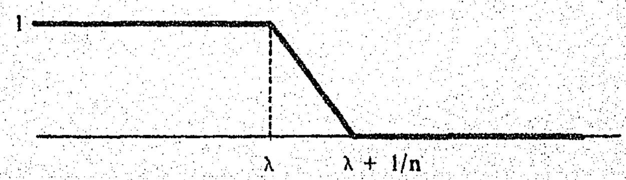


Fig 13

Pero también

$$f_n(x) + 1/n \downarrow E_\lambda(x).$$

Es posible construir, para cada n, una banda con centro en la gráfica de $f_n(x) + 1/n$, la cual no contenga a ningún otro elemento de la sucesión $f_n(x) + 1/n$. Por ejemplo se puede tomar la banda acotada por las funciones $f_n(x) + 1/n + 1/2^{n+1}$ y $f_n(x) + 1/n - 1/2^{n+1}$. Por el teorema de Weierstrass, existe un polinomio $P_n(x)$ cuya gráfica está dentro de esa banda. La fig 14 muestra dicha situación. De lo anterior se sigue que

$$P_n(x) \downarrow E_\lambda(x).$$

Lo que demuestra el teorema.

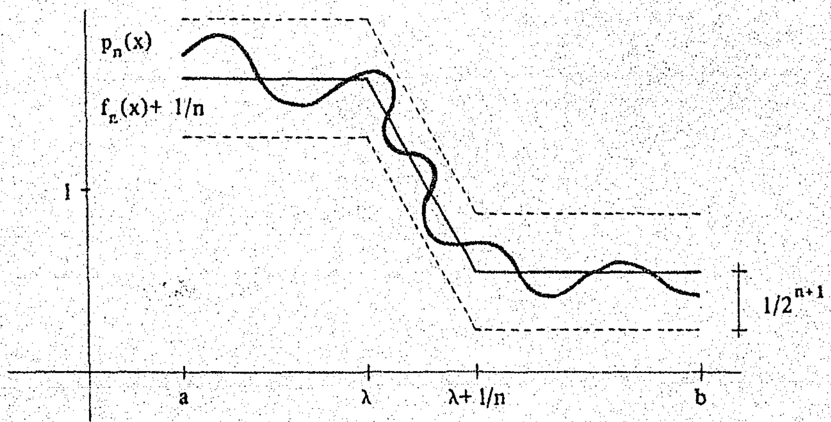


Fig 14

Sea C_1 el conjunto de funciones $f : [m, M] \to R$ semicontinuas superiormente y tales que $f \geqslant 0$ en $[m, M]$.

Por el teorema 5.5, existe $\{ P_n(x) \}$ tal que

$$P_n(x) \downarrow f(x).$$

Por los teoremas 5.3 y 5.4 tenemos que

$$P_1(A) \geqslant P_2(A) \geqslant \ldots \geqslant P_n(A) \geqslant \ldots \geqslant 0$$

y existe un operador acotado autoadjunto, al que denotaremos por f(A), tal que

$$\| P_n(A) - f(A) \| \rightarrow 0,$$

si $n \rightarrow \infty$, y $f(A) \geqslant 0$. Además si $B \in L(H)$ y $AB = BA$, entonces $Bf(A) = f(A)B$.

Estos hechos permiten definir f(A).

Definición. Si A es acotado autoadjunto y $f \in C_1$, definimos

$$f(A) = \text{lím } P_n(A), \quad n \rightarrow \infty$$

donde

$$P_n(x) \downarrow f(x).$$

Por supuesto que puede haber otra sucesión $\{q_n(x)\}$ tal que

$$q_n(x) \downarrow f(x).$$

En [1] se muestra que en este caso

$$\text{lím } P_n(A) = \text{lím } q_n(A), \quad n \rightarrow \infty$$

es decir, que f(A) está bien definido.

Sea

$$\varphi_2 : C_1 \rightarrow L(H)$$

definida como $\varphi_2(f) = f(A)$. Hemos construido φ_2 a partir de φ_1, y φ_2 tiene las siguientes propiedades:

(1') $\varphi_2(f_1 + f_2) = \varphi_2(f_1) + \varphi_2(f_2)$

(2') $\varphi_2(f_1 f_2) = \varphi_2(f_1)\, \varphi_2(f_2)$

(3') Si $a \geqslant 0$, entonces $\varphi_2(af) = a\, \varphi_2(f)$

(4') Si $f_1(x) \geqslant f_2(x)$ en $[m, M]$, entonces $f_1(A) \geqslant f_2(A)$.

La demostración de estas propiedades se siguen directamente de la definición de φ_2.

Tercer paso. Definición de $h(A)$ cuando $h = f - g$ con $f, g \in C_1$

Sean

$$C_2 = \left\{ f - g \mid f, g \in C_1 \right\}$$

y

$$\varphi_3 : C_2 \to L(H)$$

definida por $\varphi_3(h) = \varphi_2(f) - \varphi_2(g) = f(A) - g(A)$, siendo $h = f - g;\ f, g \in C_1$.

De la definición de φ_3 y de las propiedades de φ_2, se sigue fácilmente que φ_3 está bien definida y que satisface las propiedades siguientes.

i) $\varphi_3(h_1 + h_2) = \varphi_3(h_1) + \varphi_3(h_2)$

ii) $\varphi_3(h_1 h_2) = \varphi_3(h_1)\, \varphi_3(h_2)$

iii) $\varphi_3(a h) = a\, \varphi_3(h)$ para toda $a \in R$

iv) Si $h_1(x) \geqslant h_2(x)$ en $[m, M]$, entonces $h_1(A) \geqslant h_2(A)$

Ahora ya podemos definir la familia $E(\lambda)$. Sea

$$E_\lambda(x) = \begin{cases} 1 & \text{si } \lambda \leqslant x \\ 0 & \text{si } \lambda > x \end{cases}$$

Para cada $\lambda \in [m, M]$, $E_\lambda(x)$ es un elemento de C_2. Definimos

$$E(\lambda) = E_\lambda(A).$$

En términos de esta familia podemos representar a A como una integral.

5.2.3 El teorema espectral

Teorema 5.6. La familia $\{E(\lambda)\}$ constituye una familia de proyecciones ortogonales satisfaciendo las siguientes propiedades.

(1) Si $\lambda_1 \leq \lambda_2$, entonces $E(\lambda_1) \leq E(\lambda_2)$

(2) Si $\epsilon > 0$,

$$\| E(\lambda + \epsilon) - E(\lambda) \| \to 0$$

cuando $\epsilon \to 0$

(3) $E(\lambda) = 0$ para $\lambda < m$ y $E(\lambda) = I$ si $\lambda \geq M$.

(4) $AE(\lambda) = E(\lambda) A$

(5) Si $a < m$ y $b \geq M$ se tiene

$$A = \int_a^b \lambda \, d\, E(\lambda)$$

La familia $\{E(\lambda)\}$ es llamada la resolución de la identidad asociada a A.

Demostración. Para facilitar la demostración en dos partes. En la primera se demuestran las propiedades de la familia $\{E(\lambda)\}$ y en la segunda parte se ve la representación integral de A.

Primera parte. Las propiedades de $\{E(\lambda)\}$

Por construcción $E(\lambda)$ es autoadjunto y acotado, $AE(\lambda) = E(\lambda)A$ y

$E(\lambda_1)\ E(\lambda_2) = E(\lambda_2)\ E\ (\lambda_1)$. Como

$$E_\lambda(x)\ E_\lambda(x) = E_\lambda(x)$$

se sigue que

$$E(\lambda)^2 = E(\lambda)$$

Lo que muestra que $E(\lambda)$ es una proyección ortogonal para toda $\lambda \in [m, M]$.

Para demostrar (1) basta observar que si $\lambda_1 \leqslant \lambda_2$, entonces $E_{\lambda_1}(x) \leqslant E_{\lambda_2}(x)$ en $[m, M]$.

Para ver (2) demostraremos que dada una sucesión $\{\epsilon_n\}$ de reales positivos convergiendo a cero, entonces es posible construir una sucesión $\{g_n(x)\}$ de funciones continuas tal que

$$g_n(x) \downarrow E_\lambda(x) \qquad \text{y} \qquad g_n(x) \geqslant E_{\lambda+\epsilon n}(x)$$

para cada n y toda $x \in [m, M]$. En la fig 15 se muestra cómo se puede escoger $g_n(x)$. Entonces

$$E_\lambda(x) \leqslant E_{\lambda+\epsilon_n}(x) \leqslant g_n(x)$$

para toda $x \in [m, M]$. Por ser $g_n(x)$ continua, también lo es semicontinua superiormente y entonces

$$E(\lambda) \leqslant E(\lambda + \epsilon_n) \leqslant g_n(A).$$

Pero

$$\| g_n(A) - E(\lambda) \| \rightarrow 0$$

y de aquí se infiere que

$$\| E(\lambda + \epsilon_n) - E(\lambda) \| \to 0$$

cuando $\epsilon_n \to 0$.

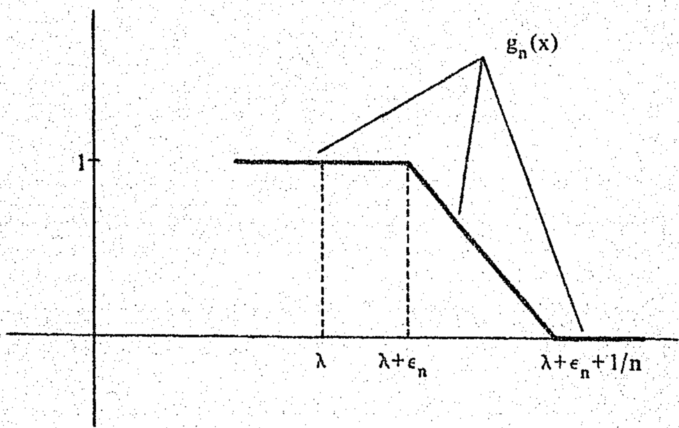


Fig 15

Para demostrar (3) supongamos primero que $\lambda < m$. En este caso $E_\lambda(x) = 0$ para toda $x \epsilon [m, M]$ y la sucesión $\{P_n(x)\}$ donde $P_n(x) \equiv 0$ es tal que $P_n(x) \downarrow E_\lambda(x)$ y por lo tanto $E(\lambda) = 0$. Ahora, si $\lambda \geqslant M$, tenemos que $E_\lambda(x) = 1$ para toda $x \epsilon [m, M]$. La sucesión de polinomios $P_n(x) \equiv 1$ es tal que $P_n(x) \downarrow E_\lambda(x)$ y entonces $E(\lambda) = I$.

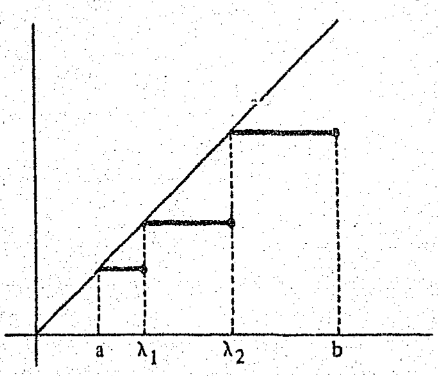
La propiedad (4) ya se hizo notar y pasamos a la segunda parte de la demostración.

Segunda parte. La representación espectral

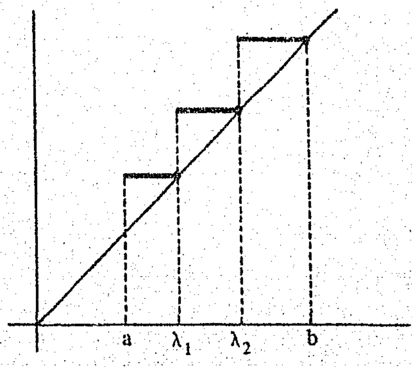
Sean $a < m$, $b \geqslant M$. Dada $\epsilon > 0$ construyamos una partición P, $a = \lambda_o < \lambda_1 < \ldots < \lambda_n = b$, de $[a, b]$ tal que

$$\text{máx}_k \ |\lambda_k - \lambda_{k\text{-}1}| < \epsilon \ .$$

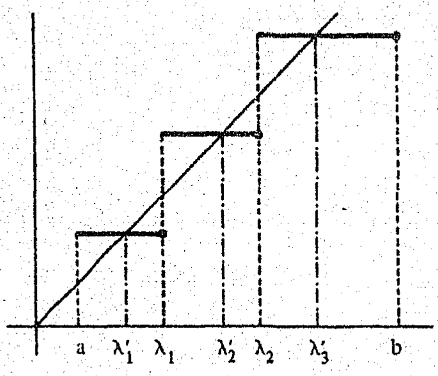
Aquí la idea es la siguiente: la función $I(x) = x$ puede ser aproximado de varias maneras. En la fig 16 se muestran tres formas de aproximación



(a)



(b)



(c)

Fig 16

La aproximación del tipo de la fig 16b es el que utilizamos en la introducción de este capítulo, aunque en un caso más particular.

La idea de la demostración es aprovechar la aproximación del tipo de la fig 16c, y hechando mano de los otros tipos de aproximación, llegar a la representación espectral.

Si $\lambda_1 < \lambda_2$

$$E_{\lambda_2}(x) - E_{\lambda_1}(x) = \begin{cases} 0 & \text{si} \quad x \leqslant \lambda_1 \\ 1 & \text{si} \quad \lambda_1 < x \leqslant \lambda_2 \\ 0 & \text{si} \quad x > \lambda_2 \end{cases}$$

De aquí se deduce que

$$\lambda_1 [E_{\lambda_2}(x) - E_{\lambda_1}(x)] \leqslant x [E_{\lambda_2}(x) - E_{\lambda_1}(x)] \leqslant \lambda_2 [E_{\lambda_2}(x) - E_{\lambda_1}(x)]$$

para cualquier x. Entonces

$$\lambda_1[E(\lambda_2) - E(\lambda_1)] \leqslant A [E(\lambda_2) - E(\lambda_1)] \leqslant \lambda_2 [E(\lambda_2) - E(\lambda_1)]$$

Lo mismo sucede para λ_{K-1} , λ_K . Entonces

$$\sum_{k=1}^{n} \lambda_{k-1} [E(\lambda_k) - E(\lambda_{k-1})] \leqslant A \leqslant \sum_{k=1}^{n} \lambda_k [E(\lambda_k) - E(\lambda_{k-1})]$$

donde se utilizó el hecho de que

$$\sum_{k=1}^{n} [E(\lambda_k) - E(\lambda_{k-1})] = I.$$

Tomemos $\lambda'_k \epsilon [\lambda_k, \lambda_{k-1}]$. Es fácil mostrar que

$$- \epsilon I \leqslant A - \sum_{k=1}^{n} \lambda'_k [E(\lambda_k) - E(\lambda_{k-1})] \leqslant \epsilon I$$

Si hacemos

$$C = A - \sum_{k=1}^{n} \lambda'_k [E(\lambda_k) - E(\lambda_{k-1})]$$

la desigualdad anterior nos dice que

$$- \epsilon \langle x, x\rangle \leqslant \langle Cx, x\rangle \leqslant \epsilon \langle x, x\rangle$$

para toda $x \epsilon H$. Por lo tanto

$$|< C x, x >| \leqslant \epsilon \; |< x, x >|$$

y de aquí que

$$\| C \| \leqslant \epsilon$$

Lo que significa que

$$\| A - \sum_{k=1}^{n} \lambda'_k [E(\lambda_k) - E(\lambda_{k-1})] \| \leqslant \epsilon .$$

Por el Apéndice A, tenemos que

$$A = \int_a^b \lambda \, d\, E(\lambda)$$

o bien

$$A = \int_{m}^{M} \lambda \, d\, E(\lambda) .$$

Así queda demostrado el teorema.

5.2.4 Consecuencias del teorema espectral

Corolario 1. Si p(x) es un polinomio con coeficientes reales, entonces

$$p(A) = \int_a^b p(\lambda)\, d\, E(\lambda)$$

Demostración. Es fácil ver que si $\lambda_j \leqslant \lambda_k$

$$E_{\lambda_j}(x) \; E_{\lambda_k}(x) = E_{\lambda_j}(x)$$

y de aquí que

$$E(\lambda_j) \; E(\lambda_k) = E(\lambda_j) .$$

Si $\lambda_j < \lambda_k$ (o sea $\lambda_j \leqslant \lambda_{k-1}$) se sigue que

$$[E(\lambda_j) - E(\lambda_{j-1})]\,[E(\lambda_k) - E(\lambda_{k-1})] = 0,$$

$$[E(\lambda_k) - E(\lambda_{k-1})]^2 = E(\lambda_k) - E(\lambda_{k-1})$$

y

$$[E(\lambda_k) - E(\lambda_{k-1})]^n = E(\lambda_k) - E(\lambda_{k-1}).$$

De las igualdades anteriores se infiere que

$$\left(\sum_{k=1}^{n} \lambda'_k\,[E(\lambda_k) - E(\lambda_{k-1})]\right)^n = \sum_{k=1}^{n} (\lambda'_k)^n\,[E(\lambda_k) - E(\lambda_{k-1})]$$

o sea que

$$A^n = \int_a^b \lambda^n \, d\,E(\lambda)$$

De donde obtenemos que

$$P(A) = \int_a^b P(\lambda)\, d\,E(\lambda)\,.$$

Corolario 2. Si $f(x)$ es continua en $[a, b]$, entonces

$$f(A) = \int_a^b f(\lambda)\, d\,E(\lambda).$$

Demostración. Sea $\{ p_n(x) \}$ una sucesión de polinomios que converje uniformemente a $f(x)$ en $[a, b]$. Entonces dada $\epsilon > 0$, existe N tal que

$$-\epsilon < f(x) - P_n(x) < \epsilon$$

para toda $x \in [a, b]$ y $n > N$. De aquí que

$$-\epsilon\, I < f(A) - P_n(A) < \epsilon\, I$$

para $n > N$, lo que implica que

$$\| f(A) - P_n(A) \| < \epsilon$$

para $n > N$, Este es

$$P_n(A) \to f(A)$$

en L(H). Por el teorema A.2 y el corolario anterior, tenemos que

$$< p_n(A)\ x,\ y> = \int_a^b p_n(\lambda)\ d < E(\lambda)\ x,\ y >$$

Ahora, por un lado tenemos que

$$\lim_{n\to\infty} < p_n(A)\, x,y > = \lim_{n\to\infty} \int_a^b p_n(\lambda)\ d < E(\lambda)\, x, y> = \int_a^b f(\lambda)\ d < E(\lambda)\ x,\ y >$$

ya que la convergencia es uniforme. Y por otro lado

$$\lim < p_n(A)\, x, y> = <\lim\ p_n(A)\, x, y> = <f(A)\, x, y> .$$

Entonces

$$<f(A)\ x,\ y> = \int_a^b f(\lambda)\ d < E(\lambda)\ x,\ y>$$

y por lo tanto

$$f(A) = \int_a^b f(\lambda)\ d\,E(\lambda)$$

Corolario 3. Si A^{-1} es acotado, entonces

$$A^{-1} = \int_{m}^{M} \lambda^{-1}\ d\,E(\lambda)$$

Demostración. Si A^{-1} existe, $0 \notin \sigma(A)$, y entonces existe $a < m$ tal que $0 \notin [a, M]$. La función $f(x) = 1/x$ es continua en $[a, M]$. Aplicando el Corolario 2 obtenemos el resultado deseado.

CAPITULO 6. CALCULO OPERACIONAL PARA OPERADORES CERRADOS Y TEOREMA ESPECTRAL PARA OPERADORES AUTOADJUNTOS NO ACOTADOS

6.0 Introducción

En vista de las complicaciones que presenta el caso en que A es cerrado, nos limitaremos a dar las ideas centrales del Cálculo Operacional y del Teorema Espectral, éste en el caso en que A es autoadjunto.

6.1 Cálculo operacional para operadores cerrados

Veremos que cuando A es cerrado podemos tener una generalización de la fórmula (4.6) y en consecuencia tendremos un Cálculo Operacional para este tipo de operadores.

Naturalmente, esta teoría se debe reducir a la ya conocida en el caso en que A sea acotado. Esto nos dice que debemos considerar, en primer lugar, una clase de funciones analíticas de manera que el dominio de cada función contenga a $\sigma(A)$. En segundo lugar, debemos definir f(A) por medio de la fórmula de Cauchy. En vista de esto último, tenemos que utilizar la resolvente y aquí tenemos que pedir que $\rho(A) \neq \phi$, es decir, que $\sigma(A)$ no sea todo el plano complejo.

Sea f una función analítica cuyo dominio Δ contenga a $\sigma(A)$. Para poder definir f(A) aplicando la fórmula de Cauchy debemos asegurarnos que podemos tomar un contorno de integración sencillo. El caso acotado sugiere que consideremos abiertos Δ para los cuales exista ω abierto tal que

$$\sigma(A) \subset \omega \subset \overline{\omega} \subset \Delta$$

y donde $\partial\omega$ conste de un número finito de curvas simples cerradas que no se intersecten entre sí. En la fig 17 se muestra un caso sencillo de esta situación, en la cual $\sigma(A)$ no es acotado

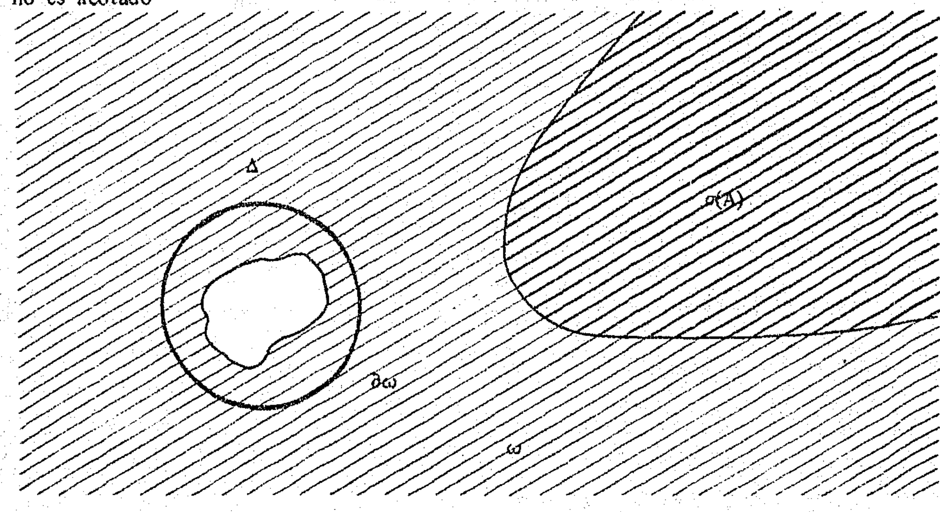


Fig 17

En este ejemplo vemos que ω^c es compacto, en consecuencia también Δ^c. Veremos a continuación que si Δ es un abierto tal que Δ^c es compacto, se puede escoger ω con las propiedades antes mencionadas.

En efecto, sea Δ conteniendo a $\sigma(A)$ y tal que Δ^c es compacto. Tenemos pues que

$$\Delta^c \subset \sigma(A)^c$$

y $\sigma(A)^c$ es abierto, ya que A es cerrado. Aplicando el Lema de la sección 4.2.3 vemos que existe un conjunto abierto U satisfaciendo

$$\Delta^c \subset U \subset \overline{U} \subset \sigma(A)^c$$

y ∂U consiste de un número finito de curvas simples cerradas. El conjunto $\omega = \overline{U}^c$ es tal que

$$\sigma(A) \subset \omega \subset \overline{\omega} \subset \Delta$$

y $\partial\omega$ tiene las propiedades de ∂U.

Consideremos, pues, la clase de funciones analíticas $f : \Delta \subset C \rightarrow C$, donde $\Delta \supset \sigma(A)$ y Δ^c es compacto. Ahora pondremos una restricción a f. Pediremos que f(z) tienda a un límite finito, $f(\infty)$, cuando $|z| \rightarrow \infty$.

Veremos aquí que la fórmula de Cauchy sigue valiendo, con ciertas variantes, para este tipo de funciones. Para esto necesitamos hacer antes un comentario sobre la orientación de $\partial\omega$.

Si γ es una curva simple cerrada que forme parte de $\partial\omega$, γ será tomada en sentido positivo o negativo según los puntos de ω cercanos a γ estén en el interior o en el exterior de ω, respectivamente.

Ahora, sea $z \in \omega$ y tomemos un círculo Γ, con centro en el origen y radio R, orientado positivamente y lo suficientemente grande para que contenga a z y a $\partial\omega$. La fig 18 muestra el caso más sencillo de esta situación.

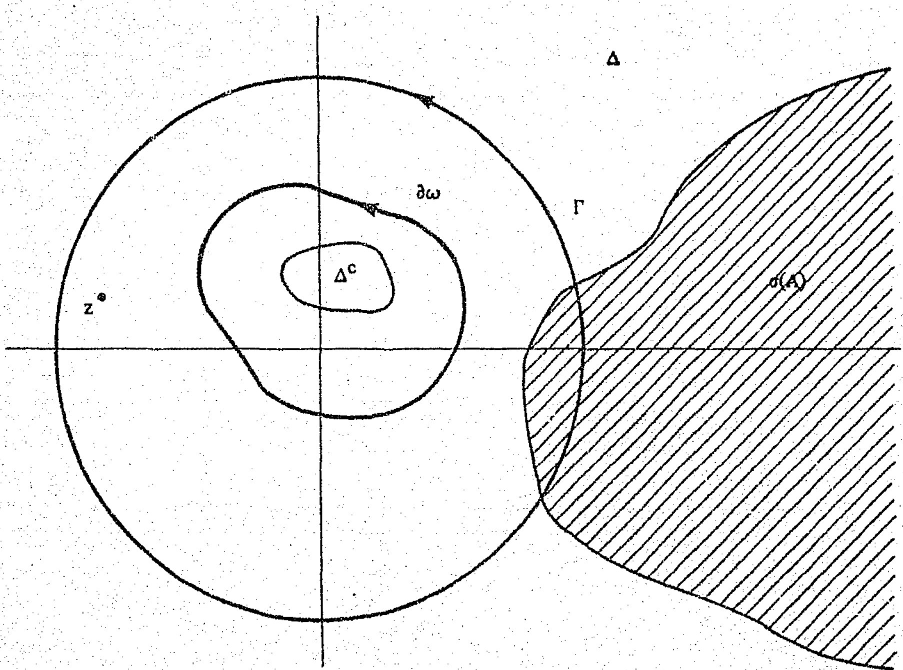


Fig 18

Por la fórmula de Cauchy

$$f(z) = \frac{1}{2\pi i}\int_{\partial\omega\cup\Gamma} \frac{f(\lambda)}{\lambda - z}\; d\lambda = \frac{1}{2\pi i}\int_{\partial\omega} \frac{f(\lambda)}{\lambda - z}\; d\lambda + \frac{1}{2\pi i}\int_{\Gamma} \frac{f(\lambda)}{\lambda - z}\; d\lambda$$

Si hacemos crecer R la integral sobre Γ no varía. Veremos ahora que cuando $R \to \infty$ la integral sobre Γ tiende a $f(\infty)$. Para esto notemos que debido a que z está en el interior de Γ,

$$\frac{1}{2\pi i}\int_{\Gamma} \frac{d\lambda}{\lambda - z} = 1$$

Entonces

$$\left| \frac{1}{2\pi i} \int_C \frac{f(\lambda)}{\lambda - z} \, d\lambda - f(\infty) \right| = \left| \frac{1}{2\pi i} \int_C \frac{f(\lambda) - f(\infty)}{\lambda - z} \, d\lambda \right| \leqslant 2\epsilon$$

siempre y cuando R sea lo suficientemente grande para que

$$|f(\lambda) - f(\infty)| < \epsilon \qquad y \qquad |\lambda - z| > R/2.$$

De aquí que

$$f(z) = f(\infty) + \frac{1}{2\pi i} \int_{\partial\omega} \frac{f(\lambda)}{\lambda - z} \, d\lambda$$

Esta fórmula nos permite definir de una manera natural f(A) como

$$f(A) = f(\infty) I + \frac{1}{2\pi i} \int_{\partial\omega} f(z) \, (zI - A)^{-1} \, dz \qquad (6.1)$$

Observación. Como $\partial\omega \subset \rho(A)$, entonces $(zI - A)^{-1} \epsilon L(H)$ y en consecuencia $f(A) \epsilon L(H)$ aunque A no sea acotado.

Se puede demostrar que en el caso en que A es acotado, la fórmula (6.1) se reduce a (4.6) y que los teoremas 4.6 y 4.7 sigue valiendo [20].

Teorema 6.1. Si f(z) es una función analítica de la clase antes especificada y A es cerrado, entonces f(A) definido por (6.1) tiene las siguientes propiedades

i) $(f + g)(A) = f(A) + g(A)$
ii) $(\alpha f)(A) = \alpha f(A)$ para toda $\alpha \epsilon C$
iii) $(fg)(A) = f(A) g(A)$
iv) Si $f(z) \neq 0$ para toda $z \epsilon \sigma(A)$ y $f(\infty) \neq 0$, entonces f(A) es invertible y

$$f(A)^{-1} = \frac{1}{f(\infty)} I + \frac{1}{2\pi i} \int_{\partial\omega} \frac{1}{f(z)} (zI - A)^{-1} \, dz$$

donde $\partial\omega$ es la frontera de un conjunto abierto ω tal que

$$\sigma(A) \subset \omega \subset \bar{\omega}$$

y $f(z)$ no se anula en $\bar{\omega}$.

Teorema 6.2. Si $p(z)$ es un polinomio y A es cerrado, entonces

$$p(\sigma(A)) = \sigma\ (p(A)).$$

Cuando A es acotado este teorema es válido aun para funciones analíticas. En el caso presente es necesario hacer algunas modificaciones para tener el correspondiente resultado.

Definición. Definimos el espectro extendido de A como

$$\sigma_e(A) = \sigma(A) \cup \{\infty\}$$

Teorema 6.3. Si $f(z)$ es analítica de la clase antes especificada y A es cerrado, entonces

$$f(\sigma_e(A)) = \sigma(f(A)).$$

6.2 Teorema espectral para operadores autoadjuntos no acotados

En [1] puede encontrarse una demostración rigurosa de este teorema. Existen varios caminos para llegar a dicho teorema. Nosotros veremos las ideas de dos de ellos. En ambos, la idea es aprovechar el teorema espectral para operadores autoadjuntos acotados.

Veamos el primer camino. La idea es encontrar una función $f(x)$ biyectiva tal que podamos definir $B = f(A)$ donde B resulte acotado. Se demuestra que la familia $\{E(t)\}$ asociada a B sirve para construir una familia $\{E(\lambda)\}$, $-\infty < \lambda < \infty$, de proyecciones artogonales, donde $\lambda = f^{-1}(t)$, y A puede ser escrito como

$$A = \int_{-\infty}^{\infty} \lambda \; d\,E(\lambda) \tag{6.2}$$

Sin embargo el sentido de esta integral no es como en el caso en que A es acotado. El sentido adecuado es el siguiente

$$< Ax, \; y> \; = \int_{-\infty}^{\infty} \lambda < d\,E(\lambda)\, x, \; y > \tag{6.3}$$

para toda x, y ϵ H. La fórmula (6.2) es sólo una manera de expresar que A tiene la propiedad (6.3). Una función que satisface las condiciones citadas anteriormente es

$$y = x\,|x|\,(1 + x^2)^{-1} .$$

Otra interpretación de (6.2) se verá a continuación en la descripción del segunda camino para llegar al teorema espectral. Es posible demostrar que existe una familia $\{E(\lambda)\}$, $-\infty < \lambda < \infty$, de proyecciones ortogonales, tal que para cada natural n, la integral

$$A_n = \int_{-n}^{n} \lambda \; d\,E(\lambda)$$

existe y es un operador autoadjunto acotado. La sucesión A_n tiene la propiedad de que

$$Ax = \lim_{n \to \infty} A_n x$$

para cada x ϵ H. Es decir

$$Ax = \lim_{n \to \infty} \int_{-n}^{n} \lambda \; d\,E(\lambda)\; x$$

para cada $x \in H$. Esto es denotado por

$$A = \int_{-\infty}^{\infty} \lambda \, d\, E(\lambda)$$

aunque la convergencia es puntual y no en norma.

Teorema 6.4 (Teorema espectral para operadores autoadjuntos no acotados) Sea $A : D \subset H \rightarrow H$ autoadjunto. Existe una familia $\{ E(\lambda) \}$, $-\infty < \lambda < \infty$, de proyecciones ortogonales tales que

1) $\lambda_1 < \lambda_2$ implica $E(\lambda_1) \leqslant E(\lambda_2)$

2) Para $\epsilon > 0$.

$$\| E(\lambda + \epsilon) - E(\lambda) \| \rightarrow 0$$

si $\epsilon \rightarrow 0$.

3) $\| E(\lambda) \| \rightarrow 0$ si $\lambda \rightarrow -\infty$; $\| E(\lambda) - I \| \rightarrow 0$ si $\lambda \rightarrow \infty$.

4) Para cada $x, y \in H$ se tiene

$$< A x, y > = \int_{-\infty}^{\infty} \lambda < d\, E(\lambda)\, x, y > \qquad (6.3)$$

5) Si f es acotada y uniformemente continua en $(-\infty, \infty)$, entonces para toda $x, y \in H$,

$$< f(A)\, x, y > = \int_{-\infty}^{\infty} f(\lambda)\, d < E(\lambda)\, x, y > \qquad (6.4)$$

Las expresiones (6.3) y (6.4) se denotan, por comodidad, como

$$A = \int_{-\infty}^{\infty} \lambda \, d\, E(\lambda) \qquad \text{y} \qquad f(A) = \int_{-\infty}^{\infty} f(\lambda)\, d\, E(\lambda),$$

respectivamente.

CAPITULO 7. APLICACIONES

7.0 Introducción

Este capítulo esta dedicado fundamentalmente a aplicar la teoría desarrollada en los capítulos previos, a la solución de los ejemplos planteados en el Cap 0.

De estos ejemplos, algunos ya se estudiaron. El ejemplo (A) fué estudiado en el capítulo 1; se vió que la solución de la ecuación diferencial

$$\dot{\overline{x}} = A\,\overline{x}$$

donde $\overline{x}(0) = \overline{x}_o$ y A es una matriz con elementos constantes, viene dada por

$$\overline{x}(t) = e^{tA}\ \overline{x}_o\ .$$

El ejemplo (D) fue estudiado en el Cap 3. Se vio que la solución de la ecuación integral

$$u(x) = f(x) + \lambda \int_0^1 k(x, t)\ u(t)\ dt,$$

que puede escribirse como

$$(I - \lambda A)\, u = f$$

donde A es un operadore integral, viene dada por

$$u = (I - \lambda A)^{-1} f$$

siempre y cuando λ no fuese un punto del espectro de A.

Nos quedan dos ejemplos por estudiar, los ejemplos (B) y (C), el de la ecuación de calor y de la cuerda vibrante, respectivamente.

Para esto es necesario estudiar un tipo particular de operadores, a saber, los operadores con espectro discreto. Los resultados acerca de este tipo de operadores pueden también tomarse como aplicaciones de la teoría previa.

7.1 Operadores con espectro discreto

7.1.1. Definición y ejemplos

Definición. Un operador $A : D \subset H \to H$ autoadjunto se dice que tiene espectro discreto si sus vectores propios forman un conjunto ortonormal completo.

Ejemplo 1. Un operador $A : D \subset L_2(0, 1) \to L_2(0, 1)$ dado por

$$A = -\frac{d^2}{dx^2}$$

con condiciones a la frontera $u(0) = u(1) = 0$. D es el conjunto de funciones $u \in L_2(0, 1)$ tales que $u'' \in L_2(0, 1)$. Los valores propios de A son de la forma

$n^2 \pi^2$ y sus vectores propios correspondientes son de la forma $u_n(x) = \sqrt{2}$ sen $n\pi x$, los cuales constituyen un conjunto ortonormal completo en $L_2(0, 1)$.

Ejemplo 2. El operador A del ejemplo anterior, pero con condiciones a la frontera $u'(0) = u'(1) = 0$ tiene espectro discreto. En efecto sus valores propios son de la forma $n^2 \pi^2$ y sus vectores propios correspondientes son de la forma $u_n(x)=\sqrt{2}\cos n\pi x$ los cuales constituyen un conjunto ortonormal completo en $L_2(0, 1)$.

Ejemplo 3. Todo operador compacto autoadjunto tiene espectro discreto. Esto es consecuencia del teorema 4.12.

Ejemplo 4. Todo operador $A : D \subset H \rightarrow H$ autoadjunto tal que A^{-1} sea compacto autoadjunto tiene espectro discreto. En efecto, sea $x \neq 0$ tal que $Ax = \lambda x$, con $\lambda \neq 0$, entonces

$$A^{-1} Ax = \lambda A^{-1} x,$$

o sea que

$$A^{-1} x = \lambda^{-1} x$$

Esto muestra que los valores propios de A^{-1} son los inversos de A y además que A y A^{-1} tienen el mismo conjunto de vectores propios. Esto muestra nuestra afirmación.

7.1.2 Los operadores con espectro discreto estan determinados por sus valores propios y vectores propios

Supongamos que A tiene espectro discreto y que $\{ e_1, e_2, \ldots \}$ es una base ortonormal completa de H que consta de vectores propios de A. Entonces cualquier $x \in H$ se puede escribir como

$$x = \sum_{k=1}^{\infty} a_k e_k$$

donde $a_k = \langle x, e_k \rangle$.

Teorema 7.1. Sea $A : D \subset H \to H$ con espectro discreto. Una condición necesaria y suficiente para que $x \in D$ es que la serie

$$\sum_{k=1}^{\infty} \lambda_k^2 \, |a_k|^2 \, , \qquad (7.1)$$

sea convergente. En este caso

$$Ax = \sum_{k=1}^{\infty} \lambda_k \, a_k \, e_k \, . \qquad (7.2)$$

Demostración. $\Rightarrow$) Sea $x \in D$. Entonces

$$< Ax, e_k > = <x, Ae_k> = <x, \lambda_k e_k> = \lambda_k <x, e_k>.$$

Esto implica que la serie (7.1) es convergente.

$\Leftarrow$) Supongamos que (7.1) es convergente y consideremos al elemento

$$x' = \sum_{k=1}^{\infty} \lambda_k \, a_k \, e_k$$

Sea

$$y_n = \sum_{k=1}^{n} a_k \, e_k ,$$

entonces $y_n \to x$ y $Ay_n \to x'$. Como A es cerrado, se sigue que $x \in D$ y $Ax = x'$. Entonces

$$Ax = \sum_{k=1}^{\infty} \lambda_k \, a_k \, e_k .$$

Corolario. Si $f(z)$ es tal que $f(A)$ está definido, A tiene espectro discreto y $x = \sum_{k=1}^{\infty} a_k \, e_k$, entonces

$$f(A)\,x = \sum_{k=1}^{\infty} f(\lambda_k)\, a_k\, e_k .$$

Demostración. Del teorema de transformación del espectro se sigue que las $f(\lambda_k)$ son los valores propios de $f(A)$. Los vectores propios son comunes a A y $f(A)$. Esto junto con el teorema anterior muestran el corolario.

7.1.3 El operador $A = -\dfrac{d^2}{dx^2}$ es positivo definido

Recordemos que un operador A es positivo definido si

$$< A\,u,\ u > \ > 0$$

para toda $u \in D$, $u \neq 0$.

Consideremos el operador $A = -\dfrac{d^2}{dx^2}$ con condiciones a la frontera $u(0) = u(1) = 0$. Ya hemos visto que los valores propios de A son de la forma $n^2\,\pi^2$ y sus vectores propios son $u_n(x) = \sqrt{2}$ sen $n\,\pi\,x$. Estos constituyen un conjunto ortonormal completo en $L_2\,(0,\ 1)$. Si $u \in D$, entonces

$$u = \sum_{k=1}^{\infty} a_k\, u_k .$$

Por ser A un operador con espectro discreto tenemos

$$Au = \sum_{k=1}^{\infty} \lambda_k\, a_k\, u_k$$

De aquí que

$$<Au,\ u> = \sum_{k=1}^{\infty} \lambda_k\, |\, a_k\,|^2 \geqslant \sum_{k=1}^{\infty} \lambda_1\, |\, a_k\,|^2 = \pi <u,\ u>$$

Esto muestra que A es positivo definido.

Si cambiamos las condiciones a la frontera $u'(0) = u'(1) = 0$, el resultado también es válido, pero ahora $u_n(x) = \sqrt{2}\cos n\pi x$.

7.2 Aplicación a los ejemplos del capítulo 0

7.2.1 Una observación importante

Observación. En los ejemplos (B) y (C), se está pensando en que primero actúa el operador $\partial/\partial t$ en la variable t y luego $\partial^2/\partial x^2$ sobre la variable x. Esto significa que podemos considerar al operador $\partial^2/\partial x^2$ como el operador d^2/dx^2.

7.2.2 La ecuación de calor

Por comodidad tomaremos $k = 1$ en la ec (I.3). En este ejemplo tenemos que considerar el operador

$$A = \frac{d^2}{dx^2}$$

con condiciones a la frontera $u'(0) = u'(1) = 0$. Los valores propios de A son $-n^2\pi^2$ y los vectores propios $u_n(x) = \sqrt{2}\cos n\pi x$.

Para cada $t \geqslant 0$, la función $g(x) = e^{tx}$, $x \leqslant 0$, satisface las condiciones para poder definir $g(A)$.

La función $f(x)$ puede ser expresada como

$$f(x) = \sum_{n=1}^{\infty} a_n u_n = \sqrt{2}\sum_{n=1}^{\infty} a_n \cos n\pi x$$

Entonces

$$g(A)\,f(x) = e^{tA} f(x) = \sum_{n=1}^{\infty} a_n\, g(\lambda_n) \cos n\pi x$$

Como $\lambda_n = - n^2 \pi^2$, entonces

$$g(\lambda_n) = e^{-n^2 \pi^2 t}$$

y por lo tanto

$$e^{tA} f(x) = \sqrt{2} \sum_{n=1}^{\infty} a_n e^{-n^2 \pi^2 t} \cos n \pi x .$$

Pero $a_n = \langle f, u_n \rangle$ y entonces la solución $u(x, t)$ de (I.3) puede ser expresada como

$$u(x,t) = e^{tA} f(x) = 2 \sum_{n=1}^{\infty} \cos n \pi x \, e^{-n^2 \pi^2 t} \int_0^1 f(x) \cos n \pi s \; ds.$$

Solución que coincide con la que se encuentra, por ejemplo, con el método de separación de variables.

7.2.3 La cuerda vibrante

Aquí consideramos el operador

$$A = - \frac{d^2}{dx^2}$$

con condiciones a la fronter $u(0) = u(1) = 0$. Ya hemos visto que sus valores propios son $n^2 \pi^2$ y sus vectores propios $u_n(x) = \sqrt{2} \operatorname{sen} n \pi x$.

La función $g(x) = \cos t\sqrt{x}$, $t \geqslant 0$, $x \geqslant 0$ satisface las condiciones para poder definir $g(A)$.

De nuevo, supondremos que en la ec (I.5) $c^2 = 1$. La función $f(x)$ puede ser expresada como

$$f(x) = \sum_{n=1}^{\infty} a_n u_n = \sqrt{2} \sum_{n=1}^{\infty} a_n \operatorname{sen} n \pi x.$$

Entonces

$$g(A)\ f(x) = \cos t\sqrt{A}\ f(x) = \sum_{n=1}^{\infty} a_n\, g\,(\lambda_n)\ u_n\ (x).$$

Pero $\lambda_n = n^2\pi^2$ y entonces $g(\lambda_n) = \cos n\pi t$. De aquí que

$$\cos t\sqrt{A}\ f(x) = \sqrt{2}\sum_{n=1}^{\infty} a_n \operatorname{sen} n\pi x \cos n\pi t$$

Por lo tanto la solución de (I.5) viene dada por

$$u(x,\ t) = 2\sum_{n=1}^{\infty} \operatorname{sen} n\pi x \cos n\pi t \int_0^1 f(s) \operatorname{sen} n\pi s\, ds$$

Nuevamente esta solución coincide con la que se encuentra por el método de separación de variables.

APENDICE A. INTEGRAL DE RIEMANN–STIELTJES PARA FUNCIONES CON VALORES EN L(H)

A.0 Introducción

La integral de Riemann–Stieltjes puede ser generalizada de dos formas

i) Cuando el integrando toma valores en L(H)

ii) Cuando la función integradora toma valores en L(H)

Para generalizar la integral necesitaremos utilizar el siguiente lema ([1]).

Lema. Si A es acotado, entonces

$$\|A\| = \sup \langle Ax,\ y\rangle, \quad \|x\| = \|y\| = 1.$$

En lo que sigue supondremos que: $F(t) : [a,b] \to L(H)$; $\alpha(t) : [a,b] \to C$; P es una partición, $a = t_o <. . .< t_n = b$, del intervalo $[a,b]$; los σ_i son puntos tales que $t_{i-1} < \sigma_i < t_i$; y $|P| = \text{máx}_i\ (t_i - t_{i-1})$.

A.1 La integral cuando el integrando toma valores en L(H)

Definición. Consideremos la suma

$$S_P (F,\alpha) = \sum_{i=1}^{n} F(\sigma_i)[\alpha(t_i) - \alpha(t_{i-1})] \qquad (A.1)$$

Si

$$\lim S_P (F, \alpha)$$

existe cuando $|P| \to 0$, definimos este límite como la integral

$$\int_a^b F(t)\ d\ \alpha(t)$$

Definición. Una función $F : \Delta \subset C \to L(H)$ se dice que es continua en $\lambda_o \in \Delta$, si dada $\epsilon > 0$ existe $\delta(\epsilon) > 0$ tal que

$$\|F(\lambda) - F(\lambda_o)\| < \epsilon \quad \text{si } |\lambda - \lambda_o| < \delta \quad \text{y} \quad \lambda \epsilon \Delta$$

F se dice que es continua en Δ si es continua en cada punto de Δ.

Teorema A.1. Si f(t) es continua y $\alpha(t)$ de variación acotada, entonces la integral

$$\int_a^b F(t)\ d\ \alpha\,(t) \qquad (A.2)$$

existe si y solo si

$$\int_a^b \langle F(t)\ x,\ y\rangle\ d\,\alpha(t) \qquad (A.3)$$

existe para toda $x, y \epsilon H$. Además

$$\langle \int_a^b f(t)d\alpha(t)\ x,\ y\rangle = \int_a^b \langle f(t)\ x,y\rangle\ d\,\alpha(t) \qquad (A.4)$$

y

$$\left(\int_a^b F(t)\, d\,\alpha(t) \right) x = \int_a^b F(t)\, x\, d\,\alpha\, (t) \qquad (A.5)$$

para toda $x \in H$.

Demostración. De (A.1) y (A.2) se sigue (A.3). Veremos ahora que (A.3) implica (A.2). Sea $g(t) = \langle F(t)\, x,\, y \rangle$. Por ser $F(t)$ continua, $g(t)$ también lo es y en consecuencia la integral

$$\int_a^b g(t)\, d\,\alpha\,(t)$$

existe. Sea

$$A_n = \sum_{i=1}^{n} F(\sigma_i)\, [\alpha(t_i) - \alpha(t_{i-1})]\ ,$$

entonces

$$\langle A_n x, y \rangle = \sum_{i=1}^{n} \langle F(\sigma_i)\, x, y \rangle\, [\alpha(t_i) - \alpha(t_{i-1})] =$$

$$= \sum_{i=1}^{n} g(\sigma_i)\, [\alpha(t_i) - \alpha(t_{i-1})]$$

y por lo tanto

$$\lim_{n \to \infty} \langle A_n\, x,\, y \rangle = \int_a^b g(t)\, d\, \alpha\, (t).$$

Pero por otro lado

$$\lim_{n \to \infty} \langle A_n\, x,\, y \rangle = \langle \lim_{n \to \infty} A_n\, x,\, y \rangle.$$

Definamos

$$Ax = \lim_{n \to \infty} A_n\, x$$

para toda $x \in H$. Veremos que

$$\|A - A_n\| \to 0$$

y esto demostrará nuestra afirmación. Para esto notemos que a $\langle A_n x, y\rangle$ lo podemos escribir como

$$\langle A_n x, y\rangle = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} g(\sigma_i)\, d\,\alpha(t)$$

y a $\langle Ax, y\rangle$ como

$$\langle Ax, y\rangle = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} g(t)\; d\;\alpha(t)$$

Ya que g(t) es continua en [a, b], en particular es uniformemente continua en [a,b] y por lo tanto dada $\epsilon > 0$ es posible escoger una partición tal que

$$|g(t) - g(\sigma_i)| < \epsilon \qquad \text{si} \qquad t \in [t_{i-1}, t_i].$$

De aquí que

$$|\langle (A - A_n)x, y\rangle| = |\sum_{k=1}^{n} \int_{a}^{b} [g(t) - g(\sigma_i)]\; d\,\alpha(t)| \leqslant \epsilon\; V(\alpha),$$

donde $V(\alpha)$ es la variación total de $\alpha(t)$. De la desigualdad anterior y del Lema se sigue que

$$\|A - A_n\| \leqslant \epsilon\; V(\alpha)$$

y entonces

$$A = \int_{a}^{b} F(t)\; d\;\alpha(t)$$

Que es lo que deseabamos demostrar. Ahora, si la integral (A.2) existe, las propiedades (A.4) y (A.5) se siguen de la definición de dicha integral.

Corolario. Sean $F : \Delta \subset C \to L(H)$ analítica y Γ una curva rectificable dada por una función continua $z(t) : [a,b] \to C$. Si Γ está contenida en Δ, entonces la integral

$$\int_\Gamma F(x)\, dz = \int_a^b F(z(t))\, dz\ (t)$$

existe y además

$$<\int_a^b F(z)\, dz\, x, y> = \int_a^b <F(z)\, x, y>\, dz$$

para toda $x, y \in H$.

A.2 La integral cuando la función integradora toma valores en L(H)

Definición. Consideremos la suma

$$\Delta_p(F, \epsilon) = \sum_{i=1}^{n} a\,(\sigma_i)\,[F(t_i) - F(t_{i-1})]. \qquad \text{(A.6)}$$

Si

$$\text{lím}\ \Delta_p\,(F, a)$$

existe cuando $|P| \to 0$, definimos este límite como

$$\int_a^b a(t)\, d\ F(t)\ .$$

Definición. Una función $F(t) : [a, b] \to L(H)$ se dice que es de variación acotada si

$$\sup_P \sum_{i=1}^{n} \|F(t_i) - F(t_{i-1})\| < \infty.$$

Teorema A.2. Si $\alpha(t) : [a, b] \to C$ es continua y $F(t) : [a, b] \to L(H)$ es de variación acotada, entonces

i) La integral

$$\int_a^b \alpha(t) \; d \; F(t)$$

existe si y solo si

$$\int_a^b \alpha(t) \; d < F(t) \; x, \; y>$$

existe para toda $x, \; y \in H$. Además

$$< \int_a^b \alpha(t) \, d \, F(t) \, x, y> = \int_a^b \alpha(t) \, d <F(t) \, x, y>$$

para toda $x, y \in H$, y

$$\left(\int_a^b \alpha(t) \, d \, F(t) \right) x = \int_a^b \alpha(t) \, dF(t) \, x$$

para toda $x \in H$.

ii) Sea $A : D \quad H \to H$ cerrado. Si $F(t) \in D$ para toda $t \in [a, b]$ y $AF(t) \in L(H)$, entonces

$$A \left(\int_a^b \alpha(t) \; d \; F(t) \right) = \int_a^b \alpha(t) \; d \; A \; F(t).$$

Los métodos de demostración son análogos a los utilizados en el caso en que $F(t)$ es el integrando.

APENDICE B. FUNCIONES ANALITICAS CON DOMINIO EN C Y VALORES EN L(H)

El objetivo de este Apéndice es precisar la noción de función analítica cuando

$$F : \Delta \subset C \rightarrow L(H)$$

y reconstruir algunos de los resultados clásicos de la teoría de funciones analíticas.

En los siguientes resultados la numeración que aparece entre paréntesis corresponde al orden en que aparece dicho resultado en el trabajo.

Teorema B.1. (Teorema 4.1). Sea $F : \Delta \subset C \rightarrow L(H)$, Δ abierto. Una condición necesaria y suficiente para que F sea analítica en Δ es que $\langle F(\lambda)\, x, y\rangle$ sea analítica en Δ para toda $x,\ y \in H$.

Demostración. Veremos primero que F analítica en Δ implica que $\langle F(\lambda)\, x, y\rangle$, entonces

$$\left| \frac{\varphi(\lambda) - \varphi(\lambda_0)}{\lambda - \lambda_0} - \langle F'(\lambda_0)\, x, y\rangle \right| = \left| \left\langle \left[\frac{F(\lambda) - F(\lambda_0)}{\lambda - \lambda_0} - F'(\lambda_0)\right] x, y \right\rangle \right|$$

$$\leqslant \left\| \frac{F(\lambda) - F(\lambda_0)}{\lambda - \lambda_0} - F'(\lambda_0) \right\| \; \|x\| \; \|y\|$$

De aquí que $\varphi'(\lambda_0)$ existe para toda $\lambda_0 \in \Delta$, toda $x, y \in H$ y además

$$\varphi'(\lambda_0) = \langle F'(\lambda_0)\, x, y\rangle .$$

Veremos ahora la otra parte de la demostración. Supongamos, pues, que $\varphi(\lambda) = \langle F(\lambda)\, x, y\rangle$ analítica en Δ. Mostraremos que para cada $\lambda_0 \in \Delta$, la expresión

$$\frac{F(\lambda) - F(\lambda_0)}{\lambda - \lambda_0} - \frac{F(z) - F(\lambda_0)}{z - \lambda_0}$$

tiende a cero cuando λ y z tienden a λ_0. Sea $r > 0$ tal que $\lambda \in \Delta$ si $|\lambda - \lambda_0| \leqslant r$. Supongamos que $0 < |\lambda - \lambda_0| < r$ y $0 < |z - \lambda_0| < r$. Por la fórmula de Cauchy tenemos que

$$\varphi(\lambda) = \frac{1}{2\pi i} \int_\Gamma \frac{\varphi(\xi)}{\xi - \lambda}\, d\xi$$

donde Γ es el círculo $|\xi - \lambda_0| = r$. Para z y λ_0 se tienen las correspondientes fórmulas. Cálculos directos muestran que

$$\left\langle \left[\frac{F(\lambda) - F(\lambda_0)}{\lambda - \lambda_0} - \frac{F(z) - F(\lambda_0)}{z - \lambda_0} \right] x, y \right\rangle = \frac{1}{2\pi i} \int_\Gamma \frac{\varphi(\xi)\,(\lambda - z)}{(\xi - \lambda)\,(\xi - \lambda_0)\,(\xi - z)}\, d\xi$$

Ahora, como $\varphi(\lambda)$ es continua y Γ compacto, debe existir un número M tal que $|\varphi(\xi)| \leqslant M$ para toda $\xi \in \Gamma$. Supongamos que $|\lambda - \lambda_0| \leqslant r/2$ y también $|z - \lambda_0| \leqslant r/2$; esto implica que $|\xi - \lambda| \geqslant r/2$. Entonces

$$\left| \left\langle \left[\frac{F(\lambda) - F(\lambda_0)}{\lambda - \lambda_0} - \frac{F(z) - F(\lambda_0)}{z - \lambda_0} \right] x, y \right\rangle \right| \leqslant 4M\, |\lambda - z| / r^2$$

y utilizando el Lema del Apéndice A tenemos

$$\left\| \frac{F(\lambda) - F(\lambda_o)}{\lambda - \lambda_o} - \frac{F(z) - F(\lambda_o)}{z - \lambda_o} \right\| \leqslant 4 M | \lambda - z | / r^2$$

Esto muestra que

$$\frac{F(\lambda) - F(\lambda_o)}{\lambda - \lambda_o}$$

tiene límite, cuando $\lambda \to \lambda_o$, para toda $\lambda_o \in \Delta$.

Teorema B.2. Si $F : \Delta \subset C \to L(H)$ es analítica en Δ, entonces es continua en Δ.

La demostración se sigue directamente de las definiciones.

A continuación veremos el teorema de Cauchy

Teorema B.3. (Teorema 4.3). Si $F : \Delta \subset C \to L(H)$ es analítica, entonces

$$\int_\Gamma F(z)\, dz = 0$$

para cualquier curva Γ simple cerrada en Δ, cuyo interior esté contenido en Δ.

Demostración. Por el corolario del Teorema A.1 y por el Teorema B.1, tenemos que

$$< \int_\Gamma F(z)\, dz\; x, y> = \int_\Gamma < F(z)\, x, y>\, dz = 0$$

para toda $x, y \in H$. Si tomamos

$$y = \int_\Gamma F(z)\, dz\; x$$

obtenemos

$$\left\| \int_\Gamma F(z)\, dz\; x \right\| = 0$$

para toda $x \in H$. De aquí que

$$\int_{\Gamma} F(z)\, dz = 0 .$$

Corolario. Sea $F : \Delta \subset C \to L(H)$ analítica y sean Γ_1, Γ_2 curvas simples cerradas que no se intersectan entre sí, una curva contenida en el interior de la otra, tales que la región acotada por ellas esté contenida en Δ, entonces

$$\int_{\Gamma_1} F(z)\, dz = \int_{\Gamma_2} F(z)\, dz .$$

Observación. Si observamos la demostración del teorema B.3, nos daremos cuenta de que no se utilizó el hecho de que Γ es simple cerrada, y que lo que es importante es que

$$\int_{\Gamma} \langle F(z)\, x, y\rangle\, dz = 0$$

De aquí que el teorema B.3 siga siendo válido en sus formas más generales. En particular, nosotros haremos uso de este teorema para el caso en el que Γ sea la unión de un número finito de curvas simples cerradas. También usaremos el Corolario en esta forma más general.

Por técnicas análogas se pueden reconstruir los resultados clásicos de la teoría de funciones analíticas tales como el teorema de Liouville, los desarrollos en series de Taylor y Laurent, el intercambio de sumatorias con la integral etc., etc.

A continuación demostraremos que la resolvente de un operador acotado es una función analítica.

Teorema B.4. Sea A acotado. Entonces si $\lambda, z \in \rho(A)$, R_λ, R_z satisfacen las siguientes propiedades:

i) $R_\lambda - R_z = (z - \lambda)\, R_\lambda\, R_z$

ii) $R_\lambda \; R_z = R_z \; R_\lambda$

iii) $\dfrac{d}{d\lambda} \; R_\lambda = - \, R_\lambda^2$

Demostración

i) Supongamos que $y \in H$ y que $x = R_z \; y$, esto es, $y = (zI - A)x$. Como

$$(zI - A) \; x - (\lambda I - A) \; x = (z - \lambda) \; x$$

tenemos

$$y - (\lambda I - A) \; R_z \; y = (z - \lambda) \; R_z \; y \; .$$

Aplicando R_λ a ambos lados de esta igualdad, obtenemos

$$R_\lambda \; y - R_z \; y = (z - \lambda) \; R_\lambda \; R_z \; y$$

para toda $y \in H$.

ii) $R_\lambda - R_z = - (R_z - R_\lambda) = - (\lambda - z) \; R_z \; R_\lambda = (z - \lambda) \; R_z \; R_\lambda$

Entonces

$$R_\lambda \; R_z = R_z \; R_\lambda$$

iii) Mostraremos que

$$\left\| \frac{R_z - R_\lambda}{z - \lambda} + R_\lambda^2 \right\| \to 0$$

Entonces

$$\frac{R_z - R_\lambda}{z - \lambda} + R_\lambda^2 = - \; R_z \, R_\lambda + R_\lambda^2 = (R_\lambda - R_z) \, R_\lambda = (z - \lambda) \; R_z \; R_\lambda^2$$

de aquí que

$$\left\| \frac{R_z - R_\lambda}{z - \lambda} + R_\lambda^2 \right\| \leq |z - \lambda| \, \|R_z\| \, \|R_\lambda\|^2$$

Por otro lado, de las desigualdades

$$\|(zI - A)\, x\| \geq \|(\lambda I - A)x\| - |z - \lambda| \, \|x\|$$

$$\|x\| = \|R_\lambda(\lambda I - A)\, x\| \leq \|R_\lambda\| \, \|(\lambda I - A)\, x\| ,$$

resulta

$$\|R_z\| \leq \|R_\lambda\| \, (1 - |z - \lambda| \, \|R_\lambda\|)^{-1}$$

siempre y cuando $|z - \lambda| < 1/\|R_\lambda\|$. De aquí que

$$\left\| \frac{R_z - R_\lambda}{z - \lambda} + R_\lambda^2 \right\| \leq |z - \lambda| \, \|R_\lambda\|^3 \, (1 - |z - \lambda| \, \|R_\lambda\|)^{-1}$$

y entonces

$$\frac{R_z - R_\lambda}{z - \lambda} \to - R_\lambda^2$$

cuando $z \to \lambda$. Esto demuestra el teorema.

Corolario. Sean $f : \Delta \subset C \to C$ analítica y $\sigma(A) \subset \Delta$. La integral

$$\frac{1}{2\pi i} \int_\Gamma f(z)\, (zI - A)^{-1}\, dz$$

está bien definida si Γ es una curva simple cerrada contenida en $\rho(A) \cap \Delta$.

BIBLIOGRAFIA

1. BACHMAN, G., NARICI, L. **Functional Analysis.** *Academic Press* (1966)

2. BUCK, R. C., Ed. **Studies in Modern Analysis.** *Prentice Hall* (1962)

3. FRIEDMAN, B. **Principles and Techniques of Applied Mathematics.** *John Wiley and Sons, Inc.* (1956)

4. FRIEDMAN, B. **An Abstract Formulation of the Method of Separtion of Variables.** *Proceedings of the Conference on Differential Equations.* University of Maryland (1956), 209-226

5. GREEN, C. D. **Integral Equations Methods.** *Nelson* (1969)

6. HALMOS, P. **Introduction to Hilbert Space.** *Chelsea* (1957)

7. HELMBERG, G. **Introduction to Spectral Theory in Hilbert Space.** *North-Holland* 1968)

8. Hille, E., PHILLIPS, R. **Functional Analysis and Semi-groups.** *American Mathematical Society* (1957)

9. HOMEISEL, G. **Integral Equations.** *Nelson* (1967)

10. HOPF, L. **Differential Equations of Physics.** *Dover* (1948)

11. LANG, S. **Analysis II.** *Addison Wesley* (1969)

12. LORCH, E. R. **Spectral Theory.** *Oxford University Press* (1962)

13. LOVITT, N. V. Linear Integral Equations. *Dover* (1950)

14. RUDIN, W. Principles of Mathematical Analysis. *McGraw-Hill Book* (1964)

15. RUDIN, W. Real and Complex Analysis. *McGraw-Hill Book* (1966)

16. SCHECHTER, . Principles of Functional Analysis. *Academic Press* (1971)

17. SHIELDS, P. S. Linear Algebra. *Addison Wesley* (1964)

18. SMIRNOV, V. I. A curse of Higher Mathematics. Vol V. *Addison Wesley* (1964)

19. TAYLOR, A. E., Spectral Theory of Closed Distributive Operators. *Acta Math.*, 84 (1950), 189-224

20. TAYLOR, A. E. Introduction to Functional Analysis. *John Wiley and Sons, Inc.* (1958)