

Métodos Matemáticos de la Física

Oscar Reula

DICIEMBRE 2008

Agradecimientos:

En primer lugar quiero agradecer a Gloria Puente, por haberme apoyado durante todo el tiempo en que escribí estas notas, las cuales tomaron bastante tiempo de nuestras horas compartidas, en general largas noches. En segundo lugar a Bernardo González Kriegel quien volcara a \LaTeX las primeras versiones de estas notas, haciendo un gran trabajo con mucho entusiasmo. También a Bob Geroch, con quien discutí varios temas de las notas y de quien también me inspiré a través de sus escritos y libros, no solo en contenido sino también en estilo. Finalmente a varias camadas de estudiantes que asimilaron estoicamente una gran cantidad del material de estos cursos en un tiempo demasiado corto.

Prefacio:

Estas notas, ahora devenidas en libro, se originaron como un intento de condensar en un solo lugar un gran conjunto de ideas, conceptos y herramientas matemáticas que considero básicas para la comprensión y el trabajo diario de un físico en nuestros días.

Usualmente sucede que si un problema es formulado desde una necesidad de origen físico, como por ejemplo la descripción de algún fenómeno natural, entonces éste está bien formulado, en el sentido de que una solución razonable al mismo existe. Esta regla ha sido en general muy fructífera y en particular les ha servido como guía a muchos matemáticos para abrirse camino en áreas desconocidas. Pero también ha servido, en particular a muchos físicos, para trabajar sin preocuparse demasiado por aspectos formales, ya sean analíticos, algebraicos o geométricos y poder así concentrarse en aspectos físicos y/o computacionales. Si bien esto permite un rápido desarrollo de algunas investigaciones, a la larga se llega a un estancamiento pues al proceder de este modo se evita enfrentar problemas que son muy ricos en cuanto a la conceptualización del fenómeno a describir. Es importante constatar que el problema formulado tiene una solución matemática y físicamente correcta.

Un ejemplo de esto ha sido el desarrollo, a mediados del siglo pasado, de la teoría moderna de las ecuaciones en derivadas parciales. Muchas de estas ecuaciones surgieron debido a que describen fenómenos de físicos: transmisión del calor, propagación de ondas electromagnéticas, ondas cuánticas, gravitación, etc. Una de las primeras respuestas matemáticas al desarrollo de estas áreas fue el teorema de Cauchy-Kowalevski que nos dice que dada una ecuación en derivadas parciales, (bajo ciertas circunstancias bastante generales) si una función analítica es dada como dato en una hipersuperficie (con ciertas características), luego existe una solución única en un entorno suficientemente pequeño de dicha hipersuperficie. Tomó mucho tiempo darse cuenta que este teorema realmente no era relevante desde el punto de vista de las aplicaciones físicas: existían ecuaciones admitidas por el teorema tales que si el dato no era analítico ¡no había solución! Y en muchos casos, si éstas existían, no dependían continuamente del dato dado, una pequeña variación del dato producía una solución totalmente distinta. Recién a mediados del siglo pasado se logró un avance sustancial al problema, encontrando que habían distinto tipo de ecuaciones, hiperbólicas, elípticas, parabólicas, etc. que se comportaban de manera distinta y esto reflejaba los distintos procesos físicos que las mismas simulaban. Debido a su relativa actualidad, este conjunto tan importante de conceptos no forman parte del conjunto de herramientas con que cuentan muchos de los físicos en actividad ni tampoco se encuentran en los libros de texto usualmente utilizados en las carreras de grado.

Como el anterior hay muchos ejemplos, en particular la teoría de ecuaciones diferenciales ordinarias y la geometría, sin la cual es imposible comprender muchas de las teorías modernas, tales como la relatividad, las teorías de partículas elementales y muchos fenómenos de la física del estado sólido. A medida que avanza nuestra comprensión de los fenómenos básicos de la naturaleza más nos damos cuenta que la herramienta más importante para su descripción es la geometría. Esta, entre otras cosas, nos permite manejar una amplia gama de procesos y teorías sin mucho en común entre sí con un conjunto muy reducido de conceptos, lográndose así una síntesis. Éstas síntesis son las que nos permiten adquirir nuevos conocimientos, ya que mediante su adopción dejamos espacio en nuestras mentes para aprender nuevos conceptos, los cuales son a su vez ordenados de manera más eficiente dentro de nuestra construcción mental del área.

Estas notas fueron originalmente pensadas para un curso de cuatro meses de duración. Pero

en realidad se adaptaban más para un curso anual o dos semestrales. Luego, a medida que se fueron incorporando más temas a las mismas, resultó más y más claro que deben darse en dos cursos semestrales o uno anual. Básicamente un curso debería contener los primeros capítulos que incluyen nociones de topología, espacios vectoriales, álgebra lineal, finalizando con la teoría de las ecuaciones en derivadas ordinarias. La tarea se simplifica considerablemente si los estudiantes han tenido previamente un buen curso de álgebra lineal. La correlación con las materias de física debería ser tal que el curso sea previo o concurrente con una mecánica avanzada. Haciendo hincapié en la misma sobre el hecho de que en definitiva uno está resolviendo ecuaciones diferenciales ordinarias con cierta estructura especial. Utilizando los conceptos del álgebra lineal para encontrar modos propios y la estabilidad de puntos de equilibrio. Y finalmente utilizando la geometría para describir aunque más no sea someramente la estructura simpléctica subyacente.

El segundo curso consiste en desarrollar las herramientas para poder discutir aspectos de la teoría de ecuaciones en derivadas parciales. Debería darse antes o concurrentemente con un curso avanzado de electromagnetismo, donde se debería hacer hincapié en el tipo de ecuaciones que se resuelven (elípticas, hiperbólicas), y el sentido de sus condiciones iniciales o de contorno, según corresponda. Usando además en forma coherente el concepto de distribución, que lejos de ser un concepto matemático abstracto es en realidad un concepto que aparece naturalmente en la física.

Nada del contenido de estas notas es material original, sí algunas formas de presentarlo, por ejemplo algunas pruebas más simples que las usuales, o la forma de integrar cada contenido con los anteriores. Mucho del material debería ser pensado como una primera lectura o una iniciación al tema y el lector interesado en profundizar debería leer los libros citados, de los cuales he extraído mucho material, siendo éstos excelentes y difíciles de superar.

Índice general

1. Conceptos Básicos de Topología	13
1.1. Introducción	13
1.1.1. Terminología	14
1.2. Conceptos Derivados	15
1.2.1. Mapas continuos	15
1.2.2. Compacidad	17
2. Álgebra Lineal	21
2.1. Espacios Vectoriales	21
2.1.1. Covectores y Tensores	22
2.1.2. Complexificación	24
2.1.3. Espacios cociente	24
2.2. Normas	25
2.2.1. Las normas inducidas en V^*	27
2.3. Teoría de Operadores Lineales	27
2.3.1. Representación Matricial	28
2.3.2. Subespacios Invariantes	29
2.3.3. Forma Canónica de Jordan	35
2.3.4. Relación de Semejanza	37
2.4. Operadores Adjuntos	38
2.5. Operadores Unitarios	43
2.6. Problemas	43
3. Geometría	47
3.1. Variedades	47
3.2. Funciones Diferenciables en M	50
3.3. Curvas en M	51
3.4. Vectores	52
3.5. Campos Vectoriales y Tensoriales	54
3.5.1. El Corchete de Lie	54
3.5.2. Difeomorfismos y la Teoría de las Ecuaciones Diferenciales Ordinarias	54
3.5.3. Campos de Covectores y Tensores	55
3.5.4. La Métrica	56
3.5.5. Notación de Índices Abstractos	57
3.6. Derivada Covariante	58
4. Ecuaciones Diferenciales Ordinarias	63
4.1. Introducción	63
4.2. El Caso de una Única Ecuación Ordinaria de Primer Orden	63
4.2.1. Ecuación Autónoma de Primer Orden	63
4.2.2. Extendiendo la Solución Local	69
4.2.3. El Caso No-autónomo	70

4.3. Reducción a Sistemas de Primer Orden	71
4.4. Sistemas de EDO	73
4.4.1. Integrales Primeras	74
4.4.2. Teorema Fundamental de los Sistemas de EDO	75
4.4.3. Dependencia en Parámetros, Ecuación de Variaciones	75
4.5. Problemas	77
5. Sistemas Lineales	81
5.1. Sistema lineal homogéneo	81
5.2. Sistema Lineal Inhomogeneo – Variación de constantes	83
5.3. Sistemas lineales homogéneos: coeficientes constantes	84
5.4. Problemas	88
6. Estabilidad	91
6.1. Problemas	96
7. Prueba del Teorema Fundamental	99
7.1. Problemas	103
8. Elementos Básicos de Análisis Funcional	105
8.1. Introducción	105
8.2. Completando un Espacio Normado	105
8.3. *Integral de Lebesgue	106
8.4. Espacios de Hilbert	111
8.5. Serie de Fourier	120
8.6. Problemas	125
8.7. Problemas de Series de Fourier	126
9. Distribuciones	129
9.1. Introducción	129
9.2. La derivada de una distribución	132
9.3. Nota sobre la completitud de \mathcal{D} y su dual \mathcal{D}'	134
9.4. Convergencia y Compacidad Débil	135
10. La Transformación de Fourier	137
10.1. Introducción	137
10.2. *Propiedades básicas de los Espacios de Sobolev	141
11. Teoría de ecuaciones en derivadas parciales	147
11.1. Introducción	147
11.2. La ecuación de primer orden	149
11.2.1. El Problema de Cauchy	150
11.3. Clasificación de ecuaciones en derivadas parciales	152
12. Ecuaciones Elípticas	157
12.1. La Ecuación de Laplace	157
12.1.1. Existencia	159
12.1.2. *Regularidad de las Soluciones	160
12.2. Teorema Espectral	162

13.Ecuaciones simétrico–hiperbólicas	167
13.1. Introducción	167
13.2. Un ejemplo	167
13.3. Desigualdad de la energía para sistemas simétrico–hiperbólicos	172
13.4. Unicidad de las soluciones	174
13.5. Dominio de dependencia	174
13.5.1. Construcción de una superficie característica	176
13.5.2. Dominio de dependencia, ejemplos	178
14.Ecuaciones Parabólicas	181
14.1. Introducción	181
14.2. Unicidad y el Teorema del Máximo	182

Índice de figuras

2.1.	Diagrama del operador dual.	25
2.2.	Diagrama del operador dual.	39
2.3.	Diagrama del operador estrella.	40
2.4.	Vectores normales y tangentes a la esfera.	42
3.1.	Un atlas de la esfera.	47
3.2.	Relación entre cartas.	48
3.3.	Sucesiones en M	48
3.4.	Ejemplo de variedad no-Hausdorff.	49
3.5.	Composición del mapa de una carta con una función.	50
3.6.	La relación entre las f_i	51
3.7.	Diferenciabilidad de curvas en M	51
3.8.	Definición de vector.	53
4.1.	Interpretación geométrica de la ecuación $f(x) = x^{1/2}$	64
4.2.	Prueba del Lema 4.1.	66
4.3.	Prueba de la unicidad de la solución.	67
4.4.	Distintas soluciones de la ecuación de crecimiento bacteriano	68
4.5.	Unicidad global.	69
4.6.	Extendiendo la solución.	70
4.7.	El péndulo Físico.	73
4.8.	La variedad Péndulo Físico.	74
6.1.	Estabilidad.	91
6.2.	Péndulo físico con fricción.	93
6.3.	La norma $\rho(x, x)$	94
7.1.	Entornos usados en la prueba del Teorema Fundamental.	102
8.1.	Una sucesión de Cauchy.	107
8.2.	Otra sucesión de Cauchy.	107
8.3.	Integral de Lebesgue.	108
8.4.	El espacio perpendicular.	115
8.5.	Ley del paralelogramo.	115
11.1.	Curvas características.	149
11.2.	Intersección de soluciones.	150
11.3.	Construyendo la solución a partir de la curva γ	151
11.4.	Construyendo la solución local.	153
12.1.	La membrana de un tambor.	157
12.2.	Problema mixto.	158

13.1. Sistema coordenado <i>nulo</i> .	168
13.2. Propagación de ondas.	169
13.3. Solución general homogénea en 1+1 dimensiones.	170
13.4. Solución general inhomogénea.	170
13.5. Desigualdad de la energía.	172
13.6. Desigualdad de la energía, vista en perspectiva.	173
13.7. Región en forma de ampolla.	175
13.8. Cono característico.	175
13.9. Construcción de S y una singularidad en S .	178
13.10. Dominio de dependencia de un fluido.	179
14.1. Condiciones de contorno para la ecuación del calor.	181
14.2. Prueba del Lema 14.1.	184

1.1. Introducción

La noción de conjunto, si bien nos dice que ciertos objetos —los elementos que lo conforman— tienen algo en común entre sí, no nos da ninguna idea de la *cercanía* de entre estos elementos, mientras que por otro lado si consideramos por ejemplo los números reales esta noción está presente. Sabemos por ejemplo que el número 2 está mucho más cercano al 1 que lo que lo está el 423. El concepto de una topología en un conjunto que definiremos a continuación trata de captar con precisión esta noción de cercanía la cual, como veremos admite muchas gradaciones.

Definición: Un **espacio topológico** consiste en un par (X, \mathcal{T}) , donde X es un conjunto y \mathcal{T} es una colección de subconjuntos de X satisfaciendo las siguientes condiciones:

1. Los subconjuntos \emptyset y X de X están en \mathcal{T} .
2. Sea O_λ , $\lambda \in I$, una familia monoparamétrica de subconjuntos de X en \mathcal{T} , luego $\bigcup_I O_\lambda$ está también en \mathcal{T} .
3. Si O y O' están en \mathcal{T} , también lo está $O \cap O'$.

Los elementos de \mathcal{T} , subconjuntos de X , son llamados los **subconjuntos abiertos** de X , el conjunto \mathcal{T} en sí es llamado una **topología** de X . La condición 2) nos dice que infinitas uniones de elementos de \mathcal{T} también están en \mathcal{T} , mientras que la condición 3) nos dice que en general solo finitas intersecciones siguen estando en \mathcal{T} . De los ejemplos siguientes se desprende el porqué de esta asimetría, ellos también nos ilustrarán de cómo dar una topología es esencialmente dar una noción de cercanía entre los puntos del conjunto en cuestión.

Ejemplo: a) Sea $\mathcal{T} = \{\emptyset, X\}$, es decir que los únicos subconjuntos abiertos de X son el subconjunto vacío y el subconjunto X . Es claro que esta colección de subconjuntos es una topología, ya que satisface las tres condiciones requeridas, a esta topología se la denomina **indiscreta**. Podemos decir que en esta topología los puntos de X están arbitrariamente cerca entre sí, ya que si un abierto contiene a uno de ellos los contiene a todos.

Ejemplo: b) Sea $\mathcal{T} = \mathcal{P}(X)$, la colección de todos los subconjuntos de X , claramente esta colección también satisface las condiciones arriba mencionadas y por lo tanto también es una topología de X , la llamada **discreta**. Podemos decir que en ésta todos los puntos están arbitrariamente separados entre sí ya que por ejemplo, dado cualquier punto de X existe un abierto que separa a éste de todos los demás, el que consiste de solo el punto en cuestión.

Ejemplo: c) Sea X el conjunto de los números reales, de ahora en más, \mathbb{R} , y sea $\mathcal{T} = \{O \mid \text{si } r \in O, \exists \varepsilon > 0 \text{ tal que si } |r - r'| < \varepsilon, r' \in O\}$, es decir la colección de abiertos en el sentido usual. Veamos que esta colección satisface las condiciones para ser una topología. Claramente $\emptyset \in \mathcal{T}$, (ya que no tiene ningún r), lo mismo que \mathbb{R} , (ya que contiene a todos los r), y así condición 1) es satisfecha.

Veamos la segunda, sea $r \in \bigcup_I O_\lambda$ luego $r \in O_\lambda$ para algún λ y por lo tanto existirá $\varepsilon > 0$ tal que todo r' con $|r - r'| < \varepsilon$ está también en O_λ , y por lo tanto en $\bigcup_I O_\lambda$. Veamos finalmente la tercera, sea $r \in O \cap O'$ luego $r \in O$ y por lo tanto existirá $\varepsilon > 0$ tal que todo r' con $|r - r'| < \varepsilon$ estará en O , como r también está en O' existirá $\varepsilon' > 0$ tal que todo r' con $|r - r'| < \varepsilon'$ estará en O' . Sea $\varepsilon'' = \min\{\varepsilon, \varepsilon'\}$ luego todo r' con $|r - r'| < \varepsilon''$ estará en O y en O' y por lo tanto en $O \cap O'$, con lo que concluimos que este último conjunto también está en \mathcal{T} . \mathbb{R} con esta topología es llamada la **línea real**.

Ejercicio: Encuentre usando el ejemplo anterior una intersección infinita de abiertos que no es abierta.

Ejemplo: d) Sea $X = \mathbb{R} \times \mathbb{R} \equiv \mathbb{R}^2$, es decir el producto cartesiano de \mathbb{R} consigo mismo –el conjunto de todos los pares (x, y) , con $x, y \in \mathbb{R}$ – y sea $\mathcal{T} = \{O \mid \text{si } (x, y) \in O, \exists \varepsilon > 0 \text{ tal que si } |x - x'| + |y - y'| < \varepsilon, (x', y') \in O\}$. Del ejemplo anterior se puede ver que este también es un espacio topológico y que esta es la topología que usualmente usamos en \mathbb{R}^2 .

Definición: Un **espacio métrico**, (X, d) es un par consistente en un conjunto X y un mapa $d : X \times X \rightarrow \mathbb{R}$, llamado usualmente distancia, satisfaciendo las siguientes condiciones:

1. $d(x, x') \geq 0, = 0 \Rightarrow x = x'$.
2. $d(x, x') = d(x', x)$.
3. $d(x, x') + d(x', x'') \geq d(x, x'')$.

Ejercicio: Pruebe que este espacio posee una topología *inducida* por su métrica en forma similar a \mathbb{R} en el ejemplo anterior.

Ejercicio: Vea que $d(x, y) = 1$ si $x \neq y$, es una distancia. ¿Qué topología nos introduce dicha distancia?

Claramente una métrica nos da una noción de cercanía entre puntos, ya que nos da un valor numérico de la distancia entre sí de estos. Una topología, al no darnos en general ningún número nos da una noción de cercanía mucho más vaga, pero de todos modos en general interesante.

1.1.1. Terminología

Damos a continuación un resumen de la terminología usual en esta área, la misma es una generalización directa de la usada comúnmente.

Definición: Llamaremos el **complemento**, O^c , del subconjunto O de X al subconjunto de todos los elementos de X que no están en O .

Definición: Diremos que un subconjunto O de X es **cerrado** si su complemento O^c es abierto.

Definición: Un subconjunto N de X es llamado un **entorno** de $x \in X$ si existe un abierto O_x , con $x \in O_x$, contenido en N .

Definición: Llamaremos el **interior** de $A \in X$ al subconjunto $\text{Int}(A)$ de X formado por la unión de todos los abiertos contenidos en A .

Definición: Llamaremos la **clausura** de $A \in X$ al subconjunto $\text{Cl}(A)$ de X formado por la intersección de todos los cerrados conteniendo a A .

Definición: Llamaremos la **frontera** de $A \in X$ al subconjunto ∂A de X formado por $Cl(A) - Int(A) \equiv Int(A)^c \cap Cl(A)$.

Ejercicio: Sea (X, d) un espacio vectorial métrico, pruebe que:

- a) $C_x^1 = \{x' | d(x, x') \leq 1\}$ es cerrado y es un entorno de x .
- b) $N_x^\varepsilon = \{x' | d(x, x') < \varepsilon\}$, $\varepsilon > 0$ es también un entorno de x .
- c) $Int(N_x^\varepsilon) = N_x^\varepsilon$
- d) $Cl(N_x^\varepsilon) = \{x' | d(x, x') \leq \varepsilon\}$
- e) $\partial N_x^\varepsilon = \{x' | d(x, x') = \varepsilon\}$.

Ejercicio: Sea (X, T) un espacio topológico y A un subconjunto de X . Pruebe que:

- a) A es abierto si y solo si cada $x \in A$ tiene un entorno contenido en A .
- b) A es cerrado si y solo si cada x en A^c (o sea no perteneciente a A) tiene un entorno que no intersecta a A .

Ejercicio: Sea (X, T) un espacio topológico, sea $A \in X$ y $x \in X$. Pruebe que:

- a) $x \in Int(A)$ si y solo si x tiene un entorno contenido en A .
- b) $x \in Cl(A)$ si y solo si todo entorno de x intersecta A .
- c) $x \in \partial A$ si y solo si todo entorno de x contiene puntos en A y puntos en A^c .

1.2. Conceptos Derivados

En las secciones anteriores hemos visto que el concepto de una topología nos lleva a una generalización de una serie de ideas y conceptos derivados que manejábamos en \mathbb{R}^n , las cuales no dependían de la métrica usual usada en estos espacios (la llamada Métrica Euclídea). Cabe entonces preguntarse si hay otras generalizaciones posibles todavía. En esta y en la próxima subsección estudiaremos dos más de ellas, éstas a su vez abren una vasta área de las matemáticas, que no trataremos en este curso pero que es muy importante en lo que respecta a la física moderna.

La primera de ellas es la de continuidad.

1.2.1. Mapas continuos

Definición: Sea $\varphi : X \rightarrow Y$ un mapa entre dos espacios topológicos. (Ver recuadro.) Diremos que el mapa φ es **continuo** si dado cualquier abierto O de Y , $\varphi^{-1}(O)$ es un abierto de X .

Definición: Un **mapa** $\phi : X \rightarrow Y$ entre un conjunto X y otro Y es una asignación a *cada* elemento de X de un elemento de Y .

Esto generaliza el concepto de función usual, note que el mapa está definido para todo elemento de X , mientras que en general su **imagen**, es decir el conjunto $\phi(X) \equiv \{y \in Y \mid \exists x \in X \text{ y } \phi(x) = y\}$, no es todo Y . En el caso que lo es, es decir que $\phi(X) = Y$, diremos que el mapa es **surjectivo**. Por otro lado si se cumple que $\phi(x) = \phi(\tilde{x}) \implies x = \tilde{x}$ diremos que el mapa es **injectivo**. En tal caso existe el mapa inverso a ϕ entre el conjunto $\phi(X) \subset Y$ y X . Si el mapa es además surjectivo entonces su inverso está definido en todo Y y en este caso se denota por $\phi^{-1} : Y \rightarrow X$. Es de utilidad considerar también los conjuntos $\phi^{-1}(O) = \{x \in X \mid \phi(x) \in O\}$

Claramente la definición anterior solo usa conceptos topológicos, ¿Tiene algo que ver con la usual *epsilon-delta* usada en \mathbb{R}^n ? La respuesta es afirmativa, como veremos más abajo en nuestro primer teorema, pero primero veamos algunos ejemplos.

Ejemplo: a) Sean X e Y cualquiera y sea la topología de X la discreta. Luego cualquier mapa entre X e Y es continuo. En efecto, para cualquier O abierto en Y , $\varphi^{-1}(O)$ es algún subconjunto en X , pero en la topología discreta todo subconjunto de X es un abierto.

Ejemplo: b) Sean X e Y cualquiera y sea la topología de Y la indiscreta. Luego también cualquier mapa entre X e Y es continuo. En efecto, los únicos abiertos en Y son \emptyset e Y , pero $\varphi^{-1}(\emptyset) = \emptyset$, mientras que $\varphi^{-1}(Y) = X$, pero cualquiera sea la topología de X , \emptyset y X son abiertos.

De los ejemplos anteriores parecería ser que nuestra definición de continuidad no es muy interesante, eso es debido a que hemos tomado casos con las topologías *extremas*, en las topologías intermedias es donde la definición se hace más útil.

Ejemplo: c) Sean X e Y líneas reales, y sea $\varphi(x) = 1$ si $x \geq 0$, $\varphi(x) = -1$ si $x < 0$. Este mapa no es continuo ya que, por ejemplo, $\varphi^{-1}((1/2, 3/2)) = \{x | x \geq 0\}$.

Teorema 1.1 *El mapa $\varphi : X \rightarrow Y$ es continuo si y solo si se cumple que: dado cualquier punto $x \in X$ y cualquier entorno M de $\varphi(x)$, existe un entorno N de x tal que $\varphi(N) \subset M$.*

Esta segunda definición está mucho más cerca del concepto intuitivo de continuidad.

Prueba: Supongamos φ continuo. Sea x un punto de X , y M un entorno de $\varphi(x)$. Luego existe un abierto O en Y contenido en M y conteniendo a $\varphi(x)$. Por continuidad $N = \varphi^{-1}(O)$ es un abierto de X , y como contiene a x , un entorno de x . Se cumple entonces que $\varphi(N) \subset O \subset M$. Supongamos ahora que dado cualquier punto $x \in X$ y cualquier entorno M de $\varphi(x)$, existe un entorno N de x tal que $\varphi(N) \subset M$. Sea entonces O un abierto cualquiera de Y , debemos mostrar ahora que $\varphi^{-1}(O)$ es un abierto de X . Sea x un punto cualquiera de $\varphi^{-1}(O)$, luego $\varphi(x) \in O$ y por lo tanto O es un entorno de $\varphi(x)$, por lo tanto existe un entorno N de x tal que $\varphi(N) \subset O$ y por lo tanto $N \subset \varphi^{-1}(O)$. Pero entonces $\varphi^{-1}(O)$ contiene un entorno de cada uno de sus puntos y por lo tanto es abierto.

Ejercicio: Sea $\phi : X \rightarrow Y$ y $\psi : Y \rightarrow Z$ mapas continuos, pruebe que $\psi \circ \phi : X \rightarrow Z$ también es continuo. (Composición de mapas preserva continuidad.)

Topología Inducida:

Sea ϕ un mapa entre un conjunto X y un espacio topológico $\{Y, \mathcal{T}\}$. Este mapa proporciona naturalmente, es decir sin la ayuda de ninguna otra estructura, una topología en X , denotada por \mathcal{T}_ϕ y llamada la **topología inducida** por ϕ en X . El conjunto de sus abiertos está dado por: $\mathcal{T}_\phi = \{O \subset X \mid O = \phi^{-1}(Q), Q \in \mathcal{T}\}$, es decir O es un abierto de X si existe un abierto Q de Y tal que $O = \phi^{-1}(Q)$.

Ejercicio: Demuestre que esta construcción realmente define una topología. No todas las topologías así inducidas son de interés y en general dependen fuertemente del mapa, como lo demuestra el siguiente ejemplo:

Ejemplo:

a) Sea $X = Y = \mathbb{R}$ con la topología usual y sea $\phi : \mathbb{R} \rightarrow \mathbb{R}$ la función $\phi(x) = 17 \forall x \in \mathbb{R}$. Esta función es claramente continua con respecto a las topologías de X e Y , las de la línea real. Sin embargo \mathcal{T}_ϕ , la topología inducida en X por este mapa es la indiscreta!

b) Sea X e Y como en a) y sea $\phi(x)$ un mapa invertible, luego \mathcal{T}_ϕ coincide con la topología de la línea real.

1.2.2. Compacidad

La otra generalización corresponde al concepto de **Compacidad**. Para ello introducimos la siguiente definición: Sea X un conjunto, A un subconjunto de éste y $\{A_\lambda\}$ una colección de subconjuntos de X parametrizados por una variable continua o discreta λ . Diremos que esta colección **cubre** A si $A \subset \bigcup_\lambda A_\lambda$.

Definición: Diremos que A es **compacto** si dada cualquier colección $\{A_\lambda\}$ de *abiertos* que lo cubren, existe un número *finito* de estos A_λ que también lo cubren.

Ejemplo: a) Sea X un conjunto infinito de puntos con la topología discreta. Luego un cubrimiento de X consiste, por ejemplo, en todos los puntos de éste, considerados cada uno de ellos como un subconjunto de del mismo. Pero la topología de X es la discreta, por lo tanto este es un cubrimiento de abiertos y ningún número finito de ellos lo cubrirá, por lo tanto X no es en este caso compacto. Claramente si X tuviese solo un número finito de elementos siempre sería compacto, cualquiera fuese su topología.

Ejemplo: b) Sea X cualquier conjunto con la topología indiscreta. Luego X es compacto. Los únicos abiertos de este conjunto son \emptyset y X , por lo tanto cualquier cubrimiento tiene a X como uno de sus miembros y éste solo alcanza para cubrir a X .

Vemos así que esta propiedad depende fuertemente de la topología del conjunto. La relación con el concepto intuitivo de compacidad queda clara de los siguientes ejemplo y ejercicio.

Ejemplo: c) Sea X la línea real y $A = (0, 1)$. Este subconjunto no es compacto pues por ejemplo el siguiente es un cubrimiento de abiertos de A tal que cualquier subconjunto finito del mismo no lo es. $A_n = (\frac{1}{n}, \frac{n-1}{n})$

Ejercicio: Sea X la línea real y $A = [0, 1]$. Pruebe que A es compacto.

Prueba:

Sea $\{A_\lambda\}$ un cubrimiento de $[0, 1]$ y $a \in [0, 1]$ la menor cota superior de los $x \in (0, 1]$ tales que $[0, x]$ está cubierto por un subcubrimiento finito. a existe pues 0 tiene un A que lo cubre. Sea A_{λ_0} un elemento del cubrimiento tal que $a \in A_{\lambda_0}$. Luego existe $b < a$ tal que $b \in A_{\lambda_0}$ y b ya está cubierto por un subcubrimiento finito. Tenemos así que a está en un subcubrimiento finito y por lo tanto, si $a \neq 1$ también algunos elementos mayores al mismo. Lo que constituiría una contradicción. ♠

Veamos ahora la relación entre los dos conceptos derivados del de Topología, es decir el de la continuidad de mapas entre espacios topológicos y el de compacidad. El hecho de que un mapa entre espacios topológicos sea continuo implica que este mapa es especial, en el sentido de que *tiene o lleva información sobre las respectivas topologías y preserva las propiedades topológicas de los conjuntos que asocia*. Esto se ve en la siguiente propiedad, la cual –como se desprende del ejemplo que sigue– es muy importante.

Teorema 1.2 Sean X e Y dos espacios topológicos y ϕ un mapa continuo entre ellos. Luego si A es un subconjunto compacto de X , $\phi(A)$ es un subconjunto compacto de Y .

Prueba: Sea O_λ una colección de abiertos en Y que cubren a $\phi(A)$. Luego la colección $\phi^{-1}(O_\lambda)$ cubre a A , pero A es compacto y por lo tanto habrá una subcolección finita $\phi^{-1}O_{\tilde{\lambda}}$ de la anterior que también lo cubre. Por lo tanto la subcolección finita $O_{\tilde{\lambda}}$ cubrirá también a $\phi(A)$. Como esto es cierto para cualquier colección de abiertos cubriendo a $\phi(A)$ concluimos que éste es compacto.

Ejemplo: Sea A compacto y sea $\phi : A \rightarrow \mathbb{R}$ continuo, es decir un mapa continuo entre A y la línea real. $\phi(A)$ es entonces un conjunto compacto de la línea real y por lo tanto un conjunto cerrado y acotado, pero entonces este conjunto tendrá un máximo y un mínimo, es decir el mapa ϕ alcanza su máximo y mínimo en A .

Finalmente otro teorema de fundamental importancia acerca de los conjuntos compactos, el cual muestra que éstos tienen otra propiedad que los hace muy interesantes. Para ello introducimos las siguientes definiciones, las cuales también solo en conceptos topológicos. Una **sucesión o secuencia** en un conjunto X $\{x_n\} = \{x_1, x_2, \dots\}$, con $x_n \in X$, es un mapa de los números enteros en este conjunto. Dada una sucesión $\{x_n\}$ en un espacio topológico X , diremos que $x \in X$ es un **punto de convergencia o límite** de esta sucesión si dado cualquier abierto O de X conteniendo a x existe un número N tal que para todo $n > N$ $x_n \in O$. Diremos que $x \in X$ es un **punto de acumulación** de esta sucesión si dado cualquier abierto O de X conteniendo a x , infinitos elementos de la sucesión también pertenecen a O .

Ejercicio: Encuentre un ejemplo de una sucesión en algún espacio topológico con diferentes puntos límites.

Teorema 1.3 *Sea A compacto. Luego toda sucesión en A tiene un punto de acumulación.*

Prueba: Supongamos –en contradicción con la afirmación del teorema– que existe una sucesión $\{x_n\}$ sin ningún punto de acumulación. Es decir, dado cualquier punto x de A existe un entorno O_x conteniéndolo y un número N_x tal que si $n > N_x$ luego $x_n \notin O_x$. Como esto es válido para cualquier x en A , la colección de conjuntos $\{O_x | x \in A\}$ cubre A , pero A es compacto y por lo tanto existirá una subcolección finita de éstos que también lo cubre. Sea N el máximo entre los N_x de esta colección finita. Pero entonces $x_n \notin A$ para todo $n > N$ lo que es absurdo.

Ejercicio: Pruebe que los conjuntos compactos en la línea real son los cerrados y acotados.

Nos podemos hacer ahora la pregunta inversa: ¿Si $A \subset X$ es tal que toda sucesión tiene puntos de acumulación, es cierto entonces que A es compacto? Una respuesta afirmativa nos daría una caracterización alternativa de la compacidad, y ésta es afirmativa para el caso de la línea real. En general la respuesta es negativa: hay topologías en las cuales toda sucesión en un conjunto tiene puntos de acumulación en él, pero este no es compacto. Sin embargo todas las topologías que nosotros veremos son **numerales de segunda especie** [Ver recuadro] y en éstas la respuesta es afirmativa.

En la línea real es cierto que si $x \in \mathbb{R}$ es un punto de acumulación de una sucesión $\{x_n\}$ entonces existe una **subsucesión**, $\{\tilde{x}_n\}$, –es decir una restricción del mapa definiendo la sucesión a un número infinito de números naturales–, que tendrá a x como punto límite. Esto tampoco es cierto en la generalidad de los espacios topológicos, pero sí lo es si consideramos solo aquellos que son **numerales de primera especie** [Ver recuadro]. Todos los espacios que veremos en este curso lo son.

***Numerabilidad de espacios topológicos.**

Definición: Diremos que un espacio topológico $\{X, \mathcal{T}\}$ es numerable de primera especie si para cada $p \in X$ existe una colección contable de abiertos $\{O_n\}$ tal que todo abierto conteniendo a p contiene también al menos uno de estos O_n .

Definición: Diremos que un espacio topológico $\{X, \mathcal{T}\}$ es numerable de segunda especie si hay una colección numerable de abiertos tal que cualquier abierto de X puede ser expresado como una unión de conjuntos de esta colección.

Ejemplo:

- a) X con la topología indiscreta es numerable de primera especie.
- b) X con la topología discreta es numerable de primera especie. Y de segunda especie si y solo si sus elementos son numerables.

Ejercicio: Demostrar que la línea real es numerable de primera y segunda especie. Ayuda: Para el primer caso use los abiertos $O_n = (p - \frac{1}{n}, p + \frac{1}{n})$ y para el segundo $O_{pqn} = (\frac{p}{q} - \frac{1}{n}, \frac{p}{q} + \frac{1}{n})$

***Separabilidad de espacios topológicos.**

Definición: Un espacio topológico X es Hausdorff si dados cualquier par de puntos de X , x e y , luego existen entornos O_x y O_y tales que $O_x \cap O_y = \emptyset$.

Ejemplo:

- a) X con la topología indiscreta es Hausdorff.
- b) X con la topología discreta no es Hausdorff.

Ejercicio: Encontrar una topología tal que los enteros sean Hausdorff y compactos.

Ejercicio: Pruebe que si un espacio es Hausdorff entonces si una sucesión tiene un punto límite este es único.

Ejercicio: Sea X compacto, Y Hausdorff y $\phi : X \rightarrow Y$ continuo. Pruebe que las imágenes de conjuntos cerrados son cerradas.

Notas bibliográficas: Este capítulo es esencialmente un condensado de los capítulos **26,27,28,29** y **30** de [1], ver también [2], [6] y [21]. La topología es una de las más apasionantes ramas de las matemáticas, ¡si profundiza quedará atrapado! De particular interés en física en la noción de Homotopía un buen lugar para entender estas ideas es el capítulo **34** de [1].

2.1. Espacios Vectoriales

Definición: Un **Espacio Vectorial Real** consiste de tres cosas – *i*) Un conjunto, V , cuyos elementos serán llamados **vectores**. *ii*) Una regla que asigna a cada par de vectores, \mathbf{v} , \mathbf{u} , un tercer vector que denotaremos por $\mathbf{v} + \mathbf{u}$ y que llamaremos su **suma**. *iii*) Una regla que asigna a cada vector, \mathbf{v} y a cada número real a , un vector que denotaremos por $a\mathbf{v}$ y que llamaremos el **producto** de \mathbf{v} con a . – sujetas a las siguientes condiciones:

- 1.a) Para cualquier par de vectores \mathbf{u} , $\mathbf{v} \in V$ se cumple que,

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \quad (2.1)$$

- 1.b) Existe en V un elemento llamado **cero** y denotado por $\mathbf{0}$, tal que $\mathbf{0} + \mathbf{v} = \mathbf{v} \forall \mathbf{v} \in V$.

- 1.c) Para cualquier vector $\mathbf{u} \in V$ existe un único vector en V , denotado $-\mathbf{u}$, tal que,

$$\mathbf{u} + (-\mathbf{u}) = \mathbf{0} \quad (2.2)$$

- 2.a) Para cualquier par de números reales a y a' y cualquier vector \mathbf{v} se cumple que,

$$a(a'\mathbf{v}) = (aa')\mathbf{v}.$$

- 2.b) Para cualquier vector \mathbf{v} se cumple que,

$$1\mathbf{v} = \mathbf{v}.$$

- 3.a) Para cualquier par de números reales a y a' y cualquier vector \mathbf{v} se cumple que,

$$(a + a')\mathbf{v} = a\mathbf{v} + a'\mathbf{v}.$$

- 3.b) Para cualquier número real a y cualquier par de vectores \mathbf{v} y \mathbf{v}' se cumple que,

$$a(\mathbf{v} + \mathbf{v}') = a\mathbf{v} + a\mathbf{v}'.$$

Las primeras son condiciones que involucran solo la regla de la suma. Las siguientes solo involucran a la regla del producto, mientras que las dos últimas tratan de la relación entre estas dos operaciones.

Ejemplo: El conjunto de todas las n -tuplas de números reales con las operaciones usuales de suma y multiplicación *tupla por tupla*. Este espacio se denomina \mathbb{R}^n .

Ejemplo: Sea S cualquier conjunto y sea V el conjunto de todas las funciones de S en los reales, $v : S \rightarrow \mathbb{R}$, con las siguientes operaciones de suma y producto: La suma de la función v con la función v' es la función (elemento de V) que al elemento s de S le asigna el valor $v(s) + v'(s)$. El producto de $a \in \mathbb{R}$ con la función v es la función que a $s \in S$ le asigna el valor $av(s)$. Este ejemplo aparecerá muy a menudo en los capítulos siguientes.

Definición: diremos que un conjunto de vectores $\{v_i\} i = 1, \dots, n$ son **linealmente independientes** si $\sum_i a^i v_i = \mathbf{0}$ $a^i \in \mathbb{R}, i = 1, \dots, n \implies a^i = 0, i = 1, \dots, n$, es decir si cualquier combinación lineal no trivial de estos vectores nos da un vector no trivial. Si un número finito de vectores linealmente independientes, n , son suficientes para **expandir** V , [es decir si cualquier vector en V puede ser obtenido como una combinación lineal de estos n vectores], entonces diremos que estos forman una **base** de V y que la **dimensión** de V es n , $\dim V = n$.

Ejercicio: Muestre que la dimensión de V es única, es decir que no depende de la base empleada para definirla.

Note que si en el ejemplo anterior S consta de un número finito de elementos, luego la dimensión de V es finita.¹ En el caso de que S tuviese un número infinito de elementos diríamos que la dimensión de V sería infinita. En lo que sigue de este capítulo solo consideraremos espacios vectoriales de dimensión finita.

Problema 2.1 Sea S un conjunto finito, $S = \{s_1, s_2, \dots, s_n\}$, encuentre una base del espacio vectorial de todas las funciones de S en \mathbb{R} . Encuentre la dimensión de este espacio.

2.1.1. Covectores y Tensores

Sea V un espacio vectorial de dimensión n . Asociado con este espacio vectorial consideremos el conjunto, $V^* = \{ \text{el espacio de mapas lineales } f : V \rightarrow \mathbb{R} \}$. Este es también un espacio vectorial, llamado el **espacio dual a V** , o **espacio de covectores**, con suma y producto dados por: $(f + \alpha g)(v) = f(v) + \alpha g(v) \quad \forall v \in V \text{ con } f, g \in V^*, \alpha \in \mathbb{R}$. ¿Cuál es su dimensión? Note que si $\{v_i\} i = 1, \dots, n$ es una base de V , es decir un conjunto linealmente independiente de vectores que expanden V , podemos definir n elementos de V^* (llamados co-vectores) por la relación

$$\theta^i(v_j) = \delta^i_j. \quad (2.3)$$

Es decir definimos la acción de θ^i sobre los $\{v_j\}$ como en la ecuación de arriba y luego extendemos su acción a cualquier elemento de V escribiendo a este elemento en la base $\{v_i\}$ y usando el hecho que la acción debe ser lineal.

Se puede ver fácilmente que cualquier $\rho \in V^*$ se puede obtener como combinación lineal de los covectores $\{\theta^j\}, j = 1, \dots, n$ y que estos son linealmente independientes, por lo tanto forman una base y por lo tanto la dimensión de V^* es también n .

Ejercicio: Vea que V^* es un espacio vectorial y que las $\{\theta^i\}$ forman realmente una base.

Como V y V^* tienen la misma dimensión son, como espacios vectoriales, la misma cosa, pero como no existe ningún mapa que los identifique los tenemos que tomar como distintos.

¿Qué pasa si ahora tomamos el dual de V^* ? ¿Obtendremos así más copias de V ? La respuesta es no, ya que existe una identificación natural entre V y su doble dual V^{**} .

En efecto a cada $v \in V$ le podemos asociar un elemento x_v de V^{**} , es decir una funcional lineal de V^* en \mathbb{R} , de la siguiente forma: $x_v(\omega) = \omega(v) \quad \forall \omega \in V^*$. Es decir el elemento x_v de V^{**} asociado a $v \in V$ es aquel que al actuar sobre un covector cualquiera ω da el número $\omega(v)$. Note que x_v actúa linealmente en los elementos de V^* y por lo tanto es un elemento de V^{**} . ¿Hay

¹Es decir un número finito de vectores linealmente independientes expanden V .

elementos de V^{**} que no provengan de algún vector en V ? La respuesta es no, ya que el mapa $\mathbf{x}_v : V \rightarrow V^{**}$ es inyectivo [$\mathbf{x}_v(\omega) = 0 \quad \forall \omega \implies v = 0$] y por lo tanto ² $\dim \mathbf{x}_V = \dim V$. Por otro lado $\dim V^{**} = \dim V^*$ ya que V^{**} es el dual de V^* y así $\dim V = \dim V^* = \dim V^{**}$, lo que indica que el mapa en cuestión es también suryectivo y por lo tanto invertible. Esto nos permite identificar V y V^{**} y concluir que *dualizando* no podremos construir más espacios vectoriales interesantes.

Ejercicio: Vea que efectivamente $\dim \mathbf{x}_V = \dim V$.

Sin embargo nada nos impide considerar también **mapas multilineales** ³ de $\underbrace{V \times V \times \cdots \times V}_{k \text{ veces}}$ en \mathbb{R} , o más generalmente de $\underbrace{V \times \cdots \times V}_{k \text{ veces}} \times \underbrace{V^* \times \cdots \times V^*}_{l \text{ veces}}$ en \mathbb{R} . El conjunto de estos mapas (para cada par (k, l) dado) es también un espacio vectorial, –con las operaciones obvias– y sus elementos son llamados **tensores de tipo** (k, l) .

Ejercicio: ¿Cuál es la dimensión de estos espacios como función del par (k, l) ?

Nota: En dimensión finita se puede demostrar que cualquier tensor de tipo (k, l) se puede escribir como combinación lineal de elementos del producto cartesiano de k copias de V^* y l copias de V –donde hemos identificado a V con V^{**} –. Por ejemplo si \mathbf{t} es de tipo $(0, 2)$, –o sea un mapa que tiene como argumentos a dos covectores–, entonces dada una base $\{\mathbf{v}^i\}$ de V habrá $n \times n$ números reales t^{ij} , $i = 1, \dots, n$ tales que

$$\mathbf{t}(\boldsymbol{\sigma}, \boldsymbol{\omega}) = \sum_{i,j=1}^n t^{ij} \mathbf{v}_i(\boldsymbol{\sigma}) \mathbf{v}_j(\boldsymbol{\omega}), \quad \forall \boldsymbol{\sigma}, \boldsymbol{\omega} \in V^*. \quad (2.4)$$

Pero el producto de combinaciones lineales de productos cartesianos de k copias de V^* y l copias de V es también un espacio vectorial, se lo llama **producto externo** de k copias de V^* y l copias de V y se lo denota por $\underbrace{V^* \otimes V^* \otimes \cdots \otimes V^*}_{k \text{ veces}} \otimes \underbrace{V \otimes \cdots \otimes V}_{l \text{ veces}}$. Por lo tanto también se los puede considerar a los tensores como elementos de estos productos externos.

Ejemplo: a) Sea \mathbf{t} de tipo $(2, 0)$, o sea $\mathbf{t} \in V^* \otimes V^*$. Este es un mapa bilineal de $V \otimes V$ en \mathbb{R} , $\mathbf{t}(v, u) \in \mathbb{R}$. Sea \mathbf{t} simétrico [$\mathbf{t}(v, u) = \mathbf{t}(u, v)$] y no degenerado [$\mathbf{t}(v, \cdot) = 0 \in V^* \implies v = 0$]. Como \mathbf{t} es no degenerado define un mapa invertible entre V y su dual.

Ejemplo: b) Sea $\boldsymbol{\varepsilon}$ un elemento de $(n, 0)$ tal que

$$\boldsymbol{\varepsilon}(\dots, \underbrace{v}_i, \dots, \underbrace{u}_j, \dots) = -\boldsymbol{\varepsilon}(\dots, \underbrace{u}_i, \dots, \underbrace{v}_j, \dots) \quad (2.5)$$

cualquiera sea la casilla i y j , es decir un tensor totalmente antisimétrico. Sea $\{\mathbf{u}_i\}$ una base de V y $\boldsymbol{\varepsilon}_{123\dots n} = \boldsymbol{\varepsilon}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$, luego cualquier otra componente de $\boldsymbol{\varepsilon}$ en esta base será o bien cero o $\boldsymbol{\varepsilon}_{123\dots n}$ o $-\boldsymbol{\varepsilon}_{123\dots n}$ de acuerdo a que algún \mathbf{u}_i se repita o sea una permutación par de la de arriba o una impar. Por lo tanto basta un número, $\boldsymbol{\varepsilon}_{123\dots n}$, para determinar el tensor $\boldsymbol{\varepsilon}$ y dado y otro tensor $\tilde{\boldsymbol{\varepsilon}}$ no idénticamente cero con las propiedades arriba mencionadas luego existirá un número α tal que $\boldsymbol{\varepsilon} = \alpha \tilde{\boldsymbol{\varepsilon}}$.

Ejercicio: Sea $\boldsymbol{\varepsilon}$ no idénticamente cero y $\{\mathbf{u}_i\}$ un conjunto de $n = \dim V$ vectores de V . Muestre que estos forman una base si y solo si

$$\boldsymbol{\varepsilon}(u_1, \dots, u_n) \neq 0. \quad (2.6)$$

²Denotando por \mathbf{x}_V la imagen por $\mathbf{x}_{(\cdot)}$ de V .

³O sea mapas separadamente lineales en cada uno de sus argumentos.

Ejemplo: c) Sea A un elemento de $(1, 1)$,

$$u \in V, v^* \in V^* \rightarrow A(u, v^*) \in \mathbb{R}. \quad (2.7)$$

Esto implica que $A(u, \cdot)$ es también un vector (identificando V con V^{**} , aquel que toma una forma $\omega \in V^*$ y da el número $A(u, \omega)$). Tenemos así un mapa lineal de V en V , o sea un operador lineal en V .

Ejercicio: Sea $\{u_i\}$ una base de V y sean a_i los vectores $A(u_i, \cdot)$, luego

$$\varepsilon(a_1, \dots, a_n) = \varepsilon(A(u_1, \cdot), \dots, A(u_n, \cdot))$$

es totalmente antisimétrico en los $\{u_i\}$ y por lo tanto proporcional a $\varepsilon(u_1, \dots, u_n)$. La constante de proporcionalidad se llama **determinante** del operador A , $\det(A)$.

Problema 2.2 Muestre que esta definición no depende del ε empleado ni de la base y por lo tanto es realmente una función del espacio de operadores de V en \mathbb{R} .

Ejercicio: Si A y B son dos operadores de V , luego $A \cdot B(v) \equiv A(B(v))$. Muestre que $\det(AB) = \det(A) \cdot \det(B)$.

2.1.2. Complexificación

Otra manera de obtener campos vectoriales a partir de uno dado, digamos V , es extendiendo el campo donde está definida la operación de multiplicación, si esto es posible. El caso más común es la **complexificación** de un espacio vectorial real en ese caso simplemente se extiende el producto a los números complejos resultando un campo vectorial de la misma dimensión. Una manera de obtenerlo, por ejemplo es tomando un conjunto de vectores linealmente independientes del espacio inicial, es decir una base, y considerando todas las combinaciones lineales con coeficientes complejos arbitrarios. El espacio así obtenido se denota por V^C . Si las componentes de los vectores en V en la base original eran n -tuplas de números reales, ahora son n -tuplas de números complejos. Como la base es la misma, la dimensión también es la misma. Estas extensiones de espacios vectoriales aparecen a menudo y más adelante veremos otras.

Ejemplo: Considere el espacio vectorial Q^n consistente de todas las n -tuplas de números racionales. En este espacio el campo es también el conjunto de los racionales. Si extendemos el campo a los reales obtenemos \mathbb{R}^n .

2.1.3. Espacios cociente

La última forma que veremos de obtener espacios vectoriales a partir de otros espacios vectoriales es la de tomar **cocientes**. Sea V un espacio vectorial y sea $W \subset V$ un subespacio del mismo. Llamaremos **espacio cociente** al conjunto de clases equivalentes en V , donde diremos que dos vectores de V son equivalentes si su diferencia es un vector en W . Este espacio se denota como V/W .

Ejercicio: Probar que esta es una relación de equivalencia.

Veamos que este es un espacio vectorial. Los elementos de V/W son clases equivalentes, las cuales denotamos como $\{v\}$, dos elementos de V , v y v' pertenecen a la misma clase equivalente si $v - v' \in W$. Sean ζ y ζ' dos elementos de V/W , es decir dos clases equivalentes de elementos de V . Definiremos las operaciones propias en los espacios vectoriales de la siguiente forma: $\zeta + \alpha\zeta'$

será la clase equivalente correspondiente un elemento \tilde{v} obtenido tomando un elemento de V en ζ , digamos v , otro de ζ' , digamos v' , y definiendo $\tilde{v} = v + \alpha v'$, tenemos $\{\tilde{v}\} = \zeta + \alpha \zeta'$.

Ejercicio: Vea que esta definición no depende de la elección de elementos en la clase equivalente tomados para hacer la operación. Es decir, considere otros dos elementos en ζ y ζ' , digamos \tilde{v} y \tilde{v}' y vea que con ellos obtiene un elemento en la misma clase que $\tilde{v} = v + \alpha v'$.

Ejemplo: Sea $V = \mathbb{R}^2$, es decir el espacio de 2-tuplas de números reales. Sea v un elemento cualquiera. Este elemento genera el espacio unidimensional W_v que consiste en todos los vectores de la forma αv , para $\alpha \in \mathbb{R}$. El espacio cociente V/W_v es el espacio compuesto por las líneas paralelas a v . Es decir, cada línea es un elemento del espacio cociente y entre ellas existe una noción de suma y multiplicación por un escalar.

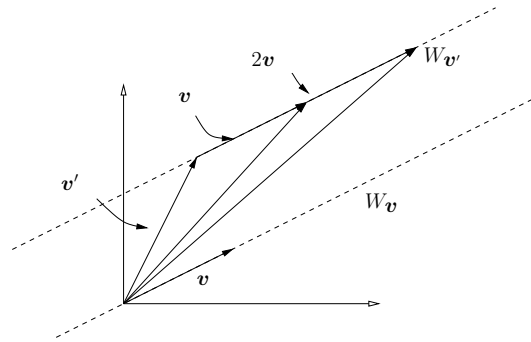


Figura 2.1: Diagrama del operador dual.

2.2. Normas

Definición: Una **norma** en un espacio vectorial V es un mapa $\|x\| : V \rightarrow \mathbb{R}^+$, satisfaciendo para todo $x, y \in V$, $\alpha \in \mathbb{R}$,

- i) $\|x\| \geq 0$ ($= \leftrightarrow x = 0$)
- ii) $\|\alpha x\| = |\alpha| \|x\|$
- iii) $\|x + y\| \leq \|x\| + \|y\|$

Ejemplos: En \mathbb{R}^2 :

- a) $\|(x, y)\| = \max\{|x|, |y|\}$;
- b) $\|(x, y)\| = \sqrt{x^2 + y^2}$, norma Euclídea;
- c) $\|(x, y)\| = |x| + |y|$.
- d) En V cualquiera sea t un tensor simétrico positivo definido de tipo $(0,2)$, es decir $t(u, v) = t(v, u)$, $t(u, u) \geq 0$ ($= \leftrightarrow u = 0$). La función $\|u\|_t = \sqrt{t(u, u)}$ es una norma. Cada tensor de este tipo genera una norma, pero hay muchas normas que no provienen de ningún tensor de este tipo. Dé un ejemplo.

Ejercicio: Pruebe que $|t(u, v)|^2 \leq \|u\|_t \|v\|_t$. Ayuda: Considere el polinomio: $P(\lambda) := t(u + \lambda v, u + \lambda v)$.

Ejercicio: Pruebe que los ejemplos dados son en realidad normas. Grafique las curvas de nivel de las tres primeras normas, es decir los conjuntos $S_a = \{(x, y) \in \mathbb{R}^2 / \|(x, y)\| = a\}$ y las “bolas de radio a ”, es decir $B_a = \{(x, y) \in \mathbb{R}^2 / \|(x, y)\| \leq a\}$.

Ejercicio: Pruebe que el mapa $d : V \times V \rightarrow \mathbb{R}^+$ dado por $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ es una métrica.

¿Qué es, geoméricamente una norma? Dado un vector $\mathbf{x} \neq 0$ de V y un número positivo cualquiera, a , existe un único número $\alpha > 0$ tal que $\|\alpha\mathbf{x}\| = a$. Esto indica que las superficies de nivel de la norma, es decir las hiper-superficies $S_a = \{\mathbf{x} \in V / \|\mathbf{x}\| = a\}$, $a > 0$ forman una familia suave de capas una dentro de la otra y cada una de ellas divide a V en tres conjuntos disjuntos, el *interior*⁴ de S_a –conteniendo el elemento $\mathbf{x} = 0$ –, S_a y el *exterior* de S_a . El *interior* de S_a es un conjunto convexo, es decir si \mathbf{x} e \mathbf{y} pertenecen al interior de S_a luego $\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}$, $\alpha \in [0, 1]$ también pertenece [ya que $\|\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}\| \leq \alpha\|\mathbf{x}\| + (1 - \alpha)\|\mathbf{y}\| \leq \alpha a + (1 - \alpha)a = a$]. Una curva de nivel caracteriza completamente una norma en el sentido que si damos un subconjunto N de V , tal que N tiene la propiedad radial, es decir dado $\mathbf{x} \neq 0$ existe $\alpha > 0$ tal que $\alpha\mathbf{x} \in N$, y su interior es convexo, luego existe una única norma tal que N es la superficie de nivel S_1 . Esta norma se define de la siguiente forma: dado \mathbf{x} habrá un $\alpha > 0$ tal que $\alpha\mathbf{x} \in N$ y entonces la norma de \mathbf{x} será $\|\mathbf{x}\| = \frac{1}{\alpha}$.

Ejercicio: Pruebe que esta es una norma.

De esta imagen se ve que dadas dos normas de un espacio vectorial de dimensión finita y una superficie de nivel de una, habrá superficies de nivel de la otra que tendrán la primera en su interior o en su exterior. En las normas a) y b) del ejemplo anterior vemos que dado un cuadrado conteniendo al cero, habrá dos círculos conteniendo al cero, uno conteniendo el cuadrado y otro contenido por éste. Esto nos lleva al siguiente Teorema.

Teorema 2.1 Sea V un espacio vectorial de dimensión finita. Luego todas sus normas son equivalentes entre sí, en el sentido que dadas $\|\cdot\|$ y $\|\cdot\|'$ existen constantes positivas M_1 y M_2 tal que para todo $\mathbf{x} \in V$ se cumple $M_1 \|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq M_2 \|\mathbf{x}\|$.

Prueba: Mostraremos que todas son equivalentes a la norma $\|\mathbf{x}\|_1 = \sum_{i=1}^n |a^i|$, donde los a^i son las componentes de \mathbf{x} , con respecto a una base $\{\mathbf{e}_i\}$ que supondremos dada, $\mathbf{x} = a^i \mathbf{e}_i$. Sea $\mathbf{y} = b^i \mathbf{e}_i$ y $\|\cdot\|$ otra norma cualquiera, luego

$$\begin{aligned} |\|\mathbf{x}\| - \|\mathbf{y}\|| &\leq \|\mathbf{x} - \mathbf{y}\| = \|(a^i - b^i) \mathbf{e}_i\| \\ &\leq |a^i - b^i| \|\mathbf{e}_i\| \leq (\max_{i=1, \dots, n} \|\mathbf{e}_i\|) \sum_{i=1}^n |a^i - b^i| \\ &= (\max_{i=1, \dots, n} \|\mathbf{e}_i\|) \|\mathbf{x} - \mathbf{y}\|_1 \end{aligned} \quad (2.8)$$

Esto demuestra que la norma $\|\cdot\|$ es una función continua con respecto a la norma $\|\cdot\|_1$. Sea S_1 la superficie de nivel de radio 1 con respecto a la métrica $\|\cdot\|_1$. S_1 es un conjunto cerrado y acotado y por lo tanto compacto⁵. Por lo tanto, por continuidad, $\|\cdot\|$ tiene un valor máximo, M_2 , y un mínimo, M_1 , que dan la desigualdad buscada.

Notas:

i) En esta prueba es crucial el hecho de que S_1 es compacto. Si V es de dimensión infinita esto no es así y hay muchas normas no equivalentes.

ii) Para nuestros propósitos cualquier norma es suficiente –ya que si por ejemplo, $f : V \rightarrow \mathbb{R}$ es diferenciable con respecto a una norma lo es también con respecto a cualquier otra equivalente a ésta– y por simplicidad de ahora en más usaremos la Euclídea.

⁴No confundir con el interior de un conjunto que definimos en el capítulo anterior, que en este caso sería S_a .

⁵Con respecto a la topología generada por $\|\cdot\|_1$.

iii) En este sentido las normas de espacios vectoriales son equivalentes a la generada por cualquier elemento simétrico-positivo del producto exterior de su dual consigo mismo.

iv) Como normas equivalentes generan una misma topología, vemos que en los espacios vectoriales de dimensión finita existe una única topología asociada con todas sus posibles normas. Esta usualmente se denomina la **topología fuerte**.

2.2.1. Las normas inducidas en V^*

Las normas definidas en V inducen naturalmente normas en su dual, V^* . Esta viene dada por:

$$\|\omega\| := \max_{\|v\|=1} \{|\omega(v)|\}. \quad (2.9)$$

Ejercicio: Vea que esta es una norma y que $|\omega(v)| \leq \|\omega\| \|v\|$.

2.3. Teoría de Operadores Lineales

Un **operador lineal** A de un espacio vectorial V es un mapa continuo⁶ de V en V tal que $\forall x, y \in V, \alpha \in \mathbb{R}, A(\alpha x + y) = \alpha A(x) + A(y)$, es decir un tensor de tipo (1,1).

El conjunto de los operadores lineales \mathcal{L} es un álgebra, es decir un espacio vectorial con producto. En efecto, si $A, B \in \mathcal{L}, \alpha \in \mathbb{R}$, luego $A + \alpha B \in \mathcal{L}$ y además $A \cdot B$ (el operador que a $x \in V$ lo envía a $A(B(x)) \in V$) también pertenece a \mathcal{L} . Debido a esta propiedad podemos además definir funciones no lineales de \mathcal{L} en \mathbb{R} y mapas de \mathcal{L} en \mathcal{L} . Para estudiar la continuidad y diferenciabilidad de estos mapas introduciremos una norma en \mathcal{L} , la más conveniente es la siguiente norma inducida de la usada en V ,

$$\|A\|_{\mathcal{L}} = \max_{\|x\|_V=1} \|A(x)\|_V. \quad (2.10)$$

Si V es de dimensión finita (lo que asumiremos de ahora en más), el espacio vectorial \mathcal{L} es también de dimensión finita y por lo tanto todas sus normas son equivalentes. El hecho que la norma de A sea finita usa nuevamente que $A : V \rightarrow V$ es continua y que $\{x \in V / \|x\|_V = 1\}$ es compacto, en el caso de dimensión infinita ninguna de estas cosas es necesariamente cierta y dentro de \mathcal{L} tenemos solo un subespacio de operadores lineales de norma finita (acotados).

Ejercicio: Muestre que $\|\cdot\|_{\mathcal{L}} : \mathcal{L} \rightarrow \mathbb{R}^+$ es una norma.

Ejercicio: Muestre que $\|A(v)\| \leq \|A\|_{\mathcal{L}} \|v\|$.

Ejercicio: Usando el resultado del ejercicio anterior muestre que $\|AB\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}} \|B\|_{\mathcal{L}}$.

Estudiamos a continuación varias funciones en el espacio de operadores.

El determinante de un operador, introducido en la sección anterior es un polinomio de grado $n = \dim V$ en A y por lo tanto diferenciable. Usando la regla de la cadena se ve que $\det(I + \varepsilon A)$ es diferenciable en ε , y de hecho un polinomio de grado n en ε . Cada uno de los coeficientes de este polinomio es una función de A . De importancia en lo que sigue es el coeficiente lineal en A que se obtiene usando la fórmula

$$\frac{d}{d\varepsilon} \det(I + \varepsilon A)|_{\varepsilon=0} = \frac{\varepsilon(A(u_1), u_2, \dots, u_n) + \dots + \varepsilon(u_1, \dots, A(u_n))}{\varepsilon(u_1, \dots, u_n)} \quad (2.11)$$

⁶Con respecto a la topología inducida por cualquiera de las normas equivalentes de V .

esta función se llama la **traza** de \mathbf{A} y se denota $tr(\mathbf{A})$.

Entre los mapas de \mathcal{L} en \mathcal{L} consideremos el mapa exponencial, definido como,

$$e^{\mathbf{A}} = \sum_{i=0}^{\infty} \frac{\mathbf{A}^i}{i!} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2} + \cdots \quad (2.12)$$

Teorema 2.2 $e^{\mathbf{A}} \in \mathcal{L}$ si $\mathbf{A} \in \mathcal{L}$ y $e^{t\mathbf{A}}$ es infinitamente diferenciable con respecto a t .

Prueba: Considere la sucesión de Cauchy $\{e_n^{\mathbf{A}}\}$, donde $e_n^{\mathbf{A}} \equiv \sum_{i=0}^n \frac{\mathbf{A}^i}{i!}$. Esta sucesión es de Cauchy ya que, tomando $m > n$, tenemos

$$\begin{aligned} \|e_m^{\mathbf{A}} - e_n^{\mathbf{A}}\|_{\mathcal{L}} &= \left\| \frac{\mathbf{A}^{m-1}}{(m-1)!} + \frac{\mathbf{A}^{m-2}}{(m-2)!} + \cdots + \frac{\mathbf{A}^{n+1}}{(n+1)!} \right\|_{\mathcal{L}} \\ &\leq \left\| \frac{\mathbf{A}^{m-1}}{(m-1)!} \right\|_{\mathcal{L}} + \left\| \frac{\mathbf{A}^{m-2}}{(m-2)!} \right\|_{\mathcal{L}} + \cdots + \left\| \frac{\mathbf{A}^{n+1}}{(n+1)!} \right\|_{\mathcal{L}} \\ &\leq \frac{\|\mathbf{A}\|_{\mathcal{L}}^{m-1}}{(m-1)!} + \frac{\|\mathbf{A}\|_{\mathcal{L}}^{m-2}}{(m-2)!} + \cdots + \frac{\|\mathbf{A}\|_{\mathcal{L}}^{n+1}}{(n+1)!} \\ &= \|e_m^{\mathbf{A}}\|_{\mathcal{L}} - \|e_n^{\mathbf{A}}\|_{\mathcal{L}} \rightarrow 0. \end{aligned} \quad (2.13)$$

Donde $e_m^{\mathbf{A}}\|_{\mathcal{L}} \equiv \sum_{i=0}^m \frac{\|\mathbf{A}^i\|_{\mathcal{L}}}{i!}$ y la última implicación sigue del hecho que la serie numérica $e^{\|\mathbf{A}\|_{\mathcal{L}}}$ converge. Pero por completitud ⁷ de \mathcal{L} toda sucesión de Cauchy converge a algún elemento de \mathcal{L} que llamaremos $e^{\mathbf{A}}$. La diferenciabilidad de $e^{t\mathbf{A}}$ sigue del hecho de que si una serie $\sum_{i=0}^{\infty} f_i(t)$ es convergente y $\sum_{i=0}^{\infty} \frac{df_i}{dt}$ es uniformemente convergente, luego $\frac{d}{dt} \sum_{i=0}^{\infty} f_i(t) = \sum_{i=0}^{\infty} \frac{d}{dt} f_i(t)$ ♠

Ejercicio: Muestre que

- a) $e^{(t+s)\mathbf{A}} = e^{t\mathbf{A}} \cdot e^{s\mathbf{A}}$,
- b) Si \mathbf{A} y \mathbf{B} conmutan, es decir si $\mathbf{AB} = \mathbf{BA}$, luego $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}} e^{\mathbf{B}}$.
- c) $\det(e^{\mathbf{A}}) = e^{tr(\mathbf{A})}$
- d) $\frac{d}{dt} e^{t\mathbf{A}} = \mathbf{A} e^{t\mathbf{A}}$.

Ayuda:

Para el punto c) use que $e^{\mathbf{A}}$ se puede definir también como $\lim_{m \rightarrow \infty} (\mathbf{I} + \frac{\mathbf{A}}{m})^m$.

2.3.1. Representación Matricial

Para describir ciertos aspectos de los operadores lineales es conveniente introducir la siguiente representación matricial.

Sea $\{\mathbf{u}_i\}$, $i = 1, \dots, n$ una base de V , es decir un conjunto de vectores linealmente independientes de V [$\sum_{i=1}^n c^i \mathbf{u}_i = 0 \implies c^i = 0$] que lo expanden [si $\mathbf{v} \in V$, existen números $\{v^i\}$, $i = 1, \dots, n$ tal que $\mathbf{v} = \sum_{i=1}^n v^i \mathbf{u}_i$]. Aplicando el operador \mathbf{A} a un miembro de la base \mathbf{u}_i obtenemos un vector $\mathbf{A}(\mathbf{u}_i)$ que a su vez puede ser expandido en la base, $\mathbf{A}(\mathbf{u}_i) = \sum_{j=1}^n A^j_i \mathbf{u}_j$. La matriz así construida, A^j_i , es una representación del operador \mathbf{A} en esa base. En este lenguaje vemos que la matriz A^j_i

⁷Todo espacio vectorial real de dimensión finita es completo.

transforma el vector de componentes $\{v^i\}$ en el vector de componentes $\{A^j_i v^i\}$. Dada una base, $\{\mathbf{u}_i\}$, y una matriz A^j_i podemos construir un operador lineal de la siguiente forma: Dada la base definimos su co-base, es decir una base en V^* como $\{\theta^i\}$, $i = 1, \dots, n$, tal que $\theta^i(\mathbf{u}_j) = \delta^i_j$, luego $\mathbf{A} = \sum_{i,j=1}^n A^j_i \mathbf{u}_j \theta^i$.

Si cambiamos de base las matrices representando los operadores cambiarán. En efecto, si tomamos otra base $\{\hat{\mathbf{u}}_i\}$ y escribimos sus componentes con respecto a la base anterior como $\hat{\mathbf{u}}_i = P_i^k \mathbf{u}_k$ entonces la relación entre las componentes del operador \mathbf{A} en ambas bases está dado por

$$\hat{A}^j_i = P_k^j A^k_l (P^{-1})^l_i \quad \text{o} \quad \hat{\mathbf{A}} = \mathbf{P} \mathbf{A} \mathbf{P}^{-1} \quad (2.14)$$

es decir que $\hat{\mathbf{A}}$ y \mathbf{A} son **matrices semejantes**.

Ejercicio: Ver, a partir de su definición (ecuación (2.11)), que $\text{tr} \mathbf{A} = \sum_{i=1}^n A^i_i$.

2.3.2. Subespacios Invariantes

Definición: Sea $\mathbf{A} : V \rightarrow V$ un operador y sea W un subespacio de V . Diremos que W es un **subespacio invariante** de \mathbf{A} si $\mathbf{A}W \subseteq W$.

Los subespacios invariantes de un operador son importantes pues nos permiten entender cual es su acción. Note que dado cualquier operador \mathbf{A} siempre existen al menos dos espacios invariantes, V y $\{\mathbf{0}\}$ y en realidad muchos más. Por ejemplo, y como veremos más adelante, dado cualquier número entre 1 y n ($= \dim V$) existe un subespacio invariante con ese número como dimensión. Los que verdaderamente codifican la acción del operador son sus subespacios invariantes **irreducibles**, es decir aquellos que no pueden ser a su vez descompuestos en subespacios invariantes.

Ejercicio: Sea el V el espacio de pares de números, es decir con elementos de la forma $\mathbf{v} = (a, b)$, $a, b \in \mathbb{R}$ y sea \mathbf{A} dado por $\mathbf{A}(a, b) = (b, 0)$. Encuentre sus subespacios invariantes irreducibles.

Estudiaremos en detalle los subespacios invariantes de dimensión 1, note que los mismos son irreducibles. Para estudiar los espacios invariantes resulta conveniente estudiar los subespacios invariantes del operador cuando su acción se extiende a V^C , es decir la **complexificación** de V .

Veamos a continuación que un operador siempre tiene al menos un subespacio invariante de dimensión 1 (y por lo tanto que siempre tiene un subespacio invariante irreducible no trivial).

Lema 2.1 Dado $\mathbf{A} : V^C \rightarrow V^C$, siempre existe un $\mathbf{u} \in V^C$ y un $\lambda \in C$ tal que,

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{u} = 0 \quad (2.15)$$

Prueba:

Una solución de esta ecuación consiste así de un escalar λ , llamado **autovalor** del operador \mathbf{A} y de un vector \mathbf{u} , llamado **autovector** del operador \mathbf{A} . El subespacio de V^C dado por $\{\alpha \mathbf{u} \mid \alpha \in C\}$ es el subespacio invariante buscado.

Es claro que el sistema tiene solución si y solo si $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$. Pero este es un polinomio en λ de orden igual a la dimensión de V y por lo tanto, por el Teorema Fundamental del álgebra tiene al menos una solución o raíz, (en general complejas), λ_1 , y por lo tanto habrá, asociada a ella, al menos un \mathbf{u}_1 solución de (2.15) con $\lambda = \lambda_1$ ♠

La necesidad de considerar todas estas soluciones es lo que nos lleva a tratar el problema para espacios vectoriales complejos.

Aplicación: Lema de triangulación de Schur

Definición: Una matriz $n \times n$ A^j_i tiene forma triangular superior si $A^j_i = 0 \ \forall j > i, j, i = 1, \dots, n$. Es decir es una matriz de la forma,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & \cdots & a_{2n} \\ 0 & 0 & \ddots & \ddots & a_{3n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & a_{nn} \end{pmatrix}. \quad (2.16)$$

Como veremos más adelante, en el capítulo 5, esta es una forma muy conveniente para poder entender cómo son las soluciones a sistemas de ecuaciones diferenciales ordinarias. Y lo más atractivo de la misma es que cualquier operador tiene una representación matricial con forma diagonal superior! Es más, si un producto escalar está presente, la base para esta representación puede ser elegida en forma orto-normal.

Lema 2.2 (Schur) Sea $A : V \rightarrow V$ un operador lineal actuando en un espacio vectorial complejo V de dimensión finita, n y sea (\cdot, \cdot) un producto escalar en V . Luego, existe una base orto-normal $\{\mathbf{u}_i\}$, $i = 1, \dots, n$ con respecto a la cual la representación matricial de A es triangular superior.

Prueba: Consideremos el problema de autovalores-autovectores para A ,

$$(A - \lambda I)\mathbf{u} = \mathbf{0}. \quad (2.17)$$

Como ya vimos este problema siempre tiene al menos una solución no trivial y por lo tanto tenemos un par $(\lambda_1, \mathbf{u}_1)$ solución del problema. Tomamos \mathbf{u}_1 de norma unidad y como primer elemento de la base a determinar. Tenemos entonces $A^j_1 = \theta^j(A(\mathbf{u}_1)) = \theta^j(\lambda_1 \mathbf{u}_1) = \lambda_1 \delta^j_1$.

Consideremos ahora el espacio $V_{n-1} = \{\mathbf{u}_1^\perp\} := \{\mathbf{u} \in V | (\mathbf{u}_1, \mathbf{u}) = 0\}$ y el operador de $V_{n-1} \rightarrow V_{n-1}$ dado por $(I - \mathbf{u}_1 \theta^1)A$. Note que como formaremos una base orto-normal conocemos ya el primer miembro de la co-base, $\theta^1 = (\mathbf{u}_1, \cdot)$. Para este operador, en este espacio, podemos plantear también la ecuación de autovalores-autovectores,

$$((I - \mathbf{u}_1 \theta^1)A - \lambda I)\mathbf{u} = \mathbf{0}. \quad (2.18)$$

Obtenemos así un nuevo par $(\lambda_2, \mathbf{u}_2)$, con \mathbf{u}_2 perpendicular a \mathbf{u}_1 y además, $A\mathbf{u}_2 = \lambda_2 \mathbf{u}_2 + \mathbf{u}_1 \theta^1(A(\mathbf{u}_2))$. Por lo tanto $A^j_2 = \theta^j(A(\mathbf{u}_2)) = \theta^j(\lambda_2 \mathbf{u}_2 + \mathbf{u}_1 \theta^1(A(\mathbf{u}_2))) = \lambda_2 \delta^j_2 + \delta^j_1 A^1_2$. El próximo paso es considerar ahora el subespacio $V_{n-2} = \{\mathbf{u}_1^\perp\} \cup \{\mathbf{u}_2^\perp\} := \{\mathbf{u} \in V | (\mathbf{u}_1, \mathbf{u}) = (\mathbf{u}_2, \mathbf{u}) = 0\}$ y allí la ecuación de autovalores-autovectores para el operador $(I - \mathbf{u}_1 \theta^1 - \mathbf{u}_2 \theta^2)A$. Prosiguiendo de esta forma generamos toda la base ♠

Problema 2.3 : Pruebe el teorema anterior usando ahora la noción de espacio cociente y prescindiendo así del producto escalar. Ayuda: en vez de usar los espacios perpendiculares a los autovectores que vaya obteniendo, use los espacios cociente por estos autovectores. Vea que un operador lineal induce una acción natural en un espacio cociente.

Continuamos ahora con el estudio de los subespacios invariantes. Si $\det(A - \lambda I)$ tiene $1 \leq m \leq n$ raíces distintas, $\{\lambda_i\}$, $i = 1, \dots, m$ habrá entonces al menos un autovector complejo \mathbf{u}_i asociado a cada una de ellas. Veamos que estos conforman subespacios invariantes distintos.

Lema 2.3 Sea $\{(\lambda_i, \mathbf{u}_i)\}$ $i = 1 \dots m$ un conjunto de pares de autovalores autovectores. Si $\lambda_i \neq \lambda_j \ \forall i \neq j, i, j = 1 \dots m$ entonces estos autovectores son linealmente independientes.

Prueba: Supongamos por contradicción que no y por lo tanto que existen constantes $c^i \in \mathbb{C}$, $i = 1, \dots, m-1$, tales que

$$\mathbf{u}_m = \sum_{i=1}^{m-1} c^i \mathbf{u}_i \quad (2.19)$$

Aplicando \mathbf{A} en ambos lados obtenemos

$$\mathbf{A}\mathbf{u}_m = \lambda_m \mathbf{u}_m = \sum_{i=1}^{m-1} c^i \lambda_i \mathbf{u}_i \quad (2.20)$$

o sea,

$$0 = \sum_{i=1}^{m-1} c^i (\lambda_m - \lambda_i) \mathbf{u}_i. \quad (2.21)$$

Concluimos así que $\{\mathbf{u}_i\}$ $i = 1, \dots, m-1$ son linealmente dependientes. Debido a 2.19 y a la hipótesis que los autovalores son distintos, al menos uno de los coeficientes tiene que ser distinto de cero y por lo tanto podemos despejar uno de los restantes autovectores en función de función otros $m-2$. Repitiendo este procedimiento $(m-1)$ veces llegamos a que necesariamente $\mathbf{u}_1 = 0$ lo que es una contradicción ya que, como hemos visto, la ecuación de autovectores siempre tiene una solución no-trivial para cada autovalor distinto ♠

Si por cada autovalor existe más de un autovector entonces éstos forman un subespacio vectorial invariante de mayor dimensión (reducible). Dentro de cada uno de estos subespacios podemos tomar una base compuesta por autovectores. El lema anterior nos asegurará entonces que todos estos autovectores así elegidos, para todos los autovalores, forman un gran conjunto linealmente independiente.

Ejercicio: Convénzase de que el conjunto de autovectores con un mismo autovalor forman un subespacio vectorial.

Si un dado operador \mathbf{A} tiene todos sus autovalores distintos entonces tenemos que los correspondientes autovectores son linealmente independientes e igualan en número a la dimensión de V , es decir generan una base de $V^{\mathbb{C}}$. En esa base la representación matricial de \mathbf{A} es diagonal, es decir $A^j_i = \delta^j_i \lambda_i$. Cada uno de sus autovectores genera un subespacio invariante irreducible y en su conjunto generan $V^{\mathbb{C}}$. En cada uno de ellos el operador \mathbf{A} actúa meramente por multiplicación por λ_i . Note que los λ_i son en general complejos y por lo tanto tal multiplicación es en realidad una rotación más una dilatación. Note, que a diferencia con la base del Lema de triangulación de Schur, ésta no es en general ortogonal con respecto a algún producto escalar dado de antemano.⁸

Ejemplo: Sea V el conjunto de 2-tuplas de números reales con elemento genérico (a, b) y sea \mathbf{A} dado por $\mathbf{A}(a, b) = (b, -a)$. Esta es una rotación de $-\pi/2$ en el plano. Tomando una base, $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$ vemos que

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \varepsilon((\mathbf{A} - \lambda \mathbf{I})\mathbf{e}_1, (\mathbf{A} - \lambda \mathbf{I})\mathbf{e}_2) / \varepsilon(\mathbf{e}_1, \mathbf{e}_2) \\ &= \varepsilon(-\mathbf{e}_2 - \lambda \mathbf{e}_1, \mathbf{e}_1 - \lambda \mathbf{e}_2) / \varepsilon(\mathbf{e}_1, \mathbf{e}_2) \\ &= 1 + \lambda^2. \end{aligned} \quad (2.22)$$

y por lo tanto los autovalores son $\lambda_1 = i$, $\lambda_2 = -i$. Los autovectores son $\mathbf{u}_1 = (1, -i)$ y $\mathbf{u}_2 = (1, i) = \bar{\mathbf{u}}_1$. Vemos entonces que la acción de \mathbf{A} en estos subespacios es multiplicación por $\pm i$ y que ambos subespacios invariantes son genuinamente complejos. En esta nueva base el espacio

⁸Note sin embargo que se puede declarar ortogonal definiendo como producto escalar $(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^n \theta^i(\bar{\mathbf{u}}) \theta^i(\mathbf{v})$

V es generado por todas las combinaciones lineales de la forma $z\mathbf{u}_1 + \bar{z}\mathbf{u}_2$ y la acción de \mathbf{A} es simplemente multiplicación por $-i$ de z .

Si la multiplicidad de alguna de las raíces $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ es mayor que uno habrá menos autovalores que la dimensión del espacio y por tanto no tendremos garantía de que habrá suficientes autovectores como para formar una base, ya que solo podemos garantizar la existencia de uno por cada autovalor. Debemos analizar por lo tanto qué pasa en estos casos. Para ello definamos, dado λ_i un autovalor de \mathbf{A} , los siguientes subespacios:

$$W_{\lambda_i p} = \{\mathbf{u} \in V \mid (\mathbf{A} - \lambda_i \mathbf{I})^p \mathbf{u} = 0\} \quad (2.23)$$

Note que estos son espacios invariantes: $\mathbf{A}W_{\lambda_i p} \subset W_{\lambda_i p}$. Además $W_{\lambda_i p} \subset W_{\lambda_i p+1}$ y por lo tanto, para un p suficientemente grande ($p \leq n$) tendremos que $W_{\lambda_i p} = W_{\lambda_i p+1}$ tomando el mínimo entre los p 's donde esto ocurre definimos $W_{\lambda_i} := W_{\lambda_i p}$. Note que si para algún λ_i , $p = 1$ entonces el subespacio W_{λ_i} está compuesto por autovectores. Estos son los máximos espacios invariantes asociados con el autovalor λ_i , en efecto tenemos:

Lema 2.4 *El único autovalor de \mathbf{A} en W_{λ_i} es λ_i .*

Prueba: Sea λ un autovalor de \mathbf{A} en W_{λ_i} . Veamos que $\lambda = \lambda_i$. Como ya hemos visto habrá entonces un autovector $\boldsymbol{\zeta} \in W_{\lambda_i}$ con λ como autovalor. Como esta en W_{λ_i} habrá algún $p \geq 1$ tal que $(\mathbf{A} - \lambda_i \mathbf{I})^p \boldsymbol{\zeta} = 0$ pero como es un autovector tenemos que $(\lambda - \lambda_i)^p \boldsymbol{\zeta} = 0$, y por lo tanto que $\lambda = \lambda_i$ ♠

Veamos ahora que estos subespacios son linealmente independientes y generan todo V^C . Probaremos nuevamente este teorema en el Capítulo 5.

Teorema 2.3 (Ver capítulo 5, Teorema 5.3) *Dado un operador $A : V \rightarrow V$, con autovectores $\{\lambda_i\}$, $i = 1..m$ el espacio V^C admite una descomposición directa en subespacios invariantes W_{λ_i} , donde en cada uno de ellos \mathbf{A} tiene solo a λ_i como autovalor.*

Prueba:

Los W_{λ_i} son independientes. Sea $\mathbf{v}_1 + \dots + \mathbf{v}_s = 0$, con $\mathbf{v}_i \in W_{\lambda_i}$, luego debemos probar que cada $\mathbf{v}_i = 0$. Aplicando $(\mathbf{A} - \lambda_2 \mathbf{I})^{p_2} \dots (\mathbf{A} - \lambda_s \mathbf{I})^{p_s}$ a la suma anterior obtenemos, $(\mathbf{A} - \lambda_2 \mathbf{I})^{p_2} \dots (\mathbf{A} - \lambda_s \mathbf{I})^{p_s} \mathbf{v}_1 = 0$, pero como λ_i , $i \neq 1$ no es un autovalor de \mathbf{A} en W_{λ_1} el operador $(\mathbf{A} - \lambda_2 \mathbf{I})^{p_2} \dots (\mathbf{A} - \lambda_s \mathbf{I})^{p_s}$ es invertible ⁹ es ese subespacio y por lo tanto $\mathbf{v}_1 = 0$. Continuando de esa forma vemos que todos los \mathbf{v}_i se deben anular.

Los W_{λ_i} generan todo V^C . Supongamos por contradicción que este no es el caso, y consideremos V^C/W , donde W es el espacio generado por todos los W_{λ_i} , es decir el espacio de todas las combinaciones lineales de elementos en los W_{λ_i} . El operador \mathbf{A} actúa en V^C/W [$\mathbf{A}\{\mathbf{u}\} = \{\mathbf{A}\mathbf{u}\}$] y por lo tanto tiene allí un par autovalor-autovector. Esto implica, que para algún elemento $\boldsymbol{\zeta}$ de V^C en alguna clase equivalente de V^C/W se cumple:

$$\mathbf{A}\boldsymbol{\zeta} = \lambda\boldsymbol{\zeta} + \mathbf{u}_1 + \dots + \mathbf{u}_s \quad (2.24)$$

donde los \mathbf{u}_i pertenecen a cada W_{λ_i} . Supongamos ahora que $\lambda \neq \lambda_i \ \forall i = 1..s$, luego $\mathbf{A} - \lambda\mathbf{I}$ es invertible en cada W_{λ_i} y por lo tanto existen vectores $\boldsymbol{\zeta}_i = (\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{u}_i \in W_{\lambda_i}$. Pero entonces $\tilde{\boldsymbol{\zeta}} := \boldsymbol{\zeta} - \boldsymbol{\zeta}_1 - \dots - \boldsymbol{\zeta}_s$ es un autovalor de \mathbf{A} ! Esto es imposible pues λ no es una raíz del polinomio característico ni $\tilde{\boldsymbol{\zeta}} = 0$, pues al pertenecer $\boldsymbol{\zeta}$ a V^C/W no es una combinación lineal de elementos en los W_{λ_i} . Tenemos así una contradicción. Supongamos ahora que $\lambda = \lambda_j$ para algún $j \in \{1..s\}$.

⁹Note que $(\mathbf{A} - \lambda_i \mathbf{I})^s|_{W_{\lambda_j}}$ es invertible si su determinante es distinto de cero. Pero $\det(\mathbf{A} - \lambda_i \mathbf{I})^s = (\det(\mathbf{A} - \lambda_i \mathbf{I}))^s = (\lambda_j - \lambda_i)^{s \dim(W_{\lambda_j})} \neq 0$

Todavía podemos definir los vectores $\zeta_i = (\mathbf{A} - \lambda_j \mathbf{I})^{-1} \mathbf{u}_i$ para todo $i \neq j$ y $\tilde{\zeta}$, donde solo sustraemos a ζ todos los ζ_i con $i \neq j$, por lo tanto tenemos que

$$(\mathbf{A} - \lambda_j \mathbf{I}) \tilde{\zeta} = \mathbf{u}_i \quad (2.25)$$

Pero aplicando $(\mathbf{A} - \lambda_j \mathbf{I})^{p_j}$ a esta ecuación, con p_j el mínimo valor para el cual $W_{\lambda_j p_j} = W_{\lambda_j p_j + 1}$ obtenemos que $\tilde{\zeta} \in W_{\lambda_j}$ y así otra contradicción, por lo tanto solo puede ser que V^C/W sea el espacio trivial y los W_{λ_i} generan todo V^C ♠

Vemos así que solo necesitamos estudiar ahora cada uno de estos subespacios W_{λ_i} para encontrar todas sus partes irreducibles (de ahora en más suprimiremos el subíndice i). Pero en estos subespacios el operador \mathbf{A} actúa en forma muy sencilla!

En efecto, sea $\Delta_\lambda : W_\lambda \rightarrow W_\lambda$ definido por $\Delta_\lambda := \mathbf{A}|_{W_\lambda} - \lambda \mathbf{I}|_{W_\lambda}$, luego Δ tiene solo al 0 como autovalor y por lo tanto es **nilpotente**, es decir, existe un entero $m \leq n$ tal que $\Delta_\lambda^m = 0$.

Lema 2.5 Sea $\Delta : W \rightarrow W$ tal que su único autovalor es 0, luego Δ es nilpotente.

Prueba: Sea $W^p := \Delta^p W$, luego tenemos que $W^p \subseteq W^q$ si $p \geq q$. Como la dimensión de W es finita deberá suceder que para algún p entero tendremos que $W^p = W^{p+1}$ vemos entonces que Δ^p actúa inyectivamente en W^p y por lo tanto no puede tener al 0 como autovalor. Pero hemos visto que todo operador tiene algún autovalor y por lo tanto tenemos una contradicción a menos que $W^p = \{0\}$. Es decir $\Delta^p = 0$ ♠

Los operadores nilpotentes tienen la importante propiedad de generar una base del espacio en que actúan a partir de su aplicación repetidas veces sobre conjunto menor de vectores linealmente independiente.

Lema 2.6 Sea $\Delta : W \rightarrow W$ nilpotente, luego existe una base de W constituida por elementos de la forma: $\{\{v_1, \Delta v_1, \dots, \Delta^{p_1} v_1\}, \dots, \{\{v_d, \Delta v_d, \dots, \Delta^{p_d} v_d\}\}$ donde p_i es tal que $\Delta^{p_i+1} v_i = 0$

Note que si $n = \dim W$ luego $n = \sum_{i=1}^d p_i$. Cada uno de estos conjuntos formados por repetidas aplicaciones de un operador se llama un **ciclo**. En este caso la base está formada por los elementos de d ciclos. Note que no necesariamente los ciclos son entes únicos, en efecto, si tenemos dos ciclos con el mismo número de elementos, entonces cualquier combinación lineal de los mismos será también un ciclo. Note que cada ciclo contiene un solo autovector.

Prueba: Lo probaremos por inducción en la dimensión de W . Si $n = 1$ tomamos cualquier vector para generar la base, ya que en este caso $\Delta = 0$. Suponemos ahora que es cierto para toda dimensión menor que n . En particular, como Δ tiene un autovalor nulo $\dim(\ker \Delta) \geq 1$ y por lo tanto tenemos que $W' (= \Delta(W))$ tiene dimensión menor que n , digamos n' y por la hipótesis inductiva una base de la forma $\{\{v'_1, \Delta v'_1, \dots, \Delta^{p'_1} v'_1\}, \dots, \{\{v'_{d'}, \Delta v'_{d'}, \dots, \Delta^{p'_{d'}} v'_{d'}\}\}$. Para formar una base de W agregaremos a estos vectores d' vectores v_i tales que $\Delta v_i = v'_i$, $i = 1, \dots, d'$. Esto siempre se puede hacer pues $v'_i \in W' = \Delta W$. Vemos así que hemos incrementado el conjunto de vectores a $\{\{v_1, \Delta v_1, \dots, \Delta^{p_1+1} v_1\}, \dots, \{\{v_{d'}, \Delta v_{d'}, \dots, \Delta^{p'_{d'}+1} v_{d'}\}\}$, es decir tenemos ahora $r = \sum_{i=1}^{d'} (p'_i + 1) = n' + d'$ vectores. Para obtener una base debemos entonces incrementar este conjunto con $n - n' - d'$ vectores. Note que este número es no-negativo, en efecto, $n - n' = \dim(\ker \Delta) \geq \dim(\ker \Delta \cap W') = d'$ y es precisamente la dimensión del subespacio de $\ker \Delta$ que no está en W' . Completamos entonces la base propuesta para W incorporando al conjunto ya obtenido $n - n' - d'$ vectores $\{z_i\}, i = 1, \dots, n - n' - d'$ del espacio nulo de Δ que sean linealmente independientes entre sí y con los otros elementos de $\ker \Delta$ en W y que además no estén en W' . Hemos obtenido así un conjunto de $d = d' + n - n' - d' = n - n'$ ciclos. Veamos que el conjunto así obtenido es una base. Como son n en número solo tenemos que ver que son linealmente independientes. Debemos entonces

probar que si tenemos constantes $\{C_{i,j}\}$, $i = 1..d$, $j = 0..p'_i + 1$ tales que

$$\mathbf{0} = \sum_{i=1}^d \sum_{j=0}^{p'_i+1} C_{ij} \Delta^{p_i} \mathbf{v}_i \quad (2.26)$$

entonces $C_{ij} = 0$. Aplicando Δ a esta relación obtenemos,

$$\begin{aligned} \mathbf{0} &= \Delta \sum_{i=1}^d \sum_{j=0}^{p'_i+1} C_{ij} \Delta^{p_i} \mathbf{v}_i \\ &= \sum_{i=1}^{d'} \sum_{j=0}^{p'_i} C_{ij} \Delta^{p'_i} \mathbf{v}'_i, \end{aligned} \quad (2.27)$$

donde hemos usado que $\Delta^{p'_i+1} \mathbf{v}'_i = 0$. Pero esta es la relación de ortogonalidad de la base de W' y por lo tanto concluimos que $C_{ij} = 0 \forall i \leq d'$, $j \leq p'_i$. La relación inicial queda entonces reducida a

$$\begin{aligned} \mathbf{0} &= \sum_{i=1}^d C_{ip'_i+1} \Delta^{p'_i+1} \mathbf{v}_i \\ &= \sum_{i=1}^{d'} C_{ip'_i} \Delta^{p'_i} \mathbf{v}'_i + \sum_{i=d'+1}^d C_{i1} \mathbf{z}_i, \end{aligned} \quad (2.28)$$

pero los miembros de la primera sumatoria son parte de la base de W' y por lo tanto linealmente independientes entre sí, mientras que los de la segunda son un conjunto de elementos fuera de W' elegidos de forma que sean linealmente independientes entre sí y con los de la primera sumatoria, y por lo tanto concluimos que todos los C_{ij} se anulan ♠

Prueba alternativa: Alternativamente el lema anterior se puede probar en forma constructiva. En efecto, si $m+1$ es la potencia para la cual Δ se anula, podemos tomar el espacio $W^m = \Delta^m(W)$ y una base $\{\mathbf{v}_i^m\}$, del mismo. Luego considerar el espacio $W^{m-1} = \Delta^{m-1}(W)$. Como $W^m = \Delta(W)^{m-1}$ por cada vector \mathbf{v}_i^m de la base $\{\mathbf{v}_i^m\}$ de W^m habrá un vector \mathbf{v}_i^{m-1} tal que $\Delta \mathbf{v}_i^{m-1} = \mathbf{v}_i^m$. Como $W^m \subset W^{m-1}$ el conjunto $\{\mathbf{v}_i^m\} \cup \{\mathbf{v}_i^{m-1}\}$ está contenido en W^{m-1} . Note que $\dim W^m = \dim \ker \Delta \cap W^m$, y $\dim W^{m-1} - \dim W^m = \dim(\ker \Delta \cap W^{m-1})$. Por lo tanto,

$$\begin{aligned} \dim W^{m-1} &= \dim W^m + \dim(\ker \Delta \cap W^{m-1}) \\ &= 2 \dim W^m + \dim(\ker \Delta \cap W^{m-1}) - \dim W^m \\ &= 2 \dim W^m + \dim(\ker \Delta \cap W^{m-1}) - \dim \ker \Delta \cap W^m \end{aligned} \quad (2.29)$$

Agregando entonces al conjunto anterior un conjunto $\{\mathbf{z}_i\}$ de $\dim(\ker \Delta \cap W^{m-1}) - \dim(\ker \Delta \cap W^m)$ vectores del espacio nulo de Δ en W^{m-1} , tales que sean linealmente independientes entre sí y con los elementos de la base de W^m , obtenemos un conjunto de $\dim W^{m-1}$ vectores. Note que la mencionada elección de los elementos $\{\mathbf{z}_i\}$ se puede hacer pues es meramente una extensión de la base $\{\mathbf{v}_i^m\}$ de W^m a una base de $\ker \Delta \cap W^{m-1}$. Probemos ahora que son linealmente independientes y por lo tanto que forman una base de W^{m-1} . Para ello habrá que probar que si

$$\mathbf{0} = \sum_i C_i^m \mathbf{v}_i^m + \sum_i C_i^{m-1} \mathbf{v}_i^{m-1} + \sum_j C_j^z \mathbf{z}_j \quad (2.30)$$

entonces cada uno de los coeficientes C_i^m , C_i^{m-1} , C_j^z debe anularse. Multiplicando la expresión anterior por Δ obtenemos,

$$\mathbf{0} = \sum_i C_i^{m-1} \Delta \mathbf{v}_i^{m-1} = \sum_i C_i^{m-1} \mathbf{v}_i^m, \quad (2.31)$$

pero entonces la independencia lineal de la base de W^m nos asegura que los $\{C_i^{m-1}\}$ son todos nulos. Tenemos entonces que,

$$\mathbf{0} = \sum_i C_i^m \mathbf{v}_i^m + \sum_j C_j^z \mathbf{z}_j. \quad (2.32)$$

Pero estos vectores fueron elegidos linealmente independientes entre sí y por lo tanto todos los coeficientes en esta suma deben anularse. Vemos así que el conjunto $\{\mathbf{v}_i^m\} \cup \{\mathbf{v}_i^{m-1}\} \cup \{\mathbf{z}_i\}$ forman una base cíclica de W^{m-1} . Continuando con W^{m-2} y así sucesivamente obtenemos una base cíclica para todo W .

Vemos así que los subespacios invariantes irreducibles de un operador están constituidos por ciclos dentro de subespacios invariantes asociados a con un dado autovector. Cada ciclo contiene un único autovalor del operador. Denotaremos a los subespacios generados por estos ciclos (y usualmente llamados también ciclos) por $C_{\lambda_i}^k$, donde el índice inferior se refiere al autovalor del ciclo y el superior indexa los distintos ciclos dentro de cada W_{λ_i} .

Ejercicio: Muestre que los ciclos obtenidos son subespacios invariantes irreducibles de \mathbf{A} .

2.3.3. Forma Canónica de Jordan

Definición: Sea $\mathbf{A} : V \rightarrow V$ un operador lineal. Diremos que \mathbf{A} es de tipo Jordan con autovalor λ si existe una base $\{\mathbf{u}_i\}$ de V tal que ¹⁰

$$\mathbf{A} = \lambda \sum_{i=1}^n \mathbf{u}_i \boldsymbol{\theta}^i + \sum_{i=2}^n \mathbf{u}_i \boldsymbol{\theta}^{i-1} \equiv \lambda I + \Delta \quad (2.33)$$

donde $\{\boldsymbol{\theta}^i\}$ es la co-base de la base $\{\mathbf{u}_i\}$.

Es decir, en esta base las componentes de \mathbf{A} forman una matriz A^j_i dada por

$$A = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ & & \lambda & 1 & 0 \\ & 0 & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \quad (2.34)$$

Note que la matriz Δ es n-nilpotente, es decir $\Delta^n = 0$.

No es cierto que todo operador sea del tipo Jordan –encuentre uno que no lo sea– pero es claro que la restricción de cualquier operador a uno de sus subespacios invariantes irreducibles (ciclos) lo es. Esto se puede ver numerando los elementos de la base generada por el ciclo convenientemente.

Ejercicio: Encuentre dicho ordenamiento de los elementos de la base.

Por lo tanto podemos resumir los resultados hallados anteriormente en el siguiente teorema sobre las representaciones matriciales de un operador cualquiera actuando en un espacio de dimensión finita.

Teorema 2.4 de Jordan Sea $\mathbf{A} : V \rightarrow V$ un operador lineal actuando en un espacio vectorial complejo V . Luego, existe una única descomposición en suma directa ¹¹ de V en subespacios $C_{\lambda_i}^k$, $V = C_{\lambda_1}^1 \oplus \cdots \oplus C_{\lambda_1}^{k_1} \oplus \cdots \oplus C_{\lambda_d}^1 \oplus \cdots \oplus C_{\lambda_d}^{k_d}$, $d \leq n$ tal que

i) Los $C_{\lambda_i}^k$ son invariantes bajo la acción de \mathbf{A} , es decir, $\mathbf{A} C_{\lambda_i}^k \subseteq C_{\lambda_i}^k$

¹⁰En el sentido que $\mathbf{A}(\mathbf{v}) = \lambda \sum_{i=1}^n \mathbf{u}_i \boldsymbol{\theta}^i(\mathbf{v}) + \sum_{i=2}^n \mathbf{u}_i \boldsymbol{\theta}^{i-1}(\mathbf{v}) \quad \forall \mathbf{v} \in V$

¹¹Recordemos que un espacio vectorial V se dice que es suma directa de dos espacios vectoriales W y Z , y lo denotamos como $V = W \oplus Z$ si cada vector en V puede ser obtenido de una única manera como la suma de un elemento en W y otro de Z .

ii) Los $C_{\lambda_i}^k$ son irreducibles, es decir no existen subespacios invariantes de $C_{\lambda_i}^k$ tales que sus sumas sea todo $C_{\lambda_i}^k$.

iii) Debido a la propiedad i) el operador \mathbf{A} induce en cada $C_{\lambda_i}^k$ un operador $\mathbf{A}_i : C_{\lambda_i}^k \rightarrow C_{\lambda_i}^k$, este es del tipo de Jordan con λ_i una de las raíces del polinomio de grado n_i ,

$$\det(\mathbf{A}_i - \lambda_i \mathbf{I}) = 0. \quad (2.35)$$

Este teorema nos dice que dado \mathbf{A} existe una base, en general compleja, tal que la matriz de sus componentes tiene forma de bloques diagonales cuadrados de $n_i \times n_i$, donde n_i es la dimensión del subespacio $C_{\lambda_i}^k$, cada uno con la forma dada en 2.34. Esta forma de la matriz se llama **forma canónica de Jordan**.

Ejercicio: Muestre que las raíces, λ_i , que aparecen en los operadores \mathbf{A}_i son invariantes ante transformaciones de semejanza, es decir $\lambda_i(\mathbf{A}) = \lambda_i(\mathbf{PAP}^{-1})$.

Ejemplo: Sea $\mathbf{A} : \mathbf{C}^3 \rightarrow \mathbf{C}^3$, luego $\det(\mathbf{A} - \lambda \mathbf{I})$ es un polinomio de 3^{er} grado, y por lo tanto tiene tres raíces. Si éstas son distintas habrá al menos tres subespacios invariantes e irreducibles de \mathbf{C}^3 , pero $\dim \mathbf{C}^3 = 3$ y por lo tanto cada uno de ellos tiene $n_i = 1$. La forma canónica de Jordan es entonces,

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad (2.36)$$

Si dos de ellas coinciden tenemos dos posibilidades, o tenemos tres subespacios, en cuyo caso la forma C. de J. será

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad (2.37)$$

o, tenemos dos subespacios, uno necesariamente de dimensión 2, la forma C. de J. será

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad (2.38)$$

Si las tres raíces coinciden entonces habrá tres posibilidades,

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \text{ y } \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad (2.39)$$

Ejemplo: Ilustraremos ahora el caso de autovalores coincidentes en dos dimensiones. Este caso no es genérico, en el sentido que cualquier perturbación en el sistema –es decir cualquier cambio mínimo en las ecuaciones– separa las raíces haciéndolas distintas.

Sea $\mathbf{A} : \mathbf{C}^2 \rightarrow \mathbf{C}^2$. En este caso el polinomio característico $\det(\mathbf{A} - \lambda \mathbf{I})$ tiene solo dos raíces las que supondremos coincidentes, $\lambda_1 = \lambda_2 = \lambda$. Nos encontramos aquí ante dos posibilidades, o existen dos autovectores linealmente independientes \mathbf{u}_1 y \mathbf{u}_2 , en cuyo caso $V = B_1 \oplus B_2$ y \mathbf{A} es diagonalizable ($\mathbf{A} = \text{diag}(\lambda, \lambda) = \lambda \mathbf{I}$), o existe solo un autovector $\tilde{\mathbf{u}}_1$. En tal caso sea $\tilde{\mathbf{u}}_2$ cualquier vector linealmente independiente de $\tilde{\mathbf{u}}_1$, luego $\mathbf{A}\tilde{\mathbf{u}}_2 = c^1 \tilde{\mathbf{u}}_1 + c^2 \tilde{\mathbf{u}}_2$ para algunos escalares c^1 y c^2 en \mathbf{C} . Calculando el determinante de $\mathbf{A} - \tilde{\lambda} \mathbf{I}$ en esta base obtenemos $(\lambda - \tilde{\lambda})(c^2 - \tilde{\lambda})$, pero λ es una raíz doble y por lo tanto $c^2 = \lambda$.

Reordenando y re-escaleando las bases $\mathbf{u}_1 = \tilde{\mathbf{u}}_1$, $\mathbf{u}_2 = c^1 \tilde{\mathbf{u}}_2$ obtenemos

$$\begin{aligned} \mathbf{A} \mathbf{u}_1 &= \lambda \mathbf{u}_1 + \mathbf{u}_2 \\ \mathbf{A} \mathbf{u}_2 &= \lambda \mathbf{u}_2 + \mathbf{u}_1, \end{aligned} \quad (2.40)$$

y por lo tanto

$$\mathbf{A} = \lambda(\mathbf{u}_1 \oplus \boldsymbol{\theta}^1 + \mathbf{u}_2 \oplus \boldsymbol{\theta}^2) + \mathbf{u}_1 \oplus \boldsymbol{\theta}^2, \quad (2.41)$$

donde $\{\boldsymbol{\theta}^i\}$ es la co-base de la base $\{\mathbf{u}_i\}$.

Note que $(\mathbf{A} - \lambda \mathbf{I}) \mathbf{u}_2 = \mathbf{u}_1$ y $(\mathbf{A} - \lambda \mathbf{I}) \mathbf{u}_1 = 0$ o sea $\Delta^2 = (\mathbf{A} - \lambda \mathbf{I})^2 = 0$.

Como veremos más adelante en las aplicaciones físicas los subespacios invariantes tienen un significado físico claro, son los llamados modos normales –caso unidimensional– y ciclos –en los otros casos.

2.3.4. Relación de Semejanza

En las aplicaciones en física es frecuente la siguiente relación de equivalencia: [Ver recuadro al final del capítulo.] Diremos que el operador \mathbf{A} es semejante al operador \mathbf{B} si existe otro operador \mathbf{P} , invertible, tal que

$$\mathbf{A} = \mathbf{P} \mathbf{B} \mathbf{P}^{-1}. \quad (2.42)$$

Es decir, si a V lo *rotamos* con un operador invertible \mathbf{P} y luego le aplicamos \mathbf{A} , obtenemos la misma acción a que si primero aplicamos \mathbf{B} y luego *rotamos* con \mathbf{P} .

Ejercicio:

- Probar que semejanza es una relación de equivalencia.
- Probar que las funciones y los mapas definidos anteriormente lo son en las distintas clases equivalentes, es decir,

$$\begin{aligned} \det(\mathbf{P} \mathbf{A} \mathbf{P}^{-1}) &= \det(\mathbf{A}) \\ \text{tr}(\mathbf{P} \mathbf{A} \mathbf{P}^{-1}) &= \text{tr}(\mathbf{A}) \\ e^{\mathbf{P} \mathbf{A} \mathbf{P}^{-1}} &= \mathbf{P} e^{\mathbf{A}} \mathbf{P}^{-1}. \end{aligned} \quad (2.43)$$

Relaciones de Equivalencia.

Definición: Una **relación de equivalencia**, \approx , entre elementos de un conjunto X es una relación que satisface las siguientes condiciones:

1. Reflexiva: Si $x \in X$, luego $x \approx x$.
2. Simétrica: Si $x, x' \in X$ y $x \approx x'$, luego $x' \approx x$.
3. Transitiva: Si $x, x', x'' \in X$, $x \approx x'$ y $x' \approx x''$, luego $x \approx x''$.

Note que la primer propiedad nos asegura que cada elemento de X cumple una relación de equivalencia con algún elemento de X , en este caso consigo mismo. Relaciones de equivalencia aparecen muy a menudo en física, esencialmente cuando usamos para describir algún proceso físico un ente matemático que tiene partes superfluas en lo que respecta a este proceso y que por lo tanto querríamos ignorarlas. Esto se logra declarando dos entes que describen el mismo fenómeno entes equivalentes.

Ejemplo: Sea X el conjunto de los números reales y sea $x \approx y$ si y solo si existe un entero n tal que $x = n + y$, esta claramente es una relación de equivalencia. Esta se usa cuando interesa describir algo usando la línea recta pero que en realidad se debería describir usando un círculo de circunferencia unitaria. Dada una relación de equivalencia en un conjunto podemos agrupar los elementos de éste en **clases equivalentes** de elementos, es decir en subconjuntos donde todos sus elementos son equivalentes entre sí y no hay ningún elemento que no esté en este subconjunto que sea equivalente a alguno de los elementos del subconjunto. (Si X es el conjunto, $Y \subset X$ es una de sus clases equivalentes, y si $y \in Y$, luego $y \approx y' \leftrightarrow y' \in Y$.)

La propiedad fundamental de las relaciones de equivalencia es la siguiente.

Teorema 2.5 *Una relación de equivalencia en un conjunto X permite reagrupar sus elementos en clases equivalentes de modo que cada elemento de X está en una y solo en una de éstas.*

Prueba: Sea $x \in X$ e Y el subconjunto de todos los elementos de X equivalentes a x . Veamos que este subconjunto es una clase equivalente. Sean y e y' dos elementos de Y , es decir $y \approx x$ e $y' \approx x$, pero por la propiedad de transitividad entonces $y \approx y'$. Si $y \in Y$ y $z \notin Y$ entonces $y \not\approx z$, ya que de otro modo z sería equivalente a x y por lo tanto estaría en Y . Por último note que por reflexividad x también está en Y . Solo resta ver que si y está en Y y también en otra clase equivalente, digamos Z , luego $Y = Z$. Como $y \in Y$ luego y es equivalente a todo elemento de Y , como $y \in Z$ luego y es equivalente a todo elemento de Z , pero por transitividad entonces todo elemento de Y es equivalente a todo elemento de Z , pero al ser estas clases equivalentes y por lo tanto cada una contener todos sus elementos equivalentes ambas deben coincidir.

Ejercicio: ¿Cuales son las clases equivalentes del ejemplo anterior?

2.4. Operadores Adjuntos

Sea A un operador lineal entre dos espacios vectoriales, $A : V \rightarrow W$, es decir $A \in \mathcal{L}(V, W)$. Debido a que V y W tienen espacios duales este operador induce naturalmente un operador lineal de W' a V' llamado su **dual**,

$$A'(\omega)(v) := \omega(A(v)) \quad \forall \omega \in W', \quad v \in V. \quad (2.44)$$

Es decir el operador que al aplicarlo a un elemento $\omega \in W'$ nos da el elemento $\mathbf{A}'(\omega)$ de V' que cuando actúa en $\mathbf{v} \in V$ da el número $\omega(\mathbf{A}(\mathbf{v}))$. Ver figura.

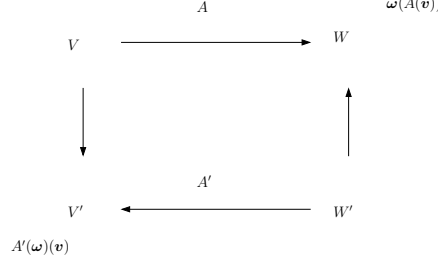


Figura 2.2: Diagrama del operador dual.

Notemos que éste es un operador lineal ya que,

$$\begin{aligned} \mathbf{A}'(\alpha\omega + \sigma)(\mathbf{v}) &= (\alpha\omega + \sigma)(\mathbf{A}(\mathbf{v})) \\ &= \alpha\omega(\mathbf{A}(\mathbf{v})) + \sigma(\mathbf{A}(\mathbf{v})) \\ &= \alpha\mathbf{A}'(\omega)(\mathbf{v}) + \mathbf{A}'(\sigma)(\mathbf{v}). \end{aligned} \quad (2.45)$$

En la representación matricial este operador está representado meramente por la misma matriz que el original, pero ahora actuando por izquierda, $\mathbf{A}'(w)_i = w_j A^j_i$, es decir, $A'^j_i = A^j_i$.

Si hay normas definidas en V y W y definimos la norma de $\mathbf{A} : V \rightarrow W$ de la manera usual,

$$\|\mathbf{A}\| := \sup_{\|\mathbf{v}\|_V=1} \{\|\mathbf{A}(\mathbf{v})\|_W\} \quad (2.46)$$

Entonces vemos que

$$\begin{aligned} \|\mathbf{A}'\| &:= \sup_{\|\omega\|_{W'}=1} \{\|\mathbf{A}'(\omega)\|_{V'}\} \\ &= \sup_{\|\omega\|_{W'}=1} \left\{ \sup_{\|\mathbf{v}\|_V=1} \{|\mathbf{A}'(\omega)(\mathbf{v})|\} \right\} \\ &= \sup_{\|\omega\|_{W'}=1} \left\{ \sup_{\|\mathbf{v}\|_V=1} \{|\omega(\mathbf{A}(\mathbf{v}))|\} \right\} \\ &\leq \sup_{\|\omega\|_{W'}=1} \left\{ \sup_{\|\mathbf{v}\|_V=1} \{\|\omega\|_{W'} \|\mathbf{A}(\mathbf{v})\|_W\} \right\} \\ &= \sup_{\|\mathbf{v}\|_V=1} \{\|\mathbf{A}(\mathbf{v})\|_W\} \\ &= \|\mathbf{A}\|. \end{aligned} \quad (2.47)$$

Vemos así que si un operador es acotado, luego su dual también lo es.

Veamos cuales son las componentes del dual de un operador en término de las componentes del operador original. Sea entonces $\{\mathbf{u}_i\}$, $\{\boldsymbol{\theta}^i\}$, $i = 1, \dots, n$ una base, respectivamente una co-base de V y sea $\{\hat{\mathbf{u}}_i\}$, $\{\hat{\boldsymbol{\theta}}^i\}$, $i = 1, \dots, m$ un par base-co-base de W . Tenemos entonces que las componentes de \mathbf{A} con respecto a estas bases son: $A^i_j := \hat{\boldsymbol{\theta}}^i(\mathbf{A}(\mathbf{u}_j))$, es decir, $\mathbf{A}(\mathbf{v}) = \sum_{i=1}^m \sum_{j=1}^n A^i_j \hat{\mathbf{u}}_i \boldsymbol{\theta}^j(\mathbf{v})$.

Por lo tanto, si \mathbf{v} tiene componentes (v^1, v^2, \dots, v^n) en la base $\{\mathbf{u}_i\}$, $\mathbf{A}(\mathbf{v})$ tiene componentes $(\sum_{i=1}^n A^1_i v^i, \sum_{i=1}^n A^2_i v^i, \dots, \sum_{i=1}^n A^m_i v^i)$ en la base $\{\hat{\mathbf{u}}_i\}$

Veamos ahora las componentes de \mathbf{A}' . Por definición tenemos,

$$(\mathbf{A}')^i_j := \mathbf{A}'(\hat{\boldsymbol{\theta}}^i)(\mathbf{u}_j) = \hat{\boldsymbol{\theta}}^i(\mathbf{A}(\mathbf{u}_j)) = A^i_j.$$

O sea las mismas componentes, pero ahora la matriz actúa por izquierda en las componentes $(\omega_1, \omega_2, \dots, \omega_m)$ de un elemento ω de W' en la co-base $\{\hat{\theta}^i\}$. Las componentes de $A'(\omega)$ en la co-base $\{\hat{\theta}^i\}$ son,

$$\left(\sum_{i=1}^m A^i_1 \omega_i, \sum_{i=1}^m A^i_2 \omega_i, \dots, \sum_{i=1}^m A^i_n \omega_i\right).$$

Un caso particularmente interesante de esta construcción es cuando $W = V$ y este es un espacio con producto interno, es decir un espacio de Hilbert. En tal caso el producto interno nos da un mapa canónico entre V y su dual V' :

$$\phi : V \rightarrow V', \quad \phi(v) := (v, \cdot), \quad (2.48)$$

y por lo tanto si $A : V \rightarrow V$ luego $A' : V' \rightarrow V'$ puede también ser considerado como un operador entre V y V que llamaremos A^* .

Note que $\phi^{-1} : V' \rightarrow V$ está definido por $(\phi^{-1}(\omega), u) = \omega(u)$, ya que

$$((\phi^{-1}(\phi(v)), u) = ((\phi^{-1}((v, \cdot)), u) = (v, u) \quad \forall u \in V.$$

Con la ayuda de este mapa definimos $A^* : V \rightarrow V$ dado por: (Ver figura)

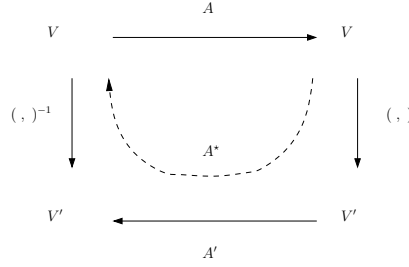


Figura 2.3: Diagrama del operador estrella.

$$A^*(v) := \phi^{-1}(A'(\phi(v))). \quad (2.49)$$

En término del producto escalar esto es:

$$(A^*(v), u) = (\phi^{-1}(A'((v, \cdot))), u) = A'((v, \cdot), (u) = (v, A(u)). \quad (2.50)$$

En su representación matricial este operador es, $A^{*j}_i = t_{il} A^l_k (t^{-1})^{kj}$, donde t_{li} es la representación del producto escalar y $(t^{-1})^{jk}$ el de su inversa ($t_{il}(t^{-1})^{lk} = \delta^k_i$). En el caso de que $t_{ik} = \delta_{ik}$, la delta de Kronecker, la representación matricial del adjunto es simplemente la matriz transpuesta, $A^{*j}_i = A^{\dagger j}_i = A^j_i$,

Un subconjunto de operadores particularmente interesante es aquel para el cual se cumple $A = A^*$. Estos operadores se llaman **Hermitianos** o **Autoadjuntos**.¹²

Los operadores autoadjuntos tienen importantes propiedades:

Lema 2.7 Sea $M = \text{Span}\{u, u_2, \dots, u_m\}$, donde $\{u_i\}$ son un conjunto de autovalores de A , un operador autoadjunto. Luego M y M^\perp son espacios invariantes de A .

Prueba: La primera afirmación es clara y general, la segunda depende de la Hermiticidad de A . Sea $v \in M^\perp$ cualquiera, veamos que $A(v) \in M^\perp$. Sea $u \in M$ arbitrario, luego

$$(u, A(v)) = (A(u), v) = 0, \quad (2.51)$$

ya que $A(u) \in M$ si $u \in M$ ♠Esta propiedad tiene el siguiente corolario:

¹²En el caso de dimensión infinita estos nombres dejan de coincidir para algunos autores.

Corolario 2.1 Sea $\mathbf{A} : H \rightarrow H$ autoadjunto. Luego los autovalores de \mathbf{A} forman una base ortonormal de H .

Prueba: $\mathbf{A} : H \rightarrow H$ tiene al menos un autovector, digamos \mathbf{u}_1 . Consideremos ahora su restricción al espacio perpendicular a \mathbf{u}_1 que denotamos también con \mathbf{A} pues por el lema anterior este es un espacio invariante, $\mathbf{A} : \{\mathbf{u}_1\}^\perp \rightarrow \{\mathbf{u}_1\}^\perp$. Este operador también tiene un autovector, digamos \mathbf{u}_2 y $\mathbf{u}_1 \perp \mathbf{u}_2$. Consideremos ahora la restricción de \mathbf{A} a $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}^\perp$ allí también tenemos $\mathbf{A} : \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}^\perp \rightarrow \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}^\perp$ y por lo tanto un autovector de \mathbf{A} , \mathbf{u}_3 con $\mathbf{u}_3 \perp \mathbf{u}_1$, $\mathbf{u}_3 \perp \mathbf{u}_2$. Continuando de esta manera llegamos a tener $n = \dim H$ autovectores todos ortogonales entre sí ♠

Este teorema tiene varias extensiones al caso donde el espacio vectorial es de dimensión infinita. Más adelante en el capítulo 12 veremos una de ellas.

Notemos que en esta base \mathbf{A} es diagonal y por lo tanto tenemos

Corolario 2.2 Todo operador autoadjunto es diagonalizable

Notemos también que si \mathbf{u} es un autovector de \mathbf{A} autoadjunto, con autovalor λ luego,

$$\bar{\lambda}(\mathbf{u}, \mathbf{u}) = (\mathbf{A}(\mathbf{u}), \mathbf{u}) = (\mathbf{u}, \mathbf{A}(\mathbf{u})) = \lambda(\mathbf{u}, \mathbf{u}) \quad (2.52)$$

y por lo tanto $\bar{\lambda} = \lambda$, es decir,

Lema 2.8 Los autovalores de un operador autoadjunto son reales

Veamos qué quiere decir la condición de hermiticidad en término de las componentes del operador en una base ortonormal. Sea entonces \mathbf{A} un operador autoadjunto y sea $\{\mathbf{e}_i\}$, $i = 1, \dots, n$ una base ortonormal del espacio donde éste actúa. Tenemos que $(\mathbf{A}(\mathbf{e}_i), \mathbf{e}_j) = (\mathbf{e}_i, \mathbf{A}(\mathbf{e}_j))$ y por lo tanto, notando que $\mathbf{I} = \sum_{i=1}^n \mathbf{e}_i \theta^i$, obtenemos,

$$\begin{aligned} 0 &= \left(\sum_{k=1}^n \mathbf{e}_k \theta^k (\mathbf{A}(\mathbf{e}_i), \mathbf{e}_j) - (\mathbf{e}_i, \sum_{l=1}^n \mathbf{e}_l \theta^l (\mathbf{A}(\mathbf{e}_j))) \right) \\ &= \sum_{k=1}^n \bar{A}^k_i (\mathbf{e}_k, \mathbf{e}_j) - \sum_{l=1}^n A^l_j (\mathbf{e}_i, \mathbf{e}_l) \\ &= \sum_{k=1}^n \bar{A}^k_i \delta_{kj} - \sum_{l=1}^n A^l_j \delta_{li}. \end{aligned} \quad (2.53)$$

de lo que concluimos que

$$\bar{A}^j_i = A^i_j \quad (2.54)$$

o sea que la matriz transpuesta es la complejo conjugada de la original. En el caso de matrices reales vemos que la condición es que en esa base la matriz sea igual a su transpuesta, lo que usualmente se denota diciendo que la matriz es simétrica.

Una propiedad interesante de los operadores autoadjuntos es que su norma es igual al supremo de los módulos de sus autovalores. Como la cuenta demostrando esto será usada más adelante en el curso damos una demostración de este hecho a continuación.

Lema 2.9 Si \mathbf{A} es autoadjunto luego $\|\mathbf{A}\| = \sup\{|\lambda_i|\}$.

Prueba: Sea $F(\mathbf{u}) := (\mathbf{A}(\mathbf{u}), \mathbf{A}(\mathbf{u}))$ definida en la esfera $\|\mathbf{u}\| = 1$. Como dicho conjunto es compacto (aquí estamos usando el hecho que el espacio es de dimensión finita) tiene un máximo que denotaremos \mathbf{u}_0 . Notemos entonces que $F(\mathbf{u}_0) := \|\mathbf{A}\|^2$. Como $F(\mathbf{u})$ es diferenciable en la esfera se debe cumplir que

$$\frac{d}{d\lambda} F(\mathbf{u}_0 + \lambda \delta \mathbf{u})|_{\lambda=0} = 0, \quad (2.55)$$

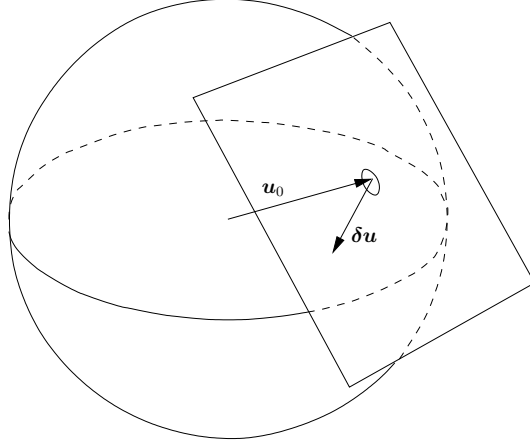


Figura 2.4: Vectores normales y tangentes a la esfera.

a lo largo de toda curva tangente a la esfera en el punto \mathbf{u}_0 , es decir, para todo $\delta \mathbf{u}$ tal que $(\mathbf{u}_0, \delta \mathbf{u}) = 0$. Ver figura.

Pero

$$\begin{aligned}
 \frac{d}{d\lambda} F(\mathbf{u}_0 + \lambda \delta \mathbf{u})|_{\lambda=0} &= \frac{d}{d\lambda} (\mathbf{A}(\mathbf{u}_0 + \lambda \delta \mathbf{u}), \mathbf{A}(\mathbf{u}_0 + \lambda \delta \mathbf{u}))|_{\lambda=0} \\
 &= (\mathbf{A}(\delta \mathbf{u}), \mathbf{A}(\mathbf{u}_0)) + (\mathbf{A}(\mathbf{u}_0), \mathbf{A}(\delta \mathbf{u})) \\
 &= 2\Re(\mathbf{A}^* \mathbf{A} \mathbf{u}_0, \delta \mathbf{u}) \\
 &= 0 \quad \forall \delta \mathbf{u}, \quad (\mathbf{u}_0, \delta \mathbf{u}) = 0
 \end{aligned} \tag{2.56}$$

Como $\delta \mathbf{u}$ es arbitrario en $\{\mathbf{u}_0\}^\perp$ esto simplemente implica

$$\mathbf{A}^* \mathbf{A}(\mathbf{u}_0) \in \{\delta \mathbf{u}\}^\perp = \{\mathbf{u}_0\}^{\perp\perp} = \{\mathbf{u}_0\}. \tag{2.57}$$

y por lo tanto que existirá $\alpha \in \mathbb{C}$ tal que

$$\mathbf{A}^* \mathbf{A}(\mathbf{u}_0) = \alpha \mathbf{u}_0. \tag{2.58}$$

Tomando producto escalar con \mathbf{u}_0 obtenemos,

$$\begin{aligned}
 (\mathbf{A}^* \mathbf{A}(\mathbf{u}_0), \mathbf{u}_0) &= \bar{\alpha}(\mathbf{u}_0, \mathbf{u}_0) \\
 &= (\mathbf{A}(\mathbf{u}_0), \mathbf{A}(\mathbf{u}_0)) \\
 &= \|\mathbf{A}\|^2
 \end{aligned} \tag{2.59}$$

y por lo tanto tenemos que $\alpha = \bar{\alpha} = \|\mathbf{A}\|^2$.

Sea ahora $\mathbf{v} := \mathbf{A} \mathbf{u}_0 - \|\mathbf{A}\| \mathbf{u}_0$, luego, usando ahora que \mathbf{A} es autoadjunto tenemos,

$$\begin{aligned}
 \mathbf{A}(\mathbf{v}) &= \mathbf{A} \mathbf{A}(\mathbf{u}_0) - \|\mathbf{A}\| \mathbf{A}(\mathbf{u}_0) \\
 &= \mathbf{A}^* \mathbf{A}(\mathbf{u}_0) - \|\mathbf{A}\| \mathbf{A}(\mathbf{u}_0) \\
 &= \|\mathbf{A}\|^2 \mathbf{u}_0 - \|\mathbf{A}\| \mathbf{A}(\mathbf{u}_0) \\
 &= \|\mathbf{A}\|(\|\mathbf{A}\| \mathbf{u}_0 - \mathbf{A}(\mathbf{u}_0)) \\
 &= -\|\mathbf{A}\| \mathbf{v},
 \end{aligned} \tag{2.60}$$

por lo tanto o \mathbf{v} es un autovector de \mathbf{A} , con autovalor $\lambda = -\|\mathbf{A}\|$, o $\mathbf{v} = 0$, en cuyo caso \mathbf{u}_0 es un autovector de \mathbf{A} con autovector $\lambda = \|\mathbf{A}\|$ ♠

2.5. Operadores Unitarios

Otra subclase de operadores lineales que aparece muy a menudo en física cuando hay un producto interno privilegiado es la de los **operadores unitarios**, es decir aquellos tales que su acción preserva el producto escalar,

$$(U(u), U(v)) = (u, v), \quad \forall u, v \in H. \quad (2.61)$$

El caso más típico de operador unitario es una transformación que envía una base ortonormal en otra. Ejemplos usuales son las rotaciones en \mathbb{R}^n .

Observemos también que $\|U\| = \sup_{\|v\|=1} \{\|U(v)\|\} = 1$.

Notemos que

$$(U(u), U(v)) = (U^*U(u), v) = (u, v), \quad \forall u, v \in H, \quad (2.62)$$

o sea,

$$U^*U = I \quad (2.63)$$

y por lo tanto,

$$U^{-1} = U^*. \quad (2.64)$$

Veamos cómo son los autovalores de un operador unitario U . Sea v_1 un autovector de U (sabemos que al menos tiene uno), luego,

$$\begin{aligned} (U(v_1), U(v_1)) &= \lambda_1 \bar{\lambda}_1 (v_1, v_1) \\ &= (v_1, v_1) \end{aligned} \quad (2.65)$$

y por lo tanto $\lambda_1 = e^{i\theta_1}$ para algún ángulo θ_1 .

Si el operador U representa una rotación entonces habrá solo un autovector real, correspondiente al eje que queda fijo en una rotación. Si tenemos más de un autovalor, luego sus correspondientes autovectores son ortogonales, en efecto, sean v_1 y v_2 dos autovectores luego

$$\begin{aligned} (U(v_1), U(v_2)) &= \bar{\lambda}_1 \lambda_2 (v_1, v_2) \\ &= (v_1, v_2) \end{aligned} \quad (2.66)$$

y por lo tanto si $\lambda_1 \neq \lambda_2$ debemos tener $(v_1, v_2) = 0$.

Ejercicio: Muestre que si A es un operador autoadjunto, luego $U := e^{iA}$ es un operador unitario.

2.6. Problemas

Problema 2.4 Sea el operador $A : V \rightarrow V$ donde $\dim V = n$, tal que $Ax = \lambda x$. Calcule $\det A$ y $\text{tr} A$.

Problema 2.5 Sea $V = \mathbb{R}^3$ y x un vector no nulo cualquiera. Encuentre geométricamente y analíticamente el espacio cociente V/W_x , donde W_x es el espacio generado por x . Tome otro vector, x' , linealmente independiente con el primero y calcule ahora $V/W_{(x, x')}$.

Problema 2.6 La norma sobre operadores se define como:

$$\|A\|_{\mathcal{L}} = \max_{\|x\|_V=1} \|A(x)\|_V. \quad (2.67)$$

Encuentre las normas de los siguientes operadores, dados por su representación matricial con respecto a una base y donde la norma en el espacio vectorial es la euclídea con respecto a dicha base.

a)

$$\begin{pmatrix} 3 & 5 \\ 4 & 1 \end{pmatrix} \quad (2.68)$$

b)

$$\begin{pmatrix} 3 & 5 & 2 \\ 4 & 1 & 7 \\ 8 & 3 & 2 \end{pmatrix} \quad (2.69)$$

Problema 2.7 Sea V un espacio vectorial cualquiera y sea $\|\cdot\|$ una norma Euclídea en dicho espacio. La norma Hilbert – Schmidt de un operador se define como:

$$\|\mathbf{A}\|_{HS}^2 = \sum_{i,j=1}^n |A^j_i|^2. \quad (2.70)$$

donde la base empleada ha sido la ortonormal con respecto a la norma Euclídea.

a) Muestre que esta es una norma.

b) Muestre que $\|\mathbf{A}\|_{\mathcal{L}} \leq \|\mathbf{A}\|_{HS}$.

c) Muestre que $\sum_{j=1}^n |A^j_k|^2 \leq \|\mathbf{A}\|_{\mathcal{L}}^2$ para cada k . Por lo tanto $\|\mathbf{A}\|_{HS}^2 \leq n\|\mathbf{A}\|_{\mathcal{L}}^2$, y las dos normas son equivalentes.

Ayuda: use que $\theta^j(\mathbf{A}(\mathbf{u})) = \theta^j(\mathbf{A})(\mathbf{u})$ y luego que $|\theta(\mathbf{u})| \leq \|\theta\|\|\mathbf{u}\|$.

Problema 2.8 Calcule los autovalores-vectores de las siguientes matrices:

a)

$$\begin{pmatrix} 3 & 6 \\ 4 & 1 \end{pmatrix} \quad (2.71)$$

b)

$$\begin{pmatrix} 3 & 6 \\ 0 & 1 \end{pmatrix} \quad (2.72)$$

c)

$$\begin{pmatrix} 2 & 4 & 2 \\ 4 & 1 & 0 \\ 3 & 3 & 1 \end{pmatrix} \quad (2.73)$$

Problema 2.9 Lleve a forma triangular superior las siguientes matrices: Nota: De la transformación de las bases.

a)

$$\begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix} \quad (2.74)$$

b)

$$\begin{pmatrix} 2 & 4 & 2 \\ 4 & 1 & 0 \\ 3 & 3 & 1 \end{pmatrix} \quad (2.75)$$

c)

$$\begin{pmatrix} 1 & 4 & 3 \\ 4 & 4 & 0 \\ 3 & 3 & 1 \end{pmatrix} \quad (2.76)$$

Problema 2.10 Muestre nuevamente que $\det e^{\mathbf{A}} = e^{\text{tr} \mathbf{A}}$. Ayuda: exprese la representación matricial de \mathbf{A} en una base donde tenga la forma canónica de Jordan. Alternativamente use una base donde \mathbf{A} sea diagonal superior y vea que el producto de matrices diagonal superior da una matriz diagonal superior y por lo tanto que la exponencial de una matriz diagonal superior es también diagonal superior.

Notas bibliográficas: Este capítulo está basado en los siguientes libros: [1], [3], [5] y [18]. También son de interés los siguientes, [11] y [12]. El álgebra lineal es una de las áreas más grandes y más prolífica de las matemáticas, sobre todo cuando trata de espacios de dimensión infinita, lo que usualmente se llama análisis real y teoría de operadores. En mi experiencia personal la mayoría de los problemas terminan reduciéndose a un problema algebraico y uno siente que ha hecho un avance cuando puede resolver dicho problema.

3.1. Variedades

Hay dos motivos principales que justifican el estudio del concepto de variedad, o más generalmente de la geometría diferencial por parte de los físicos. Uno es que en física aparecen naturalmente las variedades y por lo tanto no las podemos eludir. Solo en los cursos elementales se logra esquivar a éstas por medio del cálculo vectorial en \mathbb{R}^n . Así es que, por ejemplo, para estudiar el movimiento de una partícula restringida a moverse en una esfera la imaginamos a esta última embebida en \mathbb{R}^3 y usamos las coordenadas naturales de \mathbb{R}^3 para describir sus movimientos.

El segundo motivo es que el concepto de variedad es de gran utilidad conceptual, ya que por ejemplo en el caso de una partícula moviéndose en \mathbb{R}^3 este nos permite discernir claramente entre la posición de una partícula y su vector velocidad como entes matemáticos de naturaleza diferente. Este hecho se enmascara en \mathbb{R}^3 ya que este tipo especial de variedad tiene la estructura de un espacio vectorial.

Una variedad es una generalización de los espacios Euclídeos \mathbb{R}^n en la cual uno preserva el concepto de continuo, es decir su topología en el sentido local pero descarta su carácter de espacio vectorial. Una variedad de dimensión n es en términos imprecisos un conjunto de puntos que localmente es como \mathbb{R}^n , pero que no lo es necesariamente en su forma global.

Un ejemplo de una variedad en dos dimensiones es la esfera, S^2 . Si miramos un entorno suficientemente pequeño, U_p , de un punto cualquiera de S^2 vemos que es similar a un entorno del plano, \mathbb{R}^2 , en el sentido que podemos definir un mapa continuo e invertible entre ambos entornos. Globalmente el plano y la esfera son topológicamente distintos, ya que no existe ningún mapa continuo e invertible entre ellos. [Ver figura 3.1.]

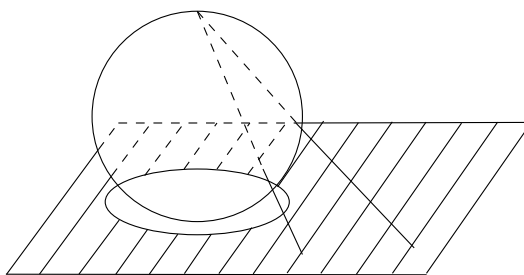


Figura 3.1: Un atlas de la esfera.

Como dijimos antes, se podría objetar la necesidad, en el ejemplo anterior, de introducir el concepto de variedad, ya que uno podría considerar S^2 como el subconjunto de \mathbb{R}^3 tal que $x_1^2 + x_2^2 + x_3^2 = 1$. La respuesta a esta objeción es que en física uno debe seguir la regla de la economía de conceptos y objetos y descartar todo lo que no sea fundamental a la descripción de un fenómeno: Si queremos describir cosas sucediendo en la esfera, ¿para qué necesitamos un espacio de más dimensiones?

Esta regla de economía nos fuerza a precisar los conceptos y descartar todo lo superfluo. Damos a continuación una serie de definiciones para desembocar finalmente en la definición de variedad de dimensión n .

Definición: Sea M un conjunto. Una **carta de M** es un par (U, φ) donde U es un subconjunto de M y φ un mapa inyectivo entre U y \mathbb{R}^n , tal que su imagen, $\varphi(U)$ es abierta en \mathbb{R}^n .

Definición: Un **atlas de M** es una colección de cartas $\{(U_i, \varphi_i)\}$ satisfaciendo las siguientes condiciones: [Ver figura 3.2.]

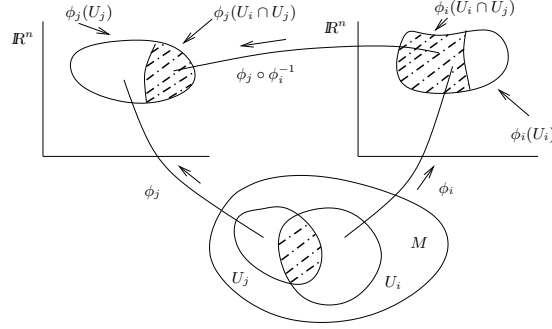


Figura 3.2: Relación entre cartas.

1. Los U_i cubren M , ($M = \bigcup_i U_i$).
2. Si dos cartas se superponen entonces $\varphi_i(U_i \cap U_j)$ es también un abierto de \mathbb{R}^n .
3. El mapa $\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \rightarrow \varphi_j(U_i \cap U_j)$ es continuo, inyectivo y suryectivo.

La condición 1 nos da una noción de *cercanía* en M inducida de la noción análoga en \mathbb{R}^n . En efecto, podemos decir que una secuencia de puntos $\{p_k\}$ en M convergen a p en U_i si existe k_0 tal que $\forall k > k_0$, $p_k \in U_i$ y la secuencia $\{\varphi_i(p_k)\}$ converge a $\varphi_i(p)$. Otra manera de ver esto es que si luego de esta construcción de una variedad imponemos que los mapas φ_i sean continuos entonces inducimos de forma unívoca una topología en M . [Ver figura 3.3.]

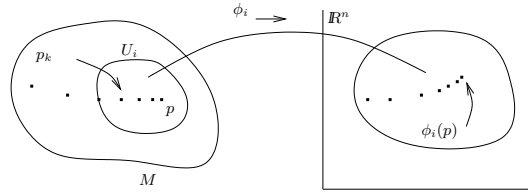


Figura 3.3: Sucesiones en M .

La condición 2 simplemente nos asegura que esta noción es consistente. Si $p \in U_i \cap U_j$ luego el hecho de que la secuencia converja es independiente de que usemos la carta (U_i, φ_i) o la (U_j, φ_j) .

La condición 3 nos permite codificar en los mapas $\varphi_j \circ \varphi_i^{-1}$ de \mathbb{R}^n en \mathbb{R}^n la información topológica *global* necesaria para distinguir, por ejemplo, si M es una esfera o un plano o un toro. Allí está,

por ejemplo la información de que no existe ningún mapa continuo e invertible entre S^2 y \mathbb{R}^2 . Pero también, si pedimos que estos mapas sean diferenciables, es lo que nos permitirá formular el cálculo diferencial en M . En efecto, note que en la condición 3 hablamos de la continuidad del mapa $\varphi_j \circ \varphi_i^{-1}$, la cual está bien definida debido a que este es un mapa entre \mathbb{R}^n y \mathbb{R}^n . Análogamente podemos hablar de la diferenciabilidad de estos mapas.

Diremos que **un atlas** $\{(U_i, \varphi_i)\}$, es C^p si los mapas $\varphi_j \circ \varphi_i^{-1}$ son p -veces diferenciables y su p -ésima derivada es continua.

Uno estaría tentado a definir la variedad M como el par que consiste del conjunto M y un atlas $\{(U_i, \varphi_i)\}$, pero esto nos llevaría a considerar como distintas variedades, por ejemplo, el plano con un atlas dado por la carta $(\mathbb{R}^2, (x, y) \rightarrow (x, y))$ y el plano con un atlas dado por la carta $(\mathbb{R}^2, (x, y) \rightarrow (x, -y))$.

Para subsanar este inconveniente introducimos el concepto de equivalencia entre atlas.

Definición: Diremos que **dos atlas son equivalentes** si su unión es también un atlas.

Ejercicio: Demuestre que esta es realmente una relación de equivalencia \approx , es decir que cumple:

- i) $A \approx A$
- ii) $A \approx B \implies B \approx A$
- iii) $A \approx B, B \approx C \implies A \approx C$.

Con esta relación de equivalencia podemos dividir el conjunto de atlas de M en distintas **clases equivalentes**. [Recuerde que cada clase equivalente es un conjunto donde todos sus elementos son equivalentes entre sí y tal que no hay ningún elemento equivalente a estos que no esté en él.]

Definición: Llamaremos **variedad M de dimensión n y diferenciabilidad p** al par que consiste del conjunto M y de una clase equivalente de atlas, $\{\varphi_i : U_i \rightarrow \mathbb{R}^n\}$, en C^p .

Se puede demostrar que para caracterizar la variedad M unívocamente es suficiente dar el conjunto M y un atlas. Si tenemos dos atlas de M , luego, o bien estos son equivalentes y así representan la misma variedad, o no lo son y entonces representan variedades distintas.

La definición de variedad que hemos introducido es todavía demasiado general para las aplicaciones físicas usuales, en el sentido de que la topología permitidas pueden aún ser patológicas desde el punto de vista de la física. Por ello en este curso impondremos a las variedades una condición extra. Supondremos que éstas son **separables o Hausdorff**. Eso es si p y $q \in M$ luego: o pertenecen al dominio de una misma carta U_i (en cuyo caso existen entornos W_p de $\varphi_i(p)$ y W_q de $\varphi_i(q)$ tales que $\varphi_i^{-1}(W_p) \cap \varphi_i^{-1}(W_q) = \emptyset$, o sea los puntos tienen entornos disjuntos) o bien existen U_i y U_j con $p \in U_i$, $q \in U_j$ y $U_i \cap U_j = \emptyset$, lo que también implica que tienen entornos disjuntos. Esta es una propiedad en la topología de M que esencialmente dice que podemos separar puntos de M . Un ejemplo de una variedad no-Hausdorff es el siguiente.

Ejemplo: M , como conjunto, consta de tres intervalos de la recta $I_1 = (-\infty, 0]$, $I_2 = (-\infty, 0]$ y $I_3 = (0, +\infty)$.

Un atlas de M es $\{(U_1 = I_1 \cup I_3, \varphi_1 = id), (U_2 = I_2 \cup I_3, \varphi_2 = id)\}$. [Ver figura 3.4.]

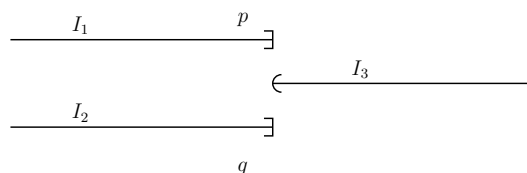


Figura 3.4: Ejemplo de variedad no-Hausdorff.

Ejercicio: Pruebe que es un atlas.

Note que dado cualquier entorno W_1 de $\varphi_1(0)$ en \mathbb{R} y cualquier entorno W_2 de $\varphi_2(0)$ tenemos necesariamente que $\varphi_1^{-1}(W_1) \cap \varphi_2^{-1}(W_2) \neq \emptyset$.

3.2. Funciones Diferenciables en M

De ahora en más asumiremos que M es una variedad C^∞ , o sea que todos sus mapas $\varphi_i \circ \varphi_j^{-1}$ son infinitamente diferenciables. Si bien matemáticamente esto es una restricción, no lo es en las aplicaciones físicas. En estas M es generalmente el espacio de posibles estados del sistema y por lo tanto sus puntos no pueden ser determinados con absoluta certeza, ya que toda medición involucra cierto error. Esto indica que por medio de mediciones nunca podríamos saber el grado de diferenciableidad de M . Por conveniencia supondremos que es C^∞ .

Una función en M es un mapa $f : M \rightarrow \mathbb{R}$, o sea un mapa que asigna un número real a cada punto de M . La información codificada en el atlas sobre M nos permite decir cuán suave es f .

Definición: Diremos que f es **p -veces continuamente diferenciable** en el punto $q \in M$, $f \in C_q^p$ si dada (U_i, φ_i) con $q \in U_i$, $f \circ \varphi_i^{-1} : \varphi_i(U_i) \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es p -veces continuamente diferenciable en $\varphi_i(q)$.

Note que esta propiedad es independiente de la carta empleada [mientras consideremos solo cartas de la clase compatible de atlas]. Diremos que $f \in C^p(M)$ si $f \in C_q^p \forall q \in M$. [Ver figura 3.5.]

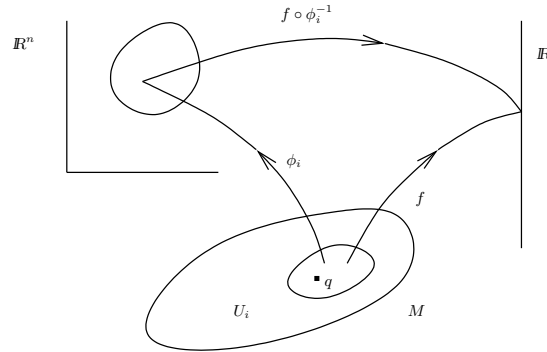
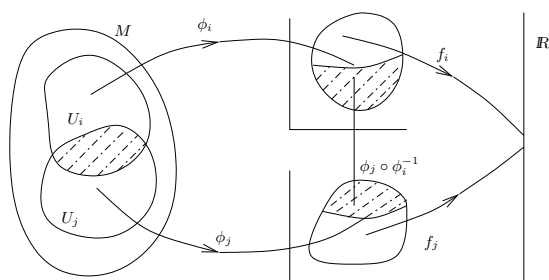


Figura 3.5: Composición del mapa de una carta con una función.

En la práctica uno define una función particular $f \in C^p(M)$ introduciendo funciones $f_i : \varphi_i(U_i) \subset \mathbb{R}^n \rightarrow \mathbb{R}$ (o sea $f_i(x^j)$ donde x^j son las coordenadas cartesianas en $\varphi_i(U_i) \subset \mathbb{R}^n$) que sean C^p en $\varphi_i(U_i)$ y tales que $f_i = f_j \circ \varphi_j \circ \varphi_i^{-1}$ en $\varphi_i(U_i \cap U_j)$. Esto garantiza que el conjunto de las f_i determinan una única función $f \in C^p(M)$. El conjunto de las $f_i (= f \circ \varphi_i^{-1})$ forman una **representación de f** en el atlas $\{(U_i, \varphi_i)\}$. [Ver figura 3.6.]

Ejercicio: El círculo, S^1 , se puede pensar como el intervalo $[0,1]$ con sus extremos identificados. ¿Cuáles son las funciones en $C^2(S^1)$?

Usando la construcción anterior también se pueden definir mapas de M en \mathbb{R}^n que sean p -veces diferenciables. Ahora realizamos la construcción inversa es decir definiremos la diferenciableidad de un mapa de \mathbb{R}^n en M . Haremos el caso $\mathbb{R} \rightarrow M$, en cuyo caso el mapa así obtenido se llama curva. El caso general es obvio.

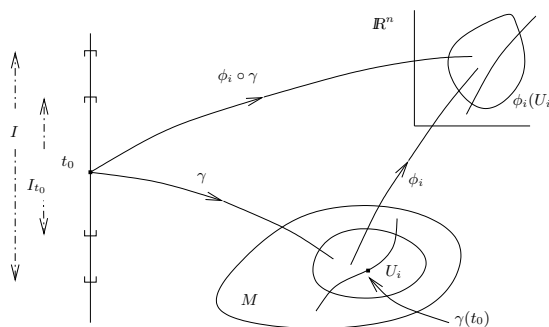
Figura 3.6: La relación entre las f_i .

3.3. Curvas en M

Definición: Una curva en M es un mapa entre un intervalo $I \subset \mathbb{R}$ y M , $\gamma : I \rightarrow M$.

Note que la curva es el mapa y no su gráfico en M , es decir el conjunto $\gamma(I)$. Se puede así tener dos curvas distintas con el mismo gráfico. Esto no es un capricho matemático sino una necesidad física : no es lo mismo que un auto recorra el camino Córdoba–Carlos Paz a 10 km/h que a 100 km/h, o lo recorra en sentido contrario.

Definición: Diremos que $\gamma \in C^p_{t_0}$ si dada una carta (U_i, φ_i) tal que $\gamma(t_0) \in U_i$ el mapa $\varphi_i \circ \gamma(t) : I_{t_0} \subset I \rightarrow \mathbb{R}^n$ es p -veces continuamente diferenciable en t_0 . [Ver figura 3.7.]

Figura 3.7: Diferenciabilidad de curvas en M .

Ejercicio: Demuestre que la definición anterior no depende de la carta usada.

Esta vez hemos usado el concepto de diferenciabilidad entre mapas de \mathbb{R} en \mathbb{R}^n .

Definición: Una curva $\gamma(t) \in C^p(I)$ si $\gamma(t) \in C^p_t \forall t \in I$.

Ejercicio: ¿Cómo definiría el concepto de diferenciabilidad de mapas entre dos variedades?

De particular importancia entre estos son los mapas de M en sí mismo $g : M \rightarrow M$ que son continuamente diferenciables e invertible. Se llaman **Difeomorfismos**. De ahora en más supondremos que todas las variedades, curvas, difeomorfismos y funciones son suaves, es decir, son

C^∞ .

3.4. Vectores

Para definir vectores en puntos de M utilizaremos el concepto de derivada direccional en puntos de \mathbb{R}^n , es decir, explotaremos el hecho de que en \mathbb{R}^n hay una correspondencia uno a uno entre vectores $(v^1, \dots, v^n)|_{x_0}$ y derivadas direccionales $\mathbf{v}(f)|_{x_0} = v^i \frac{\partial}{\partial x^i} f|_{x_0}$.

Como hemos definido funciones diferenciables en M podemos definir derivaciones, o sea derivadas direccionales, en sus puntos e identificar con ellos a los vectores tangentes.

Definición: Un **vector tangente \mathbf{v} en $p \in M$** es un mapa $\mathbf{v} : C^\infty(M) \rightarrow \mathbb{R}$ satisfaciendo:
 $\forall f, g \in C^\infty(M), a, b \in \mathbb{R}$

i) Linealidad; $\mathbf{v}(af + bg)|_p = a \mathbf{v}(f)|_p + b \mathbf{v}(g)|_p$.

ii) Leibnitz; $\mathbf{v}(fg)|_p = f(p) \mathbf{v}(g)|_p + g(p) \mathbf{v}(f)|_p$.

Note que si $h \in C^\infty(M)$ es la función constante, $h(q) = c \quad \forall q \in M$, luego $\mathbf{v}(h) = 0$. [i) $\implies \mathbf{v}(h^2) = \mathbf{v}(ch) = c \mathbf{v}(h)$ mientras que ii) $\implies \mathbf{v}(h^2) = 2h(p) \mathbf{v}(h) = 2c \mathbf{v}(h)$]. Estas propiedades muestran asimismo que $\mathbf{v}(f)$ depende solo del comportamiento de f en p .

Ejercicio: Pruebe esta última afirmación.

Sea T_p el conjunto de todos los vectores en p . Este conjunto tiene la estructura de un espacio vectorial y es llamado el espacio tangente al punto p . En efecto, podemos definir la suma de dos vectores $\mathbf{v}_1, \mathbf{v}_2$ como el vector, es decir el mapa, satisfaciendo i) y ii), $(\mathbf{v}_1 + \mathbf{v}_2)(f) = \mathbf{v}_1(f) + \mathbf{v}_2(f)$ y el producto del vector \mathbf{v} por el número a como el mapa $(a\mathbf{v})(f) = a \mathbf{v}(f)$.

Como en \mathbb{R}^n , la dimensión del espacio vectorial T_p , (es decir el número máximo de vectores linealmente independientes), es n .

Teorema 3.1 $\dim T_p = \dim M$.

Prueba: Esta consistirá en encontrar una base para T_p . Sea $\dim M = n$ y (U, φ) tal que $p \in U$ y $f \in C^\infty(M)$ cualquiera. Para $i = 1, \dots, n$ definimos los vectores $\mathbf{x}_i : C^\infty(M) \rightarrow \mathbb{R}$ dados por,

$$\mathbf{x}_i(f) = \frac{\partial}{\partial x^i} (f \circ \varphi^{-1}) \Big|_{\varphi(p)}. \quad (3.1)$$

Note que estos mapas satisfacen i) y ii) y por lo tanto las \mathbf{x}_i son realmente vectores. Note además que el lado derecho de 3.1 está bien definido ya que tenemos las derivadas parciales usuales de mapas entre \mathbb{R}^n y \mathbb{R} . Estos \mathbf{x}_i dependen de la carta (U, φ) pero esto no importa en la prueba ya que T_p no depende de carta alguna. Estos vectores son linealmente independientes, es decir si $x = \sum_{i=1}^n c^i \mathbf{x}_i = 0$ luego $c^i = 0 \quad \forall i = 1, \dots, n$. Esto se ve fácilmente considerando las funciones – en rigor definidas solamente en U – $f^j = x^j \circ \varphi$, ya que $\mathbf{x}_i(f^j) = \delta_i^j$ y por lo tanto $0 = \mathbf{x}(f^j) = c^j$. Solo resta mostrar que cualquier vector \mathbf{v} puede ser expresado como combinación lineal de los \mathbf{x}_i . Para ello usaremos el siguiente resultado cuya prueba dejamos como ejercicio.

Lema 3.1 Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}$ $F \in C^\infty(\mathbb{R}^n)$ luego para cada $x_0 \in \mathbb{R}^n$ existen funciones $H_i : \mathbb{R}^n \rightarrow \mathbb{R} \in C^\infty(\mathbb{R}^n)$ tales que $\forall x \in \mathbb{R}^n$ se cumple

$$F(x) = F(x_0) + \sum_{i=1}^n (x^i - x_0^i) H_i(x) \quad y \quad (3.2)$$

$$\text{además,} \quad \frac{\partial F}{\partial x^i} \Big|_{x=x_0} = H_i(x_0). \quad (3.3)$$

Continuamos ahora la prueba del Teorema anterior. Sea $F = f \circ \varphi^{-1}$ y $x_0 = \varphi(p)$, luego $\forall q \in U$ tenemos

$$f(q) = f(p) + (x^i \circ \varphi(q) - x^i \circ \varphi(p)) H_i \circ \varphi(q) \quad (3.4)$$

Usando $i)$ y $ii)$ obtenemos,

$$\begin{aligned} \mathbf{v}(f) &= \mathbf{v}(f(p)) + \sum_{i=1}^n (x^i \circ \varphi(q) - x^i \circ \varphi(p)) \Big|_{q=p} \mathbf{v}(H_i \circ \varphi) \\ &\quad + \sum_{i=1}^n (H_i \circ \varphi)|_p \mathbf{v}(x^i \circ \varphi - x^i \circ \varphi(p)) \\ &= \sum_{i=1}^n (H_i \circ \varphi)|_p \mathbf{v}(x^i \circ \varphi) \\ &= \sum_{i=1}^n v^i \mathbf{x}_i(f) \end{aligned} \quad (3.5)$$

donde $v^i \equiv \mathbf{v}(x^i \circ \varphi)$, y por lo tanto hemos expresado \mathbf{v} como combinación lineal de los \mathbf{x}_i , finalizando así la prueba.

La base $\{\mathbf{x}_i\}$ se llama una **base coordenada** y los $\{v^i\}$, los **componentes de \mathbf{v}** en esa base.

Ejercicio: Si $(\tilde{U}, \tilde{\varphi})$ es otra carta tal que $p \in \tilde{U}$, entonces esta definirá otra base coordenada $\{\tilde{\mathbf{x}}_i\}$. Muestre que

$$\mathbf{x}_j = \sum_{i=1}^n \frac{\partial \tilde{x}^i}{\partial x^j} \tilde{\mathbf{x}}_i$$

donde \tilde{x}^i es la i -ésima componente del mapa $\tilde{\varphi} \circ \varphi^{-1}$. Muestre también que la relación entre las componentes es $\tilde{v}^i = \sum_{j=1}^n \frac{\partial \tilde{x}^i}{\partial x^j} v^j$.

Ejemplo: Sea $\gamma : I \rightarrow M$ una curva en M . En cada punto $\gamma(t_0)$, $t_0 \in I$, de M podemos definir un vector de la siguiente forma, [Ver figura 3.8.]

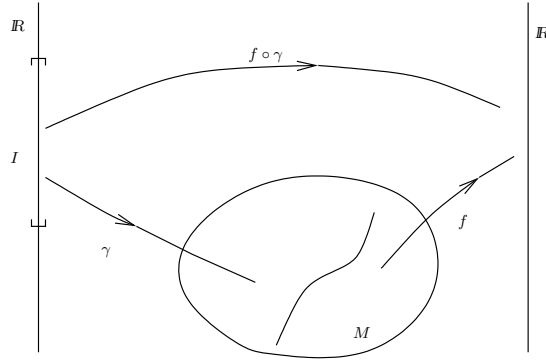


Figura 3.8: Definición de vector.

$$\mathbf{t}(f) = \frac{d}{dt} (f \circ \gamma)|_{t=t_0} \quad (3.6)$$

Sus componentes en una base coordenada se obtienen por medio de las funciones

$$x^i(t) = \varphi \circ \gamma(t) \quad (3.7)$$

$$\frac{d}{dt}(f \circ \gamma) = \frac{d}{dt}(f \circ \varphi^{-1} \circ \varphi \circ \gamma)$$

$$\begin{aligned}
&= \frac{d}{dt}(f \circ \varphi^{-1}(x^i(t))) \\
&= \sum_{i=1}^n \left(\frac{\partial}{\partial x^i}(f \circ \varphi^{-1}) \right) \frac{dx^i}{dt} \\
&= \sum_{i=1}^n \frac{dx^i}{dt} x_i(f)
\end{aligned} \tag{3.8}$$

3.5. Campos Vectoriales y Tensoriales

Si a cada punto q de M le asignamos un vector $\mathbf{v}|_q \in T_q$ tendremos un **campo vectorial**. Este estará en $C^\infty(M)$ si dada cualquier $f \in C^\infty(M)$ la función $\mathbf{v}(f)$, que en cada punto p de M le asigna el valor $\mathbf{v}|_p(f)$, esta también en $C^\infty(M)$. Al conjunto de los campos vectoriales C^∞ lo denotaremos TM y obviamente es un espacio vectorial, de dimensión infinita.

3.5.1. El Corchete de Lie

Consideremos ahora la operación en el conjunto TM de campos vectoriales, $[\cdot, \cdot] : TM \times TM \rightarrow TM$. Esta operación se llama **corchete de Lie** y dados dos campos vectoriales (C^∞) nos da un tercero:

$$[\mathbf{x}, \mathbf{y}](f) := \mathbf{x}(\mathbf{y}(f)) - \mathbf{y}(\mathbf{x}(f)). \tag{3.9}$$

Ejercicio:

- 1) Muestre que $[\mathbf{x}, \mathbf{y}]$ es en realidad un campo vectorial.
- 2) Vea que se satisface la **identidad de Jacobi**:

$$[[\mathbf{x}, \mathbf{y}], \mathbf{z}] + [[\mathbf{z}, \mathbf{x}], \mathbf{y}] + [[\mathbf{y}, \mathbf{z}], \mathbf{x}] = 0 \tag{3.10}$$

3) Sean \mathbf{x}^i y \mathbf{x}^j dos campos vectoriales provenientes de un sistema coordenado, es decir $\mathbf{x}^i(f) = \frac{\partial f}{\partial x^i}$, etc. Muestre que $[\mathbf{x}^i, \mathbf{x}^j] = 0$.

- 4) Dadas las componentes de \mathbf{x} e \mathbf{y} en una base coordenada, ¿cuáles son las de $[\mathbf{x}, \mathbf{y}]$?
Con esta operación TM adquiere el carácter de un álgebra, llamada **Álgebra de Lie**.

3.5.2. Difeomorfismos y la Teoría de las Ecuaciones Diferenciales Ordinarias

Definición: Un **grupo monoparamétrica de difeomorfismos** g^t es un mapa $\mathbb{R} \times M \rightarrow M$ tal que:

- 1) Para cada t fijo es un difeomorfismo $^1 M \rightarrow M$
- 2) Para todo par de reales, $t, s \in \mathbb{R}$ tenemos $g^t \circ g^s = g^{t+s}$ (en particular $g^0 = id$).

Podemos asociar con g^t un campo vectorial de la siguiente manera: Para un p fijo $g^t(p) : \mathbb{R} \rightarrow M$ es una curva que en $t = 0$ pasa por p y por lo tanto define un vector tangente en p , $\mathbf{v}|_p$. Repitiendo el proceso para todo punto de M tenemos un campo vectorial en M . Note que debido a la propiedad de grupo que satisface g^t el vector tangente a la curva $g^t(p)$ es también tangente a la curva $g^s(g^t(p))$ en $s = 0$.

Podemos hacernos la pregunta inversa: ¿Dado un campo vectorial suave \mathbf{v} en M existirá un grupo monoparamétrico de difeomorfismos que lo defina? La respuesta a esta pregunta, que consiste en encontrar todas las curvas integrables $g^t(p)$ que pasan por cada $p \in M$ es la teoría de las ecuaciones diferenciales ordinarias, –que será el objeto de nuestro estudio en los capítulos siguientes– ya que consiste en resolver las ecuaciones $\frac{dx^i}{dt} = v^i(x^j)$ con condiciones iniciales $x^i(0) = \varphi^i(p) \quad \forall p \in M$.

¹Es decir un mapa suave con inversa también suave.

Como veremos la respuesta es afirmativa pero solo localmente, es decir podremos encontrar solo g^t definidos en $I(\subset \mathbb{R}) \times U(\subset M) \rightarrow M$.

Ejemplo: En \mathbb{R}^1 sea el vector con componente coordenada x^2 , es decir $V(x) = x^2 \frac{d}{dx}$. La ecuación diferencial ordinaria asociada con este vector es $\frac{dx}{dt} = x^2$, cuya solución es

$$t - t_0 = \frac{-1}{x} + \frac{1}{x_0} \quad \text{ó} \quad x(t) = \frac{-1}{t - \frac{1}{x_0}} \quad (3.11)$$

donde hemos tomado $t_0 = 0$. O sea $g^t(x_0) = \frac{-1}{t - \frac{1}{x_0}}$. Note que cualquiera sea t este mapa no está definido para todo \mathbb{R} y por lo tanto no es un difeomorfismo.

Ejemplo: Sea g^t un difeomorfismo lineal en \mathbb{R} , es decir $g^t(x + \alpha y) = g^t(x) + \alpha g^t(y)$. Luego tiene la forma $g^t(x) = f(t)x$. La propiedad de grupo implica $f(t) \cdot f(s) = f(t + s)$ o $f(t) = ce^{kt} = e^{kt}$, ya que $g^0 = id$. Por lo tanto $g^t(x) = e^{kt}x$. La ecuación diferencial asociada es: $x(t) = e^{kt}x_0 \implies \dot{x} = ke^{kt}x_0 = \boxed{kx = \dot{x}}$.

Ejercicio: Grafique en un entorno del origen \mathbb{R}^2 las curvas integrales y por lo tanto g^t de los siguientes sistemas lineales.

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3.12)$$

a) $k > 1$ b) $k = 1$ c) $0 < k < 1$, d) $k = 0$, e) $k < 0$

3.5.3. Campos de Covectores y Tensores

Así como introdujimos la noción de campo vectorial podemos también introducir la de campo de co-vectores, es decir un mapa suave de M en T_p^* . Este actuará en campos vectoriales dando como resultado funciones en M . En el ejemplo que sigue vemos cómo se define el campo **diferencial de f** .

Ejemplo: Sea $f \in C_p^\infty$. Un vector en $p \in M$ es una derivación sobre funciones en C_p^∞ , $v(f) \in \mathbb{R}$. Pero dados v_1 y $v_2 \in T_p$ $(v_1 + v_2)(f) = v_1(f) + v_2(f)$ y por lo tanto f define una funcional lineal $df|_p: T_p \rightarrow \mathbb{R}$, llamada la diferencial de f , es decir un elemento de T_p^* . De esta forma la diferencial de una función df es un co-vector que al actuar sobre un vector v nos da el número *la derivada de f en el punto p en la dirección de v* .

Considere el subconjunto S de M tal que $f(M) = \text{cte}$. Esta será una subvariedad de M , es decir superficies embebidas en M , de dimensión $n - 1$. La condición $df|_p(v) = 0$ en vectores de T_p con $p \in S$ significa que estos son en realidad vectores tangentes a S , es decir elementos de $T_p(S)$. Por el contrario, si $df(v)|_p \neq 0$ entonces en ese punto v pincha a S .

Ejemplo: La función $f(x, y, z) = x^2 + y^2 + z^2$ en \mathbb{R}^3 .

$S = \{(x, y, z) \in \mathbb{R}^3 / f(x, y, z) = a^2, a^2 > 0\}$ es la esfera de radio a , y como ya hemos visto una variedad. Sea (v_x, v_y, v_z) un vector en $(x, y, z) \in \mathbb{R}$, luego la condición $df(v) = 2(xv_x + yv_y + zv_z) = 0$ implica que v es tangente a S .

Similarmente definimos **campos tensoriales** como los mapas multilineales que al actuar sobre campos vectoriales y co-vectoriales dan funciones de la variedad en los reales y que en cada punto de la variedad sólo dependen de los vectores y co-vectores definidos en ese punto. Esta última aclaración es necesaria pues de lo contrario incluiremos entre los tensores, por ejemplo, integrales lineales sobre los campos vectoriales.

3.5.4. La Métrica

Sea M una variedad n -dimensional. Hemos definido anteriormente en M las nociones de curvas, campos vectoriales y co-vectoriales, etc. pero no una noción de distancia entre sus puntos, es decir una función $d : M \times M \rightarrow \mathbb{R}$ que toma dos puntos cualquiera, p y q de M y nos da un número $d(p, q)$ satisfaciendo,

1. $d(p, q) \geq 0$.
2. $d(p, q) = 0 \leftrightarrow p = q$.
3. $d(p, q) = d(q, p)$.
4. $d(p, q) \leq d(p, r) + d(r, q)$.

Esta, y en algunos casos una noción de pseudo-distancia [donde 1) y 2) no se cumplen], es fundamental si queremos tener una estructura matemática que sea útil para la descripción de fenómenos físicos. Por ejemplo la ley de Hooke, que nos dice que la fuerza aplicada a un resorte es proporcional a la elongación (una distancia) de éste, claramente necesita de este ente. A continuación introduciremos una noción de distancia infinitesimal, es decir entre dos puntos infinitesimalmente separados, la cual responde a la noción Euclídea de distancia y permite desarrollar una noción de distancia global, es decir entre dos puntos cualesquiera de M .

La idea es entonces tener un concepto de distancia (o pseudo-distancia) entre dos puntos *infinitesimalmente cercanos*, es decir dos puntos conectados por un *desplazamiento infinitesimal*, es decir conectados por un vector. La noción que necesitamos es entonces la de norma de un vector. Como una variedad es localmente como \mathbb{R}^n , en el sentido de que el espacio de vectores tangentes a un punto p , $T_p M$ es \mathbb{R}^n , es razonable considerar allí la noción de distancia Euclídea, es decir que la distancia entre dos puntos x_0 y $x_1 \in \mathbb{R}^n$ es la raíz cuadrada de la suma de los cuadrados de las componentes (en algún sistema de coordenadas) del vector conectando estos dos puntos. El inconveniente con esto es que dicha noción depende del sistema coordenado que se esté usando y por lo tanto habrá tantas distancias como sistemas coordenados cubriendo el punto p . Esto no es más que una indicación que la estructura que hasta este momento tenemos no contiene una noción de distancia privilegiada o natural. Esta debe ser introducida como una estructura adicional. Una manera de obtener distancias infinitesimales independientes del sistema coordenado (es decir geométricas) es introduciendo en cada punto $p \in M$ un tensor de tipo $(2,0)$, simétrico $[\mathbf{g}(\mathbf{u}, \mathbf{v}) = \mathbf{g}(\mathbf{v}, \mathbf{u}) \forall \mathbf{u}, \mathbf{v} \in T_p M]$ y no degenerado $[\mathbf{g}(\mathbf{u}, \mathbf{v}) = \mathbf{0} \forall \mathbf{v} \in T_p M \Rightarrow \mathbf{u} = \mathbf{0}]$. Si además pedimos que este tensor sea definido positivo $[\mathbf{g}(\mathbf{u}, \mathbf{u}) \geq \mathbf{0} (= \Leftrightarrow \mathbf{u} = \mathbf{0})]$ se puede ver fácilmente que este define un producto escalar en $T_p M$ (o pseudoescalar si $\mathbf{g}(\mathbf{u}, \mathbf{u}) = \mathbf{0}$ para algún $\mathbf{u} \neq \mathbf{0} \in T_p M$).² Si hacemos una elección para este tensor en cada punto de M de forma suave obtendremos un campo tensorial suave que se denomina la **métrica** de M . Esta estructura extra, un campo tensorial con ciertas propiedades es lo que nos permite construir las bases matemáticas para luego edificar gran parte de la física sobre ella.

Sea \mathbf{g} una métrica en M , dado un punto cualquiera p de M existe un sistema coordenado en que sus componentes son

$$g_{ij} = \delta_{ij}$$

y por lo tanto da origen al producto escalar Euclídeo, sin embargo en general este resultado no se puede extender a un entorno del punto y en general sus componentes dependerán allí de las coordenadas. Note que esto es lo que queríamos hacer en un comienzo, pero ahora al definir esta norma vía un vector le hemos dado un carácter invariante.

Restringiéndonos ahora a métricas definidas positivas definiremos **la norma** de un vector $v \in T_p$ como $|v| = \sqrt{|g(v, v)|}$, es decir como la distancia infinitesimal dividida por ϵ entre el p y el punto $\gamma(\epsilon)$ donde $\gamma(t)$ es una curva tal que $\gamma(0) = p$, $\frac{d\gamma(t)}{dt}|_{t=0} = v$. Análogamente podemos definir la

²Posteriormente veremos que un producto escalar da origen a una distancia, correspondientemente un producto pseudo-escalar da origen a una pseudo-distancia.

longitud de una curva suave $\gamma(t) : [0, 1] \rightarrow M$ por la fórmula,

$$L(\gamma) = \int_0^1 \sqrt{g(\mathbf{v}, \mathbf{v})} dt, \quad (3.13)$$

donde $\mathbf{v}(t) = \frac{d\gamma(t)}{dt}$. Vemos entonces que definimos la longitud de una curva midiendo las longitudes infinitesimales entre puntos cercanos de ésta y luego integramos con respecto a t .

Ejercicio: Pruebe que la distancia $L(\gamma)$ es independiente del parámetro elegido.

Definimos la distancia entre dos puntos $p, q \in M$ como,

$$d_g(p, q) = \inf_{\{\gamma(t) : \gamma(0)=p, \gamma(1)=q\}} |L(\gamma)| \quad (3.14)$$

Es decir como el ínfimo de la longitud de todas las curvas conectando p con q .

Ejercicio: Encuentre un ejemplo de una variedad con dos puntos tales que el ínfimo de la definición anterior no es un mínimo. Es decir en la que no haya ninguna curva conectando los dos puntos con la mínima distancia entre ellos.

Ejercicio: a) La métrica Euclídea en \mathbb{R}^2 es $(dx)^2 + (dy)^2$, donde $\{dx, dy\}$ es la co-base asociada con $\{\partial x, \partial y\}$. ¿Cuál es la distancia entre dos puntos en este caso?

Ejercicio: b) ¿Cuál es la forma de la métrica Euclídea de \mathbb{R}^3 en coordenadas esféricas? ¿Y en cilíndricas?

Ejercicio: c) La métrica de la esfera es $(d\theta)^2 + \sin^2 \theta (d\varphi)^2$. ¿Cuál es la distancia en este caso? ¿Para qué puntos p, q existen varias curvas γ_i con $L(\gamma_i) = d(p, q)$?

Ejercicio: d) La métrica $(dx)^2 + (dy)^2 + (dz)^2 - (dt)^2$ en \mathbb{R}^4 es la métrica de Minkowski de la relatividad especial. ¿Cuál es la distancia entre el punto de coordenadas $(0, 0, 0, 0)$ y $(1, 0, 0, 1)$?

Una métrica nos da un mapa privilegiado entre el espacio de vectores tangentes a p , T_p y su dual T_p^* para cada p en M , es decir el mapa que a cada vector $\mathbf{v} \in T_p$ le asigna el co-vector $\mathbf{g}(\mathbf{v}, \cdot) \in T_p^*$. Como esto es válido para cada p obtenemos así un mapa entre campos vectoriales y co-vectoriales. Como \mathbf{g} es no degenerada este mapa es invertible, es decir existe un tensor de tipo $(2, 0)$ simétrico \mathbf{g}^{-1} tal que

$$\mathbf{g}(\mathbf{g}^{-1}(\theta, \cdot), \cdot) = \theta \quad (3.15)$$

para todo campo co-vectorial θ . Esto indica que cuando tenemos una variedad con una métrica se hace irrelevante distinguir entre vectores y co-vectores o por ejemplo entre tensores tipo $(0, 2)$, $(2, 0)$ o $(1, 1)$.

3.5.5. Notación de Índices Abstractos

Cuando se trabaja con objetos tensoriales la notación hasta ahora usada no es la más conveniente pues es difícil recordar cuál es el tipo de cada tensor, en que casilla *come* otros objetos, etc. Una solución es introducir un sistema de coordenadas y trabajar con las componentes de los tensores, donde al haber índices es fácil saber de qué objetos se trata o introducir bases generales. De esta forma, por ejemplo representamos al vector $\mathbf{l} = l^i \frac{\partial}{\partial x^i}$ por sus componentes $\{l^i\}$. Una conveniencia de esta notación es que *comer* pasa a ser *contraer*, ya que por ejemplo representamos al vector \mathbf{l}

comiéndose a una función f , por la contracción de las componentes coordenadas del vector y de la diferencial de f :

$$l(f) = \sum_{i=1}^n l^i \frac{\partial f}{\partial x^i}.$$

Pero un inconveniente grave de esta representación es que en principio depende del sistema coordinado y por lo tanto todas las expresiones que construyamos con ella tienen el potencial peligro de depender de tal sistema.

Remediamos esto introduciendo **índices abstractos** (que serán letras latinas) que indican donde irían los índices coordinados pero nada más, es decir no dependen del sistema de coordenadas y ni siquiera toman valores numéricos, es decir l^a no significa (l^1, l^2, \dots, l^n) como si fuesen índices. De esa forma l^a denotará al vector l , θ_a al co-vector θ y g_{ab} a la métrica g . Una contracción como por ejemplo $g(v, \cdot)$ se denotará $g_{ab}v^a$ y a este covector lo denotaremos por v_b , es decir la acción de g_{ab} es la de bajar el índice a v^a y dar el covector $v_b \equiv v^a g_{ab}$. Del mismo modo denotaremos a g^{-1} (la inversa de g) como g^{ab} , es decir g con *los índices subidos*.

La simetría de g es entonces el equivalente a $g_{ab} = g_{ba}$.

Ejercicio: ¿Cómo denotaría un tensor de tipo $(2, 0)$ antisimétrico?

Usando índices repetidos para la contracción vemos que $l(f)$ se puede denotar por $l^a \nabla_a f$ donde $\nabla_a f$ denota el co-vector diferencial de f , mientras el vector $\nabla^a f := g^{ab} \nabla_b f$ es llamado el **gradiente** de f y vemos que éste no solo depende de f sino también de g .

3.6. Derivada Covariante

Hemos visto que en M existe la noción de la derivada de un campo escalar f , ésta es el co-vector diferencial de f que denotamos $\nabla_a f$. ¿Existirá la noción de la derivada de un campo tensorial? Por ejemplo, ¿existirá una extensión del operador ∇_a a vectores tal que si l^a es un vector diferenciable luego $\nabla_a l^b$ sea un tensor del tipo $(1, 1)$? Para fijar ideas definamos a esta extensión del diferencial ∇_a , llamada derivada covariante, pidiendo que satisfaga las siguientes propiedades:³

i) Linealidad: Si $A_{b_1 \dots b_l}^{a_1 \dots a_k}, B_{b_1 \dots b_l}^{a_1 \dots a_k}$ son tensores de tipo (k, l) y $\alpha \in \mathbb{R}$ luego

$$\nabla_c (\alpha A_{b_1 \dots b_l}^{a_1 \dots a_k} + B_{b_1 \dots b_l}^{a_1 \dots a_k}) = \alpha \nabla_c A_{b_1 \dots b_l}^{a_1 \dots a_k} + \nabla_c B_{b_1 \dots b_l}^{a_1 \dots a_k} \quad (3.16)$$

ii) Leibnitz:

$$\nabla_e (A_{b_1 \dots b_l}^{a_1 \dots a_k} B_{d_1 \dots d_n}^{c_1 \dots c_m}) = A_{b_1 \dots b_l}^{a_1 \dots a_k} (\nabla_e B_{d_1 \dots d_n}^{c_1 \dots c_m}) + (\nabla_e A_{b_1 \dots b_l}^{a_1 \dots a_k}) B_{d_1 \dots d_n}^{c_1 \dots c_m} \quad (3.17)$$

iii) Conmutatividad con contracciones:

$$\nabla_e (\delta^i_j A_{b_1 \dots, i, \dots, b_k}^{a_1 \dots, j, \dots, a_k}) = \delta^i_j \nabla_e A_{b_1 \dots, i, \dots, b_k}^{a_1 \dots, j, \dots, a_k} \quad (3.18)$$

Es decir si primero contraemos algunos índices de un tensor y luego tomamos su derivada obtenemos el mismo tensor que si tomamos primero la derivada y luego contraemos.

iv) Consistencia con la diferencial: Si l^a es un campo vectorial y f un campo escalar, luego

$$l^a \nabla_a f = l(f) \quad (3.19)$$

v) Torsión cero: Si f es un campo escalar luego

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f \quad (3.20)$$

³Note que son una extensión de las que se pide para definir derivaciones.

Ejemplo: Sea $\{x^i\}$ un sistema coordinado global en \mathbb{R}^n y sea ∇_c el operador que cuando actúa en $A_{b_1 \dots b_l}^{a_1 \dots a_k}$ general el campo tensorial que en estas coordenadas tiene componentes.

$$\partial_j A_{j_1 \dots j_l}^{i_1 \dots i_k} \quad (3.21)$$

Por definición es un tensor y claramente satisface todas las condiciones de la definición, –ya que las satisface en este sistema de coordenadas– por lo tanto es una derivada covariante. Si tomamos otro sistema de coordenadas obtendremos otra derivada covariante, en general distinta de la anterior. Por ejemplo, hagamos actuar ∇_c en el vector l^a luego

$$(\nabla_c l^a)_j^i = \partial_j l^i. \quad (3.22)$$

En otro sistema de coordenadas $\{\bar{x}^i\}$ este tensor tiene componentes

$$(\nabla_c l^a)_k^l = \sum_{i,j=1}^n \frac{\partial \bar{x}^l}{\partial x^i} \frac{\partial x^j}{\partial \bar{x}^k} \frac{\partial l^i}{\partial x^j} \quad (3.23)$$

las cuales no son, en general, las componentes de la derivada covariante $\bar{\nabla}_c$ que estas nuevas coordenadas definen, en efecto

$$\begin{aligned} (\bar{\nabla}_c l^a)_k^l &\equiv \frac{\partial \bar{l}^l}{\partial \bar{x}^k} = \sum_{j=1}^n \left(\frac{\partial x^j}{\partial \bar{x}^k} \right) \frac{\partial}{\partial x^j} \sum_{i=1}^n \left(\frac{\partial \bar{x}^l}{\partial x^i} l^i \right) = \\ &= \sum_{i,j=1}^n \left(\frac{\partial \bar{x}^l}{\partial x^i} \right) \left(\frac{\partial x^j}{\partial \bar{x}^k} \right) \frac{\partial l^i}{\partial x^j} + \sum_{i,j=1}^n \frac{\partial x^j}{\partial \bar{x}^k} \left(\frac{\partial^2 \bar{x}^l}{\partial x^j \partial x^i} \right) l^i \\ &= (\nabla_c l^a)_j^i + \sum_{i,j=1}^n \frac{\partial x^j}{\partial \bar{x}^k} \left(\frac{\partial^2 \bar{x}^l}{\partial x^j \partial x^i} \right) l^i, \end{aligned} \quad (3.24)$$

lo que muestra claramente que son dos tensores distintos y que su diferencia es un **tensor** que depende linealmente y **no diferencialmente** de l^a . ¿Es esto cierto en general? Es decir, dada dos conexiones, ∇_c y $\bar{\nabla}_c$, ¿es su diferencia un tensor (y no un operador diferencial)? Veremos que esto es cierto.

Teorema 3.2 *La diferencia entre dos conexiones es un tensor.*

Prueba: Note que por propiedad *iii*) y *iv*) de la definición, si sabemos cómo actúa ∇_c en co-vectores sabemos cómo actúa en vectores y así por *i*) y *ii*) en cualquier tensor. En efecto, si conocemos $\nabla_c w_a$ para cualquier w_a luego $\nabla_c l^a$ es el tensor de tipo (1,1) tal que cuando contraído con un w_a arbitrario nos da el co-vector

$$(\nabla_c l^a) w_a = \nabla_c (w_a l^a) - l^a (\nabla_c w_a), \quad (3.25)$$

el cual conocemos ya que por *iv*) también sabemos cómo actúa ∇_c en escalares. Por lo tanto es suficiente ver que

$$(\bar{\nabla}_c - \nabla_c) w_a = C^b_{ca} w_b \quad (3.26)$$

para algún tensor C^b_{ca} . Probemos primero que dado cualquier $p \in M$, $(\bar{\nabla}_c - \nabla_c) w_a|_p$ depende solo de $w_a|_p$ y no de su derivada. Sea w'_a cualquier otro co-vector tal que en p coinciden, es decir $(w_a - w'_a)|_p = 0$. Luego dada una co-base suave $\{\mu_a^i\}$ en un entorno de p tendremos que $w_a - w'_a = \sum_i f^i \mu_a^i$ con f^i funciones suaves que se anulan en p . En p tenemos entonces

$$\begin{aligned} \bar{\nabla}_c (w_a - w'_a) - \nabla_c (w_a - w'_a) &= \sum_i \bar{\nabla}_c (f^i \mu_a^i) - \sum_i \nabla_c (f^i \mu_a^i) \\ &= \sum_i \mu_a^i (\bar{\nabla}_c f^i - \nabla_c f^i) = 0 \end{aligned} \quad (3.27)$$

ya que por *iv*) $\bar{\nabla}_c$ y ∇_c deben actuar del mismo modo en escalares –y en particular en las f^i –. Esto muestra que

$$(\bar{\nabla}_c - \nabla_c) w'_a = (\bar{\nabla}_c - \nabla_c) w_a \quad (3.28)$$

y por lo tanto que $(\bar{\nabla}_c - \nabla_c) w_a$ depende de solo $w_a|_p$ y obviamente en forma lineal. Pero entonces $(\bar{\nabla}_c - \nabla_c)$ debe ser un tensor de tipo (1,2) que está esperando *comerse* una forma para darnos el tensor de tipo (0,2) $(\bar{\nabla}_c - \nabla_c) w_a$. Es decir $(\bar{\nabla}_c - \nabla_c) w_a = C^b_{ca} w_b$, lo que prueba el teorema.

Note que condición *v*) nos dice que $\nabla_a \nabla_b f = \nabla_b \nabla_a f$, tomando $w_a = \nabla_a f$ obtenemos

$$\begin{aligned} \bar{\nabla}_a \bar{\nabla}_b f &= \bar{\nabla}_a \nabla_b f \\ &= \bar{\nabla}_a w_b \\ &= \nabla_a w_b + C^c_{ab} \nabla_c f \\ &= \nabla_a \nabla_b f + C^c_{ab} \nabla_c f. \end{aligned} \quad (3.29)$$

Como $\nabla_c f|_p$ puede ser un co-vector cualquiera vemos que la condición de que no haya torsión implica que C^c_{ab} es simétrico en los índices inferiores, $C^c_{ab} = C^c_{ba}$.

Ejercicio: ¿Cómo actúa $(\bar{\nabla}_c - \nabla_c)$ en vectores?

Ejercicio: Expresa el corchete de Lie en términos de una conexión cualquiera y luego pruebe explícitamente que no depende de la conexión empleada.

Ejercicio: Sea $A_{b,\dots,z}$ un tensor totalmente antisimétrico. Muestre que $\nabla_{[a} A_{b,\dots,z]}$, es decir la antisimetrización total de $\nabla_a A_{b,\dots,z}$, no depende de la derivada covariante empleada.

La diferencia entre una conexión cualquiera ∇_c y una proveniente de un sistema de coordenadas $\{x^i\}$ es un tensor llamado **símbolo de Christoffel** de ∇_c con respecto a las coordenadas $\{x^i\}$, Γ^b_{ca} ,

$$\nabla_c w_a = \partial_c w_a + \Gamma^b_{ca} w_b. \quad (3.30)$$

El conocimiento de este tensor es muy útil en la práctica, pues nos permite expresar ∇_c en término de la conexión coordenada correspondiente, ∂_c .

Como hemos visto en una variedad M existen infinitas formas de *tomar la derivada de un tensor*. ¿Hay alguna natural o privilegiada? La respuesta es no, a menos que pongamos más estructura en M . Intuitivamente la razón de esto es que en M no sabemos comparar $l^a|_p$ con $l^a|_q$ si p y q son dos puntos distintos.⁴

¿Es suficiente la presencia de una métrica en M para poder realizar esta comparación? ¡La respuesta es sí!

Teorema 3.3 *Sea g_{ab} una métrica (suave) en M , luego existe una única derivada covariante ∇_c tal que $\nabla_c g_{ab} = 0$.*

Prueba: Sea $\bar{\nabla}_c$ una conexión cualquiera y sea ∇_c tal que $\nabla_c g_{ab} = 0$, es suficiente mostrar que esta condición determina unívocamente el tensor diferencia, C^d_{ca} . Pero,

$$0 = \nabla_a g_{bc} = \bar{\nabla}_a g_{bc} - C^d_{ab} g_{dc} - C^d_{ac} g_{bd} \quad (3.31)$$

o sea

$$C_{cab} + C_{bac} = \bar{\nabla}_a g_{bc} \quad (3.32)$$

pero también

$$\begin{aligned} a \leftrightarrow b & \quad C_{cba} + C_{abc} = \bar{\nabla}_b g_{ac} \\ c \leftrightarrow b & \quad C_{bca} + C_{acb} = \bar{\nabla}_c g_{ab}. \end{aligned} \quad (3.33)$$

⁴Note que una manera de comparar vectores infinitesimalmente cercanos, dado un campo vectorial m^a , es con el corchete de Lie de m^a con l^a , $[m, l]^a$. Esto no es lo apropiado ya que $[m, l]^a|_p$ depende de la derivada de m^a en p .

Sumando estos dos últimos, sustrayendo la primera y usando la simetría de los dos últimos índices obtenemos

$$2C_{abc} = \bar{\nabla}_b g_{ac} + \bar{\nabla}_c g_{ab} - \bar{\nabla}_a g_{bc} \quad (3.34)$$

o

$$C^a_{bc} = \frac{1}{2} g^{ad} \{ \bar{\nabla}_b g_{dc} + \bar{\nabla}_c g_{db} - \bar{\nabla}_d g_{bc} \} \quad (3.35)$$

Nótese que la existencia de una métrica no es equivalente a la existencia de una conexión. Hay conexiones $\bar{\nabla}_c$ para las cuales no existe ninguna métrica g_{ab} tal que $\bar{\nabla}_a g_{bc} = 0$, es decir hay tensores C^a_{bc} para los cuales no hay ninguna g_{ab} que satisfaga (3.35).

Ejercicio: Si $\bar{\nabla}_c$ es una derivada correspondiente a un sistema de coordenadas $\bar{\nabla}_c = \partial_c$ el símbolo de Christoffel correspondiente es

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad} \{ \partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc} \}. \quad (3.36)$$

Muestre que sus componentes en ese sistema de coordenadas están dadas por,

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} \{ \partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk} \}, \quad (3.37)$$

donde g_{ij} son las componentes de la métrica en el mismo sistema de coordenadas.

Ejercicio: La métrica Euclídea de \mathbb{R}^3 en coordenadas esféricas es

$$ds^2 = (dr)^2 + r^2((d\theta)^2 + \sin^2\theta(d\varphi)^2).$$

Calcule el Laplaciano $\Delta = g^{ab} \nabla_a \nabla_b$ en estas coordenadas.

*El tensor de Riemann

¿Dada una derivada covariante en una variedad, es posible definir campos tensoriales que solo dependan de ella y por lo tanto que nos den información sobre ésta?

¡La respuesta es sí! El siguiente tensor se denomina **tensor de Riemann o de Curvatura** y solo depende de la conexión:

$$2\nabla_{[a} \nabla_{b]} l^c := (\nabla_a \nabla_b - \nabla_b \nabla_a) l^c := R^c_{dab} l^d \quad \forall l^c \in TM.$$

Ejercicio: Muestre que la definición de arriba tiene sentido, es decir que el lado izquierdo, evaluado en un punto p cualquiera de M solo depende de $l^c|_p$ y por lo tanto podemos escribir el lado derecho para algún tensor R^c_{dab} .

Ejercicio: Sea $\bar{\nabla}_a$ otra derivada covariante. Calcule la diferencia entre los respectivos tensores de Riemann en términos del tensor que aparece como diferencia de las dos conexiones.

Notas bibliográficas: Recomiendo el libro de Wald [6], sobre todo por su notación moderna, ver también [21]. El lenguaje de la intuición es la geometría, es la herramienta que nos permite visualizar los problemas, tocarlos, darlos vuelta a nuestro gusto y luego reducirlos al álgebra. Entender la geometría es la forma más eficiente de entender la física ya que ésta solo se entiende cabalmente cuando es traducida a un lenguaje geométrico. No abuse de ella, es un área muy vasta y es fácil perderse, aprenda bien lo básico y luego solo lo que le sea relevante.

4.1. Introducción

En este capítulo comenzaremos el estudio de los sistemas de ecuaciones diferenciales en derivadas ordinarias, —EDO de ahora en más—. Estos sistemas aparecen constantemente en física y conjuntamente con los sistemas de ecuaciones diferenciales en derivadas parciales, —EDP de ahora en más— que estudiaremos en la segunda parte de este curso, forman el esqueleto matemático de las ciencias exactas. Esto se debe a que la mayoría de los fenómenos físicos se pueden describir en término de estas ecuaciones. Si el parámetro libre —o variable independiente— tiene la interpretación de ser un tiempo, entonces estamos frente a **ecuaciones de evolución**. Estas nos permiten, para cada estado inicial del sistema, conocer su evolución temporal, es decir nos dan la capacidad de hacer predicciones a partir de condiciones iniciales, lo cual es el aspecto característico de toda teoría física.¹

Existen otros fenómenos físicos descriptos por EDO que no tienen un carácter evolutivo. En general en estos casos estamos interesados en estados de equilibrio del sistema. Por ejemplo si queremos ver qué figura describe la cuerda de secar la ropa que tenemos en casa. Estos casos no serán incluidos en la teoría de EDO que desarrollaremos a continuación, sino que sus casos más sencillos serán tratados como casos especiales (unidimensionales) de las EDP elípticas.

Si bien la física cuántica, o más precisamente la teoría de campos cuánticos, nos muestra que una descripción de los fenómenos físicos a esa escala microscópica no es posible por medio de los sistemas de ecuaciones diferenciales ya mencionados, usualmente nos ingeniamos para encontrar aspectos parciales de estos fenómenos, o ciertas aproximaciones en donde sí es posible una descripción en término de estos sistemas. Esto no se debe a un capricho o empecinamiento, sino al hecho de que nuestro conocimiento de las EDO, y en menor medida de las EDP, es bastante profundo, lo cual nos permite manejarlas con relativa facilidad y comodidad. Por otro lado, y es importante remarcarlo, nuestro conocimiento del tipo de ecuaciones que aparecen en las teorías de campos cuánticos no-lineales es, en comparación, prácticamente nulo.

4.2. El Caso de una Única Ecuación Ordinaria de Primer Orden

4.2.1. Ecuación Autónoma de Primer Orden

Esta tiene la forma:

$$\dot{x} = \frac{dx}{dt} = f(x) \quad (4.1)$$

o sea una ecuación para un mapa $x(t) : I \subset \mathbb{R} \rightarrow U \subset \mathbb{R}$.

Supondremos $f : U \rightarrow \mathbb{R}$ continua. Una solución de esta ecuación será un mapa diferenciable $\varphi(t)$ definido para todo I y con imagen en U . Diremos que $\varphi(t)$ satisface la **condición inicial** x_0 en

¹En algunos casos se da que la variable independiente no es el tiempo sino algún otro parámetro, pero la ecuación también puede ser considerada como describiendo evolución. Por ejemplo si queremos ver como varía la densidad de un medio al aumentar la temperatura.

$t = t_0 \in I$ si

$$\varphi(t_0) = x_0. \quad (4.2)$$

Como veremos a continuación bajo ciertas hipótesis la condición inicial que una solución cumple la determina unívocamente.

La ecuación 4.1 se puede interpretar geoméricamente de la siguiente forma. En cada punto del plano (x, t) marcamos una rayita", es decir un campo de direcciones de tal forma que el ángulo que forma con el eje $x = \text{cte}$ tenga tangente igual a $f(x)$.

Ejemplo: $\dot{x} = x^{\frac{1}{2}}$. [Ver figura (4.1).]

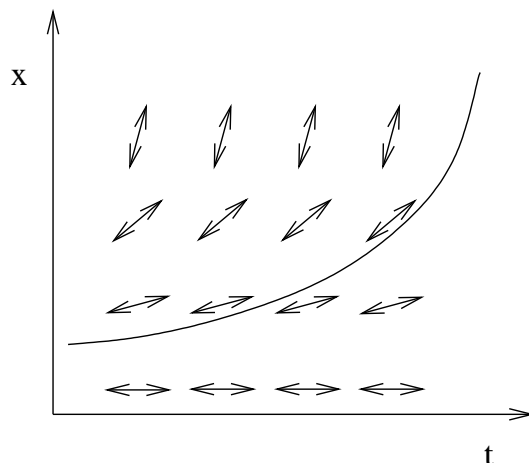


Figura 4.1: Interpretación geométrica de la ecuación $f(x) = x^{1/2}$.

Como $\varphi(t)$ tiene derivada $f(\varphi(t))$, en el punto $(\varphi(t), t)$, $\varphi(t)$ será tangente a la rayita en ese punto. Si nos paramos en algún punto (x_0, t_0) por donde ésta pase podremos encontrar la solución siguiendo las rayitas. Por lo tanto vemos que dado un punto por donde la solución pasa, siguiendo una rayita determinaremos una única solución. Esto en particular nos dice que en este gráfico la mayoría de las soluciones no se cortan. Como veremos más adelante la unicidad de la solución se da si la rayita en cada punto tiene una tangente no nula. Por otro lado, es claro del caso en la figura anterior, 4.1, que si comenzaremos con el punto $(x_0 = 0, t_0)$ tendríamos dos opciones, o bien dibujar la línea ya dibujada (t^2, t) o simplemente la línea $(0, t)$.

Ejercicio: Extender esta intuición gráfica al caso $\dot{x} = f(x, t)$.

Ejemplo: De hecho la ecuación $\dot{x} = x^{1/2}$ tiene dos soluciones que pasan por $(0, 0)$,

$$\varphi(t) = 0 \text{ y } \varphi = \frac{t^2}{4} \quad (4.3)$$

En efecto, supongamos $x(t) \neq 0$ entonces $\frac{dx}{x^{1/2}} = dt$ o $2(x^{1/2} - x_0^{1/2}) = t - t_0$, tomando $x_0 = t_0 = 0$

$$x = \left(\frac{t}{2}\right)^2 \quad (4.4)$$

La otra es trivialmente una solución.

La unicidad de las soluciones se obtiene, aun en el caso en que $f(x)$ se anula, si en vez de pedir que $f(x)$ sea continua se pide un poquito más, y eso es que, donde se anula, digamos en x_0 , se

cumpla que

$$\lim_{\epsilon \rightarrow 0} \int_{x_0+\epsilon}^x \frac{dx}{f(x)} = \infty, \quad (4.5)$$

o alternativamente que para $\epsilon > 0$ suficientemente pequeño, exista $k > 0$ tal que,

$$|f(x)| < k |x - x_0| \quad \text{si } |x - x_0| < \epsilon. \quad (4.6)$$

O sea que $f(x)$ vaya a cero cuando $x \rightarrow x_0$ a lo sumo tan rápidamente como una función lineal. Esta condición se llama condición de Lipschitz. De ahora en más supondremos que $f(x)$ es diferenciable y por lo tanto Lipschitz.

Ejercicio: Vea que esta condición es más débil que pedir que f sea diferenciable en x_0 .

Solución: La función $f(x) = x \sin(\frac{1}{x})$ es Lipschitz en x_{+0} pero no es diferenciable.

Así llegamos a nuestro primer teorema sobre ecuaciones ordinarias.

Teorema 4.1 Sea $f(x) : U \rightarrow \mathbb{R}$ continua y Lipschitz en los puntos donde se anula. Luego

i) Por cada $t_0 \in \mathbb{R}$, $x_0 \in U$ existe un intervalo $I_{t_0} \in \mathbb{R}$ y una solución $\varphi(t)$ de la ecuación 4.1 definida $\forall t \in I_{t_0}$ y la condición inicial $\varphi(t_0) = x_0$;

ii) Dos soluciones cualesquiera φ_1, φ_2 de 4.1 satisfaciendo la misma condición inicial coinciden en algún entorno de $t = t_0$

iii) la solución φ de 4.1 es tal que

$$t - t_0 = \int_{x_0}^{\varphi(t)} \frac{dx}{f(x)} \quad \text{si } f(x) \neq 0 \quad (4.7)$$

Nota: La solución cuya existencia asegura este teorema es válida en un intervalo I_{t_0} abierto, que el teorema no determina. Por lo tanto la unicidad se da solo en el máximo intervalo que esas tienen en común.

Prueba:

Si $f(x_0) = 0$ luego $\varphi(t) = x_0$ es una solución.

Si $f(x_0) \neq 0$ luego por continuidad existe un entorno de x_0 donde $f(x) \neq 0$ y por lo tanto $1/f(x)$ es una función continua, esto implica que la integral en 4.29 es una función diferenciable de sus extremos de integración, los que llamaremos $\psi(x) - \psi(x_0)$.

Por lo tanto tenemos,

$$t - t_0 = \psi(x) - \psi(x_0) \quad (4.8)$$

O sea dado x obtenemos t y además esta relación se cumple automáticamente que cuando $x = x_0$, $t = t_0$. Queremos ahora despejar x de esta relación, es decir encontrar $\phi(t)$.

Pero

$$\left. \frac{d\psi}{dx} \right|_{x=\zeta} = \frac{1}{f(\zeta)} \neq 0, \quad (4.9)$$

Por lo tanto el teorema de la función inversa nos asegura que existe un I_{t_0} y una única $\varphi(t)$ diferenciable, definida para todo $t \in I_{t_0}$ y tal que $\varphi(t_0) = x_0$ y $t - t_0 = \psi(\varphi(t)) - \psi(x_0)$. Formando su derivada obtenemos,

$$\frac{d\varphi}{dt} = \left. \frac{d\psi^{-1}}{dt} \right|_{x=\varphi(t)} = \left[\frac{1}{f(x)} \right]^{-1} \bigg|_{x=\varphi(t)} = f(\varphi(t)) \quad (4.10)$$

lo que demuestra que es una solución, necesariamente única de 4.1 con la condición inicial apropiada.

Solo resta ver que en el caso en que $f(x_0) = 0$ la solución es única. Para ello utilizaremos el siguiente Lema.

Lema 4.1 Sean f_1 y f_2 funciones reales en $U \subset \mathbb{R}$ tal que $f_1(x) < f_2(x) \quad \forall x \in U$ y sean φ_1 y φ_2 respectivas soluciones a las ecuaciones diferenciales que éstas generan, definidas en un intervalo $I \in \mathbb{R}$ y tal que $\varphi_1(t_1) = \varphi_2(t_1)$ para algún $t_1 \in I$. Luego

$$\varphi_1(t) \leq \varphi_2(t) \quad \forall t > t_1 \in I \quad (4.11)$$

Prueba: Trivial, el que corre más rápido en cada sector del camino gana! Sea el conjunto $S = \{t > t_1 \in I \mid \varphi_1(t) > \varphi_2(t)\}$. Sea $T \geq t_1$, $T \in I$ la mayor de las cotas inferiores de S . En este instante $\varphi_1(T) = \varphi_2(T)$, pero

$$\left. \frac{d\varphi_1}{dt} \right|_{t=T} < \left. \frac{d\varphi_2}{dt} \right|_{t=T} \quad (4.12)$$

Por lo tanto existe $\varepsilon > 0$ tal que para todo $t \in [T, \varepsilon)$, $\varphi_1(t) < \varphi_2(t)$ lo que es una contradicción, a menos que sea el último valor en I .

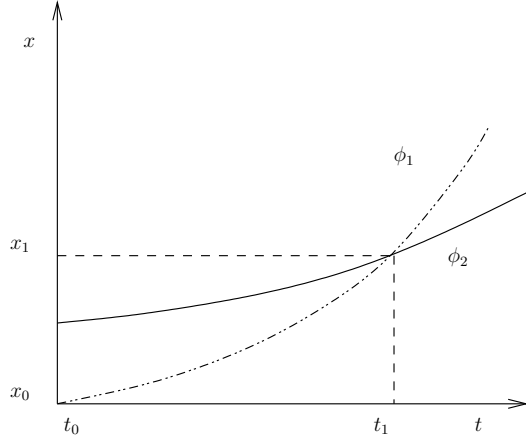


Figura 4.2: Prueba del Lema 4.1.

Con la ayuda de este Lema probaremos la unicidad de la solución. Sea $\varphi_1(t)$ una solución con $\varphi_1(t_0) = x_0$ pero distinta de la solución $\varphi(t) = x_0 \quad \forall t$. Supongamos entonces que existe $t_1 > t_0$ tal que $\varphi_1(t_1) = x_1 > x_0$, los otros casos se tratan similarmente. [Ver figura 4.3.]

Si elegimos x_1 bastante próximo a x_0 se cumple

$$f(x) < k |x - x_0| \quad \forall x_0 \leq x \leq x_1 \quad (4.13)$$

y por lo tanto que $\varphi_2(t) < \varphi_1(t)$, $\forall t_0 \leq t \leq t_1$, donde $\varphi_2(t)$ es la solución de la ecuación $\dot{x} = k(x - x_0)$ con condición inicial $\varphi_2(t_1) = x_1$. Pero $\varphi_2(t) = x_0 + (x_1 - x_0)e^{k(t-t_1)}$ que tiende a x_0 sólo cuando $t \rightarrow -\infty$. Por lo tanto $\varphi_1(t)$ no puede tomar el valor x_0 para ningún $t < t_1$ finito y por lo tanto tenemos una contradicción ♠

La solución cuya existencia y unicidad acabamos de demostrar no sólo se puede pensar como un mapa entre I_{t_0} y U sino también como un mapa g_x^t entre $U_{x_0} \times I_{t_0}$ y U (U_{x_0} entorno de x_0). El mapa g_x^t lleva (x, t) al punto $\varphi(t)$ donde φ es la solución de 4.1 con condición inicial $\varphi(0) = x$.

Estos mapas, o difeomorfismos locales, se denominan **flujos de fase locales**, (ya que en general no se pueden definir en todo $I_{t_0} \times U$) y satisfacen la siguiente importante propiedad de grupo; si $t, s \in I_0$ luego $g_x^{s+t} = g^s \circ g_x^t$.

Por la forma en que aparecen las condiciones iniciales en el Teorema anterior es fácil ver, usando nuevamente el teorema de la función inversa, que si $f(x_0) \neq 0$ luego g_x^t es diferenciable también con respecto a x .

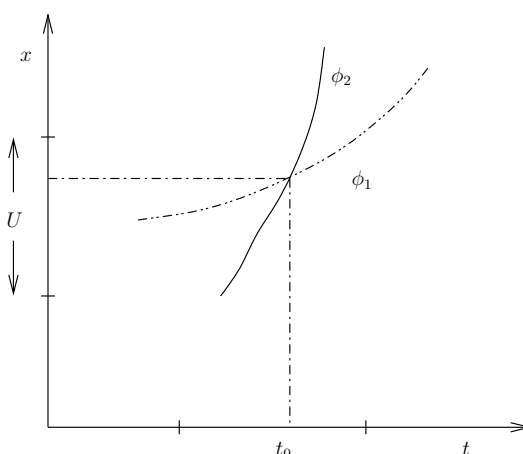


Figura 4.3: Prueba de la unicidad de la solución.

Corolario 4.1 si $f(x_0) \neq 0$ luego $g_x^t : U_{x_0} \times I_{t_0} \rightarrow U$ es diferenciable en ambos t y x .

Nota: La restricción $f(x_0) \neq 0$ es innecesaria como veremos más adelante.

Del uso del Teorema de la función inversa en el Teorema 4.1 también sigue que si consideramos el conjunto de las soluciones distinguiendo cada una por su condición inicial, $[\varphi(x_0, t_0, t)]$ es la solución que satisface $\varphi(t_0) = x_0$, luego la función $\varphi(x_0, t_0, t) : U \times I \times I \rightarrow U$ es diferenciable con respecto a todos los argumentos. Es decir que si cambiamos levemente las condiciones iniciales, luego la solución solo cambiará levemente. Más adelante daremos una prueba de todas las propiedades, incluyendo el Teorema 4.1 para los sistemas generales, o sea la ecuación 4.30.

Ejemplo: Sea $x(t)$ el número de bacterias en un vaso de cultivo al tiempo t . Si el número es lo suficientemente grande podemos pensarla como una función suave, diferenciable. Si a su vez este número no es tan grande de modo que su densidad en el vaso sea lo suficientemente baja como para que éstas no tengan que competir por aire y alimento, luego su reproducción es tal que $x(t)$ satisface la ecuación,

$$\dot{x} = ax \quad a = \text{cte.} \simeq 2,3 \frac{1}{\text{día}}, \quad (4.14)$$

o sea la velocidad de crecimiento es proporcional a la cantidad de bacterias presentes. La constante a es la diferencia entre la tasa de reproducción y la de extinción. Aplicando el Teorema 4.1 sabemos que la ecuación nos permite, dado el número de bacterias x_0 al tiempo t_0 , conocer el número de éstas para todo tiempo. En efecto la solución de 4.14 es

$$\frac{dx}{x} = a dt \quad (4.15)$$

Integrando ambas partes

$$\ln \frac{x}{x_0} = a(t - t_0) \quad \text{o; } x(t) = x_0 e^{a(t-t_0)} \quad (4.16)$$

O sea que el número de bacterias crece exponencialmente, a menos por supuesto de que $x_0 = 0$, y depende continuamente del número inicial. Se debe advertir que si bien la dependencia es continua la diferencia entre soluciones con condiciones iniciales distintas crece exponencialmente.

El flujo de fase es en este caso $g_x^t = x e^{at}$ y aquí es global. Con ese crecimiento es fácil darse cuenta que en pocos días el número de bacterias y por lo tanto su densidad será tan grande que

éstas entrarán a competir entre sí por el alimento lo cual hará disminuir la velocidad de crecimiento. Se observa experimentalmente que este hecho se puede tomar en cuenta incluyendo en la 4.14 un término proporcional al cuadrado del número de bacterias.

$$\dot{x} = ax - bx^2 \quad b = \frac{a}{375}. \quad (4.17)$$

Si x es pequeña el primer término domina y x comienza a crecer, como ya vimos, en forma exponencial hasta que el segundo término comienza a ser importante y la velocidad de crecimiento decrece tendiendo a cero cuando $x = x_s$ tal que $ax_s - bx_s^2 = 0$ o $a - bx_s = 0$ o $x_s = a/b \simeq 375$ bacterias. Como $\varphi(t) = x_s$ es una solución, debido a la unicidad, la solución que tiende a x_s que describimos arriba no puede alcanzar el valor x_s en un tiempo finito. Si el número de bacterias presentes es mayor que 375 la velocidad de crecimiento será negativa y la solución nuevamente tenderá asintóticamente a x_s .

La solución general de 4.17 puede ser obtenida de forma análoga a la de 4.14 y es,

$$x(t) = \frac{a/b}{1 + (\frac{a}{bx_0} - 1)e^{-a(t-t_0)}} \quad (4.18)$$

lo cual confirma el análisis anterior.

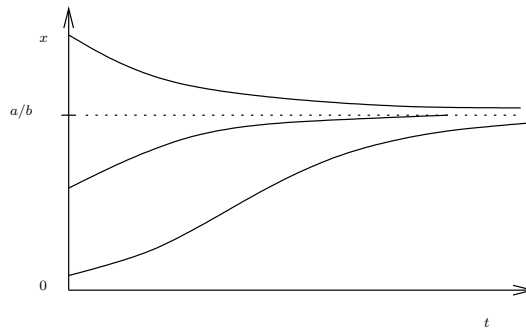


Figura 4.4: Distintas soluciones de la ecuación de crecimiento bacteriano: $\varphi_1(t_0) = 0$, $\varphi_2(t_0) = x_s = \frac{a}{b}$, $\varphi_3(t) = x_0 < x_s$ y $\varphi_4(t) = x_0 > x_s$

Este ejemplo nos permite introducir dos conceptos claves en el área. El primero es el de **solución estacionaria o de equilibrio**. Estas son las soluciones constantes o sea son tales que $\dot{x}_s = f(x_s) = 0$ o sea las raíces de la función $f(x)$. Para la ecuación 4.14 $x_s = 0$, para la ecuación 4.17 $x_s = 0$ y a/b .

Note que no siempre existen soluciones estacionarias, por ejemplo la ecuación con $f(x) = 1$ no las tiene, pero cuando existen el problema de encontrarlas se reduce a resolver una ecuación a lo sumo trascendente. Debido a la unicidad de las soluciones, esto nos permite dividir el plano (x, t) en franjas tales que si una solución tiene condiciones iniciales en esta franja ésta debe permanecer en ella.

El segundo concepto es la **estabilidad** de estas soluciones. La solución $x_s = 0$ de 4.14 y 4.17 no es estable en el sentido que si elegimos como condición inicial un punto en cualquier entorno de esta solución, la solución correspondiente se apartará inicialmente en forma exponencial de la anterior. Por lo contrario la solución de 4.17 $x_s = a/b$ es estable en el sentido que si tomamos valores iniciales próximos a ésta las soluciones correspondientes se aproximarán asintóticamente (en el futuro) a la anterior.

Las soluciones inestables no tienen interés físico ya que la más mínima perturbación del sistema originará una solución que se alejará rápidamente de la solución inestable. Esto no es así en el ejemplo anterior pues el número de bacterias es discreto y la ecuación solo representa una aproximación. Si

el vaso de cultivo está esterilizado y herméticamente cerrado luego no hay posibilidad de crecimiento bacteriano. Más adelante analizaremos en detalle el problema de la estabilidad.

4.2.2. Extendiendo la Solución Local

Si nos olvidamos del modelo físico que describimos con 4.17 y tomamos $z_o < 0$, $t_o = 0$, por ejemplo, habrá un tiempo máximo, $t_d = \frac{1}{a} \ln[\frac{a/b-z_o}{-z_o}]$, finito para el cual la solución diverge, es decir,

$$\lim_{t \rightarrow t_d} z(t) = \infty. \quad (4.19)$$

Esto muestra que el teorema anterior solo puede asegurar la existencia de un intervalo I para el cual la solución existe. Lo máximo que podemos concluir en forma general, es decir sin que se dé más información sobre f , es el siguiente Corolario del Teorema 4.1.

Corolario 4.2 Sea $U \subset \mathbb{R}$ el dominio de definición de f , U_c un intervalo compacto [es decir cerrado y acotado] de U , $z_o \in U_c$. Luego la solución $(\varphi(z_o, t_o, t), I)$ de la ecuación 4.1 puede ser extendida o bien indefinidamente o bien hasta la frontera de U_c . Esta extensión es única en el sentido que cualquier par de soluciones, $(\varphi_1(z_o, t_o, t), I_1)$, $(\varphi_2(z_o, t_o, t), I_2)$ coinciden en la intersección de sus intervalos de definición, $I = I_1 \cap I_2$.

Prueba: Primero probamos la versión global de la unicidad de la solución. Sea T la menor de las cotas superiores del conjunto $\{\tau \mid \varphi_1(z_o, t) = \varphi_2(z_o, t) \forall t_o \leq t \leq \tau\}$. Supongamos en contradicción que T es un punto interior de $I_1 \cap I_2$. Por continuidad tenemos, $\varphi_1(z_o, T) = \varphi_2(z_o, T)$, pero por el resultado del Teorema 4.1, usando condiciones iniciales $\varphi_1(z_o, T)$, en T vemos que ambas soluciones deben coincidir en un entorno de T , lo que contradice que T sea el máximo. Así es que T debe ser un extremo de alguno de los intervalos de definición por lo que las soluciones coinciden en $I_1 \cap I_2 \cap \{t \mid t \geq t_o\}$. El caso $t \leq t_o$ se trata similarmente y por lo tanto las soluciones coinciden en todo $I_1 \cap I_2$.

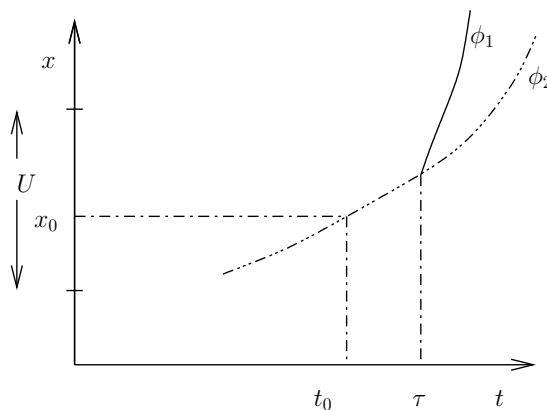


Figura 4.5: Unicidad global.

Segundo construimos la extensión. La idea es que si dos soluciones coinciden en $I_1 \cap I_2$ luego las podemos combinar y formar una en $I_1 \cup I_2$,

$$\varphi(t) = \begin{cases} \varphi_1(t) & \forall t \in I_1 \\ \varphi_2(t) & \forall t \in I_2. \end{cases} \quad (4.20)$$

Sea T la cota superior mínima de $\{\tau \mid \varphi(t) \text{ existe y está contenida en } U_c \forall t_o \leq t \leq \tau\}$. Si $T = \infty$ no hay nada que probar, por lo tanto supongamos $T < \infty$. Probaremos que existe $\varphi(t)$

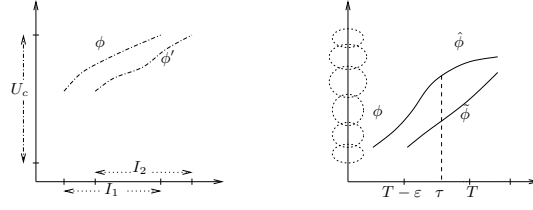


Figura 4.6: Extendiendo la solución.

definida para todo $t_o \leq t \leq T$ y tal que $\varphi(T)$ es uno de los extremos de U_c . El Corolario 4.1 nos dice que dado cualquier $\hat{z} \in U$, existen $\epsilon_{\hat{z}} > 0$ y un entorno de \hat{z} , $V_{\hat{z}} \subset U$ tal que para cualquier $z \in V_{\hat{z}}$ existe una solución con condiciones iniciales $\varphi(t_o) = z$ definida para todo t en $|t - t_o| < \epsilon_{\hat{z}}$. Como U_c es compacto podemos elegir un número finito de puntos \hat{z}_i tal que $U_c \subset \bigcup V_{\hat{z}_i}$. Sea $\epsilon > 0$ el mínimo de los $\epsilon_{\hat{z}_i}$. Como T es la cota superior mínima existe τ entre $T - \epsilon$ y T tal que $\varphi(t) \in U_c \forall t_o \leq t \leq \tau$, pero como $\varphi(\tau) \in U_c$ está en alguno de los $V_{\hat{z}_i}$ existe una solución $\hat{\varphi}(t)$ definida para todo t en $|t - \tau| < \epsilon$ con condición inicial $\hat{\varphi}(\tau) = \varphi(\tau)$. Usando $\hat{\varphi}(t)$ podemos ahora extender φ , cuya extensión llamaremos $\tilde{\varphi}$, al intervalo $t_o \leq t < \tau + \epsilon$. Pero $\tilde{\varphi}(\theta) \in U_c \forall t_o < \theta < T$ ya que $\tilde{\varphi}(\theta) = \varphi(\theta) \in U_c$, y de otro modo T no sería una cota superior mínima. Por continuidad $\tilde{\varphi}(T) \in U_c$ y como por definición de T para todo intervalo abierto, I_T , conteniendo T , $\tilde{\varphi}[I_T] \notin U_c$ concluimos que $\tilde{\varphi}(T)$ es un extremo de U_c ♠

4.2.3. El Caso No-autónomo

Este es el caso en que en vez de 4.1 tenemos

$$\frac{dz}{dt} = f(z, t). \quad (4.21)$$

La interpretación geométrica es la misma, sólo que ahora las *rayitas* tienen un ángulo que depende también de la variable t , por lo tanto esperamos que si comenzamos en un punto (t_o, z_o) y trazamos una curva tangente a las rayitas obtendremos también una única solución. Este es en realidad el caso, pero su discusión estará comprendida en la teoría general de sistemas autónomos de primer orden que enunciaremos más adelante.

Suponiendo ciertas condiciones para $f(z, t)$ –por ejemplo que sea Lipschitz con respecto a ambas variables– se puede ver que también existen flujos locales, pero éstos ahora dependen también de t_o y no solo de la diferencia entre el tiempo inicial y el final. Denotaremos a estos flujos como, $g_{t_o}^t(z)$ y la propiedad de semigrupo que cumplen es $g_{t_1}^t \circ g_{t_o}^{t_1}(z) = g_{t_o}^t(z)$.

Hay una clase especial de ecuaciones no-autónomas que se puede reducir al ya estudiado. Esta es la clase donde $f(z, t)$ se puede escribir como el cociente de dos funciones, una de z y otra de t ,

$$f(z, t) = \frac{g(z)}{h(t)}. \quad (4.22)$$

En este caso la interpretación geométrica es que el parámetro t no es el más indicado para describir este sistema. Esto se remedia eligiendo otro parámetro τ dado por, $\frac{d\tau}{dt} = \frac{1}{h(t)}$, o

$$\tau - \tau_o = \int_{t_o}^t \frac{dt}{h(t)}. \quad (4.23)$$

Una vez que se obtiene $t(\tau)$ [Ver el Teorema 4.1] debemos resolver,

$$\frac{dz}{d\tau} = g(z(\tau)), \quad (4.24)$$

[Usando nuevamente el Teorema 4.1], obteniendo $\varphi(\tau)$. La solución con respecto al parámetro original se obtiene definiendo $\varphi(t) = \varphi(\tau(t))$.

4.3. Reducción a Sistemas de Primer Orden

Definición: La **ecuación diferencial ordinaria general de orden m** es una ecuación de la forma:

$$F(x, x^{(1)}, \dots, x^{(m)}, t) = 0 \quad (4.25)$$

Es decir una ecuación implícita de x , un mapa entre $\tilde{I} \subset \mathbb{R}$ y $\tilde{U} \subset \mathbb{R}^{m \times n}$, (o sea x es un vector de n componentes), y sus derivadas (donde $x^{(i)} \equiv \frac{d^i x}{dt^i}$) denota la i -ésima derivada con respecto al parámetro independiente, t) hasta el orden n^2 .

Definición: Diremos que el mapa m -veces diferenciable $x(t) : I \rightarrow \mathbb{R}^n$ es **una solución** de la ecuación anterior si F evaluada en este mapa está bien definida y es idénticamente nula en todo el intervalo I .

En este curso supondremos que 4.25 puede ser resuelta para $x^{(m)}$, o sea que 4.25 es equivalente a la siguiente ecuación:

$$x^{(n)} = f(x, x^{(1)}, \dots, x^{(n-1)}, t) \quad (4.26)$$

en abiertos $I \subset \tilde{I} \subset \mathbb{R}$ y $U \subset \tilde{U} \subset \mathbb{R}^m$

Ejemplo: La EDO $F(x, x^{(1)}, t) = ((x^{(1)})^2 - 1) = 0$ implica una de las siguientes EDO en nuestro sentido:

$$\begin{aligned} x^{(1)} &= 1 \\ \text{o} \\ x^{(1)} &= -1 \end{aligned} \quad (4.27)$$

Si conocemos los valores de x y sus derivadas hasta el $(m-1)$ -ésimo orden en un punto t_0 entonces la ecuación 4.26 nos permite conocer la m -ésima derivada en dicho punto. Si suponemos que f es continua entonces esto nos permite conocer x y sus derivadas hasta el $(m-1)$ -ésimo orden en un punto $t_0 + \varepsilon$ suficientemente cercano (el error en que incurriremos será del orden de $\varepsilon \times$ derivadas de f), pero entonces podemos usar nuevamente la ecuación 4.26 y así conocer aproximadamente la n -ésima derivada de x en $t_0 + \varepsilon$. Prosiguiendo de este modo podemos lograr una *solución aproximada* en un intervalo dado. Al menos intuitivamente, parecería que haciendo el intervalo ε cada vez más pequeño deberíamos obtener una solución. ¡Esto en realidad es así! Y lo probaremos más adelante en el curso. Lo importante por el momento es el hecho de que para obtener una solución debemos dar en un punto t_0 el valor de todas las derivadas de la variable incógnita de orden menor de la que determina el sistema, es decir de la que aparece en el lado izquierdo de la ecuación 4.26.

Introduciendo $\tilde{u}^1 \equiv x$, $\tilde{u}^2 \equiv x^{(1)}$, \dots , $\tilde{u}^n \equiv x^{(m-1)}$, podemos escribir 4.26 de la forma:

$$\begin{aligned} \dot{\tilde{u}}^1 &= \tilde{u}^2 \\ \dot{\tilde{u}}^2 &= \tilde{u}^3 \\ &\vdots \\ \dot{\tilde{u}}^m &= f(\tilde{u}^1, \tilde{u}^2, \dots, \tilde{u}^m, t) \end{aligned} \quad (4.28)$$

donde hemos denotado con un punto sobre la función su derivada con respecto al parámetro t . Por lo tanto 4.26 es equivalente a una ecuación de primer orden para el vector de vectores $\tilde{\mathbf{u}} \subset U \subset \mathbb{R}^{n \times m}$,

$$\dot{\tilde{\mathbf{u}}} = \tilde{\mathbf{V}}(\tilde{\mathbf{u}}, t) \quad (4.29)$$

²En realidad para definir correctamente 4.25 se deben dar los abiertos \tilde{U}^i para la i -ésima derivada con respecto a x donde F está definida.

donde $\tilde{\mathbf{V}}$ es también un vector en algún subconjunto de $\mathbb{R}^{n \times m}$. Si $\tilde{\mathbf{V}}(\tilde{u}, t)$ depende explícitamente de t entonces podemos agregarle a $\tilde{\mathbf{u}}$ otra componente, la $(n \times m + 1)$ -ésima con $\tilde{u}^{n \times m + 1} = t$, y por lo tanto tenemos que $\dot{\tilde{u}}^{n \times m + 1} = 1$. La ecuación 4.29, más esta última ecuación, toman la forma ,

$$\dot{\mathbf{u}} = \mathbf{V}(\mathbf{u}) , \quad (4.30)$$

donde $\mathbf{u} = (\tilde{\mathbf{u}}, \tilde{u}^{n \times m + 1})$ y

$$V^i(\mathbf{u}) = \begin{cases} \tilde{V}^i(\tilde{\mathbf{u}}, \tilde{u}^{n \times m + 1}) & \text{si } 1 \leq i \leq m \times n \\ 1 & \text{si } i = n \times m + 1 \end{cases} , \quad (4.31)$$

un mapa entre $U \subset \mathbb{R}^{n \times m + 1}$ y $\mathbb{R}^{n \times m + 1}$.

Así es que arribamos al siguiente teorema.

Teorema 4.2 (de Reducción) *Si la función f del sistema 4.26 de m ecuaciones diferenciales ordinarias de orden n es diferenciable, luego este sistema es equivalente a un sistema de $m \times n + 1$ ecuaciones diferenciales ordinarias de primer orden que no dependen explícitamente de la variable independiente. Es decir por cada solución del sistema 4.26 podemos construir una solución del sistema 4.30 y viceversa.*

Prueba: Claramente dada una solución del sistema 4.26 suficientemente diferenciable podemos escribir con ella un vector \mathbf{u} y éste satisfará (porque así lo construimos) el correspondiente sistema 4.30. Por lo tanto toda solución del sistema original es solución del sistema de primer orden asociado a él.

Para probar el inverso, es decir que toda solución $\mathbf{u}(t)$ del sistema 4.30 da origen a una del sistema 4.26, solo es necesario tomar repetidas derivadas de las primeras componentes, $x(t) = u^1(t)$, e ir usando las ecuaciones en el correspondiente orden hasta obtener el sistema original satisfecho por la n -ésima derivada de $x(t)$ ♠

Este teorema es importante pues nos permite abarcar toda la teoría general de las ODE a partir del estudio de la ecuación 4.30, la cual como veremos tiene un sentido geométrico claro. Sin embargo se debe tener en cuenta que muchas veces en física aparecen subclases de ecuaciones muy especiales con propiedades particulares de mucha importancia física que la teoría general no contempla.

Ejemplo: Péndulo Matemático

$$\ddot{x} = kx \quad u^1 = x, \quad u^2 = \dot{x}$$

$$\frac{d}{dt} \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ k & 0 \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \quad (4.32)$$

Note que si $u = \gamma(t) : I \rightarrow U$ es una solución de 4.30 luego $\forall s \in \mathbb{R}$ tal que $t + s \in I$, $\gamma(t + s)$ es también una solución. En efecto,

$$\left. \frac{d}{dt} \gamma(t + s) \right|_{t=t_o} = \left. \frac{d}{dt} \gamma(t) \right|_{t=t_o+s} = V(\gamma(t_o + s)) = V(\gamma(t + s))|_{t=t_o} . \quad (4.33)$$

Debido a esto de ahora en más tomaremos t_o , el punto donde daremos la condición inicial, igual a cero.

Ejercicio: Si $u = \sin(t) : \mathbb{R} \rightarrow [0, 1]$, es una solución de 4.30 pruebe que $u = \cos(t)$ también lo es.

4.4. Sistemas de EDO

¿Tienen los sistemas de EDO una interpretación geométrica? Como ya vimos estos tienen la forma genérica.

$$\frac{dx^i}{dt} = v^i(x^j) \quad i = 1, \dots, n. \quad (4.34)$$

Sea $\gamma(t)$ una curva entre un intervalo $I \in \mathbb{R}$ y M una variedad de n dimensiones. Sea $p = \gamma(t_0)$ un punto de M y $(U, \varphi = (x^1, \dots, x^n))$ una carta con $p \in U$. Luego $\varphi \circ \gamma(t) = (x^1(t), \dots, x^n(t))$ es un mapa entre I y un abierto de \mathbb{R}^n , y $\frac{dx^i}{dt}$ son las componentes del vector tangente a γ en la base coordenada $\{x_i\}$ en el punto $\gamma(t)$ de M . Por lo tanto $v^i(x^j)$ son las componentes de un campo vectorial \mathbf{v} en la base coordenada $\{x_i\}$ en el punto $p \in M$ correspondiente a $\varphi(p) = (x^1, \dots, x^n)$. Vemos así que los sistemas de EDO se pueden interpretar como un campo vectorial \mathbf{v} en una variedad M , y sus soluciones como curvas $\gamma(t)$ tal que su tangente en todo punto sea \mathbf{v} .

Ejemplo: El péndulo Físico

$$\begin{cases} \ddot{\theta} &= -\sin \theta & \text{ó} \\ \dot{\theta} &= z \\ \dot{z} &= -\sin \theta \end{cases} \quad (4.35)$$

Aquí el campo vectorial es $z \frac{\partial}{\partial \theta} - \sin \theta \frac{\partial}{\partial z}$,

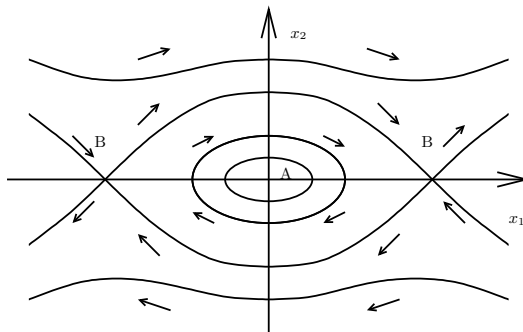


Figura 4.7: El péndulo Físico.

note que el diagrama en el plano se repite cada 2π según el eje θ . En realidad la ecuación tiene sentido físico en un cilindro, siendo θ la variable angular, o sea la variedad es un cilindro. Las **curvas integrales**, o sea las imágenes de los mapas $\gamma(t)$ que son soluciones, se construyen siguiendo los vectores de tal modo que éstas sean siempre tangentes a las curvas.

Del ejemplo queda claro que en general si damos un punto de M siguiendo los vectores, a partir de éste, podremos encontrar una solución $\gamma(t)$ que pasa por este punto. Si tenemos dos soluciones $\gamma_1(t)$, $\gamma_2(t)$ éstas no se pueden cruzar –sin ser tangentes– en los puntos donde el campo vectorial es distinto de cero [ya que de otro modo sus tangentes darían dos vectores distintos en el punto]. Si se supone que el campo vectorial es diferenciable luego se puede ver que distintas curvas nunca se pueden tocar y por lo tanto las soluciones son únicas.

Por supuesto, como se ve por ejemplo en la figura 4.8, una misma curva si se puede intersectar (note que esto solo puede suceder si el sistema es autónomo, ya que de lo contrario, la inclusión de la variable temporal hace que ninguna curva se pueda intersectar). En tal caso se puede probar que esa solución es **periódica**, es decir que existe $T > 0$ tal que $\gamma(t + T) = \gamma(t)$. Note en particular que esto implica que γ está definida para todo t en \mathbb{R} .

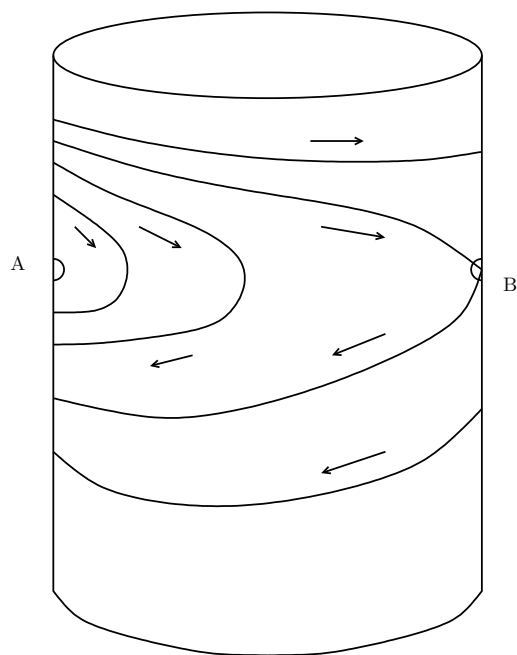


Figura 4.8: La variedad Péndulo Físico.

Problema 4.1 Si M es una variedad compacta (ejemplo: una esfera o un toro) se puede probar usando Corolario 4.2 que todas sus curvas de fase están definidas para todo t . Sin embargo existen algunas que no son cerradas, dé un ejemplo.

4.4.1. Integrales Primeras

Dado un campo vectorial \mathbf{v} en M , si existe $f : M \rightarrow \mathbb{R}$ no constante tal que $\mathbf{v}(f) = 0$ en M luego diremos que f es una **integral primera o constante de movimiento** del sistema de EDO generado por \mathbf{v} .

Se puede demostrar que si n es la dimensión de M , luego localmente, alrededor de puntos donde $\mathbf{v} \neq 0$, existen $n-1$ integrales primeras funcionalmente independientes sin embargo es fácil encontrar campos vectoriales que no tienen ninguna integral primera.

Ejemplos-Problemas

Problema 4.2 El sistema

$$\begin{cases} \dot{x}_1 &= x_1 \\ \dot{x}_2 &= x_2 \end{cases}$$

no tienen ninguna, ¿por qué? Ayuda: Dibuje las curvas integrales.

Problema 4.3 Considere el sistema

$$\begin{cases} \dot{x}_1 &= k_1 \\ \dot{x}_2 &= k_2 \end{cases}$$

en el toro (obtenido identificando los puntos (x_1, x_2) y $(x_1 + n, x_2 + m)$ de \mathbb{R}). ¿Para cuáles valores de k_1, k_2 existe una integral primera y para cuáles no? Ayuda: ídem 1).

Problema 4.4 Las ecuaciones de Hamilton de la mecánica clásica forman un sistema de EDO,

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad (4.36)$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (4.37)$$

$i = 1, \dots, n$, donde $H : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ es la función Hamiltoniano o energía. Demuestre que H es una integral primera de este sistema.

Los ejemplos anteriores muestran que hay sistemas de EDO que globalmente no son derivables de un principio variacional.

Así como el conocimiento de la energía de un sistema Hamiltoniano es una ayuda inestimable para conocer la forma cuantitativa de su movimiento, el conocimiento de una integral primera de un sistema general de EDO también brinda una gran ayuda ya que sabemos que las curvas de fase están restringidas a sus superficies de nivel ($f = \text{cte.}$) lo que nos permite reducir efectivamente el orden de la ecuación en uno: solo tenemos que considerar el problema en la variedad dada por los puntos $f = \text{cte.}$

4.4.2. Teorema Fundamental de los Sistemas de EDO

Teorema 4.3 Sea \mathbf{v} un campo vectorial continuamente diferenciable (C^1) en un entorno de $p \in M$. Luego,

i) Existe un entorno U_p de p y un abierto $I \subset \mathbb{R}$, tal que dado cualquier $q \in U_p$ ($t_0 \in I$) existe una curva diferenciable $\gamma_q(t) : I \rightarrow M$ satisfaciendo

$$\begin{aligned} \left. \frac{d\gamma_q}{dt} \right|_t &= \mathbf{v}|_{\gamma_q(t)} \\ \gamma_q(t_0) &= q. \end{aligned} \quad (4.38)$$

ii) si γ_q^1 y γ_q^2 satisfacen la condición i) luego $\gamma_q^1 = \gamma_q^2$ en $I^1 \cap I^2$.

iii) $\gamma_q(t)$ es también C^1 cuando es considerada como una función de $I \times U_p \rightarrow M$. O sea es continuamente diferenciable con respecto a la condición inicial.

iv) Sea U una región compacta (acotada) de M , $\mathbf{v} \in C^1$ y $p \in U$ luego $\gamma_p(t)$ puede ser extendida (hacia adelante y/o hacia atrás) hasta la frontera de U . En particular si M es compacta y $\mathbf{v} \in C^1(M)$, luego $\gamma_p(t)$ existe globalmente (para todo $t \in \mathbb{R}$).

Probaremos este teorema más adelante, luego de introducir las herramientas matemáticas necesarias.

4.4.3. Dependencia en Parámetros, Ecuación de Variaciones

Una consecuencia, muy importante en las aplicaciones, del teorema fundamental es el siguiente:

Corolario 4.3 Sea \mathbf{v}_λ un campo vectorial diferenciable en $U \subset M$ que depende diferenciablemente de un parámetro $\lambda \in A \subset \mathbb{R}$. Luego dados $p \in U$, $t_0 \in I$ y $\lambda_0 \in A$ existen entornos U_p , I_{t_0} y A_{λ_0} tales que para toda terna $(q, t, \lambda) \in U_p \times I_{t_0} \times A_{\lambda_0}$ existe una única curva integral de \mathbf{v}_λ , $\gamma_\lambda(t) : I_{t_0} \times A_{\lambda_0} \rightarrow U_p$ con $\gamma_\lambda(0) = q$. Esta depende diferenciablemente de q , t y λ .

Prueba: Considere el campo vectorial $(\mathbf{v}_\lambda, 0) : U \times A \rightarrow TM \times \mathbb{R}$. Por hipótesis este campo es diferenciable y por lo tanto tiene curvas integrales $(\gamma_\lambda, \lambda)$ que dependen diferenciablemente de las condiciones iniciales y por lo tanto de λ ♠

Advertencia: El intervalo $I_{t_0} \times A_{\lambda_0}$ donde la solución es diferenciable no necesariamente es el intervalo de definición de la solución. Por ejemplo no es cierto que aun cuando la solución esté definida para todo t en \mathbb{R} los límites $\lim_{T \rightarrow \infty} \gamma_\lambda(t)$ y $\lim_{\lambda \rightarrow \lambda_0} \gamma_\lambda(t)$ conmuten.

La importancia práctica de este corolario es que nos permite obtener soluciones aproximadas a las ya conocidas. En efecto supongamos que para $\lambda = 0$ conocemos alguna curva integral de $\mathbf{v}|_{\lambda=0}$, $\gamma_0(t)$ con $\gamma_0(0) = p$. Tenemos entonces, aplicando el Corolario y un desarrollo en serie de Taylor, que

$$\gamma_\lambda(t) = \gamma_0(t) + \lambda \gamma_1(t) + O(\lambda^2) \quad (4.39)$$

o sea que $\gamma_0(t) + \lambda \gamma_1(t)$ es una buena aproximación a $\gamma_\lambda(t)$ para λ suficientemente pequeño.

Resta entonces encontrar la ecuación que satisface $\gamma_1(t)$. Esta se obtiene diferenciando con respecto a λ la ecuación diferencial en $\lambda = 0$,

$$\frac{d}{d\lambda} \left(\frac{d\gamma_\lambda(t)}{dt} = \mathbf{v}_\lambda(\gamma_\lambda(t)) \right) \Big|_{\lambda=0}, \quad (4.40)$$

expandiendo y usando coordenadas

$$\begin{aligned} \frac{d}{dt} \gamma_1^j &= \frac{\partial v^j}{\partial x^i} \Big|_{(\gamma_0, \lambda=0)} \cdot \gamma_1^i + \frac{dv^j}{d\lambda} \Big|_{(\gamma_0, \lambda=0)}, \\ &= A^j_i(t) \gamma_1^i + b^j(t). \end{aligned} \quad (4.41)$$

o sea la ecuación para $\gamma_1(t)$ –usualmente llamada ecuación de variaciones– es una ecuación líneal inhomogénea no-autónoma. Como tomamos como condición inicial $\gamma_\lambda(0) = p \quad \forall \lambda$, la condición inicial que consideraremos para 4.4.3 es $\gamma_1(0) = 0$.

Problema 4.5 ¿Cuál sería la condición inicial para γ_1 si la condición para γ_λ fuese $\gamma_\lambda(0) = p(1 - \lambda) + \lambda q$? Considere solo el caso unidimensional pero piense en el genérico.

Problema 4.6 Si decidiésemos considerar una aproximación hasta segundo orden $\gamma_\lambda = \gamma_0 + \lambda \gamma_1 + \lambda^2 \gamma_2 + O(\lambda^3)$, ¿qué ecuaciones obtendríamos para γ_2 ?

Ejemplo: a) Un cuerpo cae verticalmente en un medio de baja viscosidad, la cual depende de la posición y velocidad de éste.

$$\ddot{x} = -g + \varepsilon F(x, \dot{x}) \quad \varepsilon \ll 1 \quad (4.42)$$

Calcule el efecto de esta resistencia en el movimiento.

La solución, $x(t)$, dependerá suavemente de ε por lo tanto podemos desarrollar esta en serie de Taylor con respecto a ε .

$$x(t) = x_0(t) + \varepsilon x_1(t) + O(\varepsilon^2). \quad (4.43)$$

La solución para $\varepsilon = 0$ es claramente

$$x_0(t) = x_i + v_i t - g \frac{t^2}{2}, \quad (4.44)$$

donde x_i y v_i son respectivamente la posición y la velocidad inicial, mientras que la ecuación de variaciones es

$$\ddot{x}_1 = F(x_0, \dot{x}_0), \quad x_1(0) = 0, \quad \dot{x}_1(0) = 0. \quad (4.45)$$

Integrando obtenemos

$$x_1(t) = \int_0^t \int_0^\tau F(x_0(\tilde{\tau}), \dot{x}_0(\tilde{\tau})) d\tilde{\tau} d\tau. \quad (4.46)$$

b) El segundo ejemplo es el de auto-oscilaciones.

Considere el sistema,

$$\begin{aligned}\dot{x}_1 &= x_2 + \varepsilon f_1(x_1, x_2) \\ \dot{x}_2 &= -x_1 + \varepsilon f_2(x_1, x_2), \quad \text{para } \varepsilon \ll 1.\end{aligned}\tag{4.47}$$

Sin pérdida de generalidad podemos asumir $f_i(0, 0) = 0$. Cuando $\varepsilon = 0$ obtenemos las ecuaciones del péndulo –en la aproximación de pequeñas amplitudes–, o sea un sistema conservativo. Su energía –integral primera– está dada por $E(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2^2)$ y sus curvas de fase son entonces círculos de radio $\sqrt{2E}$.

Cuando $\varepsilon > 0$ las curvas de fase no son necesariamente cerradas, pero debido al Corolario anterior a lo más pueden ser espirales con una distancia del orden de ε entre vuelta y vuelta. Estas pueden tender hacia algún punto estacionario dentro del círculo $\sqrt{2E_i}$, donde E_i = energía inicial, es decir donde $x_2 + \varepsilon f_1(x_1, x_2) = -x_1 + \varepsilon f_2(x_1, x_2) = 0$, o pueden aproximarse asintóticamente a alguna órbita cerrada o finalmente pueden diverger. Que estas son las únicas opciones es el resultado de un teorema clásico de **Poincaré–Bendixson**, el cual muestra que en general estas son las únicas posibilidades en dos dimensiones.

Para estudiar cuál es el caso entre estos tres consideramos la variación de energía entre dos vueltas consecutivas.

Usando que,

$$\dot{E}(x_1, x_2) = \varepsilon(x_1 f_1 + x_2 f_2)\tag{4.48}$$

obtenemos

$$\begin{aligned}\Delta E &= \varepsilon \oint (-f_1 dx_2 + f_2 dx_1) + O(\varepsilon^2) \\ &= \varepsilon F(E).\end{aligned}\tag{4.49}$$

donde la integral es a lo largo de un círculo de radio $\sqrt{2E}$ en la dirección del movimiento. Es decir, hemos aproximado la curva con $\varepsilon > 0$ por la curva con $\varepsilon = 0$.

Si $F(E)$ es positiva la energía aumenta y el sistema realiza oscilaciones crecientes.

Si $F(E)$ es negativa las oscilaciones disminuyen en amplitud y el sistema tiende eventualmente a un punto de equilibrio.

Si $F(E)$ cambia de signo, digamos en $E = E_0$, se puede probar entonces que cerca de este círculo hay una curva de fase Γ_ε cerrada.

Si $F'(E)|_{E=E_0}$ es negativa Γ es un ciclo límite estable –cualquier curva de fase cercana se aproxima asintóticamente a ésta–. Si $F'(E)|_{E=E_0}$ es positiva Γ es inestable.

Problema 4.7 Muestre que si hay dos ciclos límites estables luego debe haber también al menos uno que es inestable.

De los dos ejemplos anteriores queda claro la importancia práctica del Corolario. La dependencia en forma diferenciable de las condiciones iniciales también conduce a la ecuación de variaciones, por lo tanto esta ecuación, que es lineal, inhomogénea, es de suma importancia y en el próximo capítulo la estudiaremos en más detalle.

4.5. Problemas

Problema 4.8 (Kiselióv) Encuentre todas las soluciones de

$$\frac{dx}{dt} = \frac{3}{2}x^{2/3}.\tag{4.50}$$

Ayuda: hay infinitas y éstas se obtienen a partir de segmentos de algunas soluciones particulares. Grafique algunas de estas soluciones.

Problema 4.9 (Kiselióv) Aplique el teorema de existencia y unicidad para determinar las regiones donde las ecuaciones siguientes tienen solución única:

- a) $\dot{x} = x + 3x^{1/3}$.
- b) $\dot{x} = 1 - \cot x$.
- c) $\dot{x} = \sqrt{1 - x^2}$.

Problema 4.10 (Kiseliiov) Resuelva las siguientes ecuaciones, en todos los casos dé las soluciones generales en función de un dato inicial x_0 para el tiempo inicial t_0 .

- a) $t\dot{x} + x = \cos(t)$. (Use integral primera de la parte homogénea.)
- b) $\dot{x} + 2x = e^t$. (Use variación de constantes.)
- c) $(1-t^2)\dot{x} + tx = 2t$. (Resuelva primero la homogénea y luego sume una inhomogénea particular.)
- d) $x\dot{x} = t - 2t^3$.

Problema 4.11 (Kiseliiov) Resuelva las siguientes ecuaciones apelando a un cambio de variable en la variable independiente (t).

- a) $\dot{x}x = -t$. ($t = \cos(s)$.)
- b) $\dot{x}t - x = 0$.
- c) $\dot{x} + e^{\dot{x}} = t$.

Problema 4.12 (Kiseliiov) Grafique las isoclinas (líneas de igual pendiente en el plano (t, x)) y luego trace las soluciones de las siguientes ecuaciones:

- a) $\dot{x} = 2t - x$.
- b) $\dot{x} = \sin(x + t)$.
- c) $\dot{x} = x - t^2 + 2t - 2$.
- d) $\dot{x} = \frac{x-t}{x+t}$.

Problema 4.13 Resuelva la ecuación:

$$\dot{x} = A(t)x + B(t)x^n \text{ Ayuda: Use el cambio de variable } y = x^{1-n}.$$

Problema 4.14 Resuelva la ecuación

$$\frac{dx}{dt} = i\lambda x + Ae^{i\omega t} \quad \lambda, \omega \in \mathbb{R} \quad (4.51)$$

y vea cómo se comporta su parte real. Examine los casos:

- a) $A = 0$, b) ($A \neq 0$, $\lambda \neq \omega$) y c) ($A \neq 0$, $\lambda = \omega$).

Problema 4.15 La ecuación del péndulo físico.

- a) Grafique el campo vectorial de la ecuación

$$\frac{d^2\theta}{dt^2} = -\sin(\theta), \quad -\pi \leq \theta \leq \pi. \quad (4.52)$$

- b) Grafique algunas curvas integrales alrededor de $(\theta = 0, z = \frac{d\theta}{dt} = 0)$.
- c) Grafique las curvas integrales que pasan por el punto $\theta = \pm\pi, z = 0$. Infiera que el tiempo necesario para llegar por estas curvas a dicho punto es infinito.
- d) Grafique el campo vectorial a lo largo de una recta $z = z_0$. Infiera de ello que las soluciones no pueden superar nunca la velocidad adquirida cuando pasan por el punto $\theta = 0$. Ayuda: Concluya esto primero para la región $0 \leq \theta \leq \pi$, $0 \leq z \leq z_0$ y luego use la simetría de la solución.
- e) Sea $E(z, \theta) := \frac{z^2}{2} - \cos(\theta)$. Vea que $\frac{dE}{dt} = 0$. Use esta cantidad para analizar el comportamiento de las soluciones con dato inicial $(z_0, \theta_0 = 0)$. En particular determine cuáles cruzan el eje $z = 0$ y cuáles no.

Problema 4.16 Sea la ecuación:

$$\frac{dz}{dt} = \begin{cases} f(z) & |z| < 1 \\ iz & |z| \geq 1 \end{cases} \quad (4.53)$$

Donde $f(z)$ es continua con iz en $|z| = 1$. Infiera que no importando cómo sea la forma de $f(z)$, las soluciones nunca escapan de la región $\{|z(t)| \leq \max\{|z(0)|, 1\}\}$.

Problema 4.17 Sea un cuerpo afectado por una fuerza central, es decir,

$$m \frac{d^2 \vec{x}}{dt^2} = f(r) \vec{x} \quad r := |\vec{x}|. \quad (4.54)$$

- a) Encuentre un sistema equivalente de primer orden.
 b) Compruebe que dado un vector cualquiera (constante) \vec{c} , luego $F(\vec{x}, \vec{p}) := \vec{c} \cdot (\vec{x} \wedge \vec{p})$, donde $\vec{p} := \frac{d\vec{x}}{dt}$, es una integral primera.

Problema 4.18 Sea la ecuación:

$$\frac{d\vec{x}}{dt} = \vec{x} \wedge \vec{\omega}(\vec{x}). \quad (4.55)$$

- a) Vea que $R := \vec{x} \cdot \vec{x}$ es una integral primera.
 b) Concluya que las soluciones de este sistema existen para todo tiempo.

Problema 4.19 (Strichartz) Muestre que

$$J_k(t) = \sum_{j=0}^{\infty} \frac{(-1)^j (t/2)^{k+2j}}{j!(k+j)!}, \quad (4.56)$$

tiene radio de convergencia infinito y satisface la ecuación de Bessel, $x''(t) + (1/t)x'(t) + (1 - k^2/t^2)x(t) = 0$.

Problema 4.20 Resuelva el sistema:

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 + \varepsilon(x_1^2 + x_2^2) \\ x_1 \end{pmatrix} \quad (4.57)$$

Para los datos iniciales $(x_1(0) = 1, x_2(0) = 0)$. Investigue ahora la solución cerca de $\varepsilon = 0$ obteniendo los coeficientes en la serie de Taylor de la misma con respecto al parámetro ε . Encuentre qué ecuaciones satisfacen estos coeficientes.

Problema 4.21 (Arnold) Considere la ecuación del péndulo físico: $\frac{d^2 x}{dt^2} = -\sin(x)$. Vea cómo varía la frecuencia en función de la amplitud para soluciones pequeñas. Ayuda: suponga una solución en serie de Taylor en la amplitud y resuelva hasta encontrar la primera contribución no trivial a la solución linealizada. Encuentre el período ubicando los puntos de retorno (velocidad cero).

Notas bibliográficas: Para éste y los próximos tres capítulos recomiendo los siguientes libros: [3], [7] y [8]. Sobre todo el primero. Las ecuaciones diferenciales ordinarias son la base de la mecánica clásica, aunque en esta última hay además una estructura geométrica particular que es fundamental. Casi todo lo que hemos visto en este capítulo es la teoría *local*, la teoría *global*, es decir el comportamiento de las soluciones para grandes valores de la variable independiente (el tiempo en la mayoría de los casos) solo se ha comprendido en los últimos años, dando origen a conceptos nuevos como el de caos, atractores, dimensiones fractales, etc. Lamentablemente estos aspectos no han sido todavía suficientemente sintetizados y por lo tanto no puedo recomendar ningún libro al respecto. Hay demasiadas ideas interesantes pero cada una de ellas requiere de un gran bagaje de información.

5.1. Sistema lineal homogéneo

Teorema 5.1 Sea $A(t)$ continua en $I \subset \mathbb{R}$, cerrado. Luego existe una única solución $\mathbf{x}(t) : I \rightarrow V$ de la ecuación,

$$\frac{d\mathbf{x}(t)}{dt} = A(t) \mathbf{x}(t), \quad (5.1)$$

con condición inicial $\mathbf{x}(t_0) = \mathbf{x}_0 \in V$, $t_0 \in I$.

Prueba: Sean $\mathbf{x}, \mathbf{y} \in V$, luego

$$\|A(t)\mathbf{x} - A(t)\mathbf{y}\|_V \leq \|A(t)\|_{\mathcal{L}} \|\mathbf{x} - \mathbf{y}\|_V \quad (5.2)$$

y por lo tanto $A(t)\mathbf{x}$ es Lipschitz en V . El teorema fundamental nos garantiza existencia y unicidad local. El teorema quedará entonces demostrado si mostramos que $\mathbf{x}(t)$ no puede hacerse infinita en un tiempo finito. Para ello usaremos,

$$\begin{aligned} \frac{d}{dt} \|\mathbf{x}\|_V &= \lim_{t_1 \rightarrow t} \frac{\|\mathbf{x}(t_1)\|_V - \|\mathbf{x}(t)\|_V}{t_1 - t} \leq \lim_{t_1 \rightarrow t} \frac{\|\mathbf{x}(t_1) - \mathbf{x}(t)\|_V}{t_1 - t} \\ &= \left\| \frac{d\mathbf{x}}{dt} \right\|_V = \|A\mathbf{x}\|_V \leq \|A\|_{\mathcal{L}} \|\mathbf{x}\|_V. \end{aligned} \quad (5.3)$$

Integrando esta desigualdad entre t_0 y $t \in I$ obtenemos,

$$\|\mathbf{x}(t)\|_V \leq \|\mathbf{x}(t_0)\|_V e^{\int_{t_0}^t \|A(\bar{t})\|_{\mathcal{L}} d\bar{t}}, \quad (5.4)$$

lo que muestra que $\mathbf{x}(t)$ no puede escapar a infinito en un tiempo finito y por lo tanto completa la prueba.

Consideremos el conjunto de las soluciones de 5.1 definidas en un único intervalo $I \subset \mathbb{R}$ cerrado, $Sol(A, I)$. Este conjunto tiene la estructura de un espacio vectorial. En efecto, si $\mathbf{x}(t)$ e $\mathbf{y}(t)$ son dos soluciones de 5.1, luego $\mathbf{x}(t) + \alpha\mathbf{y}(t)$, $\alpha \in \mathbb{R}$, es también una solución. ¿Cuál es su dimensión? El siguiente teorema responde a esta pregunta y muestra que cualquier solución de 5.1 puede ser expresada como combinación lineal de un número finito ($n = \dim V$) de soluciones.

Teorema 5.2 $\dim Sol(A, I) = \dim V$.

Prueba: Sea $\{\mathbf{u}_i^0\}$, $i = 1, \dots, n$ una base de V , $t_0 \in I$ y $\{\mathbf{u}_i(t)\}$, $i = 1, \dots, n$, $t \in I$ el conjunto de soluciones de 5.1 con condición inicial $\mathbf{u}_i(t_0) = \mathbf{u}_i^0$. Mostraremos que las $\{\mathbf{u}_i(t)\}$ forman una base de $Sol(A, I)$, y por lo tanto $\dim Sol(A, I) = \dim V$. Primero veamos que las soluciones $\{\mathbf{u}_i(t)\}$ son linealmente independientes.

Supongamos que existen constantes c^i tales que $\mathbf{x}(t) = \sum_{i=0}^n c^i \mathbf{u}_i(t)$ se anula para algún \tilde{t} en I . Como $\mathbf{x}(t)$ es una solución de 5.1 y estas son únicas cuando un dato inicial es especificado, tomando en este caso $\mathbf{x}(\tilde{t}) = 0$, vemos que $\mathbf{x}(t) = 0 \quad \forall t \in I$. En particular tenemos que $\mathbf{x}(t_0) = \sum_{i=0}^n c^i \mathbf{u}_i^0 = 0$ y la independencia del conjunto $\{\mathbf{u}_i^0(t)\}$ implica entonces que $c^i = 0 \quad \forall i = 1, \dots, n$, probando la independencia lineal. Note que no solo hemos probado que los $\{\mathbf{u}_i(t)\}$ son linealmente independientes como elementos de $Sol(\mathbf{A}, I)$ —para lo cual solo hubiésemos tenido que probar que si $\sum_{i=0}^n c^i \mathbf{u}_i(t) = 0 \quad \forall t \in I$ luego $c^i = 0 \quad \forall i = 1, \dots, n$ lo que es trivial ya que $t_0 \in I$ — sino también que para cada $t \in I$ los $\{\mathbf{u}_i(t)\}$ son linealmente independientes como elementos de \mathbf{V} , este resultado será usado más adelante.

Para completar el teorema veamos ahora que cualquier solución de 5.1 $\mathbf{x}(t) : I \rightarrow \mathbf{V}$, es decir cualquier elemento de $Sol(\mathbf{A}, I)$ puede ser escrito como combinación lineal de los $\{\mathbf{u}_i(t)\}$. Sea $\mathbf{x}(t) \in Sol(\mathbf{A}, I)$, como los $\{\mathbf{u}_i^0\}$ forman una base de \mathbf{V} existirán constantes c^i , $i = 1, \dots, n$ tales que $\mathbf{x}(t_0) = c^i \mathbf{u}_i^0$. Consideremos entonces la solución de 5.1 dada por $\varphi(t) = c^i \mathbf{u}_i(t)$. Como $\varphi(t_0) = \mathbf{x}(t_0)$ y las soluciones son únicas tenemos,

$$\mathbf{x}(t) = \varphi(t) = c^i \mathbf{u}_i(t) \quad \forall t \in I \quad (5.5)$$

El teorema anterior nos dice que si conocemos n soluciones linealmente independientes de \mathbf{A} , $\{\mathbf{u}_i(t)\}$, conocemos todas sus soluciones, ya que éstas serán combinaciones lineales de las $\{\mathbf{u}_i(t)\}$. La dependencia con los datos iniciales es también lineal, es decir si $\mathbf{x}(t), \mathbf{y}(t) \in Sol(\mathbf{A}, I)$ y $\mathbf{x}(t_0) = \mathbf{x}_0$ e $\mathbf{y}(t_0) = \mathbf{y}_0$ luego $\mathbf{x}(t) + \alpha \mathbf{y}(t)$, $\alpha \in \mathbb{R}$, es la solución con dato inicial $\mathbf{x}_0 + \alpha \mathbf{y}_0$ y por lo tanto el mapa $g_{t_0}^t$ —que en este caso llamaremos $\mathbf{X}_{t_0}^t$ — que lleva datos iniciales dados en $t = t_0$ a soluciones con esos datos, al tiempo t es un operador lineal de \mathbf{V} en \mathbf{V} . En efecto si $\{\theta_i^0\}$ es la *co*-base de la base $\{\mathbf{u}_i\}$, *i.e.* $\theta_0^j(\mathbf{u}_i^0) = \delta_i^j$. Luego $\mathbf{X}_{t_0}^t = \sum_{i=1}^n \mathbf{u}_i(t) \theta_i^0$. Debido a la independencia lineal de los $\{\mathbf{u}_i(t)\}$ como elementos de \mathbf{V} se puede ver que el mapa $\mathbf{X}_{t_0}^t : \mathbf{V} \rightarrow \mathbf{V}$ es inyectivo y por lo tanto invertible, para cada $t \in I$. Su inversa, que denotaremos $(\mathbf{X}_{t_0}^t)^{-1}$ lleva la solución en t a su dato inicial en $t = t_0$, o sea $(\mathbf{X}_{t_0}^t)^{-1} = \mathbf{X}_t^{t_0}$.

Ejercicio: Pruebe que $(\mathbf{X}_{t_0}^t)^{-1} = \mathbf{X}_t^{t_0}$.

Ejercicio: Pruebe que $\mathbf{X}_{t_0}^t$ no depende de la base empleada.

El mapa $\mathbf{X}_{t_0}^t$ es en realidad la familia mono-paramétrica de difeomorfismos de \mathbf{V} en \mathbf{V} generados por el campo vectorial $\mathbf{A}\mathbf{x}$. En este caso estos son lineales. En efecto $\mathbf{X}_{t_0}^t \mathbf{x}_0$ es la curva que pasa por el punto $\mathbf{x}_0 \in \mathbf{V}$ en $t = t_0$ y cuyo vector tangente es $\mathbf{A} \mathbf{X}_{t_0}^t \mathbf{x}_0$.

Hemos visto entonces que si conocemos un conjunto de n soluciones linealmente independientes $\{\mathbf{u}_i(t)\}$ podemos construir cualquier solución usando el operador $\mathbf{X}_{t_0}^t$ aplicado al dato inicial de nuestro gusto.

¿Cómo sabemos en la práctica que un conjunto de n soluciones $\{\mathbf{u}_i(t)\}$ es linealmente independiente? Como ya vimos es suficiente ver que éstas son linealmente independientes, como elementos de \mathbf{V} , a un tiempo t cualquiera. Es decir que el escalar

$$w(t) = \varepsilon(\mathbf{u}_1(t), \mathbf{u}_2(t), \dots, \mathbf{u}_n(t)) \quad (5.6)$$

es distinto de cero. Esta función se llama el **Wronskiano** del sistema y satisface una ecuación particularmente simple cuya solución llamada **fórmula de Liouville**, es

$$w(t) = w(t_0) e^{\int_{t_0}^t \text{tr}(\mathbf{A}(\tilde{t})) d\tilde{t}} \quad (5.7)$$

Prueba:

$$\begin{aligned}\dot{w}(t) &= \varepsilon(\dot{\mathbf{u}}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) + \varepsilon(\mathbf{u}_1, \dot{\mathbf{u}}_2, \dots, \mathbf{u}_n) + \dots + \varepsilon(\mathbf{u}_1, \mathbf{u}_2, \dots, \dot{\mathbf{u}}_n) \\ &= \varepsilon(\mathbf{A}\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) + \varepsilon(\mathbf{u}_1, \mathbf{A}\mathbf{u}_2, \dots, \mathbf{u}_n) + \dots + \varepsilon(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{A}\mathbf{u}_n) \\ &= \operatorname{tr}(\mathbf{A}) w(t),\end{aligned}\tag{5.8}$$

cuya solución es 5.7.

La fórmula de Liouville es una demostración independiente del resultado que las $\{\mathbf{u}_i(t)\}$ forman una base de \mathbf{V} para cada $t \in I$. Note que si $\operatorname{tr}(\mathbf{A}) \equiv 0$ el Wronskiano es constante y nos puede ser de utilidad para determinar una solución en función de otras ya conocidas.

5.2. Sistema Lineal Inhomogeneo – Variación de constantes

Aquí trataremos el sistema,

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t),\tag{5.9}$$

donde $\mathbf{b}(t) : I \rightarrow \mathbf{V}$ es integrable. Veremos que si conocemos el operador $\mathbf{X}_{t_0}^t$ correspondiente al sistema homogéneo

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x},\tag{5.10}$$

Luego conocemos todas las soluciones del sistema 5.9. El método que usaremos se llama de **variación de constantes** y también es de utilidad para obtener soluciones aproximadas a sistemas no lineales. El método consiste en suponer que la solución de 5.9 será de la forma,

$$\varphi(t) = \mathbf{X}_{t_0}^t \mathbf{c}(t)\tag{5.11}$$

para algún mapa $\mathbf{c}(t) : I \rightarrow \mathbf{V}$, diferenciable. Note que si $\mathbf{c}(t) = \mathbf{c} \in \mathbf{V}$, es decir un mapa constante, luego $\varphi(t)$ satisface 5.10. Sustituyendo $\varphi(t)$ en 5.9 obtenemos,

$$\begin{aligned}\frac{d}{dt}\varphi(t) &= \left(\frac{d}{dt}\mathbf{X}_{t_0}^t\right)\mathbf{c}(t) + \mathbf{X}_{t_0}^t \frac{d}{dt}\mathbf{c}(t) \\ &= \mathbf{A}(t)\mathbf{X}_{t_0}^t \mathbf{c}(t) + \mathbf{X}_{t_0}^t \frac{d}{dt}\mathbf{c}(t) = \mathbf{A}(t)\mathbf{X}_{t_0}^t \mathbf{c}(t) + \mathbf{b}(t).\end{aligned}\tag{5.12}$$

De donde vemos que para que $\varphi(t)$ sea una solución $\mathbf{c}(t)$ debe satisfacer la ecuación $\mathbf{X}_{t_0}^t \frac{d}{dt}\mathbf{c}(t) = \mathbf{b}(t)$.

Pero como $\mathbf{X}_{t_0}^t$ es invertible y su inversa es $\mathbf{X}_t^{t_0}$ obtenemos

$$\frac{d}{dt}\mathbf{c}(t) = \mathbf{X}_t^{t_0} \mathbf{b}(t).\tag{5.13}$$

Integrando obtenemos

$$\mathbf{c}(t) = \int_{t_0}^t \mathbf{X}_{\tilde{t}}^{t_0} \mathbf{b}(\tilde{t}) d\tilde{t} + \mathbf{c}(t_0)\tag{5.14}$$

o

$$\varphi(t) = \mathbf{X}_{t_0}^t \left[\int_{t_0}^t \mathbf{X}_{\tilde{t}}^{t_0} \mathbf{b}(\tilde{t}) d\tilde{t} \right] + \mathbf{X}_{t_0}^t \varphi(t_0),\tag{5.15}$$

donde $\varphi(t_0)$ es una condición inicial cualquiera.

Debido al teorema de existencia de soluciones del sistema homogéneo sabemos que $\mathbf{X}_t^{t_0}$ existe y es diferenciable $\forall t \in I$ y por lo tanto $\varphi(t)$ también existe y es diferenciable $\forall t \in I$, ya que $\mathbf{b}(\tilde{t})$ es integrable. Como el dato inicial para $\varphi(t)$ puede darse en forma arbitraria, y las soluciones de 5.9 son únicas, concluimos que 5.15 es la solución general del sistema 5.9.

5.3. Sistemas lineales homogéneos: coeficientes constantes

La ecuación que trataremos aquí es

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}, \quad (5.16)$$

donde $\mathbf{x} : I \subset \mathbb{R} \rightarrow \mathbf{V}$ es una curva en \mathbf{V} y \mathbf{A} es un operador lineal de \mathbf{V} , $\mathbf{A} : \mathbf{V} \rightarrow \mathbf{V}$.

Ya vimos en un ejercicio que $\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0$, $\mathbf{x}_0 \in \mathbf{V}$ cualquiera, es una solución de 5.16 con condición inicial $\mathbf{x}(0) = \mathbf{x}_0$. Note que como $e^{\mathbf{A}t}$ está definido para todo $t \in \mathbb{R}$, las soluciones que hemos encontrado también están definidas en todo \mathbb{R} .

Sea $\tilde{\mathbf{x}}(t)$ una solución cualquiera de 5.16 definida en un intervalo $\tilde{I} \subset \mathbb{R}$ y sea $t_1 \in \tilde{I}$. Luego la curva $\mathbf{x}_0(t) = e^{\mathbf{A}(t_1-t)} \tilde{\mathbf{x}}(t_1)$ es también una solución de 5.16 y $\mathbf{x}(t_1) = \tilde{\mathbf{x}}(t_1)$, pero por la unicidad de las soluciones de 5.16, $\mathbf{x}(t)$ y $\tilde{\mathbf{x}}(t)$ deben coincidir en todo \tilde{I} y $\mathbf{x}(t)$ es la extensión máxima de $\tilde{\mathbf{x}}(t)$. Vemos así que por medio de la exponencial $e^{\mathbf{A}t}$ obtenemos todas las soluciones de 5.16. De hecho, hemos demostrado que la familia monoparamétrica de difeomorfismos lineales $g^t = e^{\mathbf{A}t}$ está definida globalmente [$\forall t \in \mathbb{R}, \forall \mathbf{x}_0 \in \mathbf{V}$] y es la generadora del campo vectorial $v(\mathbf{x}) = \mathbf{A}\mathbf{x}$. [Con la definición dada en 5.1 en este caso $\mathbf{X}_{t_0}^t = g^{t-t_0} = e^{\mathbf{A}(t-t_0)}$.]

¿Cuál es la forma general de esta familia? O dicho de otro modo, ¿cómo es la forma funcional del mapa exponencial? Para ello es conveniente tomar una base fija $\{\mathbf{u}_i\}$ (independiente de t) en \mathbf{V} y expresar la ecuación en término de la n -tupla de componentes $\{x^i\}$ de \mathbf{x} . De esta forma obtenemos,

$$\frac{dx^i}{dt} = \sum_{j=1}^n A^i_j x^j, \quad (5.17)$$

es decir una ecuación para la n -tupla $\{x^i\}$. La ventaja de este cambio es que podemos tomar una base conveniente, en particular la base ortogonal del Lema 2.2 (de Schur) en la cual la matriz A^i_j es triangular superior. El precio pagado por la simplificación es que, como la base de Schur es en general compleja, tanto la n -tupla $\{x^i\}$ como las componentes de \mathbf{A} , A^i_j , son complejas y por lo tanto hemos pasado a un problema en C^n , o en forma abstracta en la complexificación de \mathbf{V} .

Para esa base tenemos que la componente n -ésima satisface una ecuación particularmente simple!,

$$\frac{dx^n}{dt} = A^n_n x^n \quad (5.18)$$

y por lo tanto, $x^n(t) = e^{\lambda_n t} x_0^n$ (recuerde que la diagonal de una matriz triangular superior está compuesta por sus autovalores), donde x_0^n es la n -ésima componente de la condición inicial a $t = 0$. La ecuación para la componente $n-1$ es,

$$\frac{dx^{n-1}}{dt} = A^{n-1}_{n-1} x^{n-1} + A^{n-1}_n x^n \quad (5.19)$$

$$= \lambda_{n-1} x^{n-1} + A^{n-1}_n e^{\lambda_n t} x_0^n. \quad (5.20)$$

Usando la fórmula (5.15) para el caso unidimensional, (y \mathbf{A} constante) obtenemos,

$$x^{n-1}(t) = e^{\lambda_{n-1} t} x_0^{n-1} + e^{\lambda_{n-1} t} \int_0^t e^{-\lambda_{n-1} s} (A^{n-1}_n e^{\lambda_n s} x_0^n) ds \quad (5.21)$$

$$= e^{\lambda_{n-1} t} x_0^{n-1} + e^{\lambda_{n-1} t} \int_0^t e^{(\lambda_n - \lambda_{n-1}) s} ds A^{n-1}_n x_0^n, \quad (5.22)$$

y vemos que si $\lambda_n - \lambda_{n-1} \neq 0$ luego la solución es una suma de exponenciales, $x^{n-1}(t) = e^{\lambda_{n-1} t} x_0^{n-1} + \frac{A^{n-1}_n x_0^n}{\lambda_n - \lambda_{n-1}} (e^{\lambda_n t} - e^{\lambda_{n-1} t})$. Si los autovalores coinciden ($\lambda_n - \lambda_{n-1} = 0$) luego aparece un término

lineal en t , $x^{n-1}(t) = e^{\lambda_{n-1}t}(x_0^{n-1} + tA^{n-1}_n x_0^n)$. Conociendo $x^{n-1}(t)$ podemos ahora calcular, usando nuevamente (5.15), $x^{n-2}(t)$ y así sucesivamente todas las componentes de $\mathbf{x}(t)$ en esta base. El resultado final será que las componentes son todas sumas de exponenciales por polinomios. En efecto note que la integral de un polinomio de grado m por una exponencial es también un polinomio del mismo grado por la misma exponencial más una constante de integración (a menos que la exponencial sea nula, en cuyo caso el resultado es un polinomio de grado $m+1$).

Ejercicio: Pruebe esta última afirmación por inducción en el grado del polinomio y el uso de integración por partes, es decir, use que $e^{at}P_m(t) = \frac{d}{dt}(\frac{e^{at}}{a}P_m(t)) - \frac{e^{at}}{a}\frac{d}{dt}P_m(t)$ y que $\frac{d}{dt}P_m(t)$ es un polinomio de orden $m-1$.

Reordenando términos vemos así que la solución general será de la forma:

$$\mathbf{x}(t) = \sum_{i=1}^m e^{\lambda_i t} \mathbf{v}_i(t) \quad (5.23)$$

donde la suma es sobre los distintos autovalores y donde los vectores $\mathbf{v}_i(t)$ son polinomios en t .

Una propiedad importante de la estructura algebraica de esta solución queda de manifiesto con el siguiente lema.

Lema 5.1 *Los espacios vectoriales*

$$E_\lambda := \{f(t) : \mathbb{R} \mapsto V^C \mid f(t) = e^{\lambda t} P_m(t), \text{ con } P_m(t) \text{ un polinomio}\}$$

son linealmente independientes.

Prueba: Debemos probar entonces que si tenemos una suma finita de elementos de los E_{λ_i} , $\{\mathbf{u}_i\}$ y esta se anula, luego cada uno de estos vectores se debe anular idénticamente. Es decir

$$\sum_{i=1}^m \mathbf{u}_i = 0, \quad \mathbf{u}_i \in E_{\lambda_i} \quad \lambda_i \neq \lambda_j \Rightarrow \mathbf{u}_i = 0. \quad (5.24)$$

Lo probaremos por inducción en el número de elementos de la suma. El caso $m=1$ es trivial. Supondremos entonces que la afirmación es verdadera para $m-1$ y la probaremos para m . Supondremos además por contradicción que existe una suma de m elementos distintos de cero que se anula, es decir:

$$\sum_{i=1}^m \mathbf{u}_i = 0, \quad \mathbf{u}_i \in E_{\lambda_i}, \quad \lambda_i \neq \lambda_j, \quad \mathbf{u}_i \neq 0. \quad (5.25)$$

dividiendo por $e^{\lambda_1 t}$ obtenemos

$$0 = S(t) := e^{-\lambda_1 t} \sum_{i=1}^m \mathbf{u}_i = P_{m_1}(t) + \sum_{i=2}^m e^{(\lambda_i - \lambda_1)t} P_{m_i}(t). \quad (5.26)$$

Tomando m_1+1 derivadas de $S(t)$, donde m_1 es el orden del polinomio del primer término obtenemos:

$$0 = \frac{d^{m_1+1}}{dt^{m_1+1}} S(t) = \sum_{i=2}^m e^{(\lambda_i - \lambda_1)t} \tilde{P}_{m_i}(t). \quad (5.27)$$

donde es fácil ver que los polinomios $\tilde{P}_{m_i}(t)$ son del mismo grado que los originales. Estamos así bajo la hipótesis inductiva y por lo tanto como estos polinomios son no-nulos tenemos una contradicción



Veamos ahora que en realidad cada uno de los términos en esta suma es una solución de la ecuación 5.16. Dado $\lambda \in C$ cualquiera consideremos el espacio vectorial de funciones de $\mathbb{R} \rightarrow V^C$ de la forma $e^{\lambda t} \mathbf{w}(t)$, donde $\mathbf{w}(t)$ es un polinomio ¹ en t . Este espacio es invariante bajo la acción de $\frac{d}{dt}$, ya que tomando su derivada obtenemos una expresión similar, es decir el producto de la misma exponencial por otro polinomio. Por otro lado es invariante ante la acción de \mathbf{A} , ya que como \mathbf{A} no depende de t su acción en cualquier $e^{\lambda t} \mathbf{w}(t)$ nos da otra expresión semejante. Por lo tanto, la acción de $\frac{d}{dt} - \mathbf{A}$ también nos mantendrá en el mismo espacio. Vemos así que si tenemos una suma de términos con esa estructura y con distintos valores de λ , como es el caso en la ecuación anterior, 5.23, luego si la aplicación de $\frac{d}{dt} - \mathbf{A}$ nos da cero esto significa que la aplicación de $\frac{d}{dt} - \mathbf{A}$ a cada término de la suma debe dar cero. Es decir cada término es una solución de la ecuación 5.16! Es así que concluimos que cada término en la suma 5.16 es de la forma $e^{\mathbf{A}t} \mathbf{v}_0$ para algún $\mathbf{v}_0 \in V^C$. Consideremos ahora el subconjunto W_λ de V^C , tal que si $\mathbf{v}_0 \in W_\lambda$ luego $e^{\mathbf{A}t} \mathbf{v}_0 = e^{\lambda t} \mathbf{v}(t)$. Está claro que los únicos subespacios no triviales serán aquellos con $\lambda = \lambda_i$, donde $\{\lambda_i\}$, $i = 1..d$ son los autovalores de \mathbf{A} . Por lo tanto debemos considerar solo estos subconjuntos. Como esta relación es lineal, W_λ es un subespacio de V^C . Como $e^{\mathbf{A}t} \mathbf{A} \mathbf{v}_0 = \mathbf{A} e^{\mathbf{A}t} \mathbf{v}_0 = e^{\lambda t} \mathbf{A} \mathbf{v}(t) = e^{\lambda t} \mathbf{v}'(t)$ y $\mathbf{v}'(t)$ es también polinomial en t , vemos que W_λ es un subespacio invariante de \mathbf{A} . Mientras que usando una base de W_λ en la que la restricción a ese espacio de \mathbf{A} es triangular superior vemos que λ es el único autovalor que dicha restricción puede tener. Finalmente, como toda solución de 5.16 tiene la forma 5.23 el teorema de existencia de soluciones nos asegura que cualquier dato inicial, es decir cualquier elemento de V^C , puede ser expresado como combinación lineal de elementos en W_{λ_i} . En efecto, sea \mathbf{v}_0 un elemento cualquiera de V^C cualquiera, del teorema de existencia de soluciones, tenemos entonces una única solución de 5.16 $\mathbf{x}(t)$ con $\mathbf{x}(0) = \mathbf{v}_0$. Pero entonces

$$\begin{aligned} \mathbf{x}(t) &= \sum_{i=1}^m e^{\lambda_i t} \mathbf{v}_i(t) \\ &= \sum_{i=1}^m e^{\mathbf{A}t} \mathbf{v}_{0i}, \end{aligned} \quad (5.28)$$

para algún conjunto de vectores $\{\mathbf{v}_{0i}\} \in W_{\lambda_i}$. Evaluando esta expresión en $t = 0$ obtenemos, $\mathbf{v}_0 = \sum_{i=1}^m \mathbf{v}_{0i}$ y vemos que los W_{λ_i} generan V^C . Por otro lado la unicidad de las soluciones implica que ningún elemento de un dado W_{λ_i} puede ser escrito como una combinación lineal de elementos de los otros W_λ [de lo contrario un mismo dato inicial daría origen a dos soluciones distintas (ya que su dependencia funcional sería distinta)]. En efecto, sea $0 = \mathbf{v}_{01} + \dots + \mathbf{v}_{0s}$ una combinación lineal cualquiera de elementos de W_{λ_i} , $i = 1..s$, veamos que cada uno de ellos debe anularse. Por la unicidad de las soluciones a las ecuaciones ordinarias tenemos entonces que $0 = \sum_{i=1}^m e^{\mathbf{A}t} \mathbf{v}_{0i} = \sum_{i=1}^m e^{\lambda_i t} \mathbf{v}_i(t)$ por lo visto anteriormente, cada elemento en la última suma debe anularse y por lo tanto, evaluando a estos en $t = 0$ obtenemos $\mathbf{v}_{0i} = 0 \quad \forall i = 1..s$. Resumiendo lo anterior tenemos,

Teorema 5.3 *Dado un operador $\mathbf{A} : V^C \rightarrow V^C$ existe un conjunto de subespacios invariantes de \mathbf{A} , $\{W_{\lambda_i}\}$, donde λ_i son los autovalores de \mathbf{A} tales que:*

- a) $V^C = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \dots \oplus W_{\lambda_d}$
- b) *El único autovalor de la restricción de \mathbf{A} en $\{W_{\lambda_i}\}$ es λ_i .*

Concluimos el estudio de las soluciones de la ecuación 5.16, o sea de la función $e^{\mathbf{A}t}$ con una descripción pormenorizada de la forma de las mismas que se obtiene usando la representación matricial de \mathbf{A} en la base donde adquiere la forma canónica de Jordan. Es decir, para todo subespacio invariante W_{λ_i} la restricción de \mathbf{A} en este subespacio tiene la forma,

$$\mathbf{A} = \lambda_i + \Delta_i, \quad (5.29)$$

¹ Es decir una combinación lineal de vectores en V^C con coeficientes polinómicos en t .

donde los números λ_i son los autovalores correspondientes a \mathbf{A} —en general complejos y por lo tanto estamos considerando ahora a \mathbf{A} como un operador de C^n en C^n — y Δ_i es una matriz cuyas únicas componentes no nulas son unos y sólo pueden estar en la diagonal superior inmediata a la mayor. La matriz Δ_i tiene la importante propiedad de ser nilpotente, es decir existe un entero m_i menor o igual a la multiplicidad de λ_i , tal que $\Delta_i^{m_i} = 0$.

Como $\lambda_i I$ y Δ_i conmutan tenemos que

$$e^{t\lambda_i I + t\Delta_i} = e^{t\lambda_i I} e^{t\Delta_i}, \quad (5.30)$$

pero $e^{t\lambda_i I} = e^{t\lambda_i} I$ y $e^{t\Delta_i} = \sum_{j=0}^{m_i-1} \frac{(t\Delta_i)^j}{j!}$, es decir una suma finita.

Debido a que la exponencial de \mathbf{A} es una suma de potencias de \mathbf{A} se cumple que los espacios invariantes de ésta son los mismos que los de \mathbf{A} y por lo tanto, $e^{t\mathbf{A}}|_{B_i} = e^{t\mathbf{A}_i}$. De esto concluimos que la matriz $e^{t\mathbf{A}}$ es de la forma

$$e^{t\mathbf{A}} = \begin{pmatrix} k_0 & & & \\ & k_1 & & \\ & & \ddots & \\ & & & k_p \end{pmatrix} \quad (5.31)$$

con cada k_i una sub-matriz cuadrada, la $k_0 = \text{diag}(e^{t\lambda_1}, \dots, e^{t\lambda_n})$ y si $i \neq 0$

$$k_i = e^{t\lambda} \begin{bmatrix} 1 & t & t^2/2 & \cdots & \cdots & t^{n-1}/(n-1)! \\ & 1 & t & \cdots & \cdots & \vdots \\ & & 1 & \cdots & \cdots & \vdots \\ & & & \ddots & \cdots & t^2/2 \\ & 0 & & & \ddots & t \\ & & & & & 1 \end{bmatrix}, \quad (5.32)$$

donde el número de filas o columnas es el mismo que las de las J_i correspondiente a \mathbf{A} en la composición de Jordan, este es menor o igual a la multiplicidad de λ_i como raíz del polinomio característico.

La base en la cual \mathbf{A} tiene la forma canónica de Jordan es en general compleja, es en realidad una base en C^n . Si deseamos usar una base real, lo cual es posible ya que partimos de tener a \mathbf{A} como operador de \mathbb{R}^n en \mathbb{R}^n , y hacemos la transformación correspondiente la matriz $e^{\mathbf{A}t}$ sufrirá la transformación de semejanza correspondiente. Como esta transformación es independiente del tiempo, si bien las componentes de $e^{\mathbf{A}t}$ no tendrán la forma anterior, éstas serán sumatorias de términos exponenciales por polinomios en t . Por lo tanto cada componente de la solución general de 5.16 será de la forma,

$$x^i(t) = \sum_{p=1}^q \left\{ (e^{t\lambda_p} + e^{t\bar{\lambda}_p}) P_p^i(t) + i(e^{t\lambda_p} - e^{t\bar{\lambda}_p}) Q_p^i(t) \right\} \quad (5.33)$$

donde q es el número de autovalores distintos —contando a los pares complejos como uno solo— y P_p^i , Q_p^i son polinomios en t cuyo grado es menor o igual a la multiplicidad con que aparece el autovalor λ_p como raíz del polinomio característico. Esta información es útil en dos sentidos. En el práctico debido a que si bien el método de construir la solución usando una base en la que \mathbf{A} tiene la forma canónica de Jordan es directo, para sistemas de gran dimensión se torna engorroso. En algunos casos es conveniente calcular los autovalores y su multiplicidad y luego suponer una solución de la forma 5.33 y calcular los polinomios P_p^i , Q_p^i .

Este método también es útil en el sentido que nos permite conocer el comportamiento global de las soluciones. Por ejemplo vemos que si la parte real de los autovalores no es positiva y aquellos cuya parte real es cero no aparecen repetidos, entonces todas las soluciones son acotadas [Existe $C > 0$ tal que $\|\mathbf{x}(t)\| < C$.] Si además todos tienen parte real negativa luego todas las soluciones tienden asintóticamente a la solución trivial $\lim_{t \rightarrow \infty} \mathbf{x}(t) = 0$.

Analizaremos ahora en detalle el caso en que todos los autovalores son distintos. Note que si bien éste es el caso genérico, —en el sentido que si un sistema tiene autovalores coincidentes luego existen modificaciones arbitrariamente pequeñas de este que hacen que los autovalores sean distintos— los sistemas con autovalores coincidentes aparecen en física. Este caso contempla también la situación de un sistema general donde los datos iniciales pertenecen a uno de los subespacios unidimensionales B_i .

Como el operador \mathbf{A} es real sus autovalores serán reales o complejos conjugados [si $\det(\mathbf{A} - \lambda I) = 0$ luego $\det(\mathbf{A} - \bar{\lambda} I) = \det(\mathbf{A} - \lambda I) = 0$]. Si λ_i es real luego su autovector puede ser elegido real. En efecto si $(\mathbf{A} - \lambda I)\mathbf{u}_i = 0$, luego también $(\mathbf{A} - \lambda I)\bar{\mathbf{u}}_i = 0$, pero las raíces son simples y por lo tanto cada λ_i tiene un solo autovector—módulo un escalar complejo—, es decir $\bar{\mathbf{u}}_i = \alpha \mathbf{u}_i$ $\alpha \in \mathbb{C}$. Elijiendo $\mathbf{v}_i = \mathbf{u}_i + \bar{\mathbf{u}}_i = (1 + \alpha) \mathbf{u}_i$ obtenemos un autovalor real.

En este caso tenemos que la componente x_0^i de \mathbf{x}_0 en la auto-base en la dirección \mathbf{v}_i evoluciona como

$$x^i(t) = x_0^i e^{\lambda_i t}, \quad (5.34)$$

es decir crece o decrece exponencialmente con el tiempo de acuerdo al signo de λ_i

Si el autovalor λ_i es complejo entonces podemos elegir su autovector \mathbf{u}_i de forma que sea el complejo conjugado al elegido correspondiente a $\bar{\lambda}_i$. Este par de autovectores generan un subespacio complejo 2-dimensional. Si \mathbf{x}_0 pertenece a este subespacio y es real entonces tendrá la forma $\mathbf{x}_0 = a(\mathbf{u}_i + \bar{\mathbf{u}}_i) - ib(\mathbf{u}_i - \bar{\mathbf{u}}_i)$ con a y b reales, o sea $\mathbf{x}_1 = \mathbf{u}_i + \bar{\mathbf{u}}_i$ y $\mathbf{x}_2 = i(\mathbf{u}_i - \bar{\mathbf{u}}_i)$ forman una base real.

¿Cómo cambian estos vectores si les aplicamos el operador $e^{\mathbf{A}t}$? Es decir, ¿cuáles son las soluciones de la ecuación $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ con condiciones iniciales \mathbf{x}_1 y \mathbf{x}_2 ? Llamando a estos $\mathbf{x}_1(t)$ y $\mathbf{x}_2(t)$ respectivamente y usando que $e^{\mathbf{A}t} \mathbf{u}_i = e^{\lambda_i t} \mathbf{u}_i$ obtenemos,

$$\begin{aligned} \mathbf{x}_1(t) &= e^{\alpha_i t} (\mathbf{x}_1 \cos w_i t - \mathbf{x}_2 \sin w_i t) \\ \mathbf{x}_2(t) &= e^{\alpha_i t} (\mathbf{x}_2 \cos w_i t + \mathbf{x}_1 \sin w_i t), \end{aligned} \quad (5.35)$$

con $\lambda_i = \alpha_i + i w_i$.

Vemos que la acción del operador $e^{\mathbf{A}t}$ es en este caso la de dilatar los vectores por un factor $e^{\alpha_i t}$ y la de rotarlos un ángulo $w_i t$.

5.4. Problemas

Problema 5.1 Dado un conjunto de funciones $\{f_i(t)\}$, $i = 1, \dots, n$ se definen vectores $\mathbf{u}_i(t) := (f_i(t), f_i^{(1)}(t), \dots, f_i^{(n-1)}(t))$ y el Wronskiano del sistema como $W(\{f_i\})(t) := \varepsilon(\mathbf{u}_1(t), \mathbf{u}_2(t), \dots, \mathbf{u}_n(t))$. Si el Wronskiano de un conjunto de funciones no se anula, entonces las funciones son linealmente independientes, es decir ninguna combinación lineal no trivial de las mismas (con coeficientes constantes) se anula. La converso no es cierta. Calcule el Wronskiano de los siguientes conjuntos:

- $\{4, t\}$
- $\{t, 3t, t^2\}$
- $\{e^t, te^t, t^2 e^t\}$
- $\{\sin(t), \cos(t), \cos(2t)\}$
- $\{1, \sin(t), \sin(2t)\}$

Problema 5.2 Decida si el siguiente conjunto de funciones es linealmente dependiente o no. Luego calcule el Wronskiano.

$$f_1(t) = \begin{cases} 0, & 0 \leq x \leq 1/2 \\ (x - 1/2)^2, & 1/2 \leq x \leq 1 \end{cases} \quad (5.36)$$

$$f_2(t) = \begin{cases} (x - 1/2)^2, & 0 \leq x \leq 1/2 \\ 0, & 1/2 \leq x \leq 1 \end{cases} \quad (5.37)$$

Problema 5.3 Use la teoría de ecuaciones diferenciales ordinarias para probar las siguientes identidades:

a)

$$e^s \mathbf{A} e^t \mathbf{A} = e^{(s+t)} \mathbf{A}, \quad (5.38)$$

b)

$$e^{\mathbf{A}} e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}}, \quad \text{si y solo si } [\mathbf{A}, \mathbf{B}] := \mathbf{AB} - \mathbf{BA} = 0. \quad (5.39)$$

Problema 5.4 Grafique los campos vectoriales correspondientes a los siguientes sistemas y algún conjunto de soluciones típicas de los mismos.

a)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.40)$$

b)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.41)$$

c)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.42)$$

d)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.43)$$

e)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.44)$$

f) (compare las soluciones con las del punto b)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.45)$$

g)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.46)$$

h)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.47)$$

i)

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1+i & 0 \\ 1 & 1-i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (5.48)$$

Problema 5.5 Lleve estas ecuaciones a sistemas de primer orden y encuentre la solución general de las ecuaciones:

a) $\frac{d^3 x}{dt^3} - 2 \frac{d^2 x}{dt^2} - 3 \frac{dx}{dt} = 0$

b) $\frac{d^3 x}{dt^3} + 2 \frac{d^2 x}{dt^2} + \frac{dx}{dt} = 0$

c) $\frac{d^3 x}{dt^3} + 4 \frac{d^2 x}{dt^2} + 13 \frac{dx}{dt} = 0$

d) $\frac{d^4 x}{dt^4} + 4 \frac{d^3 x}{dt^3} + 8 \frac{d^2 x}{dt^2} + 8 \frac{dx}{dt} + 4 = 0$

Problema 5.6 (Ley de Newton) Sea la ecuación $\frac{d^2x}{dt^2} + f(x) = 0$.

- Pruebe que $\frac{1}{2}\dot{x}^2 + \int_{x_0}^x f(s)ds$ es una integral primera.
- Encuentre la integral primera de $\frac{d^2x}{dt^2} - x + x^2/2 = 0$.
- Grafique el campo vectorial correspondiente y alguna de sus soluciones. Encuentre sus soluciones estacionarias (o puntos de equilibrio) y estudie sus entornos linealizando la ecuación en estos puntos.

Problema 5.7 Estudie el siguiente sistema

$$\begin{aligned}\dot{x}_1 &= x_1 - x_1x_2 - x_2^3 + x_3(x_1^2 + x_2^2 - 1 - x_1 + x_1x_2 + x_2^3) \\ \dot{x}_2 &= x_1 - x_3(x_1 - x_2 + x_1x_2) \\ \dot{x}_3 &= (x_3 - 1)(x_3 + 2x_3x_2^2 + x_3^3)\end{aligned}\tag{5.49}$$

- Encuentre los puntos de equilibrio.
- Muestre que los planos $x_3 = 0$ y $x_3 = 1$ son conjuntos invariantes, (es decir, las soluciones que comienzan en los mismos nunca los abandonan).
- Considere el conjunto invariante $x_3 = 1$ y vea si tiene soluciones periódicas.

Las soluciones estacionarias o de equilibrio son muy importantes en física ya que muchos sistemas (especialmente aquellos donde hay disipación) se comportan de forma tal que la solución se aproxima, en su evolución, a estas soluciones.

Otra particularidad que las hace importantes es el hecho que éstas están simplemente dadas por los puntos donde el campo vectorial se anula, y por lo tanto a lo sumo se debe resolver una ecuación algebraica para encontrarlas, en contraste con el caso general donde debemos resolver una ecuación diferencial. Incluso en algunos casos es posible inferir que el campo vectorial se debe anular, tal es el caso por ejemplo si la variedad es una esfera bi-dimensional ya que cualquier campo vectorial continuo definido sobre ésta es tal que al menos existen dos puntos donde éste se anula, y por lo tanto en este caso siempre habrá soluciones estacionarias.

Ejercicio: Convénzase que esto es así.

Por lo dicho anteriormente se desprende que es importante estudiar el comportamiento de las soluciones que se originan de datos iniciales cercanos a una solución estacionaria. Para ello definimos los siguientes conceptos de estabilidad.

Definición: Sea \mathbf{v} un campo vectorial en M y sea $p \in M$ tal que $\mathbf{v}(p) = 0$. Diremos que la solución estacionaria $\gamma(t) \equiv p$ es **estable** si dado un entorno U_p de p existe otro entorno V_p de p tal que toda solución $\sigma(t)$ con $\sigma(0) \in V_p$ satisface $\sigma(t) \in U_p \quad \forall t \geq 0$. [Ver figura 6.1.]

Definición: Diremos que la solución anterior es **asintóticamente estable** si es estable y además si $\sigma(0) \in V_p$ luego $\lim_{t \rightarrow +\infty} \sigma(t) = p$.

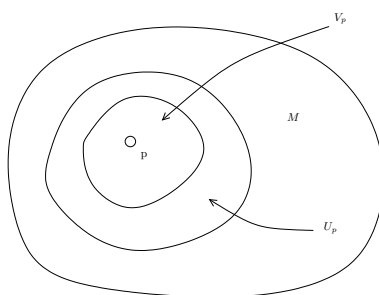


Figura 6.1: Estabilidad.

Si una solución estacionaria no es estable entonces no tiene mucho interés físico ya que la más mínima perturbación nos *aleja mucho* de ésta.

Ejemplos:

a) La ecuación del crecimiento bacteriano $\dot{x} = ax - bx^2$ tiene como soluciones estacionarias $x(t) \equiv 0$ y $x(t) \equiv \frac{a}{b}$. Como la solución general es,

$$x(t) = \frac{x(0)e^{at}}{1 + \frac{bx(0)}{a}(e^{at} - 1)}, x(0) \geq 0 \quad (6.1)$$

se ve claramente que $x(t) = 0$ no es una solución estable (a estas las llamaremos inestables) y que $x(t) \equiv \frac{a}{b}$ es asintóticamente estable. Si contaminamos un recipiente de cultivo con una sola bacteria esto es suficiente para que ésta se reproduzca ¹ hasta alcanzar (asintóticamente) la concentración $\frac{a}{b}$. Si ahora sacamos o agregamos algunas bacterias cosa de cambiar su concentración las bacterias se reproducirán o aniquilarán hasta alcanzar nuevamente la concentración estable $\frac{a}{b}$.

b) La ecuación $\dot{x} = Ax$ tiene entre sus soluciones estacionarias la dada por $x(t) \equiv 0$. ¿Cuáles otras? Estas serán estables cuando los autovalores de A , λ_i , satisfagan $\Re(\lambda_i) \leq 0$, $\lambda_i \neq \lambda_j$ o $\Re(\lambda_i) < 0$ si $\lambda_i = \lambda_j$, donde \Re indica parte real. Esto es claro ya que la solución general es $x(t) = e^{At}x(0)$ y la condición mencionada implica que $\|e^{At}\|_{\mathcal{L}} < C$, $C > 0 \forall t \geq 0$. En este caso podemos tomar como $U_{p=0} = \{x \in \mathbb{R}^n \mid |x| < \epsilon\}$ y por $V_{p=0} = \{x \in \mathbb{R}^n \mid |x| < \frac{\epsilon}{C}\}$. Si además se cumple que $\Re(\lambda_i) < 0 \forall i = 1, \dots, n$ luego $x(t) \equiv 0$ es asintóticamente estable.

El siguiente teorema nos brinda una herramienta muy práctica para conocer cuando una solución estacionaria es estable o no.

Teorema 6.1 (de Estabilidad) Sea $\gamma(t) \equiv p$ una solución estacionaria de $v \in TM$, es decir $v(p) = 0$, y sea $A : T_p M \rightarrow T_p M$ definido por,

$$Ax \equiv \frac{d}{ds} v(\sigma_x(s))|_{s=0}, \quad (6.2)$$

donde $\sigma_x(s)$ es una curva en M satisfaciendo: $\sigma_x(0) = p$ y $\frac{d}{ds} \sigma_x|_{s=0} = x \in T_p M$, es decir es una curva suave cualquiera que pasa por p cuando $s = 0$ y en ese punto tiene como tangente a $x \in T_p M$. Si $\Re(\lambda_i) < 0$, donde λ_i son los autovalores de A , luego $\gamma(t) \equiv p$ es asintóticamente estable. Si algún λ_i tiene parte real positiva luego $\gamma(t) = p$ es inestable.

Ejercicios:

a) Muestre que A es realmente un operador lineal.

b) Muestre que $Ax = [v, \tilde{x}]|_p$, donde \tilde{x} es cualquier campo vectorial tal que $\tilde{x}|_p = x$. Ayuda: no olvide que $v|_p = 0$.

Ejemplos:

a) Considere la ecuación del crecimiento bacteriano en $x = 0$ y $x = \frac{a}{b}$. Para el primer caso tomamos $\sigma_{\delta x}(s) = \delta x s$ y obtenemos,

$$A\delta x = \frac{d}{ds}(a\delta x s - b(\delta x)^2 s^2)|_{s=0} = a\delta x, \quad (6.3)$$

lo que muestra que $x = 0$ es una solución inestable ya que $a > 0$. Para el segundo caso tomamos $\sigma_{\delta x} = \frac{a}{b} + \delta x s$ luego

$$A\delta x = \frac{d}{ds}\left(\frac{a^2}{b} + a\delta x s - b\left(\frac{a}{b} + \delta x s\right)^2\right)|_{s=0} = a\delta x - 2a\delta x = -a\delta x, \quad (6.4)$$

¹Note que en el proceso real no tenemos una variable continua y por lo tanto en él las perturbaciones son como mínimo de una bacteria, esto hace que sea posible tener recipientes estériles y por lo tanto la inestabilidad matemática no se manifiesta en la realidad.

lo que muestra que es asintóticamente estable.

b) Péndulo físico con fricción: $\ddot{\theta} = -\text{sen}\theta - k\dot{\theta}$, $k > 0$. El campo vectorial en este caso está dado por,

$$\begin{aligned}\dot{\theta} &= z \\ \dot{z} &= -\text{sen}\theta - kz,\end{aligned}\tag{6.5}$$

es decir el vector con componentes $(z, -\text{sen}\theta - kz)$ en el espacio de las fases que es este caso es un cilindro, $z \in [-\infty, +\infty]$, $\theta \in [0, 2\pi]$.

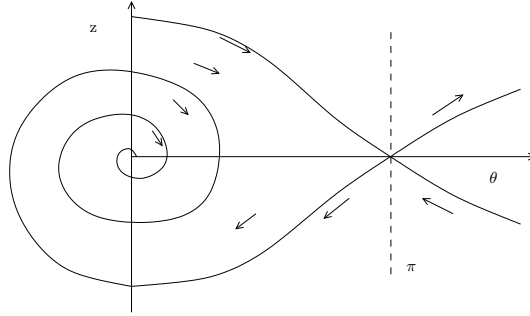


Figura 6.2: Péndulo físico con fricción.

Este vector se anula sólo en $p_1 = (\theta = 0, z = 0)$ y $p_2 = (\theta = \pi, z = 0)$. Para la primera solución estacionaria, usando $\sigma(s)$ tal que,

$$\phi \circ \sigma(s) = \begin{pmatrix} \delta\theta & s \\ \delta z & s \end{pmatrix},\tag{6.6}$$

obtenemos,

$$A \begin{pmatrix} \delta\theta \\ \delta z \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -k \end{pmatrix} \begin{pmatrix} \delta\theta \\ \delta z \end{pmatrix}.\tag{6.7}$$

En este caso la ecuación de autovalores es $\lambda^2 + k\lambda + 1 = 0$ de lo cual obtenemos, $\lambda_{\pm} = \frac{-k \pm \sqrt{k^2 - 4}}{2}$, lo que implica $\Re(\lambda_{\pm}) < 0$ y así estabilidad.

Para la segunda solución estacionaria tomamos $\sigma(s)$ tal que,

$$\phi \circ \sigma(s) = \begin{pmatrix} \delta\theta s + \pi \\ \delta z \end{pmatrix}\tag{6.8}$$

y obtenemos

$$A \begin{pmatrix} \delta\theta \\ \delta z \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -k \end{pmatrix} \begin{pmatrix} \delta\theta \\ \delta z \end{pmatrix}.\tag{6.9}$$

con autovalores $\lambda_{\pm} = \frac{-k \pm \sqrt{k^2 + 4}}{2}$, lo que implica $\Re(\lambda_{\pm}) > 0$ y así inestabilidad.

De los ejemplos se ve que este teorema es de aplicación amplia. Como estabilidad es una noción local para facilitar la demostración tomaremos $M = \mathbb{R}^n$ y un sistema coordenado cartesiano con origen en la solución estable a considerar. Usaremos varios resultados previos que probamos a continuación.

Teorema 6.2 (de Lyapunov) Sea $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un operador lineal cuyos autovalores tienen parte real negativa. Luego existe un tensor de tipo $(2,0)$, $\rho(\cdot, \cdot)$, simétrico y positivo definido [$\rho(w, v) = \rho(v, w)$ y $\rho(w, w) \geq 0$ ($=$ sii $w = 0$)] tal que,

$$Ax(\rho(x, x)) < 0 \quad \forall \quad x \neq 0,$$

es decir la derivada de $\rho(\mathbf{x}, \mathbf{x})$ en la dirección $A\mathbf{x}$ es negativa.

La interpretación geométrica de esta condición es que $\rho(\mathbf{x}, \mathbf{x})$ define una norma cuyas superficies de nivel son tales que en cada una de éstas el vector $A\mathbf{x}$ en $p = \mathbf{x}$ apunta hacia adentro. Ver figura.

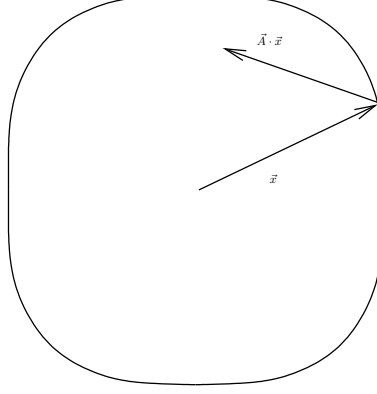


Figura 6.3: La norma $\rho(x, x)$.

Prueba: Si todos los autovalores de A son distintos luego existe una base de Jordan $\{\mathbf{u}_i\}$, $i = 1, \dots, n$ y la correspondiente co-base $\{\boldsymbol{\theta}^i\}$, $i = 1, \dots, n$ (con $\boldsymbol{\theta}^i(\mathbf{u}_j) = \delta_j^i$) tal que $A = \sum_{i=1}^n \lambda_i \mathbf{u}_i \boldsymbol{\theta}^i$. Sea en este caso $\rho = \sum_{i=1}^n \boldsymbol{\theta}^i \bar{\boldsymbol{\theta}}^i$. Si $z = \sum_{i=1}^n z^i \mathbf{u}_i$ e $y = \sum_{i=1}^n y^i \mathbf{u}_i \in C^n$, luego $\rho(z, y) = \sum_{i=1}^{n'} z^i y^i + \sum_{n'+1}^{(n-n')/2} (z^i \bar{y}^i + \bar{z}^i y^i)$ donde hemos separado la suma en la de los autovectores reales y en la de los complejos conjugados. Es fácil ver que el $\rho(\cdot, \cdot)$ obtenido al restringir z e y a \mathbb{R}^n es la norma cartesiana usual en la base real correspondiente. Calculemos ahora $A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x}))$.

$$\begin{aligned} A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x})) &= \lim_{\varepsilon \rightarrow 0} \frac{\rho(\mathbf{x} + \varepsilon A\mathbf{x}, \mathbf{x} + \varepsilon A\mathbf{x}) - \rho(\mathbf{x}, \mathbf{x})}{\varepsilon} \\ &= \rho(A\mathbf{x}, \mathbf{x}) + \rho(\mathbf{x}, A\mathbf{x}) \\ &= 2 \sum_{i=1}^n (\Re \lambda_i) x^i x^i < 0. \end{aligned} \tag{6.10}$$

Hemos probado así el teorema para el caso diagonalizable. El caso en que A no lo es es más complicado y para ello usaremos los siguientes lemas.

Lema 6.1 Dado $\varepsilon > 0$ existe una base $\{\mathbf{u}_i\}$ tal que en esa base

$$A = \text{diag.}(\lambda_1, \dots, \lambda_n) + \varepsilon \Delta,$$

con Δ una matriz con componentes distintas de cero (e igual a uno) a lo sumo en la diagonal inferior, es decir,

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & . & . & . & 0 \\ \varepsilon & \lambda_2 & 0 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & 0 \\ 0 & . & . & . & . & \varepsilon & \lambda_r \end{pmatrix}. \tag{6.11}$$

Prueba: Es un simple cambio de escala de la base de Jordan de A . Por ejemplo en C^3 si $\{\tilde{\mathbf{u}}_i\}$ es tal

que en ella,

$$A = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}. \quad (6.12)$$

definiendo $\mathbf{u}_1 = \frac{\tilde{\mathbf{u}}_1}{\epsilon^2}$, $\mathbf{u}_2 = \frac{\tilde{\mathbf{u}}_2}{\epsilon}$, y $\mathbf{u}_3 = \tilde{\mathbf{u}}_3$, vemos que en esta nueva base,

$$A = \begin{pmatrix} \lambda & 0 & 0 \\ \epsilon & \lambda & 0 \\ 0 & \epsilon & \lambda \end{pmatrix}. \quad (6.13)$$

$$[A\mathbf{u}_1 = A \frac{\tilde{\mathbf{u}}_1}{\epsilon^2} = \frac{\lambda \tilde{\mathbf{u}}_1}{\epsilon^2} = \lambda \mathbf{u}_1, A\mathbf{u}_2 = A \frac{\tilde{\mathbf{u}}_2}{\epsilon} = \frac{\tilde{\mathbf{u}}_1 + \lambda \tilde{\mathbf{u}}_2}{\epsilon} = \epsilon \mathbf{u}_1 + \lambda \mathbf{u}_2, \text{ etc.}]$$

Lema 6.2 *El conjunto de tensores simétricas positivas es abierto en el conjunto de formas simétricas, es decir si ρ_0 es simétrica y positiva y ρ_1 es simétrica luego existe $\epsilon > 0$ tal que $\rho_0 + \epsilon \rho_1$ es también positiva.*

Prueba: Sea B_1 la esfera unidad con respecto a alguna norma en \mathbb{R}^n y consideremos $\rho_0(\mathbf{x}, \mathbf{x}) : B_1 \rightarrow \mathbb{R}^+$. Como B_1 es compacta ρ_0 alcanza allí su mínimo, α_{min} . Análogamente ρ_1 alcanza allí su máximo, que llamaremos β_{max} . Tomando $\epsilon < \frac{\alpha_{min}}{\beta_{max}}$ se cumple lo requerido.

Ejemplo: En \mathbb{R}^2 sea $\rho_0((x_1, y_1), (x_2, y_2)) = x_1x_2 + y_1y_2$ y $\rho_1((x_1, y_1), (x_2, y_2)) = x_1y_2 + x_2y_1$, luego $\rho_0 + \epsilon \rho_1$ es positiva si $|\epsilon| < 2$.

Para completar la demostración del teorema de Lyapunov tomaremos

$$\rho = \sum_i \theta^i \otimes \bar{\theta}^i,$$

donde $\{\theta^i\}$ es la co-base encontrada en el lema 6.1 con $\epsilon > 0$ a determinar. Luego,

$$\begin{aligned} -A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x})) &= -\rho(A\mathbf{x}, \mathbf{x}) - \rho(\mathbf{x}, A\mathbf{x}) \\ &= -2 \sum_{i=1}^n \Re(\lambda_i) x^i x^i - 2\epsilon \sum_{i=1}^{n-1} f_i x^i x^{i+1}, \end{aligned} \quad (6.14)$$

con $f_i = 1$ ó 0 según haya ϵ o no en ese lugar de la diagonal inmediata superior. El primer término es una forma positiva definida ρ_0 evaluada en \mathbf{x} en ambas entradas. El segundo es una forma simétrica $\epsilon \rho_1$ evaluada en \mathbf{x} en ambas entradas. El segundo lema nos dice que tomando ϵ pequeño $\rho_0 + \epsilon \rho_1$ es también positiva definida. Esto concluye la prueba del teorema.

Note ahora que no solo la ρ que hemos encontrado es simétrica y positiva definida y por lo tanto una norma en \mathbb{R}^n , sino que también $-A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x}))$ es positiva definida y define una norma [ya que $-A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x})) = -\rho(A\mathbf{x}, \mathbf{x}) - \rho(\mathbf{x}, A\mathbf{x})$]. Pero ya vimos que en \mathbb{R}^n todas las normas son equivalentes y por lo tanto existirá $\gamma > 0$ tal que

$$-\frac{1}{2\gamma} \rho(\mathbf{x}, \mathbf{x}) \leq A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x})) \leq -2\gamma \rho(\mathbf{x}, \mathbf{x}). \quad (6.15)$$

Este es el resultado que utilizaremos para la prueba del teorema de estabilidad que damos a continuación.

Prueba del Teorema de Estabilidad: Hemos visto que si $\Re(\lambda_i) < 0$ luego existe una constante $\gamma > 0$ y una forma simétrica positiva definida $\rho(\cdot, \cdot)$ tal que

$$A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x})) \leq -2\gamma \rho(\mathbf{x}, \mathbf{x}). \quad (6.16)$$

Aplicando el campo vectorial que define la ecuación, $\mathbf{v}(\mathbf{x})$ a $\rho(\mathbf{x}, \mathbf{x})$ obtenemos,

$$\mathbf{v}(\mathbf{x})(\rho(\mathbf{x}, \mathbf{x})) = A\mathbf{x}(\rho(\mathbf{x}, \mathbf{x})) + O(\rho(\mathbf{x}, \mathbf{x})^{3/2}), \quad (6.17)$$

donde hemos supuesto \mathbf{v} es diferenciable y por lo tanto $\mathbf{v}(\mathbf{x}) = A\mathbf{x} + O(|\mathbf{x}|^2)$.

Ejemplo: $v(x) = (ax + bx^2) \frac{\partial}{\partial x}$ y $\rho(x, x) = \alpha x^2$ luego,

$$\begin{aligned} v(x)(\rho(x, x)) &= ax \frac{\partial}{\partial x} \rho(x, x) + 2b\alpha x^3 \\ &= ax \frac{\partial}{\partial x} \rho(x, x) + \frac{2b}{\alpha^{1/2}} \rho(x, x)^{3/2} \end{aligned} \quad (6.18)$$

Si \mathbf{x} es suficientemente pequeño, $||\mathbf{x}|| < C$ para algún $C > 0$ se cumple que $O(\rho(\mathbf{x}, \mathbf{x})^{3/2}) < \gamma(\rho(\mathbf{x}, \mathbf{x}))$ y por lo tanto tenemos que

$$\mathbf{v}(\mathbf{x})(\rho(\mathbf{x}, \mathbf{x})) \leq -2\gamma \rho(\mathbf{x}, \mathbf{x}) + \gamma \rho(\mathbf{x}, \mathbf{x}) \leq -\gamma \rho(\mathbf{x}, \mathbf{x}). \quad (6.19)$$

Sea entonces $\varphi(t)$ una solución con dato inicial suficientemente cercano a $\mathbf{x} = 0$, es decir $\dot{\varphi}(t) = v(\varphi(t))$, $|\varphi(0)| < C$. Definiendo ²

$$r(t) = \ln \rho(\varphi(t), \varphi(t)) \quad (6.20)$$

y derivando con respecto a t obtenemos,

$$\begin{aligned} \dot{r}(t) &= \frac{2 \rho(\varphi(t), \varphi(t))}{\rho(\varphi(t), \varphi(t))} = \frac{\mathbf{v}(\varphi(t))(\rho(\varphi(t), \varphi(t)))}{\rho(\varphi(t), \varphi(t))} \\ &\leq -\gamma. \end{aligned} \quad (6.21)$$

Por lo tanto $r(t) \leq r(0) - \gamma t$, he integrando concluimos que

$$\rho(\varphi(t), \varphi(t)) \leq \rho(\varphi(0), \varphi(0))e^{-\gamma t}, \quad (6.22)$$

lo que implica que $\varphi(t) \rightarrow 0$ y concluye la demostración del teorema. ³

Es instructivo observar que la prueba se basa en la construcción de una norma, $\rho(\mathbf{x}, \mathbf{x})$, especialmente adaptada al problema en el sentido que nos asegura que en sus superficies de nivel el vector $\mathbf{v}(\mathbf{x})$ (si \mathbf{x} es suficientemente pequeño) apunta hacia adentro.

6.1. Problemas

Problema 6.1 (La ecuación de Volterra-Lotka) Sea el sistema:

$$\begin{aligned} \dot{x}_1 &= (a - bx_2)x_1 \\ \dot{x}_2 &= -(c - fx_1)x_2, \end{aligned} \quad (6.23)$$

$a, b, c, f \geq 0$.

a) Haga una transformación de coordenadas y tiempo que lo lleve a la forma,

$$\begin{aligned} \dot{x}_1 &= (1 - x_2)x_1 \\ \dot{x}_2 &= -(e - x_1)x_2 \end{aligned} \quad (6.24)$$

²Por unicidad de la solución $\varphi(t) \neq 0$ y por lo tanto $\rho(\varphi(t), \varphi(t)) > 0$ y el logaritmo está bien definido.

³En realidad debemos probar además que si $|\varphi(0)| < C$ luego $\varphi(t)$ existe para todo t . Pero la última desigualdad probada nos dice que $\varphi(t)$ no puede abandonar la región compacta $\{\mathbf{x} | \rho(\mathbf{x}, \mathbf{x}) \leq \rho(\varphi(0), \varphi(0))\}$ y por lo tanto el teorema de extensión nos asegura que $\varphi(t)$ existe para todo $t \in [0, +\infty)$.

b) Grafique el campo vectorial y vea que no hay integrales primeras no triviales en los cuadrantes donde al menos una de las coordenadas es negativa.

c) Encuentre las soluciones de equilibrio y determine cuáles son estables y cuáles no.

d) Examine el cuadrante positivo mediante la transformación: $x_1 = e^{q_1}$, $x_2 = e^{q_2}$ y vea que allí la cantidad $f(q_1, q_2) := eq_1 + q_2 - (e^{q_1} + e^{q_2})$ es una integral de movimiento. Utilice esta información para inferir que en ese cuadrante las trayectorias permanecen en regiones acotadas.

e) Examine las ecuaciones linealizadas alrededor de la solución de equilibrio en el cuadrante positivo y determine cuál sería la frecuencia de oscilaciones de las variaciones cercanas a equilibrio que tendría el sistema.

Nota, esta ecuación describe la población de dos especies en competencia. Lo que acabamos de ver es que las especies no crecen indefinidamente ni desaparecen. Esto último nos dice que la aproximación no es muy buena... Note que x_2 representa un predador que depende exclusivamente para su subsistencia de la presa x_1 , ya que si ésta se extingue el predador lo sigue.

Problema 6.2 Si en el sistema anterior las especies tienen otros medios alternativos de subsistencia entonces las ecuaciones resultantes son:

$$\begin{aligned}\dot{x}_1 &= (1 - x_1 - ax_2)x_1 \\ \dot{x}_2 &= (1 - x_2 + bx_1)x_2\end{aligned}\tag{6.25}$$

donde se ha usado otra parametrización.

a) Encuentre las soluciones de equilibrio para los casos i) $0 < a < 1$, y ii) $1 < a$.

b) Se observa que $b \approx 3a$ (a nos indica cuan agresivo es el predador). ¿Cuál es el valor de a si su instinto lo lleva a maximizar la población de su especie manteniendo un equilibrio estable?

c) Mire el sistema cerca de su solución de equilibrio y determine la frecuencia de las oscilaciones en el número de especies que esperaríamos en ese entorno.

Problema 6.3 (Ciclo límite) Considere el sistema:

$$\begin{aligned}\dot{x}_1 &= x_2 + x_1(1 - (x_1^2 + x_2^2)) \\ \dot{x}_2 &= -x_1 + x_2(1 - (x_1^2 + x_2^2)).\end{aligned}\tag{6.26}$$

a) Lleve este sistema a un par de ecuaciones desacopladas,

$$\begin{aligned}\dot{r} &= f(r) \\ \dot{\theta} &= -1.\end{aligned}\tag{6.27}$$

b) Estudie los puntos de equilibrio de la primera ecuación y su estabilidad. ¿Cuál es la solución del sistema original correspondiente a este punto de equilibrio?

c) Grafique soluciones cerca de estos puntos, en el plano (r, θ) y en el plano (x_1, x_2) .

d) Utilice el método descrito al final de capítulo 4 para corroborar la estabilidad encontrada en el punto b).

Problema 6.4 (Verhulst) Encuentre las soluciones de equilibrio (y analice su estabilidad) del sistema:

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= x - 2x^3\end{aligned}\tag{6.28}$$

Grafique algunas soluciones.

Problema 6.5 (Verhulst) Encuentre las soluciones de equilibrio (y analice su estabilidad) del sistema:

$$\begin{aligned}\dot{x} &= x^2 - y^3 \\ \dot{y} &= 2x(x^2 - y)\end{aligned}\tag{6.29}$$

Grafique algunas soluciones.

Problema 6.6 (Verhulst) *Encuentre las soluciones de equilibrio (y analice su estabilidad) del sistema:*

$$\begin{aligned}\dot{x} &= -x \\ \dot{y} &= 1 - (x^2 + y^2)\end{aligned}\tag{6.30}$$

Grafique algunas soluciones.

PRUEBA DEL TEOREMA FUNDAMENTAL

Solo probaremos los puntos *i*) y *ii*), para lo cual usaremos el método de aproximaciones sucesivas de Picard, que tiene su importancia en aplicaciones. La prueba del punto *iii*) es de carácter técnico y no aporta ninguna enseñanza extra de consideración. Punto *iv*) sigue de los puntos *i*) y *ii*) y la prueba de esto es idéntica a la del Corolario 1.2 del Teorema análogo para el caso de una EDO.

Para la prueba de estos puntos deberemos desarrollar algunas ideas y resultados de la teoría matemática de espacios vectoriales de dimensión infinita. Nos restringiremos a lo mínimo que el teorema requiere ya que estos temas serán desarrollados más extensamente en la segunda parte de este curso.

Definición: Diremos que un espacio vectorial, V es de **dimensión infinita** si tiene un número infinito de vectores linealmente independientes.

Ejemplos:

a) Sea V el conjunto de todas las sucesiones $\{x_i\}$, $i = 1, \dots, \infty$ de números reales. Este es un espacio vectorial si definimos la **suma** y el **producto** de sucesiones por medio de la siguiente fórmula,

$$\{x_i\} + \alpha\{y_i\} = \{x_i + \alpha y_i\}. \quad (7.1)$$

Estos vectores se pueden escribir también como $\{x_i\} = (x_1, x_2, x_3, \dots)$ con lo que se ve que es la extensión de \mathbb{R}^n con $n \rightarrow \infty$. Claramente los $\mathbf{u}_1 = (1, 0, 0, \dots)$; $\mathbf{u}_2 = (0, 1, 0, \dots)$; $\mathbf{u}_3 = (0, 0, 1, \dots)$; son linealmente independientes e infinitos en número.

b) Sea V el conjunto de funciones continuas en el intervalo $[0, 1]$. Este es un espacio vectorial ya que si f y g son continuas en $[0, 1]$, luego $h = f + \alpha g$, $\alpha \in \mathbb{R}$, es también continua en $[0, 1]$.

El conjunto de vectores $u_n = x^n$, $n \in \mathbb{N}$, es linealmente independiente e infinito ($\sum_{n=0}^M c^n u_n = \sum_{n=0}^M c^n x^n = 0 \implies c^n = 0 \quad \forall n \leq M$) ya que un polinomio de grado M tiene a lo sumo M raíces.

A los espacios de dimensión infinita también se les puede asignar normas, pero en este caso éstas no son equivalentes y por lo tanto habrá que ser cuidadoso en no confundir las estructuras resultantes. Para ello asignaremos distintos nombres a los espacios con distintas normas.

Ejemplos: a) El espacio vectorial normado l^2 es el espacio de sucesiones infinitas con la norma

$$\|\{x_i\}\|_2 = \sqrt{\sum_{i=1}^{\infty} x_i^2} < \infty.$$

b) El espacio vectorial normado l^∞ es el espacio de sucesiones infinitas con la norma $\|\{x_i\}\|_\infty = \sup_i \{|x_i|\}$. Es decir el espacio de todas las sucesiones acotadas.

c) El espacio vectorial normado $C[a, b]$ es el espacio de funciones continuas con la norma $\|f\|_c = \sup_{x \in [a, b]} \{|f(x)|\}$. Es decir el espacio de funciones continuas acotadas en $[a, b]$.

Ejercicios:

- 1) Muestre que las normas definidas en el ejemplo anterior son en realidad normas.
- 2) Muestre que existen sucesiones en l^∞ que no pertenecen a l^2 . Ayuda: Encuentre una de ellas.

3) Muestre que $\|\{x_i\}_n\|_n := \sqrt[n]{\sum_{i=1}^n |x_i|^n} \xrightarrow{n \rightarrow \infty} \sup_i \{|x_i|\}$.

A diferencia del caso de dimensión finita, un espacio normado de dimensión infinita no necesariamente es completo. Para ilustrar esto consideremos el espacio vectorial normado l_0^∞ el cual es un subespacio de l^∞ que consiste en todas las sucesiones acotadas con solo un número finito de términos no nulos. Cada sucesión $\{x_i\}_n = (1, 1/2, 1/3, \dots, 1/n, 0, 0, \dots)$ está en l_0^∞ , la sucesión de sucesiones $\{\{x_i\}_n\}$ es de Cauchy

$$\left(\|\{x_i\}_m - \{x_i\}_n\|_\infty = \frac{1}{n+1} \xrightarrow{n \rightarrow \infty} 0 \right)$$

y converge a la sucesión $\{x_i\}_\infty = (1, 1/2, 1/3, \dots) \in l^\infty$ que no pertenece a l_0^∞ . Por lo tanto l_0^∞ no es completo.

Definición: Diremos que un espacio vectorial normado $(V, \|\cdot\|)$ es un **espacio de Banach** si es completo.

Ejemplos: \mathbb{R}^n con cualquiera de sus normas, l^2 y l^∞ son espacios de Banach.

Un importante resultado, crucial en la prueba del Teorema Fundamental es el siguiente teorema.

Teorema 7.1 *El espacio vectorial de funciones continuas acotadas, $C[a, b]$ es completo.*

Prueba: Sea $\{f_n(x)\}$ una sucesión de Cauchy en $C[a, b]$, es decir cada $f_n(x)$ es una función continua y acotada en $[a, b]$ y se cumple que dado $\varepsilon > 0$ existe $N > 0$ tal que $\forall m, n > N \quad \|f_n - f_m\|_c = \sup_{x \in [a, b]} |f_n(x) - f_m(x)| < \varepsilon$. Pero entonces para cada $x \in [a, b]$ la sucesión de números reales $\{f_n(x)\}$ es de Cauchy. Pero los números reales son completos y por lo tanto para cada $x \in [a, b]$ $\{f_n(x)\}$ converge a un número que llamaremos $f(x)$. Pero entonces dado $\varepsilon > 0$ para todo N tal que $m, n \geq N$ implique $\|f_n - f_m\|_c < \varepsilon$, tenemos que

$$\begin{aligned} \sup_{x \in [a, b]} |f(x) - f_N(x)| &= \sup_{x \in [a, b]} \lim_{n \rightarrow \infty} |f_n(x) - f_N(x)| \\ &\leq \sup_{n \geq N} \sup_{x \in [a, b]} |f_n(x) - f_N(x)| \\ &= \sup_{n \geq N} \|f_n - f_N\|_c < \varepsilon. \end{aligned} \quad (7.2)$$

Por lo tanto si pudiésemos probar que $f \in C[a, b]$ entonces tendríamos que $\|f - f_n\|_c \rightarrow 0, \quad n \rightarrow \infty$ y por lo tanto que $\{f_n\} \rightarrow f$ en $C[a, b]$ y el teorema estaría probado.

Sea $x \in [a, b]$ cualquiera y dado, probaremos que f es continua en x y así en todo $[a, b]$. Sea $\varepsilon > 0$, queremos encontrar un δ tal que $|x - y| < \delta$ implique $|f(x) - f(y)| < \varepsilon$. Tomemos N tal que $\|f - f_N\|_c < \varepsilon/3$ y δ tal que $|x - y| < \delta$ implique $|f_N(x) - f_N(y)| < \varepsilon/3$ [Esto es posible pues las $f_N(x)$ son continuas en $[a, b]$]. Luego $|x - y| < \delta$ implica

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| \\ &< \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon = \varepsilon. \end{aligned} \quad (7.3)$$

y por lo tanto la continuidad de $f(x)$. Como $[a, b]$ es un conjunto compacto (cerrado y acotado) luego f es acotada en $[a, b]$ y por lo tanto pertenece a $C[a, b]$ ♠

Definición: Sea $T : V \rightarrow V$ un mapa de un espacio de Banach V en sí mismo. Diremos que T es una **contracción** si existe $\lambda < 1$ tal que :

$$\|T(\mathbf{x}) - T(\mathbf{y})\|_V \leq \lambda \|\mathbf{x} - \mathbf{y}\|_V. \quad (7.4)$$

Ejemplos:

- a) El mapa lineal en l^2 ; $A\{x_i\} = \left\{ \frac{x_i}{i+1} \right\}$.
 b) El Mapa de \mathbb{R}^2 en \mathbb{R}^2 que manda al punto (x, y) en el punto $(x_0 + x/2, y/2)$.
 c) Una función Lipschitz cualquiera de \mathbb{R} en \mathbb{R} con módulo de continuidad (k) menor que uno.

La propiedad importante de estos mapas es el siguiente teorema

Teorema 7.2 Sea $T : V \rightarrow V$ una contracción. Luego existe un único $v \in V$ tal que $T(v) = v$, además la sucesión $T^n(u)$ tiende a v cualquiera sea u .

Prueba: Sea $\|T(u) - u\| = d$, luego $\|T^{n+1}(u) - T^n(u)\| \leq \lambda^n d$ y si $m > n$

$$\begin{aligned} \|T^m(u) - T^n(u)\| &= \|T^m(u) - T^{m-1}(u) + T^{m-1}(u) - T^{m-2}(u) + \dots + T^n(u)\| \\ &\leq \|T^m(u) - T^{m-1}(u)\| + \|T^{m-1}(u) - T^{m-2}(u)\| + \dots \\ &\leq d \sum_{n=0}^{m-n} \lambda^n. \end{aligned} \quad (7.5)$$

Como $\sum_{n=0}^{\infty} \lambda^n$ converge $\{T^n(u)\}$ es una sucesión de Cauchy. Pero V es completo y por lo tanto existe $v \in V$ tal que $\lim_{n \rightarrow \infty} T^n u = v$.

Como toda contracción es continua tenemos que

$$T(v) = T(\lim_{n \rightarrow \infty} T^n(u)) = \lim_{n \rightarrow \infty} T^{n+1}(u) = v.$$

Solo resta probar que v es único. Supongamos por contradicción que exista $w \in V$ distinto de v y tal que $T(w) = w$. Pero entonces $\|w - v\|_V = \|T(w) - T(v)\|_V \leq \lambda \|w - v\|_V$ lo que es una contradicción ya que $\lambda \neq 1$ ♠

Ejercicio: Sea $T : B_{R,x_0} \rightarrow B_{R,x_0}$, $B_{R,x_0} \in V$ una bola cerrada de radio R alrededor de x_0 , una contracción ¹. Probar para este caso las mismas afirmaciones que en el teorema anterior.

Prueba de los puntos i) y ii) del teorema fundamental. Como solo queremos ver existencia y unicidad local, es decir solo en un entorno de un punto p de M es suficiente considerar el sistema en \mathbb{R}^n . Esto nos permitirá usar la norma euclídea allí presente. Para ver esto tomemos una carta (U, φ) con $p \in U$ y por simplicidad $\varphi(p) = 0 \in \mathbb{R}^n$. Usando este mapa podemos trasladar el campo vectorial v en M a un campo vectorial \tilde{v} definido en un entorno del cero en \mathbb{R}^n . Allí trataremos el problema de encontrar sus curvas integrales $g(t, x)$ que pasan por el punto $x \in \varphi(U)$ al tiempo $t = 0$. Luego por medio del mapa φ^{-1} obtendremos en M las familias monoparamétricas de difeomorfismos $g^t(q)$, $q \in U$, que serán tangentes en todo punto al campo vectorial v .

Con esto en mente, solo resta por ver que para $R > 0$ y $\varepsilon > 0$ suficientemente pequeños existen curvas integrales, $g(t, x) : [0, \varepsilon] \times B_R = \{x \in \mathbb{R}^n \mid \|x\|_V < R\} \rightarrow \mathbb{R}^n$, del vector \tilde{v} , es decir mapas satisfaciendo

$$\frac{dg(t, x)}{dt} = \tilde{v}(g(t, x)) \quad , \quad g(0, x) = x, \quad (7.6)$$

con $g(t, x)$ continua con respecto al segundo argumento, es decir con respecto a la condición inicial. Por la suposición de que v es Lipschitz ², tenemos que existe $k > 0$ tal que $\forall x, y \in B_R$

$$\|\tilde{v}(x) - \tilde{v}(y)\|_V < k \|x - y\|_V. \quad (7.7)$$

Consideremos ahora el espacio de Banach $C([0, \varepsilon] \times B_R)$ que consiste en todos los mapas continuos (en t y x) de $[0, \varepsilon] \times B_R$ en \mathbb{R}^n con la norma

$$\|h\|_C = \sup_{\substack{x \in B_R \\ t \in [0, \varepsilon]}} \|h(t, x)\|_V, \quad (7.8)$$

¹Acomode también la definición de una contracción para este caso.

²Para probar la existencia y unicidad de soluciones locales solo es necesario suponer que v es Lipschitz.

Sea el mapa de la bola de radio R en $C([0, \varepsilon] \times B_R)$ en sí misma dado por,

$$T(h) = \int_0^t \tilde{v}(x + h(\tau, x)) d\tau . \quad (7.9)$$

Para que este mapa esté bien definido supondremos R lo suficientemente pequeño como para que \tilde{v} esté definido y satisfaga la condición de Lipschitz en B_{2R} y ε menor que R/C , donde $C = \max_{x \in B_{2R}} |\tilde{v}(x)|$, de modo que si $\|h\|_C < R$ luego $\|x + h(\tau, x)\|_V < 2R \quad \forall \tau \in [0, \varepsilon]$ y por lo tanto $\|T(h)\|_C < R$ en todo $C([0, \varepsilon] \times B_R)$. [Ver figura 7.1.]

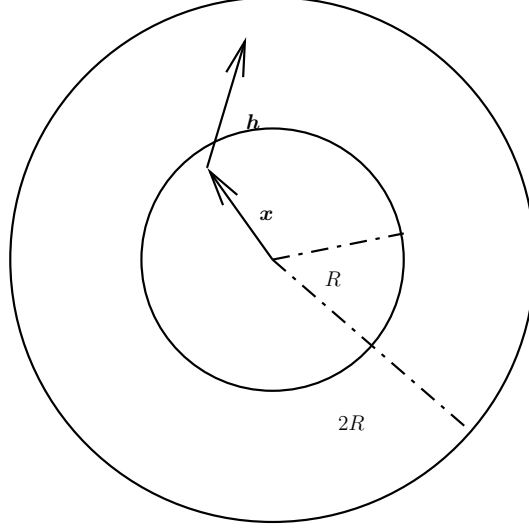


Figura 7.1: Entornos usados en la prueba del Teorema Fundamental.

Lema 7.1 Si ε es suficientemente pequeño entonces T es una contracción.

Prueba:

$$\begin{aligned} \|T(h_1) - T(h_2)\|_V &= \int_0^t \|\tilde{v}(x + h_1(\tau, x)) - \tilde{v}(x + h_2(\tau, x))\|_V d\tau \\ &\leq \int_0^t k \|h_1(\tau, x) - h_2(\tau, x)\|_V d\tau \leq k \varepsilon \|h_1 - h_2\|_C \end{aligned} \quad (7.10)$$

y por lo tanto

$$\|T(h_1) - T(h_2)\|_C \leq k \varepsilon \|h_1 - h_2\|_C \quad \forall h_1, h_2 \in C([0, \varepsilon] \times B_R). \quad (7.11)$$

Tomando $\varepsilon < 1/k$ completamos la prueba del Lema.

Este Lema y el Teorema 7.2 nos aseguran que existe un único mapa $h(t, x)$ –el punto fijo de T – satisfaciendo

$$h(t, x) = T(h(t, x)) = \int_0^t \tilde{v}(x + h(\tau, x)) d\tau. \quad (7.12)$$

Sea $g(t, x) \equiv x + h(t, x)$, esta función es continua en ambos argumentos –ya que $h(t, x) \in C([0, \varepsilon] \times B_R)$ – y por construcción continuamente diferenciable en t –ya que satisface (7.12)–. Diferenciando (7.12) con respecto a t vemos que $g(t, x)$ satisface la ecuación (7.6) y su condición inicial, lo que completa la prueba de los puntos *i*) y *ii*) del Teorema Fundamental. ♠

7.1. Problemas

Problema 7.1 *Vea que l^2 la bola unidad no es compacta. Ayuda: encuentre una sucesión infinita en la bola unidad que no tiene punto de acumulación.*

Problema 7.2 *Pruebe que la condición*

$$\|A(\mathbf{x}) - A(\mathbf{y})\| < \|\mathbf{x} - \mathbf{y}\| \quad (7.13)$$

no es suficiente para garantizar la existencia de un punto fijo. Ayuda: construya un contraejemplo. Hay algunos muy simples usando funciones de la recta en sí misma.

Problema 7.3 *Sea $f : [a, b] \rightarrow [a, b]$ una función Lipschitz con constante menor que la unidad en todo el intervalo $[a, b]$. Demuestre usando el teorema del punto fijo para contracciones que la ecuación $f(x) = x$ siempre tiene una solución. Grafique en un diagrama $(y = f(x), x)$ la sucesión dada por $x_i = f(x_{i-1})$ para el caso de f con pendiente positiva y menor que uno en todo punto. ¿Qué sucede cuando la pendiente se hace mayor que uno en algún intervalo?*

Problema 7.4 *Sea $g : [a, b] \rightarrow \mathbb{R}$ una función continuamente diferenciable tal que $g(a) < 0$, $g(b) > 0$ y $0 < c_1 < g'(x)$. Use el teorema del punto fijo de las contracciones para probar que existe una única raíz $g(x) = 0$ en el intervalo. Ayuda: defina la función $f(x) = x - \lambda g(x)$ para una constante λ convenientemente elegida y encuentre un punto fijo: $f(x) = x$. Note que la sucesión aproximante es la del método de Newton para obtener raíces.*

8.1. Introducción

Esta área de la matemática estudia los espacios funcionales, es decir los espacios cuyos elementos son ciertas funciones. Usualmente éstos son de dimensión infinita. Así es que sus resultados principales forman dos clases. Algunos son resultados generales, válidos para espacios vectoriales más generales que aquellos cuyos elementos son funciones. Éstos tienen un carácter geométrico o topológico el cual trataremos de rescatar en todo momento. Los otros son resultados particulares que nos relacionan distintos espacios funcionales entre sí. Estos son más cercanos a los resultados del análisis usual. Presentaremos aquí resultados de ambos tipos ya que de su conjunción obtendremos algunos aspectos de la teoría de las EDP. Por razones de brevedad solo consideraremos los puntos que nos serán útiles o aquellos para los que se requiera un mínimo de esfuerzo extra para obtenerlos y su importancia cultural así lo justifique.

8.2. Completando un Espacio Normado

En la primera parte de este curso introdujimos los espacios de Banach, es decir espacios vectoriales con una norma definida en ellos y que eran completos con respecto a ella. Allí también probamos que el conjunto de funciones continuas en un intervalo $[a, b] \in \mathbb{R}$ con la norma

$$\|f\|_1 = \text{Sup}_{x \in [a, b]} \{|f(x)|\} \quad (8.1)$$

era completo y por lo tanto Banach.

Cabe preguntarse si dado un espacio vectorial normado W es posible *engordarlo*, es decir agregarle vectores, y así hacerlo completo. Note que eso es lo que se hace con los racionales \mathbb{Q} (el cual es un espacio vectorial si solo permitimos la multiplicación de sus elementos por números racionales!). El espacio engordado es en este caso el de los reales. La respuesta es afirmativa y está dada por el siguiente teorema.

Teorema 8.1 *Sea W un espacio vectorial normado. Luego existe un espacio de Banach V y un mapa lineal continuo $\varphi : W \rightarrow V$ tal que $\|\varphi(w)\|_V = \|w\|_W$ y $\varphi[W]$ es un subespacio denso de V , es decir la clausura de $\varphi[W]$ es V .*

Prueba: Los detalles de esta pueden ser encontrados en, por ejemplo [Y], pag. 56. Aquí solo daremos las ideas. Si W es completo entonces tomamos $V = W$ y $\varphi = id$. Por lo tanto supondremos que W no es completo. Entonces habrán sucesiones de Cauchy $\{w_n\}$ en W que no convergen a ningún punto de W . La idea es tomar estas sucesiones como nuevos puntos con los cuales engordar W . Como muchas sucesiones distintas podrían tender a un mismo punto, a los efectos de no engordar demasiado a W , todas ellas deberán ser consideradas como un solo elemento. Esto se logra tomando clases equivalentes de sucesiones como elementos de V . Diremos que dos sucesiones, $\{w_n\}$, $\{w'_n\}$,

son equivalentes si su diferencia tiende al elemento cero,

$$\lim_{n \rightarrow \infty} \|w_n - w'_n\|_W = 0. \quad (8.2)$$

Como el conjunto de sucesiones de Cauchy de elementos de W forma un espacio vectorial, el conjunto de clases equivalentes de sucesiones también es un espacio vectorial, este será V . Este espacio hereda una norma naturalmente de W , dada por,

$$\|\{w_n\}\|_V = \lim_{n \rightarrow \infty} \|w_n\|_W, \quad (8.3)$$

la cual es claramente independiente de la sucesión particular que uno elija, para calcularla, dentro de la clase equivalente. Se puede probar fácilmente que con esta norma V es completo y por lo tanto Banach.

Ejercicio: Pruebe que la sucesión de Cauchy de sucesiones de Cauchy, $\{\{w_n\}_N\}$ converge en esta norma a la sucesión $\{\bar{w}_n\} \equiv \{\{w_n\}_n\}$.

¿Cuál es el mapa φ ? Este toma un elemento $w \in W$ y nos da un elemento en V , es decir una clase equivalente de sucesiones. Esta es la clase equivalente que converge a w y un representante es, por ejemplo,

$$\{w_n\} = (w, w, w, \dots). \quad (8.4)$$

El teorema anterior nos dice que siempre podemos completar un espacio vectorial normado y esto de una manera esencialmente única. De esta forma entonces se puede hablar de completar un espacio normado W en uno V . Como W es denso en V y el mapa φ es continuo todas las propiedades que son continuas en W valen automáticamente en V . El teorema anterior nos dice también que los elementos del espacio completado pueden tener un carácter muy diferente al de los elementos del espacio original. Debido a esto es que en la práctica uno debe ser muy cuidadoso al momento de atribuirle propiedades a los elementos de un dado espacio de Banach. Como ejemplo de lo anterior veremos la integral de Lebesgue.

8.3. *Integral de Lebesgue

La integral de Lebesgue es una extensión de la integral de Riemann a una clase de funciones más general que aquella donde la de Riemann está definida. Tenemos dos maneras de definirla, uno como la norma de un espacio de Banach completado, cuyos elementos son entonces **las funciones integrables en el sentido de Lebesgue**. La otra manera es usando un proceso de límite similar al empleado para definir la integral de Riemann. Veremos las dos.

Sea W el espacio de funciones continuas en el intervalo $[a, b]$ y sea

$$\|f\|_w = \int_a^b |f(x)| \, dx \quad (8.5)$$

su norma, donde la integral es en el sentido de Riemann. Esta definición tiene sentido ya que los elementos de W son funciones continuas y por lo tanto la integral está bien definida. Note también que esta es una norma ya que como f es continua si la integral de su módulo es cero, su módulo, y así f , es también cero. Sea $L^1([a, b])$ el espacio completado de W . Este es el espacio de **funciones integrables de Lebesgue**. ¿Qué funciones están allí?¹ Sea $\{f_n\}$ la sucesión dada por el gráfico de la fig. 8.1.

¹En realidad las funciones no son propiamente elementos de L^1 sino que éstos son la imagen por el mapa φ (definido en la sección anterior) de éstas.

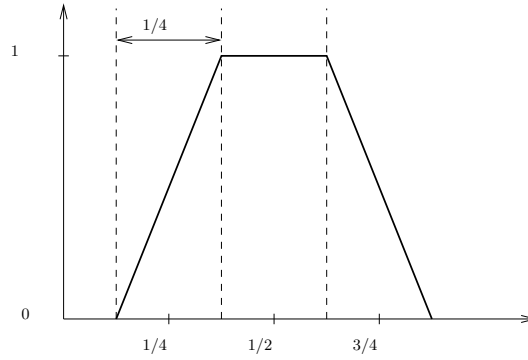


Figura 8.1: Una sucesión de Cauchy.

Esta sucesión es de Cauchy en $L^1([0, 1])$ y por lo tanto converge a un elemento de $L^1([0, 1])$, la función,

$$f(x) = \begin{cases} 0 & 0 \leq x \leq 1/4, \ 3/4 \leq x \leq 1 \\ 1 & 1/4 < x < 3/4. \end{cases} \quad (8.6)$$

Esto no es de sorprender ya que aunque esta función no es continua ésta es integrable aún en el sentido de Riemann. Sea ahora la sucesión $\{f_n\}$ dada por el gráfico de la fig. 8.2,

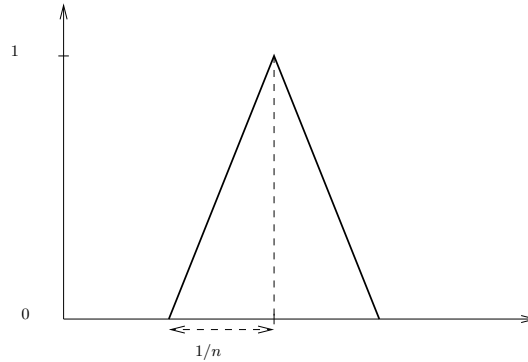


Figura 8.2: Otra sucesión de Cauchy.

esta también es de Cauchy y tiende a la función,

$$f(x) = \begin{cases} 0 & 0 \leq x < 1/2 < x \leq 1 \\ 1 & x = 1/2, \end{cases} \quad (8.7)$$

la cual sí es muy extraña. Cualquiera sea el sentido de la integral de Lebesgue la integral de esta función es de esperar que sea cero, y de hecho lo es ya que

$$\lim_{n \rightarrow \infty} \int_0^1 |f_n(x)| dx = 0. \quad (8.8)$$

Pero entonces la norma de f es cero y por lo tanto parecería ser tenemos una contradicción, a menos que f sea cero. La resolución de este problema consiste en notar que cuando completamos el espacio W tomamos como sus elementos a ciertas clases equivalentes. La función descripta arriba

está en la clase equivalente correspondiente al elemento cero. Como veremos luego los elementos de L^1 son funciones, pero solo definidas en **casi todos los puntos**.

El segundo método consiste en definir la integral de Lebesgue de forma similar a como se define la integral de Riemann. Para ello debemos definir lo que se llama una **medida** en ciertos subconjuntos de \mathbb{R} , es decir una función de estos subconjuntos en los reales positivos, que generaliza el concepto de extensión (o medida) de un abierto (a, b) , $\mu((a, b)) = b - a$.

Una vez introducido este concepto de medida la integral de Lebesgue de una función positiva $f : \mathbb{R} \rightarrow \mathbb{R}^+$ se define como,

$$\int f(x) dx = \lim_{n \rightarrow \infty} \sum_n (f) = \lim_{n \rightarrow \infty} \sum_{m=0}^{\infty} \frac{m}{n} \mu \left(f^{-1} \left[\left[\frac{m}{n}, \frac{m+1}{n} \right) \right] \right), \quad (8.9)$$

donde hemos usado el mismo símbolo para denotar la integral de Lebesgue que el normalmente usado para la integral de Riemann, esto es natural ya que como vimos en la primera definición si una función es integrable en el sentido de Riemann lo es también en el sentido de Lebesgue y el valor de las integrales coincide.

La interpretación de esta definición es la siguiente: se divide la imagen de f en intervalos regulares de longitud $\frac{1}{n}$, se considera la imagen por f^{-1} de estos intervalos, se los mide con μ y estas medidas se suman convenientemente. Ver figura. Finalmente se toma el límite n yendo a infinito, es decir los intervalos yendo a cero. Note que $\sum_{2n}(f) \geq \sum_n(f)$ y por lo tanto $\lim_{n \rightarrow \infty} \sum_n(f) = \sup_n \{\sum_n(f)\}$ existe (pudiendo ser infinito).

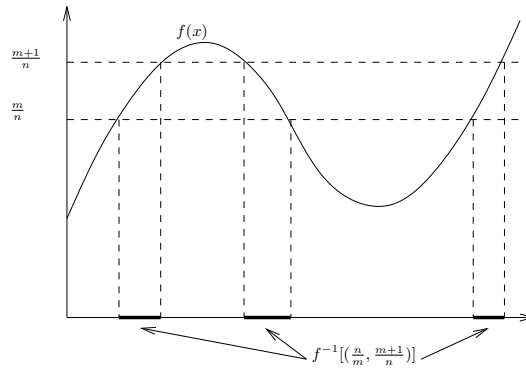


Figura 8.3: Integral de Lebesgue.

¿Para qué funciones está definida esta operación? La condición de que f fuese positiva no es realmente una restricción ya que cualquiera sea ésta siempre se la puede escribir como $f = f_+ - f_-$ con f_+ y f_- ambas positivas y la integral es una operación lineal [lo cual no es obvio de la definición dada arriba]. La condición que sí es restrictiva es que $f^{-1}(A)$, con A abierto, sea un subconjunto medible, ya que como veremos si le pedimos a la función medida que satisfaga ciertas propiedades naturales luego no todo subconjunto de \mathbb{R} puede estar en el dominio de esta función. Para estudiar esta restricción debemos tener en claro qué propiedades queremos que la medida cumpla, encontrar la colección de subconjuntos de \mathbb{R} que son medibles y finalmente definir la medida. Estas propiedades definirán la noción de medida, o más específicamente de espacio medible, ya que las propiedades que le asignaremos dependen tanto de la medida como de su dominio de definición.

Definición: Un **espacio medible** consta de una terna (X, M, μ) donde X es un conjunto (el dominio de las funciones a integrar), M una colección de subconjuntos de X llamados los **subconjuntos medibles** de X y μ es una función (llamada la **medida**) de M en $\mathbb{R}^* = \mathbb{R}^+ \cup \{\infty\}$ satisfaciendo las siguientes condiciones:

- i) $\emptyset \in M$ y $\mu(\emptyset) = 0$.
- ii) Si $A \in M$ luego A^c (el complemento de A en X) también está en M .
- iii) Si $A_i \in M, i = 1, 2, \dots$, luego $\bigcup_i A_i \in M$.
- iv) Si $A_i \in M, i = 1, 2, \dots$ y $A_i \cap A_j = \emptyset$ si $i \neq j$ luego $\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i)$.

Intuitivamente los conjuntos medibles son los conjuntos que admiten una noción de *longitud* o *área* y su medida es el valor de esta *longitud* o *área*.

Ejercicio: Muestre que:

1. $X \in M$
2. Si $A_i \in M, i = 1, 2, \dots$ luego $\bigcap_i A_i \in M$. Esta es la razón por la cual se pide ii), ya que esperaríamos que $\mu\left(\bigcap_i A_i\right) \leq \sum_i \mu(A_i)$.
3. Si $A \subset B$ luego $\mu(A) \leq \mu(B)$.

Ejemplos:

1. Sea X un conjunto cualquiera y sea M la colección de todos los subconjuntos de X . Sea μ tal que $\mu(\emptyset) = 0$ y $\mu(A) = \infty$ para todo $A \in M$ no vacío.
2. Sea $M = \{\emptyset, X\}$ y sea μ tal que $\mu(\emptyset) = 0$ y $\mu(X) = 7$.
3. Sea $X = \mathbb{Z}^+ = \{\text{números enteros positivos}\}$ y sea $\{x_i\}, x_i \in \mathbb{R}^+$, una sucesión. Sea M la colección de todos los subconjuntos de \mathbb{Z}^+ y $\mu(A) = \sum_{i \in A} (x_i)$.
El siguiente no es un espacio medible (¿por qué?) pero nos será útil para construir luego el que deseamos.
4. Sea $X = \mathbb{R}$ y $M = \mathcal{J}$ el conjunto de todas las uniones (contables) de intervalos abiertos disjuntos, es decir un elemento, I , de \mathcal{J} es un subconjunto de \mathbb{R} de la forma,

$$I = \bigcup_i (a_i, b_i), \quad (8.10)$$

con $a_1 \leq b_1 < a_2 \leq b_2 < a_3 \dots$. Sea $\mu \equiv m : \mathcal{J} \rightarrow \mathbb{R}^*$ definida como,

$$m(I) = \sum_i (b_i - a_i). \quad (8.11)$$

Ejercicio: Muestre que el tercer ejemplo es un espacio medible.

Como se ve de estos ejemplos la noción de espacio medible es amplia. Distintos ejemplos de espacios medibles aparecen en diversas ramas de la física. De ahora en más nos restringiremos a un espacio medible y éste es el de Lebesgue, el cual construiremos a partir del cuarto ejemplo.

Sea $X = \mathbb{R}$, \bar{M} todo subconjunto de X y $\bar{\mu} : \bar{M} \rightarrow \mathbb{R}^*$ dada por,

$$\begin{aligned} \bar{\mu}(A) &= \inf_{\substack{I \in \mathcal{I} \\ A \subset I}} (m(I)) \end{aligned} \quad (8.12)$$

Es decir, dado un subconjunto de \mathbb{R} , A , consideramos todos los elementos de \mathcal{J} tales que A está contenido en éstos, calculamos su medida y tomamos el ínfimo sobre todos los elementos de \mathcal{J} .

Ejemplos:

1. Sea $A = (0, 1)$, luego un candidato es $I_1 = (-1, 1) \cup (3, 5)$, $m(I_1) = 2 + 2 = 4$, pero también tenemos $I_2 = (0, 1)$ con $m(I_2) = 1$ y claramente $\bar{\mu}(A) = 1$.
2. Sea $B = 1 \cup 2 \cup 3 \cup \dots$ luego $\bar{\mu}(B) = 0$ ya que podemos cubrir B con $I = (1 - \varepsilon, 1 + \varepsilon) \cup (2 - \varepsilon, 2 + \varepsilon) \cup \dots$.

Como vemos esta terna $(\mathbb{R}, \bar{M}, \bar{\mu})$ parece tener las condiciones deseadas para ser la medida que buscamos, pero en realidad ni siquiera es un espacio medible!

Contraejemplo: Sea (S^1, M, μ) , donde S^1 es el círculo, un espacio medible con μ una medida finita e invariante ante translaciones, es decir $\mu(A) = \mu(A_r)$ donde $A_r = \{a + r | a \in A\}$ -note que nuestra $\bar{\mu}$ lo es, pero este resultado es mucho más general-. Luego M no puede ser la colección de todos los subconjuntos de S^1 .

La idea es encontrar un subconjunto A de S^1 tal que si suponemos que sea medible obtenemos una contradicción. Para ello pensamos a S^1 como \mathbb{R} con sus extremos identificados ($0=1$). Introducimos ahora una relación de equivalencia: Diremos que dos puntos de S^1 son equivalentes $a \approx b$ si $a - b$ es un número racional.

Ejercicio: Muestre que es una relación de equivalencia.

Construimos A tomando exactamente un elemento de cada clase equivalente -note que hay infinitas maneras de elegir un conjunto A - Supongamos que A es medible con $\mu(A) = \alpha \in \mathbb{R}^+$ y sea $A_r = \{a | (a + r) \pmod{1} \in A\}$, con r racional, es decir una translación por $-r$ de A . Luego es fácil ver que si $r \neq r'$ luego $A_r \cap A_{r'} = \emptyset$, es decir los A_r son disjuntos [Si $x \in A_r$ y $x \in A_{r'}$ luego $x = a + r = b + r'$, con $a, b \in A$, pero entonces $a - b = r - r' \in \mathbb{Q}$ lo que es una contradicción.] y que $S^1 = \bigcup_{r \in \mathbb{Q}} A_r$ [Sea $x \in S^1$, luego x pertenece a alguna de las clases equivalentes en que hemos

separado a S^1 , pero entonces existe $a \in A$ y $r \in \mathbb{Q}$ tal que $a + r = x$, o sea $x \in A_r$]. Como por hipótesis $\mu(A_r) = \mu(A) = \alpha$ entonces,

$$1 = \mu(S^1) = \mu\left(\bigcup_{r \in \mathbb{Q}} A_r\right) = \sum_{r \in \mathbb{Q}} \mu(A_r) = \sum_{r \in \mathbb{Q}} \alpha, \quad (8.13)$$

lo cual lleva a una contradicción ya que si $\alpha = 0$ luego la sumatoria da cero y si $\alpha \neq 0$ la sumatoria da infinito.

Este contraejemplo nos dice entonces que debemos restringir M a ser un subconjunto de \bar{M} si queremos tener una medida. Hay muchas maneras de caracterizar esta restricción una de ellas está dada por el siguiente teorema (que no probaremos).

Teorema 8.2 Sea M el subespacio de \bar{M} tal que si $A \in M$ luego

$$\bar{\mu}(E) = \bar{\mu}(A \cap E) + \bar{\mu}((X - A) \cap E) \quad \forall E \in \bar{M}. \quad (8.14)$$

Sea μ la restricción de $\bar{\mu}$ a M , luego (X, M, μ) es el espacio medible de Lebesgue.

Intuitivamente vemos que los conjuntos medibles son aquellos que cuando usados para separar otros conjuntos en dos partes dan una división que es **aditiva** con respecto a $\bar{\mu}$.

Como en general solo tenemos que $\bar{\mu}(E) \leq \bar{\mu}(A \cap E) + \bar{\mu}((X - A) \cap E)$ vemos entonces que los conjuntos no-medibles son aquellos cuyos puntos están distribuidos en X de tal forma que cuando uno trata de cubrir $A \cap E$ y $(X - A) \cap E$ con abiertos éstos se superponen tanto que en realidad uno puede obtener un ínfimo menor cubriendo directamente E con abiertos.

Existe un gran número de teoremas que nos dan información acerca de cuáles son los conjuntos medibles. Bastará decir que todos los conjuntos abiertos de \mathbb{R} (y muchos más) están en M y que si $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función continua luego $f^{-1}[M] \in M$.

Retornamos ahora a la integral de Lebesgue. Naturalmente diremos que f es **medible** si $f^{-1}[(a, b)] \in M$ para todo intervalo abierto $(a, b) \in \mathbb{R}$. Las funciones medibles tienen las siguientes propiedades:

Teorema 8.3 a) Si f y g son medibles y $\lambda \in \mathbb{R}$ luego $f + \lambda g$ es medible.

b) También son medibles fg , $\max\{f, g\}$ y $\min\{f, g\}$. Note que $|f| = \max\{f, -f\}$ y $f_{\pm} = \max_{\min}\{f, 0\}$.

c) Sea $\{f_n(x)\}$ una sucesión de funciones medibles que convergen punto a punto a $f(x)$, es decir $\lim_{n \rightarrow \infty} f_n(x) = f(x)$, luego $f(x)$ es también medible.

La primera parte de este teorema nos dice que el conjunto de funciones medibles es un espacio vectorial. Esto sigue siendo cierto si nos restringimos al espacio de funciones integrables en el sentido de Lebesgue, es decir f medible y $\int |f| dx < \infty$. Denotaremos a este espacio como \mathcal{L}_1 o $\mathcal{L}_1(\mathbb{R})$. Similarmente definiremos $\mathcal{L}_1[a, b]$ al espacio de funciones integrables $f : [a, b] \rightarrow \mathbb{R}$ donde en este caso la integral se define como antes pero extendiendo la función f a todo \mathbb{R} con valor cero fuera del intervalo $[a, b]$.

Como ya vimos antes para hacer de \mathcal{L}_1 un espacio normado debemos tomar como sus elementos las clases equivalentes de funciones donde diremos que $f, g \in \mathcal{L}_1$ son **equivalentes** si $\int |f - g| dx = 0$. Esto es equivalente a decir que el subconjunto de \mathbb{R} donde f es distinta de g es de medida cero, o en otras palabras que f es igual a g en **casi todo punto**.

Denotaremos al espacio de clases equivalentes de funciones integrables (Lebesgue) con L_1 y sus elementos con letras tildadas. Note que un elemento \tilde{f} de L_1 es una clase equivalente de funciones y por lo tanto en general el valor de \tilde{f} en x , $\tilde{f}(x)$, no tiene sentido alguno ya que podemos tener funciones en la clase equivalente \tilde{f} , f_1 y f_2 con $f_1(x) \neq f_2(x)$ para algún $x \in \mathbb{R}$. Por lo tanto debemos ser precavidos al respecto.

¿Es este espacio L_1 el mismo que habíamos obtenido anteriormente? La respuesta es sí y sigue trivialmente de los siguientes teoremas.

Teorema 8.4 (Riez-Fisher) : L_1 es completo.

Teorema 8.5 $C^1[a, b]$ es denso en $L^1[a, b]$, es decir $L^1[a, b]$ es el espacio completado de $C^1[a, b]$.

8.4. Espacios de Hilbert

Los espacios de Hilbert son espacios de Banach² cuya norma proviene de un producto escalar, es decir de un mapa (\cdot, \cdot) de $V \times V \rightarrow \mathbb{C}$ satisfaciendo:

$$\begin{aligned} i) \quad (x, y + cz) &= (x, y) + c(x, z), \\ (x + cy, z) &= (x, z) + \bar{c}(y, z) \\ \text{para cualquier } x, y, z \text{ en } V \text{ y } c \text{ en } \mathbb{C}. \end{aligned}$$

$$ii) \quad \overline{(x, y)} = (y, x)$$

$$iii) \quad (x, x) \geq 0 \quad (0 \text{ si } x = 0).$$

²De ahora en más y esencialmente por la misma razón que dimos en el caso del teorema de la forma canónica de Jordan, consideraremos espacios vectoriales complejos.

La primera parte de la condición *i)* indica que el mapa producto escalar es lineal con respecto a su segundo argumento. La segunda parte indica que es lineal en el primer argumento, excepto por el hecho de que se toma el complejo conjugado del escalar. Se dice que este mapa es **anti-lineal** con respecto al primer argumento. La condición *ii)* nos dice que el mapa es todo lo simétrico que puede ser, dado que en el primer argumento es anti-lineal. Esta condición nos garantiza que (x, x) es un número real y, junto con *iii)*, que es no-negativo.

Ejercicio: Pruebe que dando un tensor de tipo $(2, 0)$, $t(\cdot, \cdot)$, simétrico, real y positivo definido, entonces $(x, y) \equiv t(\bar{x}, y)$ es un producto escalar.

La norma que este producto escalar induce es simplemente la función, $\|\cdot\|_V : V \rightarrow \mathbb{R}^+$ dada por,

$$\|x\|_V = \sqrt{(x, x)}. \quad (8.15)$$

Que esta es en realidad una norma sigue de los siguientes lemas:

Lema 8.1 (Desigualdad de Schwarz) : $|(x, y)| \leq \|x\| \|y\|$.

Prueba: Para cualquier $x, y \in V$, $\lambda \in \mathbb{R}$ tenemos por *iii)*,

$$\begin{aligned} 0 \leq (y + \lambda(x, y)x, y + \lambda(x, y)x) &= \|y\|^2 + \lambda^2 |(x, y)|^2 \|x\|^2 + \\ &\quad + \lambda \overline{(x, y)}(x, y) + \lambda (x, y)(y, x) \\ &= \|y\|^2 + 2\lambda |(x, y)|^2 + \lambda^2 |(x, y)|^2 \|x\|^2. \end{aligned} \quad (8.16)$$

Como esta relación debe valer para todo $\lambda \in \mathbb{R}$ entonces el discriminante del polinomio en λ de la derecha debe ser no-positivo, es decir,

$$4 |(x, y)|^4 - 4 \|x\|^2 \|y\|^2 |(x, y)|^2 \leq 0. \quad (8.17)$$

lo que nos da la desigualdad buscada.

Lema 8.2 (Desigualdad triangular) : $\|x + y\| \leq \|x\| + \|y\|$.

Prueba:

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2 \operatorname{Re}(x, y) \\ &\leq \|x\|^2 + \|y\|^2 + 2 |(x, y)| \\ &\leq \|x\|^2 + \|y\|^2 + 2 \|x\| \|y\| \\ &\leq (\|x\| + \|y\|)^2, \end{aligned}$$

Lo que nos da la desigualdad buscada.

Ejemplos:

1. C^n , el espacio vectorial de n-tuplas de números complejos.

Sea $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$, luego

$$(x, y) = \sum_{j=1}^n \bar{x}_j y_j.$$

2. l^2 , el espacio vectorial de sucesiones de números complejos $\{x_i\}$ tales que

$$\|\{x_j\}\|_2 \equiv \sqrt{\sum_{j=1}^{\infty} |x_j|^2} < \infty,$$

con el producto escalar,

$$(\{x_j\}, \{y_j\}) = \sum_{j=1}^{\infty} \bar{x}_j y_j.$$

Note que la desigualdad de Schwarz nos garantiza que el producto escalar está bien definido para todo par de vectores de l^2 .

Para asegurarnos que l^2 es un espacio de Hilbert debemos probar que es completo, es decir que toda sucesión de Cauchy (con respecto a la norma de l^2) converge a un elemento de l^2 .

Lema 8.3 l^2 es un espacio de Hilbert.

Prueba: Sea $\{\{x_i\}_N\}$ una sucesión de sucesiones. Que esta sea de Cauchy significa que dado $\varepsilon > 0$ existe \bar{N} tal que

$$\|\{x_i\}_N - \{x_i\}_M\|^2 = \sum_{i=1}^{\infty} |x_i^N - x_i^M|^2 < \varepsilon^2 \quad \forall \quad N, M > \bar{N}, \quad (8.18)$$

pero eso implica que para cada i

$$|x_i^N - x_i^M| < \varepsilon \quad (8.19)$$

es decir la sucesión de números complejos (en N, i fijo) $\{x_i^N\}$ es Cauchy. Pero el plano complejo es completo y por lo tanto para cada i , $\{x_i^N\}$ converge a un número complejo que denotaremos \bar{x}_i

Sea $\{\bar{x}_i\}$ la sucesión (en i) de estos números, probaremos ahora que $\{\bar{x}_i\} \in l^2$ y que $\{\{x_i\}_N\} \longrightarrow \{\bar{x}_i\}$ para $N \rightarrow \infty$.

Tomando el límite $N \rightarrow \infty$ vemos que si $M > \bar{N}$ luego

$$\sum_{j=1}^k |x_j^M - \bar{x}_j|^2 < \varepsilon^2$$

Pero esta es una sucesión (en k) de números reales que es acotada (por ε^2) y por lo tanto, tomando ahora el límite $k \rightarrow \infty$ vemos que $\{\{x_i\}_N\} - \{\bar{x}_i\} \in l^2$ y que si $\{\bar{x}_i\} \in l^2$ luego $\{\{x_i\}_N\} \rightarrow \{\bar{x}_i\}$ en norma. Pero

$$\|\{\bar{x}_i\}\|^2 = \|\{\{x_i\}_N\} + (\{\bar{x}_i\} - \{\{x_i\}_N\})\|^2 \leq \|\{\{x_i\}_N\}\|^2 + \|\{\bar{x}_i\} - \{\{x_i\}_N\}\|^2,$$

y por lo tanto que $\{\bar{x}_i\} \in l^2$ ♠

Este ejemplo y el que lo sigue son ejemplos clásicos a tener en cuenta, esencialmente todo espacio de Hilbert que trataremos es alguna de estas variantes.

3. L^2 (o H^0), el espacio de funciones medibles con cuadrado integrable en \mathbb{R} e identificadas entre sí, si su diferencia está en un conjunto de medida cero ($f \sim g$ si $\int |f - g|^2 dx < 0$). El producto escalar es $(f, g) = \int \bar{f} g dx$ y su norma obviamente $\|f\|_{H^0} = \sqrt{\int |f|^2 dx}$.

4. (Espacios de Sobolev) Sea la norma

$$\|f\|_{H^m}^2 = \int_{\Omega} \{|f|^2 + \sum_{i=1}^n |\partial_i f|^2 + \sum_{i,j=1}^n |\partial_i \partial_j f|^2 + \cdots + \underbrace{\sum_{i,j,\dots,k=1}^n |\partial_i \partial_j \dots \partial_k f|^2}_m\}, \quad (8.20)$$

donde las derivadas parciales son con respecto a un sistema cartesiano de coordenadas en \mathbb{R}^n . Definiremos el espacio de Sobolev de orden m como, $H^m(\Omega) = \{ \text{Completamiento del espacio de funciones } m \text{ veces diferenciables en } \Omega \subset \mathbb{R}^n \text{ con respecto a la norma } \|\cdot\|_{H^m} \}$

Es obvio cuál es el producto escalar correspondiente, además note que por definición H^m es completo. H^0 coincide con el definido en el ejemplo anterior, ya que las funciones continuas son densas en L^2 .

Como vemos de estos ejemplos los espacios de Hilbert son la generalización inmediata de \mathbb{R}^n (o C^n) a dimensión infinita donde hemos preservado la noción no solo de la magnitud de un vector sino también la del ángulo entre dos vectores. Esto hace que los espacios de Hilbert tengan propiedades más interesantes que las que tienen los espacios de Banach en general. La más interesante se refiere a los subespacios de H . Sea M un subespacio cerrado de H . Este subespacio hereda el producto escalar definido en H (simplemente restringiendo el mapa (\cdot, \cdot) a actuar solo en elementos de M) y por ser cerrado es completo, por lo tanto es también un espacio de Hilbert.

Ejemplos:

- Sea M el subespacio generado por el vector $(1, 0, 0)$ en C^3 , es decir todos los vectores de la forma $(c, 0, 0)$ con $c \in C$.
- Sea M el subespacio de H^0 que consiste de todas las funciones que se anulan en el intervalo $(0, 1)$ excepto en un subconjunto de medida nula. M es cerrado, ya que si una sucesión de Cauchy de funciones se anula en $(0, 1)$ luego la función límite (que existe pues H es completo) también se anula en dicho intervalo y por lo tanto está en M .
- Sea $M = C[a, b]$ el subespacio de funciones continuas de $H^0([a, b])$. Este no es cerrado y por lo tanto no es un espacio de Hilbert. [Muestre, encontrando una sucesión de Cauchy de funciones continuas que no tiene como límite una función continua, que este subespacio no es cerrado.]
- Sea K un subconjunto cualquiera de H y considere la intersección de todos los subespacios cerrados de H que contienen a K . Esta intersección nos da el subespacio más pequeño que contiene a K y se llama el subespacio generado por K . En el ejemplo a) $K = \{(1, 0, 0)\}$ genera el plano complejo que contiene a este vector. En el ejemplo c) $K = C[a, b]$ genera todo $H^0([a, b])$.

El concepto definido anteriormente no usa del producto escalar y por lo tanto también es válido para espacios de Banach en general (un subespacio cerrado de un espacio de Banach es también un espacio de Banach).

El producto escalar nos permite introducir el concepto de ortogonalidad y así definir el complemento ortogonal de M , es decir el conjunto,

$$M^\perp = \{x \in H \mid (x, y) = 0 \forall y \in M\}, \quad (8.21)$$

que tiene la siguiente propiedad,

Teorema 8.6 *Sea M un subespacio cerrado de un espacio de Hilbert H . Luego M^\perp es también un espacio de Hilbert. Además M y M^\perp son complementarios, es decir todo vector en H puede ser escrito de una única manera como la suma de un vector en M y otro en M^\perp .*

$$H = M \oplus M^\perp \quad (8.22)$$

Prueba:

Es inmediato que M^\perp es un subespacio vectorial y que es cerrado. [Si $\{x_i\}$ es una sucesión en M^\perp convergiendo a x en H luego $|(x, y)| = |(x - x_i, y)| \leq \|x - x_i\| \|y\| \rightarrow 0 \quad \forall y \in M$ y por lo

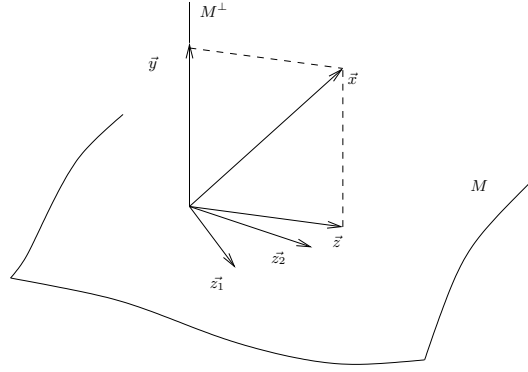


Figura 8.4: El espacio perpendicular.

tanto $x \in M^\perp$.] Solo debemos probar la complementaridad. Para ello, dado $x \in H$, buscaremos el elemento z en M más próximo a x , ver figura.

Sea $d = \inf_{w \in M} \|x - w\|_H$ y elijamos una sucesión $\{z_n\}$ en M tal que $\|x - z_n\|_H \rightarrow d$. Esto es posible por la definición de ínfimo.

Luego

$$\begin{aligned}
 \|z_n - z_m\|_H^2 &= \|(z_n - x) - (z_m - x)\|_H^2 \\
 &= 2\|z_n - x\|_H^2 + 2\|z_m - x\|_H^2 - \|(z_n - x) + (z_m - x)\|_H^2 \\
 &= 2\|z_n - x\|_H^2 + 2\|z_m - x\|_H^2 - 4\left\|\frac{z_n + z_m}{2} - x\right\|_H^2 \\
 &\leq 2\|z_n - x\|_H^2 + 2\|z_m - x\|_H^2 - 4d^2 \longrightarrow \\
 &\quad \xrightarrow[m \rightarrow \infty]{n \rightarrow \infty} 2d^2 + 2d^2 - 4d^2 = 0
 \end{aligned} \tag{8.23}$$

La segunda igualdad proviene de la llamada ley del paralelogramo, ver figura.

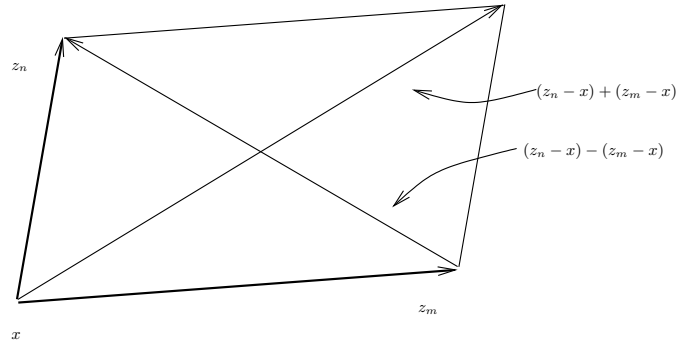


Figura 8.5: Ley del paralelogramo.

Esto nos dice que $\{z_n\}$ es Cauchy y por lo tanto que converge a un único z en M .

Sea $y = x - z$, solo resta ver que y está en M^\perp , es decir que $(y, v) = 0 \ \forall v \in M$. Sea $v \in M$ y $t \in \mathbb{R}$, luego

$$\begin{aligned} d^2 &\equiv \|x - z\|^2 \leq \|x - (z + tv)\|^2 = \|y - tv\|^2 \\ &= d^2 - 2t \operatorname{Re}(y, v) + t^2 \|v\|^2 \end{aligned} \quad (8.24)$$

y por lo tanto $-2t \operatorname{Re}(y, v) + t^2 \|v\|^2 \geq 0 \ \forall t$ lo que implica que $\operatorname{Re}(y, v) = 0$. Tomando it se obtiene que $\operatorname{Im}(y, v) = 0$ y se completa la prueba ♠

Problema 8.1 En la prueba solo usó la ley del paralelogramo,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (8.25)$$

Muestre que una norma la satisface si y solo si ésta proviene de un producto escalar. Ayuda: Utilice la denominada identidad de polarización para definir un producto interno a partir de la norma,

$$(x, y) = \frac{1}{4} \{ [\|x + y\|^2 - \|x - y\|^2] - \Im[\|x + \Im y\|^2 - \|x - \Im y\|^2] \} \quad (8.26)$$

Este teorema tiene importantes corolarios que veremos a continuación.

Corolario 8.1 $(M^\perp)^\perp = M$

Prueba:

Probaremos éste mostrando que $M \subset (M^\perp)^\perp$ y que $(M^\perp)^\perp \subset M$.

La primera inclusión es obvia ya que $(M^\perp)^\perp$ es el conjunto de vectores ortogonales a M^\perp el cual a su vez es el conjunto de vectores ortogonales a M .

La segunda inclusión sigue de que si descomponemos cualquier vector $x \in (M^\perp)^\perp$ en su parte en M y en parte en M^\perp , $x = x_1 + x_2$, $x_1 \in M$, $x_2 \in M^\perp$, la primera inclusión nos dice que $x_1 \in (M^\perp)^\perp$ pero $(M^\perp)^\perp$ y M^\perp son también complementarios y por lo tanto $x_2 = 0$ ♠

Este nos dice que tomando complementos no obtendremos más que un solo espacio de Hilbert extra.

Para formular el siguiente corolario es necesario introducir un nuevo concepto, el de espacio dual de un espacio de Hilbert. Podríamos definir el dual de un espacio de Hilbert de la misma forma que lo hicimos para espacios vectoriales de dimensión finita, es decir como el conjunto de mapas lineales de H en \mathbb{C} . El espacio vectorial así obtenido no hereda H ninguna propiedad interesante por ser éste un espacio demasiado grande. Para lograr un espacio más pequeño y con propiedades atractivas restringiremos los mapas lineales a aquellos que sean continuos. Es decir el **espacio dual** a H , H' será el conjunto de mapas lineales continuos de H en \mathbb{C} . ¿Qué mapas están en H' ? Note que si $y \in H$ luego el mapa $\varphi_y : H \rightarrow \mathbb{C}$ definido por $\varphi_y(x) = (y, x)$ es lineal y como $|\varphi(x)| \leq \|y\|_H \|x\|_H$ también es continuo. Este mapa nos da una correspondencia inyectiva (no canónica, ya que depende del producto escalar) entre los elementos de H y los de su dual. Vemos así que H está contenido naturalmente en H' .

Ejercicio: Medite sobre cuál es la norma natural en H' . Pruebe que con esa norma tenemos $\|\phi_y\|_{H'} = \|y\|_H$.

Problema 8.2 Sea $\phi : H \rightarrow \mathbb{C}$ un mapa lineal. Muestre que la continuidad del mapa en el origen asegura la continuidad del mapa en todo punto.

Si la dimensión de H es finita luego sabemos que H' tendrá la misma dimensión y entonces tendremos una correspondencia (no canónica) uno a uno entre vectores y covectores. En principio esto no tiene por qué ser así en el caso de dimensión infinita y efectivamente si usamos una definición análoga y definimos el dual de un espacio de Banach en general no habrá ninguna relación entre éste y el espacio original. En el caso de espacios de Hilbert todo es más simple, como muestra el siguiente corolario.

Corolario 8.2 (Teorema de representaciones de Riez) : Sea $\varphi \in H'$, luego existe un único $y \in H$ tal que $\varphi(x) = (y, x) \forall x \in H$, es decir $H \approx H'$ en el sentido que existe un mapa natural invertible entre H y H' .

Prueba: Sea $M = \{x \in H \text{ tal que } \varphi(x) = 0\}$. Como φ es lineal, M es un subespacio de H , como φ es continuo este es cerrado. Por el teorema anterior M^\perp es un espacio de Hilbert y complementa a M . Si M^\perp es cero luego $M = H$ y φ es el mapa cero e $y = 0$ será su representante en H . Supongamos entonces $M^\perp \neq \{0\}$ y elijamos $w \in M^\perp$ tal que $\varphi(w) = 1^3$. Para cualquier $v \in M^\perp$ $v - \varphi(v)w \in M^\perp$, pero $\varphi(v - \varphi(v)w) = 0$ y por lo tanto también $v - \varphi(v)w \in M$. Así concluimos que $v - \varphi(v)w = 0 \forall v \in M^\perp$ o sea que $v = \varphi(v)w \forall v \in M^\perp$, es decir que M^\perp es unidimensional. Veamos ahora que $\varphi(x) = (\frac{w}{\|w\|^2}, x) \forall x \in H$. En efecto por el teorema anterior $x = \alpha w + y$ con $\alpha \in C$ e $y \in M$ y por lo tanto, $\varphi(x) = \varphi(\alpha w + y) = \alpha \varphi(w) + \varphi(y) = \alpha$, pero por otro lado, $(\frac{w}{\|w\|^2}, x) = \alpha$ ♠

Para enunciar el tercer corolario necesitamos el concepto de base ortonormal. Una **base ortonormal** de H es un subconjunto de vectores de H que tienen norma uno, que son mutuamente ortogonales y que generan H (es decir que dado $x \in H$ y $\varepsilon > 0$ existe una combinación lineal **finita** de elementos de esta base y tal que $\|x - y\|_H < \varepsilon$).

Ejemplo: En l^2 sea $e_1 = (1, 0, 0, \dots)$, $e_2 = (0, 1, 0, \dots)$, etc.

Ejercicio: Muestre que si $a \in l^2$ luego $\|a - \sum_{n=1}^i (a, e_n) e_n\|_{l^2_{i \rightarrow \infty}} \longrightarrow 0$.

Corolario 8.3 *Todo espacio de Hilbert tiene una base ortonormal.*

Este resultado, como el anterior, es muy poderoso ya que nos dice que siempre podemos aproximar distintos elementos de H usando una dada base. No probaremos este Corolario pues para ello se necesitan herramientas que no se verán en este curso. Pero sí otro resultado más simple, para lo cual introducimos la siguiente definición.

Definición: Diremos que un espacio de Banach (y en particular de Hilbert) es **separable** si tiene un subconjunto denso con una cantidad numerable de elementos.

Ejemplos:

a) El subconjunto S de sucesión l^2 que tienen un número finito de elementos *racionales* es denso en l^2 [ya que como veremos más adelante cualquier elemento de l^2 se puede expresar como el límite de una sucesión S] y numerable [ya que los racionales lo son].

b) El subconjunto $S(\mathbb{R})$ que consiste de las funciones $f_{r,s,t}(x) = r e^{-s|x-t|}$ (note que estas son suaves), donde r, s, t son números racionales y $s > 0$, es denso y numerable en $L^2(\mathbb{R})$.

La mayoría de los espacios de Hilbert que aparecen en la práctica son separables y por lo tanto tienen la siguiente propiedad:

Teorema 8.7 *Un espacio de Hilbert H es separable si y solo si tiene una base ortonormal numerable S . Si S tiene un número finito de elementos, N , luego H es isomorfo a C^N (es decir existe un mapa $\varphi : H \rightarrow C^N$, en este caso lineal, con $\|\varphi(x)\|_{C^N} = \|x\|_H \forall x \in H$ que es continuo e invertible y su inversa es también continua). Si S tiene un número infinito de elementos luego H es isomorfo a l^2 .*

Este teorema nos dice que entre los espacios de Hilbert separables esencialmente no hay más estructura que la ya presente en C^N o l^2 .

³El lector debe convencerse que si φ no es el mapa nulo entonces siempre existe $w \in H$ tal que $\varphi(w) = 1$.

Prueba: Para obtener la base ortonormal (numerable) primero descartamos de un subconjunto denso y numerable de H , S , algunos de sus elementos hasta obtener una subcolección de vectores (por construcción numerable) linealmente independientes y que todavía expanden S . Luego aplicamos a esta subcolección el procedimiento de Gram-Schmidt para obtener la base ortonormal. Para probar el resto del teorema necesitamos los siguientes lemas.

Lema 8.4 (Pitágoras): Sea $\{x_n\}_{n=1}^N$ un conjunto ortonormal de H , no necesariamente una base. Luego,

$$\|x\|_H^2 = \sum_{n=1}^N |(x_n, x)|^2 + \|x - \sum_{n=1}^N (x_n, x)x_n\|^2 \quad \forall x \in H. \quad (8.27)$$

Prueba: Sea $v = \sum_{n=1}^N (x_n, x)x_n$ y $w = (x - \sum_{n=1}^N (x_n, x)x_n)$. Es fácil de ver que $v \in V$, el subespacio de Hilbert generado por los $\{x_n\}_{n=1}^N$ y $w \in V^\perp$, pero entonces

$$\|x\|_H^2 = (x, x) = (v + w, v + w) = (v, v) + (w, w) = \sum_{n=1}^N |(x_n, x)|^2 + \|w\|_H^2. \quad (8.28)$$



Lema 8.5 (Desigualdad de Bessel) :

Sea $\{x_n\}_{n=1}^N$ un conjunto ortonormal de H , no necesariamente una base. Luego,

$$\|x\|_H^2 \geq \sum_{n=1}^N |(x, x_n)|^2 \quad \forall x \in H \quad (8.29)$$

Prueba: Obvia conclusión del lema anterior ♠

Lema 8.6 Sea $S = \{x_n\}$ una base ortonormal numerable⁴ de H . Luego $\forall y \in H$,

$$y = \sum_n (x_n, y)x_n \quad (8.30)$$

y

$$\|y\|^2 = \sum_n |(x_n, y)|^2, \quad (8.31)$$

donde la primera igualdad significa que la suma converge con respecto a la norma de H a $y \in H$. La última igualdad es llamada **relación de Parseval** y los coeficientes (x_n, y) los **coeficientes de Fourier** de y . Inversamente si $\{c_n\} \in l^2$ luego $\sum_n c_n x_n \in H$.

Prueba: De la desigualdad de Bessel sigue que dado cualquier subconjunto finito $\{x_n\}_{n=1}^N$,

$$a_N = \sum_{n=1}^N |(x_n, y)|^2 \leq \|y\|^2 \quad (8.32)$$

lo que implica que la sucesión $\{a_N\}$, que es monótonamente creciente, es acotada y por lo tanto converge a un valor límite finito. Esto a su vez implica que $\{a_N\}$ es Cauchy. Sea $y_N = \sum_{n=1}^N (x_n, y)x_n$, luego para $N > M$

$$\|y_N - y_M\|_H^2 = \left\| \sum_{j=M+1}^N (x_j, y)x_j \right\|_H^2 = \sum_{j=M+1}^N |(x_j, y)|^2 = a_N - a_M \quad (8.33)$$

⁴Un resultado similar es válido aún en el caso de que la base no sea numerable, o sea aún cuando H no sea separable.

La convergencia de $\{a_N\}$ garantiza así que $\{y_N\}$ es una sucesión de Cauchy y así que esta converge a un elemento y' de H . Probemos entonces que $y' = y$. Pero para cualquier elemento de S , x_n

$$\begin{aligned}(y - y', x_n) &= \lim_{N \rightarrow \infty} (y - \sum_{j=1}^N (x_j, y) x_j, x_n) \\ &= (y, x_n) - (y, x_n) = 0\end{aligned}\tag{8.34}$$

y por lo tanto $y - y'$ es perpendicular al espacio generado por S , que es todo H y por lo tanto tiene que ser el elemento cero. Solo resta ver que si $\{c_n\} \in l^2$ luego $\sum_{n=1}^{\infty} c_n x_n \in H$. Un cálculo idéntico al anterior muestra que $\{y_N = \sum_{n=1}^N c_n x_n\}$ es una sucesión de Cauchy y por lo tanto que $\lim_{N \rightarrow \infty} y_N \in H$ ♠

Continuamos ahora la prueba del Teorema 8.7 El último argumento del Lema 8.6 nos muestra que dada una base numerable $\{x_n\}_{n=1}^{\infty}$ de H el conjunto numerable $\text{Span}_{\mathbb{Q}}\{S\} = \{x | x = \sum_{n=1}^{\infty} c_n x_n \text{ con } \{c_n\} \text{ una sucesión finita de racionales}\}$ es denso en H y por lo tanto que H es separable. Esto concluye la parte *sii* de la prueba, solo resta encontrar un isomorfismo entre H y C^N o l^2 . Sea $\{x_n\}$ una base numerable de H y sea $\varphi : H \rightarrow C^N$ o l^2 el mapa

$$\varphi(y) = \{(x_n, y)\}.\tag{8.35}$$

Si la base es finita, con dimensión N este es claramente un mapa en C^N . Si la base es infinita la imagen de H por φ está incluida en el espacio de sucesiones complejas infinitas, pero usando el Lema 8.6 vemos que ,

$$\|\varphi(y)\|_{l^2}^2 = \sum_{n=1}^{\infty} |(x_n, y)|^2 = \|y\|_H^2\tag{8.36}$$

y por lo tanto que dicha imagen está contenida en l^2 . La continuidad e invertibilidad del mapa quedan como ejercicio para el lector ♠

El teorema anterior, entre otras cosas sirve para dar una caracterización que diferencia los espacios de Hilbert de dimensión finita de los de dimensión infinita. (En realidad esta caracterización es válida para espacios de Banach en general).

Teorema 8.8 *Sea H un espacio de Hilbert y sea $B_1(H) = \{x \in H / \|x\|_H \leq 1\}$ la bola de radio uno en H . Luego $B_1(H)$ es compacta sii H es de dimensión finita. [Se puede ver que en este caso un subconjunto B de H es compacto sii toda sucesión $\{x_n\}$ de elementos de H en B tiene una subsucesión que converge a un elemento de B .]*

Prueba: Veremos solo el caso en que el espacio es separable. Si H es de dimensión finita es isomorfo a C^N , pero $B_1(C^N)$ es un subconjunto cerrado y acotado, y por lo tanto compacto. Si H es de dimensión infinita (y separable) entonces es isomorfo a l^2 , pero la sucesión en $B_1(l^2)$, $\{x_n\}$ con

$\{x_n = (0, \dots, 0, \underset{n}{1}, 0, \dots)\}$ claramente no tiene ninguna subsucesión convergente ♠

Ejercicio: Ver que ninguna subsucesión de la sucesión anterior es de Cauchy.

Ejemplo: Sea $H = L^2([0, 2\pi])$ y $S = \{\frac{1}{\sqrt{2\pi}} e^{inx}, n = 0, \pm 1, \pm 2, \dots\}$. El desarrollo de este ejemplo será nuestra ocupación hasta el final de este capítulo.

Ejercicio: Muestre que los elementos de $L^2(0, 2\pi)$ de la forma $f_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$, $n = 0, \pm 1, \pm 2, \dots$, forman un conjunto ortonormal.

8.5. Serie de Fourier

Teorema 8.9 $S = \left\{ \frac{e^{inx}}{\sqrt{2\pi}}, n = 0, \pm 1, \pm 2, \dots \right\}$ es un conjunto completo de vectores ortonormales (es decir una base ortonormal) de $L^2[0, 2\pi]$.

Note que por Lema 8.6 esto es equivalente a asegurar que si $f \in L^2[0, 2\pi]$ luego $f_N(x)$ converge (en la norma de $L^2[0, 2\pi]$) a f cuando $N \rightarrow \infty$.

Prueba: La ortogonalidad es trivial y forma parte de un ejercicio anterior, por lo tanto solo nos concentraremos en la completitud. El espacio de funciones Lipschitz en $[0, 2\pi]$, $Lip[0, 2\pi]$ es denso [esencialmente por definición! ya que contiene a todas las funciones suaves] en $L^2[0, 2\pi]$, pero entonces su subespacio consistente de funciones periódicas $Lip_p[0, 2\pi]$ también lo es [ya que los elementos de $L^2[0, 2\pi]$ son clases equivalentes de funciones y en cada clase siempre hay una que es periódica]. Supongamos ahora que

$$S_N(g)(x) := \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^N g_n e^{inx} \rightarrow g(x) \quad (8.37)$$

punto a punto y uniformemente si $g(x) \in Lip_p([0, 2\pi])$. Luego por densidad, dada cualquier función $f(x) \in L^2([0, 2\pi])$ y $\varepsilon > 0$ existirá $f_\varepsilon \in Lip_p([0, 2\pi])$ tal que $\|f - f_\varepsilon\|_{L^2} < \varepsilon/3$ y por lo tanto tendremos,

$$\begin{aligned} \|f - S_N(f)\|_{L^2} &= \|f - S_N(f) - f_\varepsilon + S_N(f_\varepsilon) + f_\varepsilon - S_N(f_\varepsilon)\|_{L^2} \\ &\leq \|f - f_\varepsilon\|_{L^2} + \|f_\varepsilon - S_N(f_\varepsilon)\|_{L^2} + \|S_N(f - f_\varepsilon)\|_{L^2} \end{aligned} \quad (8.38)$$

pero

$$\begin{aligned} \|S_N(f - f_\varepsilon)\|_{L^2} &= \sum_{n=-N}^N |(f - f_\varepsilon)_n|^2 \\ &\leq \|f - f_\varepsilon\|^2 \\ &\leq \varepsilon/3 \end{aligned} \quad (8.39)$$

y por lo tanto eligiendo N tal que $|f_\varepsilon(x) - S_N(f_\varepsilon)(x)| < \frac{\varepsilon}{6\pi}$, lo que sigue de nuestra suposición anterior (todavía no probada), tenemos que $\|f_\varepsilon - S_N(f_\varepsilon)\|_{L^2} < \frac{\varepsilon}{3}$ y por lo tanto concluimos que para dicho N , $\|f - S_N(f)\|_{L^2} < \varepsilon$ ♠

Vemos así que solo resta probar nuestra suposición de convergencia uniforme de $S_N(f)$ a f para funciones Lipschitz periódicas. La prueba de esta afirmación es muy instructiva y nos muestra como trabaja la serie de Fourier. Preparatoriamente para este resultado probamos una serie de lemas:

Lema 8.7 Si f es integrable ($f \in L^1$) luego,

$$\sup_n \{|f_n|\} \leq \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} |f(x)| dx. \quad (8.40)$$

Prueba:

$$\begin{aligned} |f_n| &= \left| \left(\frac{e^{inx}}{\sqrt{2\pi}}, f \right) \right| \\ &\leq \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} |e^{-inx} f(x)| dx \\ &\leq \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} |f(x)| dx \end{aligned} \quad (8.41)$$

♠

Lema 8.8 Si f tiene derivada integrable, luego $f_n \rightarrow 0$ cuando $n \rightarrow \infty$.

Prueba: Integrando por partes tenemos,

$$\begin{aligned}
 f_n &= \left(\frac{e^{inx}}{\sqrt{2\pi}}, f \right) \\
 &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{-inx} f(x) dx \\
 &= \frac{1}{in\sqrt{2\pi}} \int_0^{2\pi} e^{-inx} f'(x) dx - \frac{1}{in\sqrt{2\pi}} e^{-inx} f(x) \Big|_0^{2\pi} \\
 &= \frac{1}{in\sqrt{2\pi}} \int_0^{2\pi} e^{-inx} f'(x) dx - \frac{1}{in\sqrt{2\pi}} [f(2\pi) - f(0)]
 \end{aligned} \tag{8.42}$$

y por lo tanto

$$|f_n| \leq \frac{1}{n\sqrt{2\pi}} [\|f'\|_{L^1} + |f(2\pi) - f(0)|], \tag{8.43}$$

con lo que el lema queda demostrado ♠

Este lema y su generalización a un número mayor de derivadas nos indica que cuanto más diferenciable es una función más rápido decaen sus coeficientes de Fourier (asintóticamente).

Ejercicio: Pruebe un lema similar que dé una mejor cota para f_n si la función es periódica y es m veces diferenciable.

Lema 8.9 (Riemann-Lebesgue) Si f es una función integrable, es decir $f \in L^1([0, 2\pi])$. Luego

$$\lim_{n \rightarrow \infty} f_n = 0 \tag{8.44}$$

Prueba: Si $f \in L^1([0, 2\pi])$, luego es aproximable por funciones suaves, en particular, dado cualquier $\varepsilon > 0$ existe $f_\varepsilon : [0, 2\pi]$, suave, tal que $\|f - f_\varepsilon\|_{L^1} < \frac{\varepsilon}{2\sqrt{2\pi}}$. Pero como,

$$\begin{aligned}
 |f_n - f_{\varepsilon n}| &= \left| \left(\frac{e^{inx}}{\sqrt{2\pi}}, f - f_\varepsilon \right) \right| \\
 &\leq \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} |f(x) - f_\varepsilon(x)| dx \\
 &\leq \frac{1}{\sqrt{2\pi}} \|f - f_\varepsilon\|_{L^1} \\
 &\leq \varepsilon/2
 \end{aligned} \tag{8.45}$$

tenemos que

$$|f_n| \leq |f_n - f_{\varepsilon n}| + |f_{\varepsilon n}| \leq \varepsilon/2 + |f_{\varepsilon n}|. \tag{8.46}$$

Aplicando el lema anterior y notando que f_ε es diferenciable vemos que $f_{\varepsilon n} \rightarrow 0$ y por lo tanto, dado $\varepsilon > 0$ puedo elegir N tal que para todo $n > N$ $|f_{\varepsilon n}| < \varepsilon/2$ con lo que $|f_n| < \varepsilon$ para todo $n > N$ y por lo tanto concluimos que $f_n \rightarrow 0$ cuando $n \rightarrow \infty$ ♠

Ahora estamos en condiciones de probar el teorema sobre convergencia puntual de funciones Lipschitz.

Teorema 8.10 Sea $f : [0, 2\pi] \rightarrow \mathbb{C}$ periódica y Lipschitz. Luego

$$S_N(f)(x) := \frac{1}{\sqrt{2\pi}} \sum_{n=-N}^N f_n e^{inx} \tag{8.47}$$

converge puntualmente y uniformemente a $f(x)$.

Prueba: Probaremos solo la convergencia puntual, la uniforme sigue fácilmente y la prueba de ello no agrega nada nuevo. Comenzamos con el siguiente cálculo,

$$\begin{aligned}
 S_N(f)(\theta) &:= \frac{1}{\sqrt{N}2\pi} \sum_{n=-N}^N f_n e^{in\theta} \\
 &= \frac{1}{2\pi} \sum_{n=-N}^N \int_0^{2\pi} e^{in\theta'} f(\theta') e^{in\theta} d\theta' \\
 &= \frac{1}{2\pi} \int_0^{2\pi} f(\theta') \left(\sum_{n=-N}^N e^{in(\theta-\theta')} \right) d\theta' \\
 &= \int_0^{2\pi} f(\theta') D_N(\theta - \theta') d\theta'
 \end{aligned} \tag{8.48}$$

donde hemos definido el **núcleo de Dirichlet**

$$D_N(\theta - \theta') := \frac{1}{2\pi} \sum_{n=-N}^N e^{in(\theta-\theta')}. \tag{8.49}$$

Estudiamos ahora algunas propiedades del núcleo de Dirichlet. Notemos primero que,

$$\begin{aligned}
 \int_0^{2\pi} D_N(\theta - \theta') d\theta' &= \frac{1}{2\pi} \sum_{n=-N}^N \int_0^{2\pi} e^{in(\theta-\theta')} d\theta' \\
 &= \frac{1}{2\pi} \left[\sum_{n=-N, n \neq 0}^N \frac{-i}{n} [e^{in(\theta-2\pi)} - e^{in\theta}] + 2\pi \right] \\
 &= 1.
 \end{aligned} \tag{8.50}$$

Por otro lado tenemos que

$$\begin{aligned}
 2\pi D_N(\theta) &= \sum_{n=-N}^N e^{in\theta} \\
 &= \sum_{n=-N}^N (e^{i\theta})^n
 \end{aligned} \tag{8.51}$$

y llamando $q := e^{i\theta}$ para simplificar la escritura tenemos,

$$\begin{aligned}
 &2\pi D_N(\theta) \\
 &= \sum_{n=-N}^N q^n \\
 &= \sum_{n=0}^N q^n + \sum_{n=0}^N q^{-n} - 1 \\
 &= \frac{q^{N+1} - 1}{q - 1} + \frac{q^{-(N+1)} - 1}{q^{-1} - 1} - 1 \\
 &= \frac{(q^{N+1} - 1)(q^{-1} - 1) + (q^{-(N+1)} - 1)(q - 1) - (q - 1)(q^{-1} - 1)}{(q - 1)(q^{-1} - 1)} \\
 &= \left[(q^{N+1/2} - q^{-1/2})(q^{-1/2} - q^{1/2}) + (q^{-(N+1/2)} - q^{1/2})(q^{1/2} - q^{-1/2}) \right. \\
 &\quad \left. - (q^{1/2} - q^{-1/2})(q^{-1/2} - q^{1/2}) \right] / (q^{1/2} - q^{-1/2})(q^{-1/2} - q^{1/2})
 \end{aligned}$$

$$\begin{aligned}
&= \frac{q^{N+1/2} - q^{-N-1/2}}{q^{1/2} - q^{-1/2}} \\
&= \frac{e^{i\theta(N+1/2)} - e^{-i\theta(N+1/2)}}{e^{i\theta/2} - e^{-i\theta/2}} \\
&= \frac{\sin((N+1/2)\theta)}{\sin(\theta/2)}.
\end{aligned} \tag{8.52}$$

Por lo tanto

$$D_N(\theta) = \frac{1}{2\pi} \frac{\sin((N+1/2)\theta)}{\sin(\theta/2)}. \tag{8.53}$$

Tenemos por lo tanto,

$$\begin{aligned}
S_N(f)(\theta) &= \int_0^{2\pi} f(\theta') D_N(\theta - \theta') d\theta' \\
&= \int_0^{2\pi} f(\theta - \theta') D_N(\theta) d\theta' \\
&= \frac{1}{2\pi} \int_0^{2\pi} [f(\theta - \theta') - f(\theta)] \frac{\sin((N+1/2)\theta')}{\sin(\theta'/2)} d\theta' + f(\theta)
\end{aligned} \tag{8.54}$$

donde en la segunda igualdad hemos usado que la integral de una función periódica sobre su período es independiente del punto donde comienza el intervalo de integración, y en la última hemos restado y sumado $f(\theta)$ y usado que la integral del núcleo de Dirichlet es uno (8.50).

Si $f(\theta)$ es Lipschitz, $|f(\theta - \theta') - f(\theta)| \leq k|\theta'|$ y por lo tanto

$$g_\theta(\theta') := \frac{|f(\theta - \theta') - f(\theta)|}{\sin(\theta'/2)} \tag{8.55}$$

es continua en $(0, 2\pi)$ y acotada en los extremos, por lo tanto integrable. Aplicando ahora el Lema de Riemann-Lebesgue concluimos entonces que la integral tiende a cero con N y por lo tanto que $\lim_{N \rightarrow \infty} S_N(f)(\theta) = f(\theta)$ en todo punto donde $f(\theta)$ es Lipschitz. ♠

Para probar que S es completo solo tenemos que ver que si $(e^{inx}, g) = 0 \forall n$ luego $g = 0$. Sea $f \in C_p^1[0, 2\pi]$ y c_n sus coeficientes de Fourier con respecto a S , luego

$$(f, g) = \lim_{M \rightarrow \infty} \left(\sum_{-M}^M c_n \frac{e^{inx}}{\sqrt{2\pi}}, g \right) = 0 \tag{8.56}$$

Concluimos entonces que g es ortogonal a toda $f \in C_p^1[0, 2\pi]$, pero como vimos este espacio es denso en $L^2[0, 2\pi]$ y por lo tanto por continuidad debemos tener entonces que $(f, g) = 0 \forall f \in H$, o sea $g = 0$ ♠

Ejemplo[Aplicación de la serie de Fourier] Sea S un aro de metal de circunferencia 2π y sea $T_0(\theta)$ una distribución de temperatura que supondremos de cuadrado integrable ($\in L^2(S)$). La evolución temporal de $T(\theta, t)$, (despreciando pérdidas al medio circundante por conducción o por radiación) está dada por,

$$\frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial \theta^2} \tag{8.57}$$

con k una constante positiva. Supondremos $T(\theta, 0) = T_0(\theta)$. Suponiendo además que $T(\theta, t)$ admite una descomposición en serie de Fourier,

$$T(\theta, t) = \sum_{n=-\infty}^{\infty} T_n(t) \frac{e^{in\theta}}{\sqrt{2\pi}} \tag{8.58}$$

y que se puede pasar las derivadas dentro de la sumatoria infinita obtenemos,

$$\begin{aligned} 0 &= \frac{\partial T}{\partial t} - k \frac{\partial^2 T}{\partial \theta^2} \\ &= \sum_{n=-\infty}^{\infty} \left(\frac{dT_n(t)}{dt} + kn^2 T_n(t) \right) \frac{e^{in\theta}}{\sqrt{2\pi}} \end{aligned} \quad (8.59)$$

y tomando el producto escalar con $\frac{e^{in\theta}}{\sqrt{2\pi}}$ vemos que se debe cumplir,

$$\frac{dT_n(t)}{dt} = -kn^2 T_n(t) \quad \forall n \quad (8.60)$$

o sea,

$$T_n(t) = T_n(0) e^{-kn^2 t}. \quad (8.61)$$

Si la temperatura inicial era

$$T_0(\theta) = \sum_{n=-\infty}^{\infty} T_n^0 \frac{e^{in\theta}}{\sqrt{2\pi}} \quad (8.62)$$

vemos entonces que $T_n(t) = T_n^0 e^{-kn^2 t}$ y

$$T(\theta, t) = \sum_{n=-\infty}^{\infty} T_n^0 e^{-kn^2 t} \frac{e^{in\theta}}{\sqrt{2\pi}} \quad (8.63)$$

nos da la solución del problema.

Notemos que: a) La distribución de temperatura en la barra solo depende de la distribución inicial. b) El hecho de haber podido reducir un problema en derivadas parciales en un problema en derivadas ordinarias se debe a que hemos elegido representar a las funciones en una base de autovectores del operador derivada. En efecto, lo que hemos usado fundamentalmente es que $\frac{\partial}{\partial \theta} e^{in\theta} = in e^{in\theta}$. c) No importa cuán mala (no diferenciable) es la distribución inicial $T_0(\theta)$, mientras esté en $L^2(S)$, la solución se suaviza para tiempos positivos. En efecto si por ejemplo la distribución inicial es solo Lipschitz luego para todo $t > 0$ la solución es infinitamente diferenciable, tanto en t como en θ . Para ver esto, por ejemplo tomemos,

$$\frac{\partial^p T(\theta, t)}{\partial t^p} = \sum_{n=-\infty}^{\infty} T_n^0 (-kn^2)^p e^{-kn^2 t} \frac{e^{in\theta}}{\sqrt{2\pi}}, \quad (8.64)$$

pero como para $t > 0$, $n^{2p} e^{-kn^2 t} \rightarrow 0$ rápidamente cuando $n \rightarrow \infty$ la serie converge absolutamente. Por el contrario, para dato inicial genérico, aun infinitamente diferenciable, la solución no existe para tiempos negativos, ya que en ese caso los coeficientes de la serie crecen rápidamente con n .

Ejercicio: Pruebe que un conjunto ortogonal $\{\mathbf{x}_n\}$ es completo si y solo si $(\mathbf{x}_n, \mathbf{g}) = 0 \quad \forall n \Rightarrow \mathbf{g} = 0$.

Problema 8.3 : Use Gram-Schmidt para obtener una base ortonormal a partir de los monomios

$$1, x, x^2, \dots, x^n, \dots, \quad (8.65)$$

con respecto a los espacios de Hilbert obtenidos a partir de los siguientes productos escalares:

1. $(f, g) = \int_{-1}^1 \bar{f}g \, dx$ (En este caso obtendrá los polinomios de Legendre.)
2. $(f, g) = \int_{-\infty}^{\infty} \bar{f}g e^{-x^2} \, dx$ (En este caso obtendrá los polinomios de Hermite.)
3. $(f, g) = \int_0^{\infty} \bar{f}g e^{-x} \, dx$ (En este caso obtendrá los polinomios de Laguerre.)

Estas bases polinómicas reciben el nombre genérico de sistemas de polinomios de Tchebyshev.

El hecho de que los polinomios de Legendre son una base sigue del Teorema de Aproximación de Weirstrass y del hecho de que las funciones continuas son densas en L^2 .

***Método de Gram – Schmidt.**

Dado un conjunto numerable de elementos linealmente independientes de H , $\{\mathbf{x}_i\}$ generamos recursivamente los siguientes conjuntos:

$$\mathbf{y}_i = \mathbf{x}_i - \sum_{l=1}^{i-1} (\mathbf{u}_l, \mathbf{x}_i) \mathbf{u}_l, \quad \mathbf{u}_i = \frac{\mathbf{y}_i}{\|\mathbf{y}_i\|}.$$

Note que la segunda operación está bien definida ya que $\mathbf{y}_i \neq 0$. Esta aseveración sigue del hecho de que el lado derecho en la definición de \mathbf{y}_i es una combinación lineal de los \mathbf{x}_j , $j = 1, \dots, i$ y hemos supuesto que estos son linealmente independientes.

Veamos ahora que $(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$. Para ello probaremos por inducción que dado i , $(\mathbf{u}_i, \mathbf{u}_j) = 0 \quad \forall j < i$ positivo. Esto es cierto para $i = 1$ [ya que no hay j]. Supongamos entonces que es cierto para $i - 1$ y veamos que también lo es para i . Pero, dado $j < i$ tenemos,

$$(\mathbf{u}_j, \mathbf{y}_i) = [(\mathbf{u}_j, \mathbf{x}_i) - \sum_{l=1}^{j-1} (\mathbf{u}_l, \mathbf{x}_i)(\mathbf{u}_j, \mathbf{u}_l)] = [(\mathbf{u}_j, \mathbf{x}_i) - (\mathbf{u}_j, \mathbf{x}_i)] = 0.$$

Si el conjunto de partida es un conjunto completo, es decir un conjunto que no es subconjunto de un conjunto mayor de vectores linealmente independientes, luego el conjunto resultante es una base ortogonal. En particular, si $(\mathbf{x}, \mathbf{u}_i) = 0 \quad \forall i$, luego $\mathbf{x} = 0$. Caso contrario tomamos $\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$ y tendríamos otro elemento de la base, lo que sería una contradicción con respecto a la completitud de los $\{\mathbf{u}_i\}$.

8.6. Problemas

Problema 8.4 Sea $\phi : H \rightarrow C$ un mapa lineal. Muestre que ϕ es continuo si y solo si es acotado.

Problema 8.5 Muestre que el mapa $I : C[a, b] \rightarrow \Re$ dado por

$$I(f) \equiv \int_a^b f(x) dx, \quad (8.66)$$

es un mapa lineal y continuo.

Problema 8.6 Sea V un espacio de dimensión finita y sea $\{\mathbf{u}_i\}$, $i = 1..n$ una base y $\{\boldsymbol{\theta}^i\}$, $i = 1..n$ la correspondiente co-base. Sea $\mathbf{x} = \sum_{i=1}^n x^i \mathbf{u}_i$ un vector de V cualquiera y $\boldsymbol{\omega} = \sum_{i=1}^n \omega_i \boldsymbol{\theta}^i$ una funcional lineal cualquiera, es decir un elemento de V' . Considere en V la norma

$$\|\mathbf{x}\|_p \equiv \left(\sum_{i=1}^n |x^i|^p \right)^{\frac{1}{p}}. \quad (8.67)$$

Vea que esta es una norma y pruebe que la norma inducida en V' por ésta está dada por,

$$\|\boldsymbol{\omega}\|_q \equiv \left(\sum_{i=1}^n |\omega_i|^q \right)^{\frac{1}{q}}, \quad (8.68)$$

donde

$$\frac{1}{p} + \frac{1}{q} = 1 \quad (p, q \geq 1); \quad (8.69)$$

Ayuda: Expresa $\omega(\mathbf{x})$ en componentes con respecto a la base/co-base dada y luego use (pruebe) la desigualdad:

$$\left| \sum_{i=1}^n x^i \omega_i \right| \leq \left(\sum_{i=1}^n |x^i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |\omega_i|^q \right)^{\frac{1}{q}}. \quad (8.70)$$

Problema 8.7 Sea c_0 el espacio de sucesiones $\{x\} = (x_1, x_2, \dots)$ convergiendo a cero con la norma

$$\|\{x\}\|_{c_0} \equiv \sup_i \{|x_i|\}. \quad (8.71)$$

Probar que el dual del espacio c_0 es el espacio l_1 de las sucesiones absolutamente sumables $\{\omega\} = (\omega_1, \omega_2, \dots)$ con la norma

$$\|\{\omega\}\|_{l_1} \equiv \sum_{i=1}^{\infty} |\omega_i|. \quad (8.72)$$

Ayudas: Note que dado un elemento de l_1 , $\{\omega\} = (\omega_1, \omega_2, \dots)$, tenemos un funcional lineal dado por,

$$\omega(\{x\}) \equiv \sum_{i=1}^{\infty} x_i \omega_i. \quad (8.73)$$

Pruebe que este cumple

$$\|\omega\| \leq \|\{\omega\}\|_{l_1}. \quad (8.74)$$

Luego encuentre un elemento de norma igual o menor que la unidad en c_0 y con su ayuda vea que

$$\|\omega\| \geq \|\{\omega\}\|_{l_1}. \quad (8.75)$$

de lo cual se concluye que las normas son las mismas. Solo resta ver que para cada elemento del dual de c_0 , ω , existe un elemento de l_1 , $\{\omega\} = (\omega_1, \omega_2, \dots)$ tal que la ecuación 8.73 valga. Para ello construya una base de c_0 y la respectiva base de su dual. Atención: en algún punto tendrá que usar que las funcionales lineales consideradas son continuas.

8.7. Problemas de Series de Fourier

Problema 8.8 Sea f una función integrable de período T . Muestre que

$$\int_0^T f(x) dx = \int_a^{T+a} f(x) dx, \quad \forall a \in \mathbb{R} \quad (8.76)$$

Problema 8.9 a.- Encuentre la serie de Fourier de la función $f(x) := x$ en el intervalo $[-\pi, \pi]$.

b.- Use la relación de Parseval para probar que

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \pi^2/6 \quad (8.77)$$

Problema 8.10 a.- Encuentre la serie de Fourier de la función $f(x) := e^{sx}$ en el intervalo $[-\pi, \pi]$.

b.- Use la relación de Parseval para probar que

$$\pi \coth(\pi s)/s = \sum_{n=-\infty}^{\infty} \frac{1}{s^2 + n^2} \quad (8.78)$$

Problema 8.11 Sea $S_n : L^2 \rightarrow L^2$ el mapa que envía $f \in L^2$ en la serie parcial de Fourier,

$$S_n(f) := \sum_{m=-n}^n c_m e^{imx}, \quad c_m := \frac{1}{2\pi} (e^{imx}, f(x)). \quad (8.79)$$

Muestre que los S_n son proyecciones ortogonales y que $S_n S_m = S_m S_n = S_m$ si $m \leq n$.

Problema 8.12 En este problema se intenta probar que la serie de Fourier de una función continua es sumable en el sentido de Césaro en todo punto. Sea $f(\theta)$ una función periódica en L^2 , $f(\theta) \in L^2[0, 2\pi]$, $c_m := \frac{1}{2\pi} (e^{im\theta}, f(\theta))$ y $S_n(f) := \sum_{m=-n}^n c_m e^{imx}$

a.- Probar que

$$S_n(f)(\theta) = \frac{1}{2\pi} \int_0^{2\pi} f(\theta + x) \frac{\sin((n+1/2)x)}{\sin(x/2)} dx \quad (8.80)$$

b.- Sea $SS_n(f)(\theta) = \frac{1}{n+1} \sum_0^n S_m(f)(\theta)$ (suma de Césaro), pruebe que:

$$SS_n(f)(\theta) = \frac{1}{2\pi(n+1)} \int_0^{2\pi} f(\theta + x) \frac{\sin^2((n+1)x/2)}{\sin^2(x/2)} dx \quad (8.81)$$

c.- Sea $K_n(x) = \frac{\sin^2((n+1)x/2)}{2\pi(n+1)\sin^2(x/2)}$ pruebe que para todo $\delta > 0$, $K_n(x) \rightarrow 0$ uniformemente en $[\delta, 2\pi - \delta]$.

d.- Pruebe que $SS_n(f)(\theta_0) \rightarrow f(\theta_0)$ si f es acotada y continua en θ_0 .

e.- Probar que si f es continua y periódica, entonces $SSL_n(f)(\theta) \rightarrow f(\theta)$ uniformemente en θ . Ayuda: recuerde que si f es continua en $[0, 2\pi]$ entonces f es uniformemente continua.

f.- Mostrar que $\|f - S_n(f)\| \leq \|f - SSL_n(f)\|$ y concluir que $S_n(f) \rightarrow f$ en L^2 si f es continua.

Problema 8.13 Suponga que $f \in C_p^1([0, 2\pi])$ y sea $c_n = (e^{in\theta}, f)$ y $b_n = (e^{in\theta}, f')$.

a.- Vea que $\sum_{n=-\infty}^{\infty} |b_n|^2 < \infty$ y concluya que $\sum_{n=-\infty}^{\infty} n^2 |c_n|^2 < \infty$.

b.- Probar que $\sum_{n=-\infty}^{\infty} |c_n| < \infty$.

c.- Probar que $\sum_{m=-n}^n c_m e^{mi\theta}$ es uniformemente convergente para $n \rightarrow \infty$.

d.- Utilice el ítem f.- del problema anterior para concluir que $\sum_{m=-n}^n c_m e^{mi\theta} \rightarrow 2\pi f(\theta)$ uniformemente.

Problema 8.14 Considere la expansión en serie de Fourier para la siguiente función:

$$f(x) := \begin{cases} -1 & 0 < x < \pi \\ +1 & \pi < x < 2\pi \end{cases} \quad (8.82)$$

$f(\theta + 2\pi) = f(\theta)$. La suma de los primeros n términos produce una función que tiene un máximo absoluto en las cercanías positivas de 0 de altura $1 + \delta_n$. Mostrar que $\lim_{n \rightarrow \infty} \delta_n \approx 0,18$. Esto se conoce como fenómeno de Gibbs.

Notas bibliográficas: Estas notas están basadas en los siguientes libros: [9], [10], [13], [1] y [14]. Ésta es una de las áreas mas bellas y útiles de las matemáticas, básica para casi todo, en particular la mecánica cuántica. No deje de profundizar un poquito en ella, sobre todo recomendando los libros [9] y [1] de lectura placentera.

9.1. Introducción

Para hacer del espacio de funciones de cuadrado integrable $L^2(\mathbb{R})$ un espacio normado fue necesario generalizar el concepto de función (como mapa de \mathbb{R} en \mathbb{R}) en el sentido que los elementos de L^2 son solo funciones definidas en casi todos los puntos, es decir clases equivalentes de funciones de cuadrado integrable ante la relación de equivalencia $f \approx g$ si $\int |f - g|^2 dx = 0$.

Si bien esta generalización es útil ya que entre otras cosas nos permite agrupar las funciones en espacios de Hilbert y así usar la poderosa estructura geométrica que éstos tienen, es conveniente considerar una generalización aun mayor la cual, como veremos más adelante, nos proveerá de una importante herramienta en lo que hace al cálculo formal en física matemática.

Las funciones generalizadas que definiremos a continuación, llamadas distribuciones, tienen muchas propiedades interesantes, entre ellas que la operación diferenciación es cerrada en este espacio, es decir la derivada de una distribución es otra distribución. Esto es particularmente sorprendente si se tiene en cuenta que entre las distribuciones hay funciones que ni siquiera son continuas!

¿Cuál es la idea detrás de esta generalización? El teorema de representación de Riez nos mostró que el dual de $L^2(\mathbb{R})$ es ese mismo espacio. Ahora bien, si en vez del dual de $L^2(\mathbb{R})$ consideramos el dual de un subespacio de $L^2(\mathbb{R})$, obtendremos un espacio mayor que $L^2(\mathbb{R})$ y que contiene a éste de una manera natural. Este espacio lineal, que contiene al de las funciones usuales es un espacio de funciones generalizadas.

Está claro que existen muchos espacios de funciones generalizadas ya que no solo podemos considerar distintos subespacios de $L^2(\mathbb{R})$, sino también podemos considerar distintas nociones de continuidad más débiles que la continuidad proveniente de la norma de $L^2(\mathbb{R})$ para definir los espacios duales.

¿Cuál de ellos estudiar? La respuesta es: El que más convenga al tratamiento del problema para el cual se las quiere usar. Aquí trataremos aquellas obtenidas a partir de un subespacio bastante pequeño con lo cual se obtiene una generalización lo suficientemente amplia como para abarcar con ella la mayoría de los problemas de la Física. Es necesario remarcar que el concepto de distribución que introduciremos no es una necesidad física, en el sentido que las teorías físicas se pueden enunciar usando simplemente funciones infinitamente diferenciables¹, pero sí es una herramienta muy útil que permite una formulación más “condensada” de algunas de estas leyes. El subespacio de $L^2(\mathbb{R})$ que usaremos es de las funciones infinitamente diferenciables y de soporte compacto $C_0^\infty(\mathbb{R})$. [Recuerde que el soporte de una función clásica $f(x)$ está dado por el subconjunto de \mathbb{R} ,

$$Cl\{f^{-1}[\mathbb{R} - \{0\}]\},$$

donde Cl significa tomar la clausura. Como el soporte de una función es automáticamente cerrado, que sea compacto (como subconjunto de \mathbb{R} o \mathbb{R}^n) significa meramente que es acotado.]

Ejercicio: Muestre que $C_0^\infty(\mathbb{R})$ es realmente un espacio vectorial.

¹Aquí nos referimos a las teorías fundamentales. Existen aproximaciones, como la teoría de los fluidos, donde las distribuciones aparecen naturalmente.

Ejercicio: Muestre que además es un álgebra con respecto al producto usual. ¿Cuál es el soporte de $f \cdot g$?

¿Qué noción de continuidad introduciremos en las funcionales de $C_0^\infty(\mathbb{R})$ para definir el dual? Desgraciadamente no hay en este espacio ninguna norma que sea *natural*, en particular no hay ninguna en que éste sea completo ².

Hay sí, en este espacio, una topología conveniente. La correspondiente noción de continuidad de esta topología se obtiene del siguiente criterio de convergencia:

Definición: Diremos que una sucesión $\{\varphi_n\}$, $\varphi_n \in C_0^\infty(\mathbb{R})$ **converge** a $\varphi \in C_0^\infty(\mathbb{R})$ si:

1. Existe $K \subset \mathbb{R}$ compacto tal que $\text{soporte}(\varphi_n) \subset K \quad \forall n$.
2. Las sucesiones $\{\varphi_n^{(p)}\}$ de sus derivadas de orden p convergen uniformemente en K a $\varphi^{(p)}$ para todo $p = 0, 1, 2, \dots$, es decir dado p y $\varepsilon > 0$ existe N tal que para todo $n > N$ se cumple

$$\sup_{x \in K} |\varphi_n^{(p)}(x) - \varphi^{(p)}(x)| < \varepsilon. \quad (9.1)$$

Note que la primera condición nos restringe a considerar como sucesiones convergentes a aquellas que solo pueden converger a una función de soporte compacto. Esto es fundamental para la completitud del espacio y para que la convergencia uniforme en la segunda condición tenga sentido usando el supremo. Con este criterio de convergencia asociamos la siguiente noción de continuidad sobre las funcionales de $C_0^\infty(\mathbb{R})$ en \mathbb{R} .

Definición: Diremos que la funcional $F : C_0^\infty(\mathbb{R}) \rightarrow \mathbb{R}$ (o \mathbb{C}) es **continua** en $\varphi \in C_0^\infty(\mathbb{R})$ si dada cualquier sucesión convergente $\{\varphi_n\}$ en $C_0^\infty(\mathbb{R})$ a φ se cumple,

$$F(\varphi_n) \longrightarrow F(\varphi). \quad (9.2)$$

Esta noción proviene de la topología antes mencionada.

Ejercicio: Sea B un espacio de Banach con la noción de convergencia dada por su norma. Muestre que en ese caso la noción de continuidad definida arriba coincide con la usual (ε, δ) .

Con esta noción de continuidad el espacio $C_0^\infty(\mathbb{R})$ es llamado el **espacio de funciones de prueba** de \mathbb{R} y denotado por $\mathcal{D}(\mathbb{R})$.

Definición: El espacio dual al espacio de funciones de prueba, \mathcal{D}' , es decir el espacio de funcionales lineales continuas $T : C_0^\infty(\mathbb{R}) \rightarrow \mathbb{R}$ es llamado el espacio de las **distribuciones**.

Ejemplos:

a) Sea f continua y sea la funcional lineal

$$T_f(\varphi) = \int_{\mathbb{R}} f \varphi \, dx. \quad (9.3)$$

²Aún en el caso que, por ejemplo, usásemos como norma

$$\|f\| = \sum_{n=0}^{\infty} \frac{1}{n!} \sup_{x \in \mathbb{R}} \{|f^{(n)}(x)|\}$$

no obtendríamos un espacio completo ya que en este hay sucesiones de Cauchy que tienden a funciones infinitamente diferenciables pero cuyo soporte no es compacto.

Como

$$|T_f(\varphi)| \leq \left(\int_K |f| dx \right) \sup_{x \in K} |\varphi| \quad (9.4)$$

donde K es cualquier compacto conteniendo el soporte de φ vemos que T_f es continua y por lo tanto una distribución. Vemos así que las funciones continuas dan origen a distribuciones, es decir están incluidas naturalmente en el espacio de las distribuciones.

Ejercicio: Muestre que si $f \neq g$ luego $T_f \neq T_g$.

b) Sea f integrable (en el sentido de Lebesgue) es decir un elemento de $\mathcal{L}^1(\mathbb{R})$ y sea

$$T_f(\varphi) = \int_{\mathbb{R}} f \varphi dx \quad \forall \varphi \in C_0^\infty(\mathbb{R}). \quad (9.5)$$

Pero $|T_f(\varphi)| \leq \sup_{x \in K} |\varphi| \int_{\mathbb{R}} |f| dx$ y por lo tanto si $\{\varphi_n\} \rightarrow 0$ entonces $T_f(\varphi_n) \rightarrow 0$ lo que nos asegura (por linealidad) que T_f es continua y así una distribución. Note que si $f = g$ en casi todo punto ($f \sim g$) luego $T_f = T_g$ por lo que concluimos que son en realidad los elementos de L^1 los que definen estas distribuciones. Las distribuciones obtenidas de esta forma se denominan **regulares**.

c) Sea $T_a : C_0^\infty(\mathbb{R}) \rightarrow \mathbb{R}$, $a \in \mathbb{R}$, dada por $T_a(\varphi) = \varphi(a)$, este mapa es claramente lineal,

$$T_a(\varphi + \alpha\psi) = \varphi(a) + \alpha\psi(a) = T_a(\varphi) + \alpha T_a(\psi),$$

y continuo

$$|T_a(\varphi)| = |\varphi(a)| \leq \sup_{x \in \mathbb{R}} |\varphi(x)|$$

y por lo tanto una distribución. Esta es llamada la delta de Dirac en el punto a . ¿Hay alguna función continua, f , tal que $T_a = T_f$? Supongamos que sí, y que $f(r) \neq 0$ con $r \neq a$. Eligiendo φ distinto de cero solo en un entorno suficientemente pequeño de r que no contenga a a y con el mismo signo que $f(r)$ obtenemos $\varphi(a) = 0$ y $T_f(\varphi) \neq 0$ lo que implica que $f(r) = 0 \quad \forall r \neq a$ pero por continuidad concluimos entonces que $f \equiv 0$ y por lo tanto que $T_f(\varphi) = 0 \quad \forall \varphi \in C_0^\infty(\mathbb{R})$. Vemos entonces que esta distribución no proviene de ninguna función continua y se puede ver que en realidad no proviene de ningún elemento de $L^1(\mathbb{R})$, es así una **distribución irregular**. Estos son los elementos **extra** que nos da la generalización definida. Usualmente, en manipulaciones formales se pretende que esta distribución proviene de una función, denominada delta de Dirac y denotada por $\delta(x - a)$. Con ella se escriben cosas como

$$\int_{\mathbb{R}} \delta(x - a) \varphi(x) dx = \varphi(a). \quad (9.6)$$

Como hemos visto en realidad no existe ninguna función para la cual esta expresión tenga sentido por lo tanto ésta es solo **formal** y debe considerársela con precaución, es decir siempre como un mapa lineal y continuo del espacio de funciones de prueba.

¿Qué estructura tiene $\mathcal{D}'(\mathbb{R})$? Por ser un espacio dual ésta es un espacio vectorial con la suma de sus elementos y el producto de éstos por números reales definidos en la manera obvia, es decir si $T, \tilde{T} \in \mathcal{D}$, $\alpha \in \mathbb{R}$, luego $(T + \alpha\tilde{T})(\varphi) = T(\varphi) + \alpha(\tilde{T}(\varphi))$. Estas operaciones generalizan las operaciones definidas sobre funciones integrables ya que $T_f + \alpha T_g = T_{f+\alpha g}$.

¿Está definido el producto de distribuciones? La respuesta es que en general no lo está —así como tampoco lo está entre funciones integrables—. Sí lo está si una de ellas proviene de una función de prueba, es decir

$$T_\varphi \tilde{T}(\psi) := \tilde{T}(\varphi\psi) \quad \forall \psi \in \mathcal{D}, \quad (9.7)$$

donde hemos usado que los elementos de \mathcal{D} forman un álgebra. Note que ésto generaliza la operación definida sobre funciones integrables:

$$T_\varphi T_f(\psi) = \int_{\mathbb{R}} f \varphi \psi = T_f(\varphi\psi) = T_{\varphi f}(\psi) \quad (9.8)$$

ya que si $f \in L^1$ luego $f \varphi \in L^1$ si $\varphi \in \mathcal{D}$.

9.2. La derivada de una distribución

Si f es continuamente diferenciable luego su derivada también da origen a una distribución $T_{f'}$. ¿Qué relación hay entre T_f y $T_{f'}$? Note que

$$T_{f'}(\varphi) = \int_{\mathbb{R}} f' \varphi \, dx = - \int_{\mathbb{R}} f \varphi' \, dx = -T_f(\varphi') \quad \forall \varphi \in \mathcal{D} \quad (9.9)$$

donde al integrar por partes se usó que las φ en \mathcal{D} tienen soporte acotado. Esto sugiere que la posibilidad de extender la noción de derivada a toda distribución por medio de la fórmula

$$T'(\varphi) := -T(\varphi'), \quad \forall \varphi \in \mathcal{D}, \quad (9.10)$$

donde el lado derecho está bien definido ya que si $\varphi \in \mathcal{D}$ luego $\varphi' \in \mathcal{D}$. Note que esta operación satisface las condiciones que uno esperaría ya que son las mismas que satisface la derivada usual, teniendo en cuenta que ahora estamos en este espacio más amplio donde el producto de funciones no está definido en general. En efecto esta operación es lineal.

$$\begin{aligned} (T + \alpha \tilde{T})'(\varphi) &= -(T + \alpha \tilde{T})(\varphi') = -(T(\varphi') + \alpha \tilde{T}(\varphi')) \\ &= T'(\varphi) + \alpha \tilde{T}'(\varphi) \end{aligned} \quad (9.11)$$

y satisface la regla de Leibniz **tanto como es posible** ya que como vimos el producto de distribuciones no está definido en general. Cuando si lo está, es decir cuando una de ellas es un elemento de \mathcal{D} , tenemos que

$$\begin{aligned} (T_{\varphi} \tilde{T})'(\psi) &= -(T_{\varphi} \tilde{T})(\psi') = -\tilde{T}(\varphi \psi') \\ &= -\tilde{T}((\varphi \psi)' - \varphi' \psi) \\ &= -\tilde{T}((\varphi \psi)') + \tilde{T}(\varphi' \psi) \\ &= \tilde{T}'(\varphi \psi) + \varphi' \tilde{T}(\psi) \\ &= \tilde{T}_{\varphi} \tilde{T}'(\psi) + T_{\varphi'} \tilde{T}(\psi) \\ &= (T_{\varphi} \tilde{T}' + T_{\varphi'} \tilde{T})(\psi). \end{aligned} \quad (9.12)$$

Pero estas son las condiciones que definen una derivación. Vemos por lo tanto que ésta es una extensión de la derivada de una función en \mathcal{D} . Note que al generalizar la noción de función en la de distribución la hemos ampliado tanto que ahora objetos como la derivada de funciones discontinuas (incluso en todos sus puntos!) están incluidas entre éstas e incluso son a su vez infinitamente diferenciables $[(T)^{(n)}(\psi) \equiv (-1)^n T(\psi^{(n)})]$.

Ejemplo: T'_a , la derivada de la función de Dirac en a es la distribución tal que $T'_a(\varphi) = -\varphi'(a) \quad \forall \varphi \in \mathcal{D}$.

Comenzamos con el espacio de funciones infinitamente diferenciables de soporte compacto y finalizamos también con un espacio infinitamente diferenciable [con una noción distinta de derivada], cabe preguntarse si existe una noción de soporte de una distribución. Obviamente no podemos utilizar la misma noción que para funciones continuas y tendremos que proceder de una manera indirecta.

Sea O un abierto acotado, de \mathbb{R} y $\mathcal{D}(O)$ el espacio de funciones de prueba con soporte en la clausura de O , \bar{O} . Diremos que una distribución T se **anula en O** si $T[\mathcal{D}(O)] = 0$, es decir si se anula para toda función de prueba con soporte en O . Llamaremos el **soporte de T** al complemento de la unión de todos los abiertos donde T se anula. Por ser el complemento de un conjunto abierto este conjunto es cerrado.

Ejemplo: El soporte de T_o es $\{0\}$. Sea $O_n = (1/n, n) \cup (-n, -1/n)$, luego $\mathbb{R} - \bigcup_n O_n = \{0\}$.

Ejercicio: Sea f continua. Pruebe que $\text{soporte}\{f\} = \text{soporte}\{T_f\}$

Ejercicio: ¿Cómo extendería las nociones de funciones pares ($f(x) = f(-x)$) y el de impares ($f(x) = -f(-x)$). ¿Qué propiedades conserva esta extensión?

Ejercicio: Sea

$$g(x) = \begin{cases} x, & x \geq 0 \\ 0, & x \leq 0 \end{cases}$$

$x \in \mathbb{R}$. Claramente $g(x)$ es continua pero no diferenciable (en el sentido clásico). Encuentre las 3 primeras derivadas de g en el sentido de las distribuciones.

Ejercicio: La parte principal, en el sentido de Cauchy, de una función,

$$\mathcal{P}(1/x)(f) = \lim_{\varepsilon \rightarrow 0} \int_{|x| \geq \varepsilon} \frac{1}{x} f(x) dx$$

es una distribución. ¿Cómo debe interpretarse la fórmula?

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{x - x_0 + i\varepsilon} = \mathcal{P}\left(\frac{1}{x - x_0}\right) - i\pi \delta(x - x_0)$$

Hemos visto que una distribución es diferenciable, es decir dada $T \in \mathcal{D}'$ existe $S \in \mathcal{D}'$ tal que $T' = S$. Cabe preguntarse lo opuesto, es decir si dada $S \in \mathcal{D}'$ existe T tal que la fórmula de arriba valga, es decir

$$-T(\varphi') = S(\varphi) \quad \forall \varphi \in \mathcal{D}. \quad (9.13)$$

Esta es una generalización a distribuciones de la más simple de las ecuaciones diferenciales ordinarias ya estudiadas y la respuesta al problema planteado es afirmativa. Note que dada T la ecuación (9.13) nos define S , es decir su derivada, pero si damos S luego (9.13) no define T completamente ya que esta fórmula nos dice solamente cómo actúa T sobre funciones de prueba (φ') cuya integral (φ) es también una función de prueba. Esto es de esperar ya que en el caso de funciones la primitiva de una función está sólo determinada hasta una constante. Esta indeterminación se remedia dando **valores iniciales generalizados** lo cual se logra pidiendo que $T(\theta)$ tenga un dado valor, T_θ , para alguna $\theta \in \mathcal{D}$ que no sea la primitiva de otra función en \mathcal{D} , es decir que $\varphi(x) = \int_{-\infty}^x \theta(\tilde{x}) d\tilde{x}$ no tenga soporte acotado (o sea que $\int_{\mathbb{R}} \theta(\tilde{x}) d\tilde{x} \neq 0$).

Teorema 9.1 Dada $S \in \mathcal{D}'$, $\theta \in \mathcal{D}$ tal que $\int_{\mathbb{R}} \theta dx \neq 0$ y $T_\theta \in \mathbb{R}$ existe una única T satisfaciendo

$$\begin{aligned} -T(\varphi') &= S(\varphi) \quad \forall \varphi \in \mathcal{D} \\ T(\theta) &= T_\theta \end{aligned} \quad (9.14)$$

Prueba: Solo tenemos que conocer la acción de T sobre una arbitraria $\psi \in \mathcal{D}$. Sin pérdida de generalidad tomemos θ tal que $\int_{\mathbb{R}} \theta = 1$. Vemos que dada $\psi \in \mathcal{D}$ existe un único $\lambda \in \mathbb{R}$ y una única $\varphi \in \mathcal{D}$ tal que, $\psi - \lambda\theta = \varphi'$, es decir es una función de prueba con primitiva. En efecto, sea

$$\varphi(x) = \int_{-\infty}^x (\psi - \lambda\theta) d\tilde{x}$$

luego la condición para que φ tenga soporte compacto es que

$$0 = \int_{-\infty}^{\infty} (\psi - \lambda\theta) d\tilde{x} = \int_{-\infty}^{\infty} \psi dx - \lambda = 0 \quad (9.15)$$

o sea

$$\lambda = \int_{-\infty}^{\infty} \psi \, dx. \quad (9.16)$$

Sea entonces

$$T(\psi) = \lambda T_{\theta} - S(\varphi), \quad (9.17)$$

esta distribución satisface las ecuaciones del enunciado del teorema. λT_{θ} se puede considerar como la distribución constante, ya que $T_{\theta}\lambda(\psi) = T_{\theta} \int_{\mathbb{R}} \psi \, dx$. De esto concluimos que como en el caso de funciones dos soluciones cualesquiera de (9.13) difieren por una constante ♠

9.3. Nota sobre la completitud de \mathcal{D} y su dual \mathcal{D}'

Usando la noción de convergencia introducida en \mathcal{D} podemos definir un concepto análogo al de sucesión de Cauchy:

Definición: diremos que una sucesión de funciones de prueba $\{\varphi_n\}$, $\varphi_n \in \mathcal{D}$ es **convergente** si:

- 1) Existe $K \in \mathbb{R}$ compacto tal que $\text{sup}(\varphi_n) \subset K \, \forall n$.
- 2) Dado p y $\varepsilon > 0$ existe N tal que para todo $n, m > N$ se cumple

$$\sup_{x \in K} |f_n^{(p)}(x) - f_m^{(p)}(x)| < \varepsilon$$

Con esta noción de convergencia el espacio \mathcal{D} es completo, es decir toda sucesión convergente converge a un elemento de \mathcal{D} . Para discutir completitud de \mathcal{D}' debemos introducir nociones similares en este espacio. La noción de convergencia apropiada es la siguiente.

Definición: Diremos que la sucesión $\{T_n\}$, $T_n \in \mathcal{D}'$ **converge** a $T \in \mathcal{D}'$ si $T_n(\varphi) \rightarrow T(\varphi)$ para todo $\varphi \in \mathcal{D}$ ³.

Ejemplos:

- a) Sea T_n la distribución asociada con la función $e^{-|x-n|^2}$. Luego T_n converge a la distribución cero. Esto muestra lo débil que es este tipo de convergencia.
- b) Sea T_n la distribución asociada con alguna función f_n satisfaciendo

1. $f_n(t) \geq 0$ si $|t| < 1/n$ y cero si $|t| \geq 1/n$.

2. $\int_a^b f_n(t) \, dt = 1,$

Luego $T_n \rightarrow T_0$, la función de Dirac con soporte en el cero.

Análogamente podemos definir la noción de convergencia de distribuciones.

Definición: $\{T_n\}$, $T_n \in \mathcal{D}'$ es **convergente** si para cada $\varphi \in \mathcal{D}$ y $\varepsilon > 0$ existe N tal que si $n, m > N$ luego

$$|T_n(\varphi) - T_m(\varphi)| < \varepsilon$$

Con esta noción de convergencia el espacio \mathcal{D}' es completo.

³Nuevamente estamos introduciendo una topología de manera indirecta, esta vez en \mathcal{D}' .

9.4. Convergencia y Compacidad Débil

En un espacio normado, H , tenemos la noción de convergencia con respecto a la norma, la cual llamaremos convergencia fuerte $\{x_n\} \xrightarrow{f} x$ si $\lim_{n \rightarrow \infty} \|x_n - x\|_H = 0$. En estos espacios existe otra noción de convergencia, la llamada convergencia débil y que usa la existencia del espacio dual de H , H' .

Definición: Diremos que $\{x_n\}$ **converge débilmente** a x , $\{x_n\} \xrightarrow{d} x$, si $f(x_n) \rightarrow f(x) \forall f \in H'$.

Si H es un espacio de Hilbert (lo que supondremos de ahora en más) el Teorema de Representación de Riez nos dice que $\{x_n\} \xrightarrow{d} x$ sii $(x_n, y) \rightarrow (x, y) \forall y \in H$.

Claramente esta noción de convergencia es la más débil tal que los elementos de H' son funcionales continuos [En el sentido que f es continua en x si **dada cualquier sucesión** $\{x_n\}$ **convergiendo a** x luego $\lim_{n \rightarrow \infty} f(x_n) = f(x)$], donde decimos que una noción de convergencia es más débil que otra si toda sucesión que converge con respecto a la segunda también lo hace con respecto a la primera y hay sucesiones que convergen con respecto a la primera pero no con respecto a la segunda. Veamos, a manera de ejemplo, que la convergencia en norma, o fuerte, es en realidad más fuerte que la llamada débil. Supongamos entonces que $\{x_n\} \xrightarrow{f} x$ es decir $\lim_{n \rightarrow \infty} \|x - x_n\|_H = 0$, luego como los elementos de H' son funcionales lineales acotadas (\iff continuas) se cumple que

$$|f(x) - f(x_n)| \leq \|f\|_{H'} \|x - x_n\| \quad \forall f \in H', \quad (9.18)$$

y por lo tanto,

$$\lim_{n \rightarrow \infty} |f(x) - f(x_n)| = 0, \quad (9.19)$$

o sea $\{x_n\} \xrightarrow{d} x$. El siguiente es un ejemplo de una sucesión que convergen débilmente y no fuertemente.

Ejercicio: Muestre que la sucesión $\{x_n = (0, \dots, 0, \underset{n}{1}, 0, \dots)\}$ converge débilmente en l^2 pero no fuertemente.

El anterior fue también el ejemplo que utilizamos para mostrar que la bola de radio unidad en l^2 no era compacta con respecto a la convergencia fuerte. ¿Será ésta débilmente compacta? Es decir, dada una sucesión $\{x_n\} \in B_1(l^2)$ cualquiera ¿existirá una subsucesión que converja débilmente? La respuesta es afirmativa y es una de las herramientas más útiles del análisis funcional.

Teorema 9.2 B_1 es débilmente compacta.

Prueba: Sólo probaremos el caso en que H es separable. Sea $\{x_n\}$ con $\|x_n\|_H \leq 1$ y $S = \{y_m\}$ una base ortonormal numerable de H . Construiremos, usando inducción, una subsucesión $\{x_n^\infty\}$ tal que,

$$(x_n^\infty, y_m) \xrightarrow{n \rightarrow \infty} \alpha_m \quad \forall y_m \in S. \quad (9.20)$$

Sea $m = 1$, luego $|(x_n, y_1)| \leq \|x_1\|_H \|y_1\|_H \leq 1$. Vemos entonces que $\{(x_n, y_1)\}$ es una sucesión acotada en \mathbb{C} . Pero la bola de radio unidad en \mathbb{C} es compacta y por lo tanto habrá alguna subsucesión $\{x_n^1, y_1\}$ convergiendo a algún α_1 en \mathbb{C} . Supongamos ahora que tenemos una subsucesión $\{x_n^{m-1}\}$ tal que

$$(x_n^{m-1}, y_p) \rightarrow \alpha_p \quad \forall 1 \leq p \leq m-1. \quad (9.21)$$

De la misma forma que como hicimos para el caso $m = 1$, considerando en este caso $\{(x_n^{m-1}, y_m)\}$ obtenemos una subsucesión $\{x_n^m\}$ de $\{x_n^{m-1}\}$ que satisface,

$$(x_n^m, y_m) \xrightarrow{n \rightarrow \infty} \alpha_m, \quad (9.22)$$

lo que completa la inducción. Tenemos así un mapa $f : \{y_m\} \rightarrow \mathbf{C}$ dado por $f(y_m) = \alpha_m$, como $\{y_m\}$ es una base podemos extender este mapa linealmente a todo H . Como $\{x_n^\infty\}$ es acotada,

$$|f(y)| = \lim |(x_n^\infty, y)| \leq \|y\|_H \quad (9.23)$$

y f es también acotada, por lo tanto continua. Usando el Teorema de Representación de Riez sabemos entonces que existe $x \in H$ tal que

$$(x, y) = f(y) = \lim (x_n^\infty, y) \quad \forall y \in H. \quad (9.24)$$

Concluimos así que $\{x_n^\infty\} \xrightarrow{d} x \spadesuit$

Notas bibliográficas: Recomendando leer: [10], [1] [9]. A pesar de que fue un físico, Dirac, el que introdujo el concepto de distribución muchos físicos las desdeñan como algo *matemático* y las usan como una abreviación útil para hacer cálculos. Usualmente la persona que las manipula conoce lo que está haciendo y no comete errores, pero es bastante fácil cometerlos si no se siguen cuidadosamente las reglas y se pierde de vista lo que son. Esto por ejemplo lleva a errores tales como asignarle significado al producto de dos distribuciones arbitrarias. No es difícil entender el concepto básico de distribución ni tampoco que no hay que salirse de las reglas operativas, sígalas siempre y no se equivocará.

10.1. Introducción

Consideremos el siguiente problema en el círculo S^1 cuya circunferencia es 2π :
Dada ρ continua en S^1 encuentre f en S^1 tal que

$$\frac{\partial^2 f}{\partial \theta^2} = \rho. \quad (10.1)$$

Una manera de resolver este problema es utilizando la base ortogonal de Fourier, $\left\{ \frac{1}{\sqrt{2\pi}} e^{in\theta} \right\}$ de $L^2(S^1)$. Sea $F : L^2(S^1) \rightarrow l^2$ el mapa entre estos espacios generado por esta base, es decir,

$$F(f) = \left\{ \left(\frac{1}{\sqrt{2\pi}} e^{in\theta}, f \right) \right\} \equiv \{c_n\}. \quad (10.2)$$

Luego

$$\begin{aligned} F\left(\frac{\partial^2 f}{\partial \theta^2} - \rho\right) &= \left\{ \left(\frac{1}{\sqrt{2\pi}} e^{in\theta}, \frac{d^2 f}{d\theta^2} - \rho \right) \right\} \\ &= \{-n^2 c_n - a_n\} \end{aligned} \quad (10.3)$$

donde $\{a_n\} = F(\rho)$.

Como $F(0) = \{0\}$ vemos que (10.1) implica una ecuación algebraica en l^2 ,

$$-n^2 c_n = a_n. \quad (10.4)$$

Si ρ es L^2 ortogonal a $f = cte.$, es decir a las únicas soluciones de (10.1) con $\rho = 0$, lo que implica $a_0 = 0$, la sucesión con c_0 arbitrario y $c_n = -\frac{a_n}{n^2}$, $n \neq 0$ satisface (10.4). El mapa inverso

$F^{-1} : l^2 \rightarrow L^2(S^1)$ nos define $f(\theta) = \sum_{n=-\infty}^{\infty} \frac{c_n}{\sqrt{2\pi}} e^{in\theta}$, el cual al menos formalmente¹ es una solución de (10.1).

Note que en esta aplicación no hemos usado directamente el hecho de que las funciones $\left\{ \frac{1}{\sqrt{2\pi}} e^{in\theta} \right\}$ forman una base ortogonal sino solo que $\frac{d}{d\theta} e^{in\theta} = in e^{in\theta}$ y ciertas propiedades del mapa F generado por dicha base. En particular esta propiedad de la base se induce en el mapa en el sentido que si $f \in L^2$ es diferenciable y $F(f) = \{c_n\}$ luego $F\left(\frac{df}{d\theta}\right) = \{+in c_n\}$.

Estas observaciones son muy útiles pues existen casos, como el de $L^2(\mathbb{R})$, en los cuales no se conocen bases ortogonales interesantes, pero sí mapas similares a F con propiedades interesantes.

¹El mapa define un elemento de L^2 pues $\rho \in L^2$ y por lo tanto $\sum_{n=-\infty}^{\infty} |a_n|^2 < \infty$ lo que implica que $\{c_0 \text{ arbitrario}, c_n = -\frac{a_n}{n^2}, n \neq 0\}$ es también un elemento de l^2 . Restaría ver que $f = F^{-1}\{c_n\}$ es dos veces diferenciable.

Uno de ellos es la transformación de Fourier que estudiaremos ahora. El problema en $L^2(\mathbb{R})$ es que desearíamos tener una base $\{\varphi_n\}$ cuyas funciones satisfagan $\frac{d}{dx}\varphi_n = i c_n \varphi_n$, pero las soluciones a esta ecuación son $\varphi_n = a_n e^{i c_n x}$, las cuales no son funciones de cuadrado integrable cualquiera sea c_n o $a_n \in \mathbb{R}$ (excepto por supuesto que tomemos $a_n \equiv 0$). Sin embargo aunque estas funciones no formen una base ortogonal si generan un mapa \mathcal{F} , esta vez de $L^2(\mathbb{R})$ en otra copia (considerada como distinta) de $L^2(\mathbb{R})$, con propiedades similares a las del mapa F considerado anteriormente.

Teorema 10.1 (de Fourier) *La transformación de Fourier*

$$\hat{f}(\lambda) := \mathcal{F}(f)(\lambda) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix \cdot \lambda} f(x) dx, \quad \text{donde } x \cdot \lambda = \sum_{i=1}^n x^i \lambda_i \quad (10.5)$$

Satisface:

a) $\mathcal{F} : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ es un mapa lineal, continuo e invertible de $L^2(\mathbb{R}^n)$ en sí mismo que preserva la norma (es decir unitario), $\|\hat{f}\|_{L^2} = \|\mathcal{F}(f)\|_{L^2} = \|f\|_{L^2}$ (identidad de Plancherel).

b) Su inversa está dada por

$$\check{g}(x) = \mathcal{F}^{-1}(\hat{g}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix \cdot \lambda} \hat{g}(\lambda) d\lambda. \quad (10.6)$$

c) Si la derivada de $f \in L^2(\mathbb{R}^n)$ en el sentido distribucional está también en $L^2(\mathbb{R}^n)$ se cumple entonces que

$$\mathcal{F}\left(\frac{\partial f}{\partial x_j}\right)(\lambda) = i \lambda_j \mathcal{F}(f)(\lambda). \quad (10.7)$$

Este teorema nos dice que esta transformación tiene todas las propiedades que tenía la generada por la base de Fourier y por lo tanto será tan útil como aquella en aplicaciones similares, pero ahora en \mathbb{R}^n .

La prueba de este teorema usa varias de las técnicas más comunes y poderosas del análisis funcional y por lo tanto es recomendable una lectura atenta y no automática de la misma.

Prueba: El mapa \mathcal{F} es obviamente lineal y claramente está definido para cualquier función en $C^\infty(\mathbb{R}^n)$, es decir infinitamente diferenciable y de soporte compacto. Veamos primero que \mathcal{F} preserva la norma $L^2(\mathbb{R}^n)$ para estas funciones. Sea $f \in C^\infty(\mathbb{R}^n)$ cualquiera y tomemos un n-cubo C_ε de volumen $(\frac{\varepsilon}{2})^n$ con $\varepsilon > 0$ suficientemente pequeño como para que $\text{supp}(f) \subset C_\varepsilon$. Sea $K_\varepsilon = \{k \in \mathbb{R}^n \mid k_i = p\varepsilon \text{ para algún entero } p\}$, [por ejemplo en \mathbb{R}^3 el vector $(\pi\varepsilon 5, \pi\varepsilon 17, 0)$ es un elemento de K_ε]. Luego el conjunto de funciones $\left\{ \left(\frac{\varepsilon}{2}\right)^{n/2} e^{ik \cdot x} \mid k \in K_\varepsilon \right\}$ forman una base ortogonal de $L^2(C_\varepsilon)$.

Pero $f \in L^2(C_\varepsilon)$ y es continuamente diferenciable, por lo tanto

$$\sum_{k \in K_\varepsilon} \left(\left(\frac{\varepsilon}{2}\right)^{n/2} e^{ik \cdot \tilde{x}}, f(\tilde{x}) \right) \left(\frac{\varepsilon}{2}\right)^{n/2} e^{ik \cdot x} = \sum_{k \in K_\varepsilon} \frac{\hat{f}(k)}{(2\pi)^{n/2}} e^{ik \cdot x} (2\pi\varepsilon/2)^n, \quad (10.8)$$

es la representación en serie de Fourier de f la cual converge uniformemente a f en C_ε y por lo tanto en \mathbb{R}^n . Tenemos entonces que

$$\begin{aligned} \|f\|_{L^2}^2 &= \int_{\mathbb{R}^n} |f(x)|^2 dx = \int_{C_\varepsilon} |f(x)|^2 dx = \sum_{k \in K_\varepsilon} \left| \left(\frac{\varepsilon}{2}\right)^{n/2} e^{ik \cdot x}, f(x) \right|^2 \\ &= \sum_{k \in K_\varepsilon} |\hat{f}(k)|^2 (\pi\varepsilon)^n. \end{aligned} \quad (10.9)$$

Como \mathbb{R}^n es la unión de n -cubos de lados $\pi\varepsilon$ (o sea de volumen $(\pi\varepsilon)^n$ alrededor de los puntos de K_ε el lado derecho de (10.9) es simplemente la serie de Riemann de la función $|\hat{f}(k)|^2$. Si esta función fuese continua luego la serie de Riemann convergería a $\|\hat{f}\|_{L^2}^2$ y habríamos probado que la transformada de Fourier de una función en $C^\infty(\mathbb{R}^n)$ preserva la norma $L^2(\mathbb{R}^n)$. Pero

$$\begin{aligned} \sup_k \left| \frac{\partial \hat{f}(k)}{\partial k_j} \right| &= \sup_k \left| \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} -i x^j e^{-ik \cdot x} f(x) d^n x \right| \\ &\leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |x^j f(x)| d^n x < \infty, \end{aligned} \quad (10.10)$$

ya que $f \in C^\infty(\mathbb{R}^n)$. Vemos entonces que las derivadas de $\hat{f}(k)$ están acotadas y por lo tanto que \hat{f} es continua. Este resultado no solo completa la demostración de que \mathcal{F} preserva la norma sino también, ya que $\hat{f}(k)$ es continua, que la serie de Riemann en el lado derecho de la ecuación (10.8) converge a la integral y por lo tanto que $\mathcal{F}^{-1}(\mathcal{F}(f)) = f(x) \forall f \in C^\infty(\mathbb{R}^n)$.

Si $f \in C^\infty(\mathbb{R}^n)$ luego la identidad del punto c) se muestra trivialmente. Solo resta entonces extender estos resultados para f arbitraria en $L^2(\mathbb{R}^n)$, pero $C^\infty(\mathbb{R}^n)$ es un subespacio denso de $L^2(\mathbb{R}^n)$, es decir todo elemento $f \in L^2(\mathbb{R}^n)$ puede obtenerse como límite (con respecto a la norma $L^2(\mathbb{R}^n)$) de una sucesión $\{f_n\}$ de funciones en $C^\infty(\mathbb{R}^n)$, esto nos permite extender la acción de \mathcal{F} a todo elemento de $L^2(\mathbb{R}^n)$. En efecto, sea $f \in L^2(\mathbb{R}^n)$ arbitraria y sea $\{f_n\} \rightarrow f \in C^\infty(\mathbb{R}^n)$ luego la sucesión $\hat{f}_n(k) = \mathcal{F}(f_n)$ satisface,

$$\|\hat{f}_n(k) - \hat{f}_m(k)\|_{L^2(\mathbb{R}^n)} = \|f_n(x) - f_m(x)\|_{L^2(\mathbb{R}^n)}. \quad (10.11)$$

Como $\{f_n\}$ es convergente es de Cauchy y por lo tanto $\{\hat{f}_n\}$ es también de Cauchy en $L^2(\mathbb{R}^n)$, pero $L^2(\mathbb{R}^n)$ es completo y por lo tanto existe $\hat{f} \in L^2(\mathbb{R}^n)$ tal que $\{\hat{f}_n\} \rightarrow \hat{f}$. Extenderemos el mapa \mathcal{F} a $L^2(\mathbb{R}^n)$ definiendo $\mathcal{F}(f) \equiv \hat{f}$. Esta extensión es claramente lineal y acotada por lo tanto es continua. Usando el mismo razonamiento extendemos \mathcal{F}^{-1} a todo $L^2(\mathbb{R}^n)$ y vemos que $\mathcal{F}^{-1}(\mathcal{F}(f(x))) = f(x) \forall f(x) \in L^2(\mathbb{R}^n)$ lo que demuestra que \mathcal{F} es invertible, que su inversa es \mathcal{F}^{-1} y es continua. El mismo argumento muestra que la fórmula en c) también es válida para todo f tal que su gradiente también está en $L^2(\mathbb{R}^n)$. Esto completa la prueba del teorema ♠

Note que la estrategia ha sido primero tomar un espacio ($C^\infty(\mathbb{R}^n)$) donde todas las propiedades valen obviamente, luego tomar un espacio que contiene al anterior y donde este es denso, ver que el mapa es allí continuo (con respecto a la noción de continuidad del espacio mayor) y finalmente tomar la extensión que esta continuidad nos brinda.

¿Hay otras extensiones posibles? La respuesta es sí y ahora veremos una de ellas que nos permitirá usar la transformación de Fourier en distribuciones. Lo haremos en forma de un problema dividido en varios ejercicios.

Notación de Schwartz: I_+^n denotará el conjunto de n -tuplas de enteros no-negativos $\alpha = \langle \alpha_1, \dots, \alpha_n \rangle$

y $|\alpha| = \sum_{i=1}^n \alpha_i$. $D^\alpha \equiv \frac{\partial^\alpha}{\partial x^\alpha} \equiv \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$, es decir D^α es el operador diferencial que toma α_1 , derivadas parciales con respecto a x_1 , α_2 derivadas parciales con respecto a x_2 , etc. El grado de D^α es $|\alpha|$.

Definición: Llamaremos **espacio de Schwartz** $\mathcal{S}(\mathbb{R}^n)$ al espacio vectorial de funciones infinitamente diferenciables que decaen asintóticamente más rápido que la inversa de cualquier polinomio, es decir que

$$\|\varphi\|_{\alpha, \beta} \equiv \sum_{\substack{|\tilde{\alpha}| \leq |\alpha| \\ |\tilde{\beta}| \leq |\beta|}} \sup_{x \in \mathbb{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty \quad (10.12)$$

$\forall \alpha, \beta \in I_+^n$, con la siguiente noción de convergencia: diremos que $\{\varphi_n\}$, $\varphi_n \in \mathcal{S}$ converge a $\varphi \in \mathcal{S}$ si dado $\varepsilon > 0$ para cada par $\alpha, \beta \in I_+^n$ existe $N_{\alpha, \beta}$ tal que si $n > N_{\alpha, \beta}$ luego $\|\varphi - \varphi_n\|_{\alpha, \beta} < \varepsilon$.

Ejercicios:

- 1) ¿Cómo definiría la noción de sucesión convergente en este espacio?
- 2) Muestre que si $\{\varphi_n\} \rightarrow \varphi$ luego $\{\varphi_n\}$ converge en el sentido anterior.
- 3) Muestre que \mathcal{D} está estrictamente contenido en \mathcal{S} . Ayuda: Encuentre $\varphi \in \mathcal{S}$ con $\text{sop}(\varphi) = \mathbb{R}^n$.

Las propiedades de este espacio que nos interesan son las siguientes:

Lema 10.1 *Con la noción de convergencia correspondiente el espacio $\mathcal{S}(\mathbb{R}^n)$ es completo.*

Lema 10.2 *El espacio $C^\infty(\mathbb{R}^n)$ es denso en \mathcal{S} , es decir cualquier elemento $\varphi \in \mathcal{S}$ puede ser obtenido como límite de una sucesión convergente con cada uno de sus miembros en $C^\infty(\mathbb{R}^n)$.*

Ejercicio: Pruebe el Lema 10.2.

Lema 10.3 $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$ es continuo e invertible con inversa continua.

Ejercicio: Pruebe el Lema 10.3. Ayuda: Pruebe que dado $\alpha, \beta \in I_+^n$ existen $\tilde{\alpha}, \tilde{\beta} \in I_+^n$ y $c > 0$ tales que para todo $\varphi \in \mathcal{S}$ se cumple

$$\|\hat{\varphi}\|_{\alpha, \beta} \leq c \|\varphi\|_{\tilde{\alpha}, \tilde{\beta}}. \quad (10.13)$$

Ejercicio: Encuentre $\mathcal{F}(e^{-\alpha x^2/2})$.

Definición: El espacio dual a \mathcal{S} , \mathcal{S}' se llama el espacio de las **distribuciones temperadas**.

Ejercicio: Muestre que \mathcal{S}' está estrictamente contenido en \mathcal{D}' . Ayuda: Encuentre f tal que $T_f \in \mathcal{D}'$ y no a \mathcal{S}' .

¿Cómo extender \mathcal{F} a \mathcal{S}' ? Usando la identidad de Plancherel y la de polarización² tenemos que $(\varphi, \psi) = (\hat{\varphi}, \hat{\psi})$. Pero entonces

$$T_{\hat{\sigma}}(\hat{\psi}) = \int \hat{\sigma} \hat{\psi} dk = \int \sigma \psi dx = T_{\sigma}(\psi) := \hat{T}_{\sigma}(\hat{\psi}). \quad (10.14)$$

Esto nos induce a definir en general

$$\hat{T}(\hat{\varphi}) := T(\varphi). \quad (10.15)$$

Ejercicio: Calcule $\mathcal{F}(\delta_a)$ y $\mathcal{F}(\delta'_a)$.

¿Qué otras propiedades tiene la transformación de Fourier? Note que de la identidad del punto c) del Teorema de Fourier sigue que si $x^\alpha D^\beta f \in L^2(\mathbb{R}^n)$ luego $k^\beta D^\alpha \hat{f} \in L^2(\mathbb{R}^n)$. Esto nos dice esencialmente que diferenciabilidad en x es equivalente a decaimiento en k y viceversa. En particular esto nos dice que la transformada de Fourier de una función de soporte compacto es infinitamente diferenciable.³ Más aún se puede probar que es analítica.

²Es decir que $(x, y) = \frac{1}{4}\{||x+y||^2 - ||x-y||^2 + i||x-iy||^2 - i||x+iy||^2\}$

³En realidad el argumento anterior nos dice que si $\text{sop}(f)$ es compacto luego $D^\alpha \hat{f} \in L^2(\mathbb{R}^n) \quad \forall \alpha \in I_+^n$. Como veremos más adelante esto implica que es infinitamente diferenciable.

Otra propiedad importante de la transformación de Fourier es que lleva producto de funciones en convolución de funciones y viceversa.

Definición: Sea $f, g \in \mathcal{S}(\mathbb{R}^n)$. La **convolución** de f con g denotada $f * g$ es la función, también en $\mathcal{S}(\mathbb{R}^n)$, dada por

$$(f * g)(y) = \int_{\mathbb{R}^n} f(y - x) g(x) d^n x. \quad (10.16)$$

Teorema 10.2 a) Para cada $f \in \mathcal{S}(\mathbb{R}^n)$, el mapa lineal $g \rightarrow f * g$ de $\mathcal{S}(\mathbb{R}^n)$ en $\mathcal{S}(\mathbb{R}^n)$ es continuo.

$$b) \widehat{f * g} = \frac{1}{(2\pi)^{n/2}} \widehat{f} * \widehat{g} \quad y \quad \widehat{f * g} = (2\pi)^{n/2} \widehat{f} \widehat{g}.$$

$$c) f * g = g * f \quad y \quad f * (g * h) = (f * g) * h.$$

Prueba: Una vez probado el punto b) los otros siguen trivialmente. Asegúrese! Probaremos entonces b). Como ya vimos $(\varphi, \psi) = (\widehat{\varphi}, \widehat{\psi}) \quad \forall \varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$. Aplicando esta identidad a $e^{iy \cdot x} \bar{f}(x)$ y $g(x)$ obtenemos $(e^{iy \cdot x} \bar{f}, g) = (\widehat{e^{iy \cdot x} \bar{f}}, \widehat{g})$. pero

$$(e^{iy \cdot x} \bar{f}, g) = \int_{\mathbb{R}^n} e^{-iy \cdot x} f(x) g(x) dx = (2\pi)^{n/2} \widehat{f} \widehat{g} \quad (10.17)$$

y

$$\begin{aligned} (\widehat{e^{iy \cdot x} \bar{f}}, \widehat{g}) &= \int_{\mathbb{R}^n} \left((2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\lambda \cdot x + iy \cdot x} \bar{f}(x) d^n x \right) \widehat{g}(\lambda) d^n \lambda \\ &= \int_{\mathbb{R}^n} \widehat{f}(y - \lambda) \widehat{g}(\lambda) d^n \lambda \\ &= \widehat{f} * \widehat{g}. \end{aligned} \quad (10.18)$$

La otra fórmula se obtiene aplicando \mathcal{F}^{-1} a la anterior ♠

Otra propiedad interesante de la convolución se obtiene al notar que como $\mathcal{S}(\mathbb{R}^n)$ es cerrado con respecto a la operación de convolución ($f, g \in \mathcal{S}(\mathbb{R}^n) \implies f * g \in \mathcal{S}(\mathbb{R}^n)$) podemos definir la convolución de una distribución temperada con una función en $\mathcal{S}(\mathbb{R}^n)$ como,

$$(T * f)(\varphi) = T(\tilde{f} * \varphi) \quad \text{donde } \tilde{f}(x) := f(-x). \quad (10.19)$$

Ejercicio: Aplique (10.19) a $T = T_g$ para $g \in \mathcal{S}(\mathbb{R}^n)$ y vea que esta definición tiene sentido. La propiedad interesante se ve en el siguiente lema.

Lema 10.4 $T * f$ es equivalente a una función infinitamente diferenciable.

10.2. *Propiedades básicas de los Espacios de Sobolev

Como una aplicación de la transformación de Fourier veremos a continuación las propiedades básicas de los espacios de Sobolev. Éstas serán usadas más adelante cuando estudiemos la teoría de las ecuaciones diferenciales parciales. El primer paso será una generalización de los espacios de Sobolev. Como vimos los espacios de Sobolev son espacios de Hilbert con norma proveniente del siguiente producto escalar.

$$(f, g)_{H^m} = \sum_{k=0}^m \sum_{i,j,p=1}^n \left(\underbrace{\partial_i \partial_j \dots \partial_p}_{k \text{ veces}} f, \partial_i \partial_j \dots \partial_p g \right)_{L^2}, \quad (10.20)$$

donde f y g son funciones definidas en un abierto Ω cualquiera de \mathbb{R}^n . Si $\Omega = \mathbb{R}^n$ luego de las propiedades de la transformación de Fourier obtenemos

$$\begin{aligned} (f, g)_{H^m} &= \sum_{k=0}^m \sum_{i,j,p=1}^n (\lambda_i \lambda_j \dots \lambda_p \hat{f}(\lambda), \lambda_i \lambda_j \dots \lambda_p \hat{g}(\lambda))_{L^2} \\ &= (\hat{f}(\lambda) \sqrt{\sum_{k=0}^m |\lambda|^{2k}}, \hat{g}(\lambda) \sqrt{\sum_{k=0}^m |\lambda|^{2k}})_{L^2}, \end{aligned} \quad (10.21)$$

es decir las funciones $f(x) \in H^m(\mathbb{R}^n)$ son aquellas que su transformada de Fourier \hat{f} decae lo suficientemente rápido como para que $\hat{f}(\lambda) \sqrt{\sum_{k=0}^m |\lambda|^{2k}}$ sea de cuadrado integrable.

Esto nos sugiere generalizar los espacios permitiendo índices no enteros y aún negativos por medio del siguiente producto escalar,

$$(f, g)_{H^s} = (\hat{f}(\lambda)(1 + |\lambda|^2)^{s/2}, \hat{g}(\lambda)(1 + |\lambda|^2)^{s/2})_{L^2}. \quad (10.22)$$

Ejercicio: Muestre que este es un producto escalar.

Note que para $s = m$ en vez del polinomio $\sum_{k=0}^m |\lambda|^{2k}$, ahora tenemos $(1 + |\lambda|^2)^m$. Como existen constantes positivas c_1 y c_2 , tales que

$$c_1(1 + |\lambda|^2)^m \leq \sum_{k=0}^m |\lambda|^{2k} \leq c_2(1 + |\lambda|^2)^m \quad (10.23)$$

este cambio en la norma es trivial en el sentido de que estas normas son equivalentes entre sí. En realidad, como veremos en un caso especial, incluso el 1 en el polinomio se puede ignorar y aún así obtener una norma equivalente. La primera propiedad que veremos es el siguiente lema:

Lema 10.5 Si $s' > s$ luego $H^{s'} \subset H^s$.

Prueba: Trivial ♠

En particular si $s > 0$ luego $H^s \subset H^0 = L^2$, si $s < 0$ luego $L^2 \subset H^s$. ¿Qué son los elementos de H^s para s negativo?

Dada $g \in H^{-s}$ puedo definir el siguiente mapa de H^s en \mathbf{C} ,

$$\Psi_g(f) := (\hat{g}, \hat{f})_{H^0} = \left(\frac{\hat{g}}{(1 + |\lambda|^2)^{s/2}}, \hat{f}(1 + |\lambda|^2)^{s/2} \right)_{H^0}.$$

Este mapa es lineal, está bien definido $\forall f, g \in C_0^\infty$ y puede ser extendido a todo H^s (y H^{-s}) por continuidad, ya que,

$$|\Psi_g(f)| \leq \|g\|_{H^{-s}} \|f\|_{H^s}.$$

Por lo tanto tenemos un mapa entre H^{-s} y el dual de H^s , que preserva la norma, es más, se puede probar que este mapa es un isomorfismo (o sea que además es invertible) entre estos espacios.⁴ Por lo tanto podemos identificar a H^{-s} con el dual de H^s .

Ejercicio: Usando que $C^\infty(\mathbb{R}^n)$ es denso en $H^s(\mathbb{R}^n)$ muestre que la delta de Dirac está en $H^s(\mathbb{R}^n)$ para todo $s < -n/2 - 1$.

Quizás la más importante propiedad de $H^s(\mathbb{R}^n)$ es la siguiente,

⁴Por el teorema de Representación de Riez cada $\Psi : H^s \rightarrow \mathbf{C}$ puede además ser escrito como $\Psi(f) = (\tilde{g}, f)_{H^s}$, $\tilde{g} \in H^s$. Si tomamos $g = \mathcal{F}^{-1}((1 + |\lambda|^2)^s \mathcal{F}(\tilde{g})) \in H^{-s}$ entonces $\Psi(f) = \Psi_g(f)$.

Lema 10.6 (de Sobolev) Si $m < s - n/2$ luego $H^s(\mathbb{R}^n) \subset C^m(\mathbb{R}^n)$.

Prueba: Basta con probarlo para $m = 0$, el resto sigue por inducción. Basta también probar la desigualdad. $\|f\|_{C^0} < C \|f\|_{H^s}$, $s > \frac{n}{2}$ para alguna constante $C > 0$ independiente de f y para toda $f \in C^\infty(\mathbb{R}^n)$, el resto sigue por la continuidad de la norma. Pero

$$\begin{aligned}
 \|f\|_{C^0} = \sup_{x \in \mathbb{R}^n} |f(x)| &= \sup_{x \in \mathbb{R}^n} \left| \int_{\mathbb{R}^n} e^{i\lambda \cdot x} \hat{f}(\lambda) d^n \lambda \right| \\
 &\leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |\hat{f}(\lambda)| d^n \lambda \\
 &\leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{|\hat{f}(\lambda)| (1 + |\lambda|^2)^{s/2}}{(1 + |\lambda|^2)^{s/2}} d^n \lambda \\
 &\leq \frac{1}{(2\pi)^{n/2}} \|\hat{f}(\lambda)(1 + |\lambda|^2)^{s/2}\|_{L^2} \left\| \frac{1}{(1 + |\lambda|^2)^{s/2}} \right\|_{L^2} \\
 &\leq C \|f\|_{H^s}.
 \end{aligned} \tag{10.24}$$

donde hemos usado que $s > n/2$ para probar que $\left\| \frac{1}{(1 + |\lambda|^2)^{s/2}} \right\|_{L^2} < \infty$ ♠

Este lema nos dice que si $f \in H^m(\mathbb{R}^n)$ para m suficientemente grande luego f puede ser identificada con una función ordinaria, continua e incluso diferenciable.

Sea f una función continua en $\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x_n \geq 0\}$ y sea τf la restricción de esa función al hiperplano $x_n = 0$ es decir $(\tau f)(x_1, x_2, \dots, x_{n-1}) = f(x_1, x_2, \dots, x_{n-1}, 0)$. Claramente τ es un mapa lineal y si f es continua en \mathbb{R}_+^n luego τf es continua en $\mathbb{R}^{n-1} = \{x \in \mathbb{R}^n \mid x_n = 0\}$. ¿Se puede extender este mapa a funciones más generales? La respuesta es el siguiente lema que nos muestra además el porqué de la necesidad de extender los espacios de Sobolev a índices no enteros.

Lema 10.7 (Traza) Sea $m > 0$ entero, luego:

- i) $\tau : H^m(\mathbb{R}_+^n) \rightarrow H^{m-1/2}(\mathbb{R}^{n-1})$ es continuo.
- ii) es suryectivo.

Prueba: Por el mismo argumento que en el lema anterior para probar i) basta probar que existe $c > 0$ tal que

$$\|\tau f\|_{H^{1/2}(\mathbb{R}^{n-1})} < C \|f\|_{H^1(\mathbb{R}_+^n)} \quad \forall f \in C_0^\infty(\mathbb{R}_+^n) \tag{10.25}$$

Sea $\hat{f}(\lambda', x_n)$ la transformada de Fourier de $f(x)$ con respecto a las coordenadas $(x_1, x_2, \dots, x_{n-1})$ de \mathbb{R}_+^n , luego

$$|\hat{f}(\lambda', 0)|^2 = -2 \operatorname{Re} \int_0^\infty \frac{\partial \hat{f}}{\partial x_n}(\lambda', t) \overline{\hat{f}(\lambda', t)} dt. \tag{10.26}$$

Multiplicando ambos lados de esta igualdad por $(1 + |\lambda'|^2)^{1/2}$ e integrando con respecto a λ' obtenemos,

$$\begin{aligned}
 \|\tau f\|_{H^{1/2}(\mathbb{R}^{n-1})} &= \int_{\mathbb{R}^{n-1}} |\hat{f}(\lambda', 0)|^2 d^{n-1} \lambda' \\
 &= 2 \left| \int_{\mathbb{R}^{n-1}} \int_0^\infty (1 + |\lambda'|^2)^{1/2} \frac{\partial \hat{f}(\lambda', t)}{\partial x_n} \overline{\hat{f}(\lambda', t)} dt d^{n-1} \lambda' \right| \\
 &\leq 2 \left\{ \int_{\mathbb{R}^{n-1}} \int_0^\infty \left| \frac{\partial \hat{f}}{\partial x_n}(\lambda', t) \right|^2 dt d^{n-1} \lambda' \right\}^{1/2} \times \\
 &\quad \times \left\{ \int_{\mathbb{R}^{n-1}} \int_0^\infty (1 + |\lambda'|^2) |\hat{f}(\lambda', t)|^2 dt d^{n-1} \lambda' \right\}^{1/2}.
 \end{aligned} \tag{10.27}$$

Usando que $|2ab| \leq a^2 + b^2$ y la identidad de Plancherel obtenemos

$$\begin{aligned} \|\tau f\|_{H^{1/2}(\mathbb{R}^{n-1})}^2 &\leq \int_{\mathbb{R}_+^n} \left\{ |f(x)|^2 + \sum_{k=0}^{n-1} |\partial_k f|^2 + |\partial_n f|^2 \right\} dx \\ &= \|f\|_{H^1(\mathbb{R}_+^n)}. \end{aligned} \quad (10.28)$$

Para probar suryectividad debemos ver que dada $g \in H^{1/2}(\mathbb{R}^{n-1})$ existe al menos una (en realidad infinitas) $f \in H^1(\mathbb{R}_+^n)$ tal que $\tau f = g$. Nuevamente basta definir una anti-traza K en $C^\infty(\mathbb{R}^n)$ y probar que existe $C > 0$ tal que

$$\|Kg\|_{H^1(\mathbb{R}_+^n)} < C \|g\|_{H^{1/2}(\mathbb{R}^{n-1})} \quad \forall g \in C_0^\infty. \quad (10.29)$$

Sea $K(g) = \mathcal{F}^{-1}(e^{-(1+|\lambda'|^2)^{1/2}x^n} \mathcal{F}(g))$ y notando que el argumento de \mathcal{F}^{-1} está en $\mathcal{S}(\mathbb{R}^{n-1})$ dejamos la prueba de la desigualdad anterior como ejercicio.

En las aplicaciones tendremos que considerar funciones definidas solo en abiertos de \mathbb{R}^n , Ω y sus bordes $\partial\Omega$. Supondremos que Ω es tal que su borde, $\partial\Omega = \mathbb{C}l\Omega - \Omega$, es una variedad suave.

Sea $H^m(\Omega)$ el espacio de Sobolev obtenido tomando la integral en el producto escalar simplemente sobre Ω y completando el espacio $C_0^\infty(\Omega)$ con respecto a su norma y sea $H_0^m(\Omega)$ el obtenido con la misma norma pero completando el espacio $C^\infty(\Omega)$. Si $m = 0$ o si $\Omega = \mathbb{R}^n$ luego estos espacios coinciden. Pero si $m \geq 1$ y $\partial\Omega$ es no vacío entonces son distintos (obviamente $H_0^m \subset H^m$) ya que como se controlan derivadas en la norma las sucesiones de funciones de soporte compacto no pueden converger una función no nula en el borde. ¿Cómo extenderemos los resultados obtenidos en \mathbb{R}^n a Ω ? El lema clave para ello es el siguiente.

Lema 10.8 Sea γ la restricción de funciones en \mathbb{R}^n a funciones en Ω , luego

$$\gamma : H^m(\mathbb{R}^n) \rightarrow H^m(\Omega)$$

es continuo y suryectivo.

Prueba: La continuidad es clara, solo resta ver la suryectividad. Probaremos suryectividad para el caso $\Omega = \{x \in \mathbb{R}^n, x_n > 0\}$, $\partial\Omega = \{x \in \mathbb{R}^n / x_n = 0\}$. Sea $f \in H^m(\Omega)$ luego por el Lema anterior $\tau(\partial_n^k f) \in L^2$ para $k = 0, 1, \dots, m-1$. Continuamos f para x_n negativo como,

$$f(x) \equiv \left[\sum_{k=1}^{m-1} \frac{(-x_n)^k}{k!} \tau(\partial_n^k f) \right] \left(1 - e^{1/x_n^2} \right), \quad x_n < 0 \quad (10.30)$$

La sumatoria hace que las derivadas de f sean continuas en $x_n = 0$ y la exponencial hace que la norma H^m sea acotada. Pero esta norma solo dependerá de las normas de $\tau(\partial_n^k f)$ las cuales a su vez están acotadas por $\|f\|_{H^m(\Omega)}$ y por lo tanto existirá $c > 0$ tal que $\|f(x)\|_{H^m(\mathbb{R}^n)} < C \|f\|_{H^m(\Omega)} \forall f \in C_0^\infty$. El caso general es técnicamente más engorroso la idea básica es la de usar cartas tales que mapeen regiones conteniendo parte del borde de Ω en \mathbb{R}^{n+} mapeando borde con borde. En cada una de estas cartas sabemos cómo probar la desigualdad correspondiente. La desigualdad global se prueba suponiendo, por el absurdo, que ésta no vale ♠

Este lema nos permite inmediatamente generalizar los lemas anteriores.

Definición: Diremos que $f \in H^s(\Omega)$ si admite una extensión \tilde{f} a \mathbb{R}^n tal que $\tilde{f} \in H^s(\mathbb{R}^n)$ [Note que esto concuerda con lo probado para s entero positivo].

Es inmediato entonces que si $s' > s$ luego $H^{s'}(\Omega) \subset H^s(\Omega)$, que si $m < s - n/2$ luego $H^s(\Omega) \subset C^m(\Omega)$ y que si $m > 0$ entero $\tau : H^m(\Omega) \rightarrow H^{m-1/2}(\partial\Omega)$ es continuo y suryectivo. En la última afirmación usamos $H^{m-1/2}(\partial\Omega)$ donde en general $\partial\Omega \neq$ abierto en \mathbb{R}^{n-1} y por lo tanto no se encuadra en la definición anterior. En el caso en que $\partial\Omega$ sea compacto diremos que $f \in H^{m-1/2}(\partial\Omega)$ si, dado cualquier cubrimiento abierto $\{U_i\}$ de $\partial\Omega$ lo suficientemente pequeño como para que exista φ_i , tal que (U_i, φ_i) sea una carta de $\partial\Omega$ entonces $f \in H^{m-1/2}(U_i) \forall i$.

Notas bibliográficas: Recomiendo los libros: [15] y [9]. Los espacios de Sobolev permitieron comprender gran parte de la teoría de ecuaciones no lineales y previamente la teoría de ecuaciones lineales pero a coeficientes no suaves. No son ideas complejas, pero muy útiles e imprescindibles para la investigación en ecuaciones en derivadas parciales.

11.1. Introducción

Definición: Una **ecuación diferencial en derivadas parciales de orden m** en M es una ecuación de la forma

$$F \left(p, u, \nabla_a u, \nabla_a \nabla_b u, \dots, \overbrace{\nabla_a \cdots \nabla_c}^{m \text{ veces}} u \right) = 0, \quad (11.1)$$

donde ∇_a es alguna conexión en M . Más generalmente u puede ser una tupla de campos tensoriales y F tener rango en alguna otra tupla de campos tensoriales.

Ejemplos:

a) La ecuación de Laplace en \mathbb{R}^3 con respecto a una métrica g_{ab} ,

$$\Delta u \equiv g^{ab} \nabla_a \nabla_b u = 0 \quad (11.2)$$

Si g_{ab} es la métrica Euclídea, luego en coordenadas cartesianas

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}. \quad (11.3)$$

b) La ecuación de Poisson,

$$\Delta u - \rho = 0, \quad (11.4)$$

donde ρ es una función dada.

c) La ecuación de onda en \mathbb{R}^{n+1} . Esta tiene la forma de la ecuación de Laplace pero para una métrica de la forma $-(dx_0)^2 + \sum_{i=1}^n (dx^i)^2$. Por ejemplo en \mathbb{R}^2 , $g_{ab} = -(dt)_{ab}^2 + (dx)_{ab}^2$,

$$g^{ab} \nabla_a \nabla_b u = -\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2}. \quad (11.5)$$

d) Las ecuaciones de Maxwell,

$$\begin{aligned} \nabla_a F^{ab} &= j^b \\ \nabla_{[a} F_{bc]} &= 0 \end{aligned} \quad (11.6)$$

donde M es \mathbb{R}^4 , la métrica (usada tanto para subir los índices de F_{ab} como para definir ∇_c) es la de Minkowski, $g_{ab} = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2$, F_{ab} es un campo tensorial antisimétrico en M , y j^b es un campo vectorial (la cuadracorriente) en M .

e) Las ecuaciones de elasticidad en $\mathbb{R}^3(\text{Euclídeo}) \times \mathbb{R}$

$$\rho \frac{\partial^2 u^a}{\partial t^2} = \mu \Delta u^a + (\lambda + \mu) \nabla^a (\nabla_c u^c), \quad (11.7)$$

donde u^a es el vector desfaseamiento (en \mathbb{R}^3), ρ la densidad del medio elástico y λ y μ las constantes de Lamé del medio.

f) La ecuación de conducción del calor en $\mathbb{R}^3(\text{Euclídeo}) \times \mathbb{R}$

$$\frac{\partial u}{\partial t} = k \Delta u, \quad k > 0 \quad (11.8)$$

g) La ecuación de Schrödinger en $\mathbb{R}^3(\text{Euclídeo}) \times \mathbb{R}$,

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi + V \psi, \quad (11.9)$$

donde ψ es una función compleja y V un potencial.

h) La ecuación de Navier-Stokes para un fluido viscoso e incompresible (ejemplo agua) en $\mathbb{R}^3(\text{Euclídeo}) \times \mathbb{R}$

$$\frac{\partial u^a}{\partial t} + u^b \nabla_b u^a + \frac{1}{\rho} \nabla^a p - \gamma \Delta u^a = 0 \quad (11.10)$$

$$\nabla_a u^a = 0, \quad (11.11)$$

donde u^a es el vector velocidad del fluido, p su presión, ρ su densidad (constante) y γ la viscosidad cinemática.

De los ejemplos citados, de los cuales solo el último no es lineal, vemos la tremenda importancia física que tiene tener una teoría general de estas ecuaciones y es así que ésta es una de las ramas más activas de la matemática. Por su complejidad, en el gran número de casos distintos que presenta, la teoría general dista mucho de estar completa, sin embargo hay casos o clases de ecuaciones donde esto se ha logrado. Uno de éstos es el caso de una única ecuación de primer orden, donde como veremos el problema se reduce al de las ecuaciones diferenciales ordinarias.

Otro de estos casos es el de las ecuaciones lineales con coeficientes constantes (es decir para los cuales existe un sistema de coordenadas en el cual todos los coeficientes son constantes -ejemplo el Laplaciano para una métrica Euclídea). Esto se debe fundamentalmente al uso de transformadas como la de Fourier. Hago notar sin embargo que hay trabajos recientes mostrando nuevos resultados, aún en el caso del Laplaciano! Si permitimos a los coeficientes variar el problema se complica, sin embargo ciertas subclases de éstos han sido estudiadas completamente. Si a esto le agregamos no-linealidad en una forma no demasiado drástica -lo que se conoce como ecuaciones cuasi-lineales- el conocimiento se reduce bruscamente, aunque algunas subclases han sido doblegadas y algunas ecuaciones particulares completamente estudiadas. El caso donde la no-linealidad es drástica no ha sido todavía atacado con éxito -de ninguna clase-. Por fortuna los problemas físicos que hasta el momento hemos podido modelar o describir por ecuaciones en derivadas parciales tienen ecuaciones a lo sumo del tipo cuasi-lineal. En este curso veremos en detalle solo la teoría de algunas de las ecuaciones más simples [esencialmente las ecuaciones en los ejemplos a), b), c) y f)], siempre tratando de usar métodos que pueden ser aplicados a ecuaciones similares pero más complejas. Esto no solo se debe a la simplicidad de estas ecuaciones, lo que permite su conocimiento completo y detallado sino a que por un lado representan los “*ejemplos canónicos*” de distintas clases de ecuaciones. Las soluciones de las ecuaciones en cada una de estas clases se comportan en forma muy similar mientras que lo hacen en forma radicalmente diferente a las soluciones a ecuaciones en las otras clases. Por otro lado estas son las ecuaciones más usadas en física y aparecen en multitud de problemas diferentes, incluso como casos particulares de las ecuaciones en los ejemplos d), e), f) y g)!

11.2. La ecuación de primer orden

Esta es una ecuación de la forma

$$F(p, u, \nabla_a u) = 0. \quad (11.12)$$

donde u es una función en alguna variedad M . Esta ecuación puede ser atacada con mucho éxito y resulta en ecuaciones ordinarias, cuya teoría ya conocemos. Por simplicidad aquí consideraremos solo el caso cuasi-lineal y en \mathbb{R}^2 , es decir una ecuación de la forma,

$$a(x, y, u) u_x + b(x, y, u) u_y = c(x, y, u), \quad (11.13)$$

donde $u_x = \frac{\partial u}{\partial x}$ y $u_y = \frac{\partial u}{\partial y}$.

Por motivos geométricos es útil representar las soluciones de esta ecuación en \mathbb{R}^3 o más precisamente en una región Ω de \mathbb{R}^3 donde a , b y c estén definidos, eso es asociar una solución $u(x, y)$ de (11.13) con la hiper-superficie de \mathbb{R}^3 dada por $\tau = z - u(x, y) = cte$. Estas hiper-superficies se denominan superficies integrales de la ecuación (11.13). El gradiente de τ es en estas coordenadas es $(\nabla_a \tau) = (-u_x, -u_y, 1)$, por lo que vemos que la ecuación (11.13) es simplemente la condición que τ sea constante a lo largo del campo vectorial $(l^a) = (a(x, y, z), b(x, y, z), c(x, y, z))$, es decir $l^a \nabla_a \tau = 0$, lo que es equivalente a decir que l^a es **tangente** a las superficies integrales. Note que si $l^a \nabla_a \tau = 0$ luego $(f l^a) \nabla_a \tau = 0$ o sea que lo que determina la ecuación (11.13) no es l^a sino su dirección. Este campo de direcciones se llama de **direcciones características**, y sus curvas integrales **curvas características**.¹ La teoría de EDO nos dice entonces que por cada punto de Ω pasa una única curva característica. El conocimiento de estas curvas es fundamental ya que si formamos una superficie S tomando la unión de ciertas curvas características luego claramente S será una superficie integral de (11.13), pero por otro lado, dada una superficie integral S y cualquier $p \in S$ la curva característica que pasa por p será tangente a S en todo punto y por lo tanto será una subvariedad de S con lo que se ve que S estará formada por la unión de curvas características.

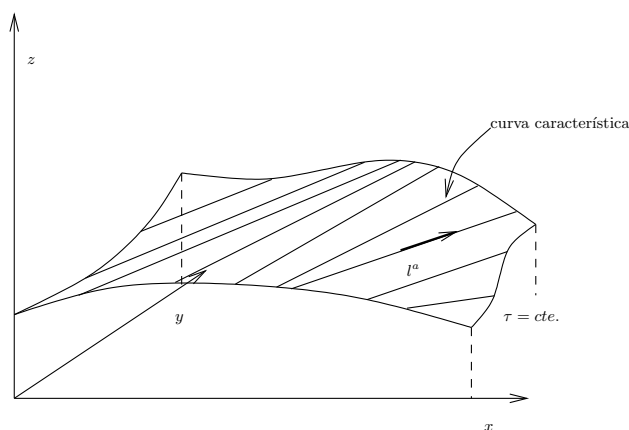


Figura 11.1: Curvas características.

En particular note que si dos superficies integrales S, S' , es decir dos soluciones de (11.13), u y u' , tienen un punto p en común luego que tener toda una curva característica en común, ya

¹Recuerde que una curva integral es la “imagen” de una curva $\gamma(t)$ (en este caso $= (x(t), y(t), z(t))$) solución de una EDO, en este caso

$$\frac{d\gamma(t)}{dt} = \frac{d}{dt} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a(x, y, z) \\ b(x, y, z) \\ c(x, y, z) \end{pmatrix}. \quad (11.14)$$

que por p solo pasa una de tales curvas y ésta no puede abandonar ninguna de las dos superficies. Por otro lado si S y S' se intersectan en una curva γ ésta debe ser integral. Para ver esto tomemos un punto p de γ y consideremos $T_p S$ y $T_p S'$; como las superficies se intersectan en una curva estos dos sub-espacios de $T_p \mathbb{R}^3$ se intersectan en una línea, como ambos deben contener la dirección dada por l^a ésta será la línea. Pero esto es cierto para todo punto de γ y por lo tanto el vector tangente a γ es proporcional a l^a y γ es característica.

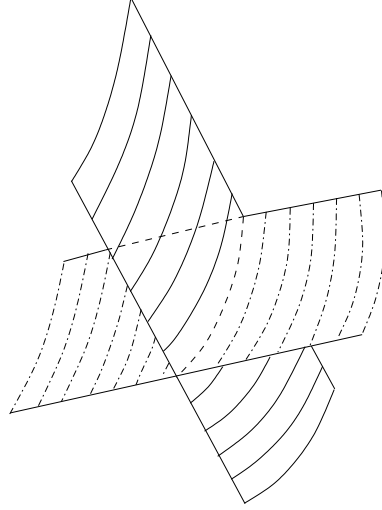


Figura 11.2: Intersección de soluciones.

11.2.1. El Problema de Cauchy

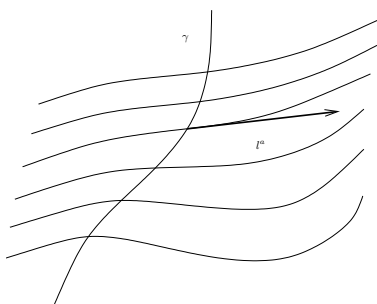
Se llama **problema de Cauchy** de una ecuación al problema de encontrar ciertos datos tales que dando éstos exista una única solución de esta ecuación. Es decir el de encontrar un cierto conjunto cuyos elementos son llamados datos y un mapa entre este espacio y el conjunto de soluciones de la ecuación. Por ejemplo en el caso de los EDO, el problema consiste en dada un campo vectorial suave v^a en M encontrar algún conjunto tal que a cada elemento de éste le corresponda una curva integral de v^a . Claramente podríamos tomar este conjunto a los puntos de M donde $v^a \neq 0$ ya que por cada uno de estos puntos pasa una única curva integral. Claramente también podríamos tomar como conjunto de datos –al menos localmente– una hiper-superficie s de M tal que en ninguno de sus puntos v^a sea tangente a ésta. En tal caso tenemos además la muy deseable propiedad de que cada punto de S determina una única solución, es decir *no contamos cada solución más que una vez*.

¿Cuáles serán estos datos en el caso de la ecuación (11.13)? Sea $\gamma(s) = (x_0(s), y_0(s), z_0(s))$, $s \in [0, 1]$, una curva en \mathbb{R}^3 . Buscaremos una solución tal que su superficie integral contenga a γ , es decir una $u(x, y)$ tal que se cumpla,²

$$z_0(s) = u(x_0(s), y_0(s)) \quad \forall s \in [0, 1]. \quad (11.15)$$

Consideraremos primero el caso en que $\gamma(s)$ no es una curva característica. Tomando cada punto $\gamma(s)$ como punto inicial para la ecuación diferencial ordinaria que determina l^a y resolviendo esta obtenemos para cada s la curva característica que pasa por $\gamma(s)$. (Ver figura.)

²Solo interesa la imagen de la curva y no su parametrización, por lo tanto tomaremos una en que el rango de S sea el intervalo $[0, 1]$.

Figura 11.3: Construyendo la solución a partir de la curva γ .

Obtenemos así un mapa $\gamma(s, t) : I_s \times I_t \rightarrow \mathbb{R}^3$ dado por,

$$\gamma(s, t) = (x(s, t), y(s, t), z(s, t)) \quad (11.16)$$

con $x(s, 0) = x_0(s)$, $y(s, 0) = y_0(s)$ y $z(s, 0) = z_0(s)$ y donde a s fijo se cumple,

$$\frac{d\gamma(s, t)}{dt} = l^a(\gamma(s, t)). \quad (11.17)$$

Si pudiésemos invertir las funciones $x(s, t)$ e $y(s, t)$ y así obtener s y t como funciones de x e y luego

$$u(x, y) \equiv z(s(x, y), t(x, y)) \quad (11.18)$$

sería la solución buscada, ya que tal u satisface por construcción (11.13) y (11.15).

No siempre es posible tal inversión. La razón es que en general habrá valores de s y t tales que el plano tangente a $\gamma(s, t)$ en ese punto contiene al eje z . Veamos si hay condiciones que nos aseguren que tal inversión es posible al menos localmente, es decir en un entorno de algún punto $(x(s_0, 0), y(s_0, 0))$ sobre $\gamma(s)$.

El teorema de la función implícita nos dice que esto será posible si el diferencial de la transformación en ese punto $(s_0, 0)$ es invertible, es decir si su determinante (el Jacobiano de la transformación) no se anula. En este punto tenemos,

$$J = \begin{vmatrix} \frac{\partial x}{\partial s} \big|_{(s_0, 0)} & \frac{\partial y}{\partial s} \big|_{(s_0, 0)} \\ \frac{\partial x}{\partial t} \big|_{(s_0, 0)} & \frac{\partial y}{\partial t} \big|_{(s_0, 0)} \end{vmatrix} = \frac{\partial x}{\partial s}(s_0) b_0 - \frac{\partial y}{\partial s}(s_0) a_0 \neq 0, \quad (11.19)$$

donde $a_0 = a(x(s_0), y(s_0), z(s_0))$ y $b_0 = b(x(s_0), y(s_0), z(s_0))$. Esta es entonces la condición para la existencia local de soluciones y nos dice que $\gamma(s)$ debe ser elegida de modo tal que su proyección en el plano (x, y) tenga vector tangente que no sea proporcional a la proyección en dicho plano del vector (a, b, c) .

Ejemplo: En algunas aplicaciones la coordenada y es el tiempo. En tal caso es natural especificar u en un instante de tiempo, digamos $y = 0$, es decir dar su **valor inicial**. Así es que el **problema de valores iniciales** consiste simplemente en elegir $\gamma(s) = (s, 0, h(s))$. La ecuación (11.15) resulta entonces $h(s) = u(s, 0)$ o $h(x) = u(x, 0)$, es decir $h(s)$ será el valor inicial que tendrá la solución a $y = 0$. en este caso habrá solución local siempre y cuando b , el coeficiente de $\frac{\partial u}{\partial y}$, no se anule en $(x, 0, h(x))$.

Si γ fuese una curva característica luego existirán infinitas soluciones (locales) ya que dado un punto $\gamma(s)$ de γ y una curva no característica $\gamma^*(r)$ pasando por este punto podremos construir,

usando el procedimiento anterior, pero ahora con $\gamma^*(r)$, una solución (una superficie) $\gamma^*(r, t)$ que necesariamente contendrá a γ .

Ejercicio: Resuelva por el método descripto las ecuaciones:

a)

$$\frac{\partial u}{\partial y} + c \frac{\partial u}{\partial x} = 0$$

$$u(x, 0) = h(x).$$

b)

$$\frac{\partial u}{\partial y} + u \frac{\partial u}{\partial x} = 0$$

$$u(x, 0) = -x.$$

c) ¿Por cuánto tiempo ($y = t$) se pueden extender las soluciones de b)?

11.3. Clasificación de ecuaciones en derivadas parciales

Para facilitar la clasificación procederemos en forma similar a como lo hicimos cuando estudiamos las ecuaciones en derivadas ordinarias, es decir reduciremos los sistemas de ecuaciones a sistemas de primer orden. Para ello tomaremos como variables independientes todas las derivadas de orden inferior al mayor orden en que aparece cada una de las variables.

Ejemplo: Sea $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfaciendo

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = \rho, \quad (11.20)$$

es decir el Laplaciano en una variedad dos dimensional plana. Definamos $u^1 := \phi$, $u^2 := \frac{\partial \phi}{\partial x}$ y $u^3 := \frac{\partial \phi}{\partial y}$, luego esta ecuación es equivalente al siguiente sistema

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \partial_x \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \partial_y \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix} = \begin{pmatrix} \rho \\ u^2 \\ u^3 \\ 0 \end{pmatrix} \quad (11.21)$$

El motivo por el cual agregamos la cuarta ecuación se verá más adelante, pero adelantamos que ella nos permite deducir, sin recurrir a las ecuaciones de las filas segunda y tercera, de que las componentes u^2 y u^3 son las componentes de la diferencial de alguna función. Si conocemos una solución de 11.21, (u^1, u^2, u^3) podemos probar que u^1 satisface también el Laplaciano. En efecto, tomando una derivada con respecto a x de la segunda fila y una derivada con respecto a y de la tercera, sumándolas y luego usando la primer fila obtenemos el resultado buscado. De todos los sistemas de primer orden (y básicamente también por la misma razón que dimos cuando consideramos sistemas de ecuaciones ordinarias) solo consideraremos aquellos de la forma,

$$M_{A'B}^a \nabla_a u^B = I_{A'}, \quad (11.22)$$

con $M_{A'B}^a$ y $I_{A'}$ funciones suaves del punto en la variedad (campos) y de u^A .

Los índices que estamos usando son índices abstractos y puede denotar no solo un conjunto de escalares, sino también un gran vector hecho de diversos campos vectoriales. Veremos esto en ejemplos. También se pueden tomar sistemas coordinados y bases y entonces plantear el problema en componentes, como hicimos en el ejemplo anterior, en cuyo caso entonces podemos pensar a u^A

como en gran arreglo vectorial de campos escalares. Hemos usado índices con y sin primar para dejar en claro que el espacio (co-vectorial) de los índices primados no tiene la misma dimensión que el espacio vectorial de los índices sin primar.

El tipo de sistema que acabamos de definir se llama cuasi-lineal, pues las derivadas aparecen en forma lineal. Esta no es una pérdida de generalidad desde el punto de vista de la física:

¡Todos los sistemas físicos conocidos son de esta forma!

Históricamente la clasificación que daremos a continuación nace del intento de encuadrar todas las ecuaciones en el Problema de Cauchy, es decir, tomar una hiper-superficie Σ de M , dar allí como dato u y su derivada en la dirección normal y tratar de construir soluciones en un entorno de Σ en M (soluciones locales). Este intento no fue en general exitoso, ya que en general esta no es la manera correcta de dar datos, pero sí lo fue la clasificación de estas ecuaciones así obtenida, ya que las propiedades de las soluciones a las ecuaciones en cada una de estas clases en las que las clasificaremos son muy similares y en cada una de estas clases los datos se prescriben también de manera similar.

Para fijar ideas sea M de dimensión n y sea $p \in M$. Queremos encontrar las soluciones en un entorno de p dando como dato u^A en alguna hiper-superficie Σ de M que contiene a p . Por simplicidad haremos las cuentas necesarias en un sistema de coordenadas adaptado al problema en el sentido que elegiremos a la coordenada x^n de tal forma que la subvariedad $x^n = 0$ sea la superficie Σ y p sea el origen de coordenadas. Los datos serán entonces $\Phi^A(x^1, \dots, x^{n-1})$, que luego corresponderá a la solución u^A restringida a Σ , es decir $\Phi^A = u^A|_{\Sigma}$.

Como conocemos lo que será u^A en Σ conocemos lo que serán todas sus derivadas (de cualquier orden) con respecto a las coordenadas x^i , $i = 1, \dots, n-1$. La idea es ahora usar la ecuación para encontrar $\partial_n u^A|_{\Sigma}$ y así sucesivamente encontrar todas las derivadas de u en p (o cualquier otro punto de Σ). Si lo lográsemos podríamos, al menos formalmente y en un entorno de p , construir u^A como su serie de Taylor alrededor de p . Si pudiésemos encontrar las derivadas normales a Σ , si Φ^A fuesen datos analíticos, y si los coeficientes de la ecuación fuesen también analíticos, luego se puede demostrar (Teorema de Cauchy-Kowalevski) que la solución u^A construida formalmente arriba en realidad existe y es analítica en un entorno de p en M .

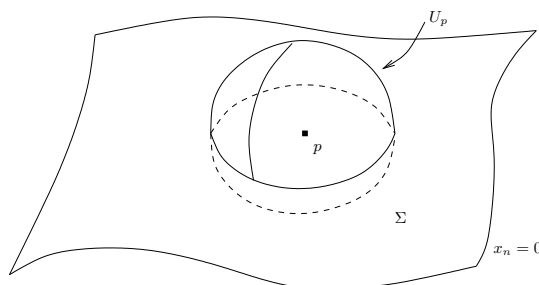


Figura 11.4: Construyendo la solución local.

¿Qué requisitos debe satisfacer la ecuación para que esto sea posible? Usando el sistema de coordenadas mencionado se puede ver que

$$M(\Phi^C, q)_{A'B}^n \partial_n u^B|_{\Sigma} = \text{Términos en } (\Phi^A, \partial_i \Phi^C, q)_{A'}, \quad (11.23)$$

Está claro entonces que podremos resolver para $\partial_n u^A$ sólo si $M(\Phi^C, q)_{A'B}^n$ es invertible. En general (como en el ejemplo que dimos) el espacio de llegada del mapa $M_{A'B}^n$ no tiene la misma dimensión que el espacio de partida y por lo tanto el mapa no es en general invertible, por lo tanto solo pediremos que **el rango de dicho mapa tenga la dimensión máxima** (es decir la dimensión

del espacio de partida –la de los vectores u^A –). En particular esto implica que estamos suponiendo que la dimensión del espacio de llegada es mayor o igual que la del espacio de partida. Estamos pidiendo la posibilidad mínima de tener soluciones.

Si esto no sucede entonces es claro que la ecuación será una relación entre términos determinados a partir de los datos dados y en general no se satisfará. En ese caso los datos a dar libremente serán un número menor al de la dimensión del espacio de partida, solo se podrán dar algunas componentes de u^A .

¿Qué es geoméricamente la condición de maximalidad del rango de $M_{A'B}^a$ en lo que hace a las coordenadas elegidas?

Si la superficie $x^n = 0$ es una superficie suave entonces podemos suponer que $\nabla_a x^n$ existe y es distinta de cero, en este caso la condición anterior es simplemente la condición que $M_{A'B}^a \nabla_a x^n$ tiene rango máximo, pero esta es independiente del sistema coordenado y solo depende de Σ , ya que aquí x^n es simplemente una función en M que es constante en Σ y cuyo gradiente no se anula. Las superficies en donde la condición anterior se viola en todos sus puntos se llaman **superficies características**. Clasificaremos a las ecuaciones de acuerdo al número de superficies características que se intersecten en un dado punto. Note que la clasificación solo depende del tensor $M_{A'B}^a$ y no de los términos de menor orden. $M_{A'B}^a \nabla_a u^B$ es llamada la **parte principal** de la ecuación. Note también que como $M_{A'B}^a$ es un campo tensorial no necesariamente constante, la condición definiendo las superficies características puede ser muy distinta de punto a punto, por lo tanto la clasificación que introduciremos vale en general solo para el punto en cuestión. Por fortuna en las aplicaciones muy pocas veces aparecen ecuaciones donde su tipo cambia de región en región. Note también que para sistemas no-lineales la condición depende también de la solución que uno esté considerando! Como nuestra clasificación solo se basa en la parte principal de la ecuación ésta debe tener toda la información sobre la ecuación, es por eso que en el ejemplo anterior agregamos la última fila en el sistema de ecuaciones, sin ella y considerando el sistema principal no podríamos saber que las dos últimas componentes de u^A debían ser las componentes de la diferencial de una función.

Diremos que una ecuación es **elíptica en** $p \in M$ si p no es intersectado por ninguna superficie característica. Es decir si $\text{Rango } M_{A'B}^a k_a$ no es maximal entonces $k_a = 0$ [como estamos en un punto la condición de que k_a sea un gradiente es vacía].

Ejercicio: El ejemplo canónico de ecuación elíptica es el Laplaciano en M , $\Delta u := g^{ab} \nabla_a \nabla_b \phi = \rho$ donde g^{ab} es una métrica Riemanniana. Muestre que el ejemplo dado anteriormente es un sistema elíptico.

Diremos que una ecuación es **parabólica en** $p \in m$ si p es intersectado por una única superficie característica. Es decir existe un único $k_a \neq 0$ –a menos de un factor multiplicativo– tal que $\text{Rango } M_{A'B}^a k_a$ no es maximal.

El ejemplo canónico de ecuación parabólica (o de difusión) es la ecuación del calor en $\mathbb{R} \times \mathbb{R}^{n-1}$,

$$\partial_t u = \Delta u. \quad (11.24)$$

Ejercicio: Considere la ecuación anterior en $\mathbb{R} \times \mathbb{R}$, $\partial_t u = \frac{\partial^2 u}{\partial x^2}$. Reduzca dicho sistema a primer orden y encuentre las superficies características de esta ecuación.

Diremos que una ecuación es **hiperbólica en** $p \in M$ si es intersectado por más de una superficie característica.

El ejemplo canónico en este caso es la ecuación de onda en M

$$\Delta u = g^{ab} \nabla_a \nabla_b u, \quad (11.25)$$

donde g^{ab} es una métrica tal que dado cualquier punto $p \in m$ existe un sistema coordenado en donde g_{ab} toma la forma,

$$g_{\mu\nu}|_p = \{-(dt)^2 + \sum_i (dx^i)^2\}|_p. \quad (11.26)$$

Ejercicio: Considere en \mathbb{R}^2 la métrica $ds^2 = -(dt)^2 + (dx)^2$ (en todo punto) y la ecuación $g^{ab}\nabla_a\nabla_b u = \rho$.

- Reduzca la ecuación a primer orden.
- Encuentre las características de la ecuación.
- Haga lo mismo en $\mathbb{R} \times S^1$ (un cilindro espacio temporal) con $ds^2 = -(dt)^2 + (d\theta)^2$ y con $ds^2 = -(dt)^2 + t^2(d\theta)^2$.

Notas bibliográficas: Lectura recomendada para éste y los capítulos que siguen: [15], [16], [17] y [19]. Lo expuesto en estos capítulos es lo mínimo e imprescindible para tener una noción de esta área. Se conoce muchísimo más y a la vez, como siempre, hay muchísimo más que desconocemos, soluciones débiles, existencia global, ondas de choque, condiciones de contorno, estabilidad, etc, etc. Esta es probablemente el área más activa, con más gente trabajando y con el mayor número de aplicaciones de toda la matemática. La mayoría de estas aplicaciones han sido tradicionalmente en el área de la ingeniería y en particular en la de fluidos, lo que ha hecho de que solo se trataran cierto tipo específico de ecuaciones, bastante difíciles por cierto, y no las más usadas en otras áreas de la física, esto ha evolucionado en los últimos años y ahora hay un cambio considerable de atención hacia muchos de los problemas de la física moderna.

12.1. La Ecuación de Laplace

Tomaremos como modelo de ecuación elíptica a la ecuación de Laplace en \mathbb{R}^n con métrica Euclídea $ds^2 = \sum_{i=1}^n (dx^i)^2$, o una región abierta Ω de \mathbb{R}^n con la misma métrica. Los resultados que obtendremos que no dependan de funciones especiales se pueden generalizar para ecuaciones elípticas de la forma (12.51) si se asume que $c \leq 0$.

Como ya mencionamos, el programa de Cauchy, es decir formular el problema de obtener soluciones dando como dato u^A en una hipersuperficie Σ no funciona en general, solo lo hace para ecuaciones hiperbólicas. Si consideramos datos no-analíticos, y los datos físicos son en general no-analíticos, no hay en general ni siquiera soluciones locales. ¿Cuál es la manera apropiada entonces de dar datos para la ecuación de Laplace? Obtendremos la respuesta a esto considerando un fenómeno físico descrito por esta ecuación.

Consideremos un tambor e imprimamos una fuerza (pequeña) sobre su membrana (o parche) en forma perpendicular a ésta.

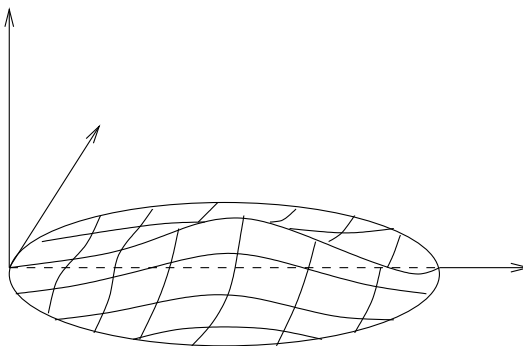


Figura 12.1: La membrana de un tambor.

La membrana se moverá de su posición plana (de reposo) y adquirirá una nueva posición de equilibrio generando una fuerza elástica que cancele exactamente a la aplicada. ¿Cuál será esta nueva forma de la membrana? Si denotamos por $u(x, y)$ al desplazamiento de ésta en la dirección del eje vertical (z) y f a la fuerza por unidad de área aplicada (en la dirección vertical) se puede ver que u debe satisfacer la ecuación,

$$\Delta u = f, \text{ en } \Omega = \{(x, y) | \sqrt{x^2 + y^2} < \text{radio del tambor}\}. \quad (12.1)$$

Ejercicio: Convénzase de que (12.1) es la ecuación de equilibrio. Ayuda: Haga primero el caso unidimensional.

Como el borde del tambor sujeta la membrana allí tendremos la siguiente **condición de contorno**,

$$u|_{\partial\Omega} = 0, \quad \partial\Omega = \{(x, y) | \sqrt{x^2 + y^2} = \text{radio del tambor}\}. \quad (12.2)$$

Tenemos así el siguiente problema: Dado f en Ω encuentre u satisfaciendo (12.1) y (12.2). Por nuestra experiencia física sabemos que este problema se debe poder resolver! Más aún, se debe poder resolver el problema donde el tambor no es circular pero tiene forma arbitraria, permitiendo que $\partial\Omega$ sea un borde arbitrario pero suave y también que $\partial\Omega$ no esté en el plano $x = 0$ o equivalentemente que esté en tal plano pero que u pueda tomar un valor cualquiera –pero suave– ϕ en $\partial\Omega$.

Llegamos así al siguiente **Problema de Dirichlet**: Dada Ω con $\partial\Omega$ suave, $f : \Omega \rightarrow \mathbb{R}$ suave y $\phi_0 : \partial\Omega \rightarrow \mathbb{R}$, también suave, encuentre $u : \Omega \rightarrow \mathbb{R}$ satisfaciendo,

$$1. \Delta u = f \quad \text{en } \Omega,$$

$$2. u|_{\partial\Omega} = \phi_0.$$

Más adelante reafirmaremos nuestra intuición viendo que este problema siempre se puede resolver.

Supongamos ahora que permitimos que los bordes de la membrana se pueda deslizar verticalmente pero en el borde ponemos resortes verticales con una constante de Hooke que depende de la posición, $k : \partial\Omega \rightarrow \mathbb{R}$ y dispuestos de tal forma que la posición de equilibrio antes de aplicar f sea $u = 0$. Cuando apliquemos f la membrana se moverá hasta que nuevamente en el interior tengamos,

$$\Delta u = f \quad \text{en } \Omega \quad (12.3)$$

y en el borde

$$(ku + n^a \nabla_a u)|_{\partial\Omega} = 0, \quad (12.4)$$

donde n^a es la normal a $\partial\Omega$. Este **Problema Mixto** también se puede resolver.

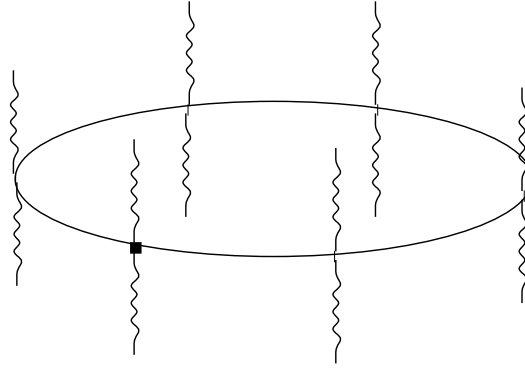


Figura 12.2: Problema mixto.

Un caso particular de éste es cuando para ejercer la fuerza de borde no usamos resortes sino simplemente una fuerza dada $\phi_1 : \partial\Omega \rightarrow \mathbb{R}$, pero teniendo la precaución de que su contribución total, $\int_{\partial\Omega} \phi_1 dS$, cancele exactamente la contribución total de f –de no ser así tendríamos una membrana acelerándose–. En tal caso tenemos el **Problema de Neumann**:

$$\Delta u = f \quad \text{en } \Omega \quad (12.5)$$

$$n^a \nabla_a u|_{\partial\Omega} = \phi_1, \quad \text{en el borde}, \quad (12.6)$$

$$\text{con } \int_{\partial\Omega} \phi_1 dS = \int_{\partial\Omega} n^a \nabla_a u dS = \int_{\Omega} f dV = \int_{\Omega} \Delta u dV. \quad (12.7)$$

12.1.1. Existencia

A continuación veremos que el problema de Dirichlet siempre tiene una única solución (suponiendo que f , ϕ_0 y $\partial\Omega$ son lo suficientemente suaves). Los otros problemas se resuelven en forma similar.

Supongamos que existe $u \in H^2(\Omega)$ satisfaciendo, $\Delta u = f$ en Ω y $u|_{\partial\Omega} = 0$ –lo que implica $u \in H_0^1(\Omega)$ –. Luego usando el teorema de la divergencia obtenemos la **primera identidad de Green**, $\forall v \in C^\infty(\Omega)$.

$$\int_{\Omega} v \bar{f} d^n x = \int_{\Omega} v \Delta \bar{u} d^n x = - \int_{\Omega} e^{ab} \nabla_a v \nabla_b \bar{u} d^n x + \int_{\partial\Omega} v (n^a \nabla_a \bar{u}) d^{n-1} x, \quad (12.8)$$

Si suponemos que $v \in H_0^1(\Omega)$ la identidad todavía es válida y en este caso se reduce a,

$$\int_{\Omega} e^{ab} \nabla_a v \nabla_b \bar{u} d^n x = - \int_{\Omega} v \bar{f} d^n x, \quad \forall v \in H_0^1(\Omega). \quad (12.9)$$

Pero note que esta identidad todavía es válida si suponemos simplemente que $u \in H_0^1(\Omega)$ –no necesariamente en $H^2(\Omega)$ – y que $f \in H^{-1}(\Omega)$.

Tenemos así el **Problema Débil de Dirichlet** (con $\phi_0 = 0$): Encuentre $u \in H_0^1(\Omega)$ tal que dada $f \in H^{-1}(\Omega)$ se satisfaga (12.9).

Si el lado derecho (12.9) fuese un producto escalar este daría origen (por completamiento) a un espacio de Hilbert H y entonces (12.9) tendría la forma

$$(u, v)_H = \Phi_f(v) \equiv - \int_{\Omega} \bar{f} v d^n x, \quad \forall v \in H_0^1(\Omega). \quad (12.10)$$

Si $H = H_0^1(\Omega)$ y si $\Phi_f : H \rightarrow \mathbb{C}$ fuese continua luego el teorema de representación de Riez nos diría que existe un único u en H satisfaciendo (12.9) y por lo tanto (12.10). Como veremos a continuación (lema de Poincaré-Hardy) en este caso la parte derecha de (12.9) es un producto escalar equivalente al de $H_0^1(\Omega)$ y por lo tanto $H = H_0^1(\Omega)$. Pero entonces Φ_f es claramente continua ya que $H^{-1}(\Omega)$ es el dual de $H^1(\Omega) \supset H_0^1(\Omega)$ y $u \in H_0^1(\Omega)$. Hemos probado entonces:

Teorema 12.1 (de existencia y unicidad) Dada $f \in H^{-1}(\Omega)$ existe una única $u \in H_0^1(\Omega)$ satisfaciendo el problema débil de Dirichlet.

Corolario 12.1 El mapa $(\Delta, \tau) : H^1(\Omega) \rightarrow H^{-1}(\Omega) \times H^{1/2}(\Omega)$ dado por,

$$\begin{aligned} \Delta u &= f \in H^{-1}(\Omega), \\ \tau u &= \phi_0 \in H^{1/2}(\partial\Omega), \end{aligned} \quad (12.11)$$

es un isomorfismo¹.

Prueba: Es claro que el mapa es continuo –ya que es lineal y acotado–. Es también inyectivo ya que si $\phi_0 = 0$ luego $u \in H_0^1(\Omega)$ y si $f = 0$ el teorema anterior (unicidad) nos dice entonces que $u = 0$. Veamos que es suryectivo. Sea $\phi_0 \in H^{1/2}(\Omega)$, luego existe $w \in H^1(\Omega)$ tal que su restricción a $\partial\Omega$, $\tau w = \phi_0$ y sea $f \in H^{-1}(\Omega)$. Como también $\Delta w \in H^{-1}(\Omega)$ el teorema anterior nos garantiza que existe $\bar{u} \in H_0^1(\Omega)$ tal que

$$\Delta \bar{u} = f - \Delta w \text{ en } \Omega. \quad (12.12)$$

Pero entonces $u = \bar{u} + w$ satisface $\Delta u = f$ en Ω y $\tau u = \phi_0$ en $\partial\Omega$.

Solo resta probar entonces,

¹Siempre entendiendo estas ecuaciones en su forma débil o distribucional

Lema 12.1 (de Poincaré-Hardy) Existe $C > 0$ tal que para todo $u \in H_0^1(\Omega)$, se cumple,

$$\int_{\Omega} |u|^2 dx \leq C \int_{\Omega} e^{ab} \nabla_a \bar{u} \nabla_b u dx. \quad (12.13)$$

Esto nos dice que el producto escalar en $H_0^1(\Omega)$ es equivalente al producto escalar usado anteriormente.

Prueba: Como $C_0^\infty(\Omega)$ es denso en $H_0^1(\Omega)$ es suficiente probar la desigualdad para estas funciones. Sea Γ_d un n -cubo de lado d conteniendo a Ω . Extendiendo las funciones en $C_0^\infty(\Omega)$ como cero en $\Gamma_d - \Omega$ obtenemos,

$$\begin{aligned} |u|^2(x) &= \left| \int_{-d/2}^{x^1} \partial_1 u(\xi^1, x^2, \dots, x^n) d\xi^1 \right|^2 \\ &\leq (x^1 + \frac{d}{2}) \int_{-d/2}^{x^1} |\partial_1 u|^2 d\xi^1 \\ &\leq d \int_{-d/2}^{d/2} |\partial_1 u|^2 d\xi^1. \end{aligned} \quad (12.14)$$

Por lo tanto,

$$\int_{-d/2}^{d/2} |u(x)|^2 dx^1 \leq d^2 \int_{-d/2}^{d/2} |\partial_1 u|^2 d\xi^1, \quad (12.15)$$

y

$$\int_{\Gamma_d} |u|^2(x) dx \leq d^2 \int_{\Gamma_d} |\partial_1 u|^2 dx \leq C \int_{\Gamma_d} \nabla_a \bar{u} \nabla^a u dx, \quad (12.16)$$

con $C = d^2$. Esto prueba el lema y concluye la prueba de existencia y unicidad ♠♠

Tanto el teorema de existencia y unicidad como el de regularidad (que damos a continuación) pueden ser generalizados a ecuaciones elípticas con coeficientes no constantes mientras éstos sean suaves, se cumpla que $c \leq 0$ y que $a^{ab} l_a l_b > k g^{ab} l_a l_b$, donde k es una constante y g^{ab} es una métrica positiva definida tal que el volumen de Ω con respecto a esta métrica es finito. La prueba que dimos es válida solo si $Vol_g(\Omega) < \infty$, ya que en esta usamos el lema de Poincaré-Hardy. Si $\Omega = \mathbb{R}^n$ entonces la prueba es todavía válida si sustituimos al espacio H_0^1 por el espacio H_1^1 , es decir el espacio de las funciones que decaen al infinito, con norma

$$\|f\|_{H_1^1(\mathbb{R}^n)}^2 = \int_{\mathbb{R}^n} \frac{|f|^2}{r^2} + e^{ab} \nabla_a \bar{f} \nabla_b f.$$

En este caso se puede probar que la solución del problema no solo será suave sino también que decaerá asintóticamente como $1/r$.

12.1.2. *Regularidad de las Soluciones

La solución que hemos encontrado lo es solo en el sentido débil. Si suponemos ahora que f y ϕ_0 tienen cierta regularidad podemos concluir además que,

Teorema 12.2 (de Regularidad) Sea $u \in H^1(\Omega)$ una solución débil de la ecuación $\Delta u = f$ en Ω con condición de contorno $u|_{\partial\Omega} = \phi_0$ y sea $f \in H^k(\Omega)$, $\phi_0 \in H^{k+\frac{3}{2}}(\partial\Omega)$, $k \geq -1$ luego $u \in H^{k+2}(\Omega)$. En particular si $f \in C^\infty(\Omega)$ y $\phi_0 \in C^\infty(\partial\Omega)$, luego $u \in C^\infty(\Omega)$.

Antes de proceder con la prueba definiremos el operador diferencia finita y veremos sus propiedades. Sea $\{x^i\}$ un sistema coordenado en un abierto de Ω .

Definición: $\Delta_i^h u(x^1, \dots, x^i, \dots, x^n) \equiv \frac{u(x^1, \dots, x^i + h, \dots, x^n) - u(x^1, \dots, x^i, \dots, x^n)}{h}$.

Lema 12.2 Si $u \in H^1(\Omega)$ y $\Omega' \subset\subset \Omega$ (contenido estrictamente, $\partial\Omega \cap \Omega' = \emptyset$) luego,

$$\|\Delta_i^h u\|_{H^0(\partial\Omega)} \leq \|\partial_i u\|_{H^0(\Omega)}. \quad (12.17)$$

Prueba: Es suficiente considerar el caso $u \in C^1(\Omega)$. Si $h < \text{dist}(\partial\Omega, \Omega')$, luego

$$\Delta_i^h u(x) = \frac{1}{h} \int_0^h \partial_i u(x^1, \dots, x^i + \xi, \dots, x^n) d\xi \quad \forall x \in \Omega', \quad (12.18)$$

por lo tanto, usando la desigualdad de Schwarz vemos que,

$$|\Delta_i^h u(x)|^2 \leq \frac{1}{h} \int_0^h |\partial_i u(x^1, \dots, x^i + \xi, \dots, x^n)|^2 d\xi. \quad (12.19)$$

Integrando sobre Ω' obtenemos,

$$\int_{\Omega'} |\Delta_i^h u(x)|^2 d^n x \leq \frac{1}{h} \int_0^h \int_{\Omega'} |\partial_i u|^2 d^n x d\xi \leq \int_{\Omega} |\partial_i u|^2 d^n x. \quad (12.20)$$

♠

Lema 12.3 Sea $u \in H_0(\Omega)$ (y por lo tanto $\Delta_i^h u \in H_0(\Omega')$) y supongamos existe $k < \infty$ tal que $\|\Delta_i^h u\|_{H_0(\Omega')} \leq k \quad \forall h > 0$ y $\Omega' \subset\subset \Omega$, con $h < \text{dist}(\partial\Omega, \Omega')$. Luego la derivada débil $\partial_i u$ existe y satisface, $\|\partial_i u\|_{H_0(\Omega)} \leq k$.

Prueba: Por la compacidad débil de los conjuntos cerrados y acotados en $H_0(\Omega')$ existirá una sucesión $\{h_m\} \rightarrow 0$ y $v_i \in H_0(\Omega)$ con $\|v_i\|_{H_0(\Omega')} \leq k$ tal que $\Delta_i^{h_m} u \xrightarrow{d} v_i$, es decir

$$\int_{\Omega} \phi \Delta_i^{h_m} u d^n x \rightarrow \int_{\Omega} \phi v_i d^n x, \quad \forall \phi \in C_0^1(\Omega). \quad (12.21)$$

Por otro lado, si $h_m < \text{dist}(\text{sop.}\phi, \partial\Omega)$, luego

$$\int_{\Omega} \phi \Delta_i^{h_m} u d^n x = - \int_{\Omega} u (\Delta_i^{-h_m} \phi) d^n x \rightarrow - \int_{\Omega} u \partial_i \phi d^n x. \quad (12.22)$$

Concluimos así que,

$$\int_{\Omega} \phi v_i d^n x = - \int_{\Omega} u \partial_i \phi d^n x, \quad (12.23)$$

o sea que v_i es la derivada débil (distribucional) de u en la dirección x^i ♠

Prueba: (Teorema de Regularidad): Nos contentaremos con probar que u es regular en cualquier Ω' contenido estrictamente en Ω , la extensión de la prueba a todo Ω no agrega ningún concepto nuevo. Probaremos el enunciado para $k = 0$, el resto sigue por inducción en k . Como u es una solución débil tenemos que

$$\int_{\Omega} \nabla^a \bar{u} \nabla_a v d^n x = \int_{\Omega} \bar{f} v d^n x, \quad \forall v \in C_0^1(\Omega). \quad (12.24)$$

Reemplazando v por $-\Delta_i^{-h} v$ con $|2h| < \text{dist}(\text{sop.}v, \partial\Omega)$, tenemos,

$$\int_{\Omega} (\Delta_i^h \nabla^a \bar{u}) \nabla_a v d^n x = - \int_{\Omega} \bar{f} \Delta_i^h v d^n x, \quad \forall v \in C_0^1(\Omega) \quad (12.25)$$

y por lo tanto (Usando el lema (12.2)),

$$\left\| \int_{\Omega} (\Delta_i^h \nabla^a \bar{u}) \nabla_a v d^n x \right\| \leq \|f\|_{H_0} \|\partial_i v\|_{H_0}. \quad (12.26)$$

Tomando ahora $v = \Delta_i^h u$ obtenemos,

$$\left\| \int_{\Omega} (\Delta_i^h \nabla^a \bar{u}) \Delta_i^h \nabla_a u d^n x \right\| \leq \|f\|_{H_0} \|\partial_i \Delta_i^h u\|_{H_0}, \quad (12.27)$$

o sea,

$$\|\Delta_i^h \nabla^a u\|_{H_0}^2 \leq \|f\|_{H_0} \|\Delta_i^h \nabla_a u\|_{H_0}, \quad (12.28)$$

lo que finalmente implica,

$$\|\Delta_i^h \nabla^a u\|_{H_0} \leq \|f\|_{H_0} < k. \quad (12.29)$$

El lema (12.3) nos dice entonces que $\partial_i u \in H^1(\partial\Omega)$ y completa la primera parte de la prueba.

Veamos ahora, para completar la prueba, que si $f \in C^\infty(\Omega)$ luego $u \in C^\infty(\Omega)$. Pero esto es obvio ya que si $u \in H^p(\Omega)$, $\Omega \subset \mathbb{R}^n$, luego por el Lema de Sobolev $u \in C^{p-\frac{n}{2}-\varepsilon}(\Omega) \quad \forall \varepsilon > 0$ y en particular si $u \in H^p(\Omega) \quad \forall p$ luego $u \in C^\infty(\Omega)$ ♠

12.2. Teorema Espectral

Sea una barra de un material de conductividad térmica $q = 1$ y longitud $L = 2\pi$ y supongamos que queremos describir la evolución de su distribución de temperatura $T(t, x)$. Para ello supondremos una distribución inicial $T_0(x)$ y que los extremos de la barra están conectados con un reservorio infinito de calor a cero grado. Debemos resolver entonces el problema matemático,

$$\begin{aligned} \frac{d}{dt} T &= -\frac{d^2}{dx^2} T, \quad t \geq 0 \\ T(t_0, x) &= T_0(x), \\ T(t, 0) &= T(t, 2\pi) = 0 \quad t \geq 0. \end{aligned} \quad (12.30)$$

Para resolverlo usaremos un desarrollo en serie de Fourier, es decir plantearemos una solución de la forma,

$$T(t, x) = \sum_{n=1}^{\infty} C_n(t) \sin\left(\frac{nx}{2}\right), \quad (12.31)$$

la cual satisface claramente las condiciones de contorno. Aplicándole la ecuación y usando la ortogonalidad de las funciones obtenemos,

$$\frac{d}{dt} C_n(t) = -\frac{n^2}{4} C_n(t), \quad (12.32)$$

la cual tiene como solución,

$$C_n(t) = C_n(0) e^{-\frac{n^2}{4}t}. \quad (12.33)$$

Por lo tanto si damos como condición inicial $T_0 \in L^2(0, 2\pi)$, ésta determinará una sucesión $\{C_n^0 = (T_0, \sin(\frac{nx}{2}))_{L^2}\}$ en l^2 que usaremos como condición inicial y así obtendremos $T(t, x)$. Como

$$\sum_{n=1}^{\infty} |C_n(t)|^2 = \sum_{n=1}^{\infty} |C_n^0|^2 e^{-\frac{n^2}{2}t} \leq \sum_{n=1}^{\infty} |C_n^0|^2 < \infty \quad (12.34)$$

la solución formal –ya que no sabemos si es diferenciable– estará en $L^2(0, 2\pi)$ para todo $t \geq 0$. Para cualquier $t > 0$ y $q \in \mathbb{N}$ tenemos que $n^{2q} e^{-\frac{n^2}{2}t}$ tiende en forma exponencial a cero cuando n tiende a infinito y por lo tanto $T(t, x) \in H^q(0, 2\pi)$ lo que implica $T(t, x) \in C^\infty((0, +\infty) \times [0, 2\pi])$ ².

²Se puede demostrar además que $T(t, x)$ es analítica en ambas variables en $(0, +\infty) \times [0, 2\pi]$

Hemos resuelto así completamente este problema.

¿Qué hemos usado para construir estas soluciones? Hemos usado que cualquier función que se anula en los extremos puede ser expandida en término de su serie de Fourier, es decir de las funciones $\text{sen}(\frac{nx}{2})$ y además que ,

$$\frac{d^2}{dx^2} \text{sen}(\frac{nx}{2}) = -\frac{n^2}{4} \text{sen}(\frac{nx}{2}) \quad (12.35)$$

En analogía con la teoría de operadores entre espacios vectoriales de dimensión finita llamaremos a la ecuación de arriba la ecuación de autovectores-autovalores del operador lineal $\frac{d^2}{dx^2}$. Dado un operador lineal $L : D \subset L^2 \rightarrow L^2$ el problema de encontrar sus autovalores y autovectores con condiciones de contorno dadas se llama problema de Sturm-Liouville. Como veremos a continuación existe una gran cantidad de operadores para los cuales este problema tiene solución. ¿Es una casualidad que el conjunto de autovalores del operador en el ejemplo sea una base de L^2 ?³ El siguiente teorema y su corolario nos dicen que no lo es, si el operador en cuestión satisface ciertas condiciones.

Teorema 12.3 (Espectral para el Laplaciano) *Sea $\Omega \subset \mathbb{R}^n$ acotado. Luego el Laplaciano tiene un conjunto numerable y discreto de autovalores, $\Sigma = \{\lambda_i\}$, cuyas autofunciones (autovectores) expanden $H_0^1(\Omega)$.*

Para probar este teorema necesitamos de los siguientes lemas:

Lema 12.4 *Sea Γ_d un cubo en \mathbb{R}^n con lados de longitud $d > 0$. Si $u \in H^1(\Gamma_d)$ luego,*

$$\|u\|_{H^0(\Gamma_d)}^2 \leq d^{-n} \left| \int_{\Gamma_d} u d^n x \right|^2 + \frac{nd^2}{2} \sum_{j=1}^n \|\partial_j u\|_{H^0(\Gamma_d)}^2. \quad (12.36)$$

Prueba: Es suficiente considerar $u \in C^1(\Gamma_d)$. Para cualquier x e y en Γ_d tenemos,

$$u(y) - u(x) = \sum_{j=1}^n \int_{x^j}^{y^j} \partial_j u(y^1, \dots, y^{j-1}, s, x^{j+1}, \dots, x^n) ds. \quad (12.37)$$

Tomando su cuadrado y usando la desigualdad de Schwarz obtenemos,

$$\begin{aligned} & |u|^2(x) + |u|^2(y) - 2\Re(u(x)u(y)) \\ & \leq nd \sum_{j=1}^n \int_{-d/2}^{d/2} |\partial_j u(y^1, \dots, y^{j-1}, s, x^{j+1}, \dots, x^n)|^2 ds. \end{aligned} \quad (12.38)$$

Integrando con respecto a todos los x^j e y^j obtenemos,

$$2d^n \|u\|_{H^0(\Gamma_d)}^2 \leq 2 \left| \int_{\Gamma_d} u d^n x \right|^2 + nd^{n+2} \sum_{j=1}^n \|\partial_j u\|_{H^0(\Gamma_d)}^2, \quad (12.39)$$

de la cual sigue trivialmente la desigualdad buscada ♠

Lema 12.5 *Sea $\Omega \subset \mathbb{R}^n$ acotado. Si una sucesión $\{u_p\}$ en $H_0^1(\Omega)$ es acotada, luego existe una subsucesión que converge (fuertemente) en $H^0(\Omega)$. Es decir el mapa natural $H_0^1(\Omega) \rightarrow H^0(\Omega)$ es compacto.*

³Para obtener convergencia puntual (y no simplemente en L^2) en el caso de condiciones de contorno no nula hay que agregar los autovectores (con cero autovalor) $f_0(x) = 1$ y $f_1(x) = x$.

Prueba: Sea $k = \sup\{\|u_p\|_{H_0^1(\Omega)}\}$. Como Ω es acotado lo podemos encerrar en un cubo Γ_D y extender cada u_n como cero en $\Gamma_D - \Omega$. Sea $\varepsilon > 0$ y M lo suficientemente grande como para que $\frac{2nk^2D^2}{M^2} < \varepsilon$. Descomponiendo Γ_D en M^n cubos Γ_d^j con $d = D/M$ y usando que como $\{u_n\}$ es también acotada en $H^0(\Omega)$ existe una subsucesión $\{\tilde{u}_p\}$ y $u \in H^0(\Omega)$ tal que esta converge débilmente a u . Vemos entonces que existe un entero N tal que si p y $q > N$,

$$\left| \int_{\Gamma_d^j} (\tilde{u}_p - \tilde{u}_q) d^n x \right|^2 < \frac{\varepsilon}{2} \left(\frac{D}{M} \right)^{2n} \frac{1}{D^n}. \quad (12.40)$$

Si aplicamos a cada Γ_d^j el lema (12.4) y sumamos sobre j obtenemos que $\forall p$ y $q > N$,

$$\begin{aligned} \|\tilde{u}_p - \tilde{u}_q\|_{H^0(\Omega)}^2 &\leq \left(\frac{D}{M} \right)^n \left(\sum_{j=1}^{M^n} \frac{\varepsilon}{2} \left(\frac{D}{M} \right)^{2n} \right) + \frac{n}{2} \left(\frac{D}{M} \right)^2 (2k^2) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &\leq \varepsilon. \end{aligned} \quad (12.41)$$

Pero entonces $\{\tilde{u}_p\}$ es una sucesión de Cauchy en $H^0(\Omega)$ y por lo tanto converge fuertemente a u en $H^0(\Omega)$ ♠

Prueba: [Teorema Espectral] Considere la funcional $\mathcal{J} : H_0^1(\Omega) \rightarrow \mathbb{R}^+$ dada por,

$$\mathcal{J}(u) = \frac{\int_{\Omega} \nabla_a \bar{u} \nabla^a u d^n x}{\int_{\Omega} \bar{u} u d^n x}. \quad (12.42)$$

Como $\mathcal{J}(u)$ es no negativa existirá $\lambda_0(\Omega) \in \mathbb{R}^+$ tal que

$$\lambda_0 = \inf_{u \in H_0^1(\Omega)} \mathcal{J}(u). \quad (12.43)$$

Ejercicio: Relacione este λ_0 con la constante en el Lema de Poincaré-Hardy.

Como veremos a continuación este λ_0 es el menor de los autovalores del Laplaciano y su autofunción u_0 es la que minimiza a \mathcal{J} . En efecto supongamos que existe u_0 en $H_0^1(\Omega)$ que minimiza a \mathcal{J} , como \mathcal{J} es una funcional diferenciable debemos tener que

$$\frac{d}{ds} \mathcal{J}(u_0 + sv)|_{s=0} = 0 \quad \forall v \in H_0^1(\Omega). \quad (12.44)$$

Pero

$$\frac{d}{ds} \mathcal{J}(u_0 + sv)|_{s=0} = \frac{1}{\int_{\Omega} \bar{u} u d^n x} \left[\int_{\Omega} (\nabla^a \bar{v} \nabla_a u_0 + \nabla^a \bar{u}_0 \nabla_a v) d^n x - \lambda_0 \left[\int_{\Omega} (\bar{v} u_0 + \bar{u}_0 v) d^n x \right] \right] \quad (12.45)$$

y por lo tanto (12.44) es equivalente a que u_0 satisfaga la versión débil de la ecuación de autovalores-autovectores. Veamos ahora que u_0 existe. Como λ_0 es el ínfimo de \mathcal{J} existe una sucesión $\{u_p\} \in H_0^1(\Omega)$ tal que $\|u_p\|_{H^0(\Omega)} = 1$ y $\mathcal{J}(u_p) \rightarrow \lambda_0$. Pero tal sucesión es acotada en $H_0^1(\Omega)$ y por lo tanto por el lema (12.5) existe una subsucesión $\{\tilde{u}_p\}$ que converge fuertemente en $H^0(\Omega)$ a una función u_0 con $\|u_0\|_{H^0(\Omega)} = 1$. Como $Q(u) \equiv \int_{\Omega} \nabla^a \bar{u} \nabla_a u d^n x$ es una norma proveniente de un producto escalar se cumple la regla del paralelogramo,

$$Q\left(\frac{\tilde{u}_p - \tilde{u}_q}{2}\right) + Q\left(\frac{\tilde{u}_p + \tilde{u}_q}{2}\right) = \frac{1}{2}(Q(\tilde{u}_p) + Q(\tilde{u}_q)), \quad (12.46)$$

la que implica

$$Q\left(\frac{\tilde{u}_p - \tilde{u}_q}{2}\right) \leq \frac{1}{2}(Q(\tilde{u}_p) + Q(\tilde{u}_q)) - \lambda_0 \left\| \frac{\tilde{u}_p + \tilde{u}_q}{2} \right\|_{H^0(\Omega)}^2 \xrightarrow{p, q \rightarrow \infty} \lambda_0 - \lambda_0 = 0. \quad (12.47)$$

Pero como la norma $Q(u)$ es equivalente a la de $H_0^1(\Omega)$ vemos que $\{\tilde{u}_p\}$ es también de Cauchy en $H_0^1(\Omega)$ y por lo tanto que $u_0 \in H_0^1(\Omega)$. Usando el teorema de regularidad con $f = \lambda_0 u_0$ vemos que $u_0 \in C^\infty(\Omega) \cap H_0^1(\Omega)$ y por lo tanto que u_0 es un autovalor en el sentido clásico ($\Delta u_0 + \lambda_0 u_0 = 0$). Probemos ahora la existencia de los otros autovalores-autovectores. Sea $H(1) = \{u \in H_0^1(\Omega) | (u, u_0)_{H^0(\Omega)} = 0\}$. Este es un subespacio vectorial y es cerrado, por lo tanto es un espacio de Hilbert ⁴, por lo tanto existirá λ_1 , tal que

$$\lambda_1 = \inf_{u \in H(1)} \mathcal{J}(u). \quad (12.48)$$

Repitiendo el argumento anterior vemos que λ_1 , será un autovalor, es decir que existirá una autofunción u_1 , con autovalor λ_1 .

Definiendo $H(2) = \{u \in H_0^1(\Omega) | (u, u_0)_{H^0(\Omega)} = (u, u_1)_{H^0(\Omega)} = 0\}$, etc. podemos continuar indefinidamente y obtener Σ y un conjunto ortonormal ⁵, en $H^0(\Omega)$, de autovectores. Para completar la prueba veamos que este conjunto expande $H_0^1(\Omega)$. Como u_i satisface $\Delta u_i + \lambda_i u_i = 0$ y $(u_i, u_j)_{H^0(\Omega)} = 0$ si $i \neq j$ tenemos que

$$0 = (u_j, \Delta u_i + \lambda_i u_i)_{H^0(\Omega)} = - \int_{\Omega} \nabla^a \bar{u}_j \nabla_a u_i d^n x \quad (12.49)$$

y por lo tanto también que

$$(u_j, u_i)_{H_0^1(\Omega)} = 0. \quad (12.50)$$

Vemos entonces que este conjunto es ortogonal en $H_0^1(\Omega)$ y que por construcción su subespacio perpendicular es $\{0\}$ lo que implica que este conjunto es una base de $H_0^1(\Omega)$ ♠

Ejercicio: Encuentre los autovalores y autovectores de Laplaciano en $H_0^1(\Omega)$ cuando: a) $\Omega \subset \mathbb{R}^2$ es un cuadrado de lado L . b) $\Omega \subset \mathbb{R}^3$ es una esfera de radio R . Construya en ambos casos usando estas autofunciones la función de Green del problema en cuestión.

¿Para qué ecuaciones se puede generalizar el teorema anterior? Para la prueba se usaron propiedades específicas del Laplaciano solo para afirmar que \mathcal{J} era acotada por debajo –para concluir que el ínfimo existía– y que $Q(u)$ era una norma proveniente de un producto escalar –para concluir que la ley del paralelogramo valía–. Si

$$L(u) = a^{ab} \nabla_a \nabla_b u + b^a \nabla_b u + cu, \quad (12.51)$$

con $a^{ab} l_a l_b \geq k g^{ab} l_a l_b$, $k > 0$, para todo campo vectorial l_a en Ω (elipticidad) y con a^{ab} , b^a y c campos suaves en Ω , luego existen constantes positivas c_1 y c_2 tales que

$$(u, -L(u))_{H^0(\Omega)} = - \int_{\Omega} \bar{u} L(u) d^n x \leq c_1 \int_{\Omega} g^{ab} \nabla_a \bar{u} \nabla_b u d^n x - c_2 \int_{\Omega} |u|^2 d^n x. \quad (12.52)$$

Ejercicio: Pruebe esto.

⁴Note que $H(1) \neq$ espacio perpendicular a u_0 en la norma $H_0^1(\Omega)$.

⁵Luego de normalizarlos convenientemente.

Por lo tanto en este caso también tenemos que

$$\mathcal{J}(u) \equiv \frac{(u, -L(u))_{H^0(\Omega)}}{(u, u)_{H^0(\Omega)}}, \quad (12.53)$$

es acotada por debajo. La condición de que $Q(u) \equiv (u, -L(u))_{H^0(\Omega)}$ satisfaga la *regla del paralelogramo*. Aunque no sea positiva definida es mucho más restrictiva, y es equivalente a pedir que L satisfaga

$$(v, L(u))_{H^0(\Omega)} = (L(v), u)_{H^0(\Omega)} \quad \forall u, v \in H_0^1(\Omega). \quad (12.54)$$

Los operadores que satisfacen esta relación se denominan **auto-adjuntos** o **hermitianos**.

Ejercicio: Muestre que

$$L(u) = \nabla_a(a^{ab}\nabla_b u + b^a u) - b^a \nabla_a u + cu, \quad (12.55)$$

con a^{ab} , b^a y c campos tensoriales reales es auto-adjunto.

Ejercicio: Muestre que si L es autoadjunto luego los autovectores son reales.

Ejercicio: Encuentre los autovalores y autovectores en $H_0^1(\Omega)$ de

$$L(u) = \frac{d^2}{dx^2} u + cx^2 u. \quad (12.56)$$

Llegamos así a la siguiente generalización:

Teorema 12.4 (Espectral) *Sea Ω acotado y L un operador elíptico autoadjunto con coeficientes a^{ab} , b^a y c en $C^\infty(\Omega)$ [Condición que se puede debilitar considerablemente]. Luego el problema de autovalores $L(u_i) = \lambda_i u_i$, $u_i \in H_0^1(\Omega)$, tiene un conjunto numerable y discreto de autovalores reales, cuyas autofunciones $u_i \in C^\infty(\Omega)$ expanden $H_0^1(\Omega)$.*

Ejercicio:

a) Pruebe el siguiente **corolario**:

Si L elíptico y autoadjunto es tal que sus autovalores son distintos de cero luego el problema de Dirichlet

$$L(u) = f, \quad (12.57)$$

$f \in H^0(\Omega)$, $u \in H^1(\Omega)$, tiene una única solución.

b) Si algún $\lambda_i = 0$ entonces el problema anterior tiene solución sii $(u_i, f)_{H^0(\Omega)} = 0$ para toda autofunción con autovalor cero.

13.1. Introducción

En este capítulo estudiaremos sistemas de ecuaciones hiperbólicas bajo la siguiente restricción:

Definición: Diremos que un sistema es **simétrico-hiperbólico** si:

- a.) El espacio de llegada del mapa lineal $M_{A'B}^a k_a$ es de la misma dimensión que el espacio de partida para todo $k_a \neq 0$. Por lo tanto de ahora en más usaremos índices sin primar.
- b.) El mapa $M_{AB}^a k_a$ es simétrico para todo $k_a \neq 0$.
- c.) En un entorno de cada punto existe una función τ tal que si llamamos a su diferencial por t_a ($= \nabla_a \tau$), el mapa $H_{AB} := M_{AB}^a t_a$ es positivo definido. (Es decir $H_{AB} u^A u^B \geq 0$ ($= 0$ si $u^A = 0$).)

Note que esta última condición implica que H_{AB} es una métrica en el espacio de las variables independientes. Esta y su inversa, que denotaremos por H^{AB} serán usadas para subir y bajar índices.

Esta tampoco es una restricción importante desde el punto de vista de la física, ya que todos los sistemas físicos que conocemos son simétrico-hiperbólicos.

También, pero solo por simplicidad en la exposición ya que así evitaremos algunas complicaciones técnicas, solo consideraremos sistemas lineales.

Comenzaremos este capítulo con un ejemplo simple que ilustra las características básicas de esta clase de ecuaciones.

13.2. Un ejemplo

Sea una cuerda infinita en el plano x, y y sea $y = u(x, t)$ la posición de la cuerda al tiempo t en dicho plano. Ajustando las dimensiones (de longitud o de tiempo) se puede ver que $u(x, t)$ satisface la ecuación,

$$-\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = f(x, t), \quad (13.1)$$

donde $f(x, t)$ es la densidad de fuerza que se ejerce sobre la cuerda y que suponemos no depende de la posición de la cuerda con respecto a la coordenada y .¹ Tenemos así el problema matemático de encontrar las soluciones de la ecuación,

$$\square u = g^{ab} \nabla_a \nabla_b u = f, \quad (13.2)$$

¹De otro modo deberíamos considerar $f(x, t, u)$ lo que complica el problema.

en $M = \mathbb{R}^2$ con métrica pseudo-euclídea $g_{ab} = -(dt)^2 + (dx)^2$. Para tratar esta ecuación es conveniente introducir un sistema coordenado apropiado a esta métrica, es decir uno que tiene sus líneas características como ejes,

$$\begin{aligned}\xi &= x + t = cte., \\ \eta &= x - t = cte.\end{aligned}\tag{13.3}$$

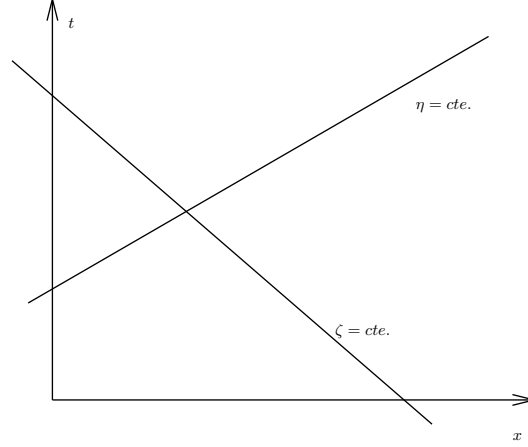


Figura 13.1: Sistema coordenado *nulo*.

tenemos entonces que

$$\begin{aligned}x &= \frac{\xi + \eta}{2} \\ t &= \frac{\xi - \eta}{2}\end{aligned}\tag{13.4}$$

y

$$\begin{aligned}g_{ab} = -(dt)^2 + (dx)^2 &= \frac{1}{4}[\{-(d\xi)^2 - (d\eta)^2 + d\xi \otimes d\eta + d\eta \otimes d\xi\} \\ &\quad + \{(d\xi)^2 + (d\eta)^2 + d\xi \otimes d\eta + d\eta \otimes d\xi\}] \\ &= \frac{1}{2}[d\xi \otimes d\eta + d\eta \otimes d\xi]\end{aligned}\tag{13.5}$$

Notando que $g^{ab} = 2[\partial\xi \otimes \partial\eta + \partial\eta \otimes \partial\xi]$ y que los símbolos de Christoffel se anulan debido a que la métrica tiene componentes constantes tenemos,²

$$\square u = \frac{\partial^2 u}{\partial\eta\partial\xi} = 4f(\eta, \xi).\tag{13.6}$$

Esta ecuación puede ser integrada inmediatamente, obteniéndose

$$\begin{aligned}\frac{\partial u}{\partial\eta}(\eta, \xi) &= 4 \int_{\xi_0}^{\xi} f(\eta, \tilde{\xi}) d\tilde{\xi} C(\eta) \\ u(\eta, \xi) &= 4 \int_{\eta_0}^{\eta} \int_{\xi_0}^{\xi} f(\tilde{\eta}, \tilde{\xi}) d\tilde{\xi} d\tilde{\eta} + u_I(\xi) + u_{II}(\eta),\end{aligned}\tag{13.7}$$

donde $u_I(\xi)$ y $u_{II}(\eta)$ son funciones arbitrarias. Consideremos primero el caso $f \equiv 0$, es decir la ecuación homogénea. Sus soluciones son entonces suma de una función cualquiera de ξ y otra cualquiera de η . Volviendo a las coordenadas x, t obtenemos

$$u(x, t) = u_I(x + t) + u_{II}(x - t).\tag{13.8}$$

²Note además que $g(\partial_\eta, \partial_\eta) = g(\partial_\xi, \partial_\xi) = 0$, es decir estos vectores coordenados tienen norma nula.

Por ejemplo,

$$u_I(x+t) = \begin{cases} e^{\frac{1}{(x+t)^2-1}} & x+t \in [-1, 1] \\ 0 & x+t \in (-\infty, -1] \cup [1, +\infty) \end{cases} \quad (13.9)$$

es una solución que representa una **onda** (\equiv solución de la ecuación homogénea) moviéndose hacia la izquierda sin cambiar de forma y con velocidad 1.

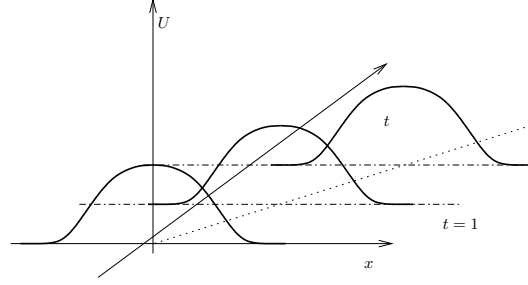


Figura 13.2: Propagación de ondas.

Similarmente u_{II} representa una onda moviéndose hacia la derecha. Ver figura. Veamos ahora que el problema de Cauchy en este caso tiene solución. Esto es extremadamente importante en física: nos dice que si tomamos una superficie no característica (por ejemplo la $t = 0$) y damos allí como dato u y su derivada temporal obtendremos una única solución para tiempos futuros. Esto es lo que nos permite, si conocemos el presente, predecir el futuro, es decir, si preparamos un experimento, predecir el resultado. Este hecho es el que distingue a la física de las otras ciencias naturales. Supongamos entonces que a $t = 0$ ($\xi = \eta = x$) damos $u(x, 0) = u_0(x)$ y su derivada $\frac{\partial u}{\partial t}(x, 0) = u_1(x)$. Tenemos entonces que

$$u_0(x) = u(x, 0) = u_I(x) + u_{II}(x), \quad (13.10)$$

$$u_1(x) = \frac{\partial u}{\partial t}(x, 0) = u'_I(x) - u'_{II}(x). \quad (13.11)$$

Diferenciando (13.10) con respecto a x y resolviendo el sistema lineal así obtenido tenemos,

$$\begin{aligned} u'_I(x) &= \frac{u'_0(x) + u_1(x)}{2} \\ u'_{II}(x) &= \frac{u'_0(x) - u_1(x)}{2}, \end{aligned} \quad (13.12)$$

e integrando,

$$\begin{aligned} u_I(x) &= \frac{u_0(x)}{2} + \frac{1}{2} \int_0^x u_1(\tilde{x}) d\tilde{x} + C_I, \\ u_{II}(x) &= \frac{u_0(x)}{2} - \frac{1}{2} \int_0^x u_1(\tilde{x}) d\tilde{x} + C_{II}. \end{aligned} \quad (13.13)$$

Para que (13.10) se satisfaga debemos tener $C_I = -C_{II}$ y por lo tanto

$$u(x, t) = \frac{1}{2}(u_0(x+t) + u_0(x-t)) + \frac{1}{2} \int_{x-t}^{x+t} u_1(\tilde{x}) d\tilde{x}. \quad (13.14)$$

Vemos entonces que si damos como dato $u_0(x) \in C^2(\mathbb{R})$ y $u_1(x) \in C^1(\mathbb{R})$ obtenemos una solución $u(x, t) \in C^2(\mathbb{R} \times \mathbb{R})$. Por construcción esta solución es única.

Ejercicio: Muestre explícitamente que (13.14) satisface (13.1) con $f \equiv 0$.

La ecuación (13.14) nos dice que a $u(x, t)$ contribuyen *solo* el promedio de los valores de u_0 en $x - t$ y $x + t$ y la integral de u_1 , entre estos dos valores. [Ver figura 13.3.]

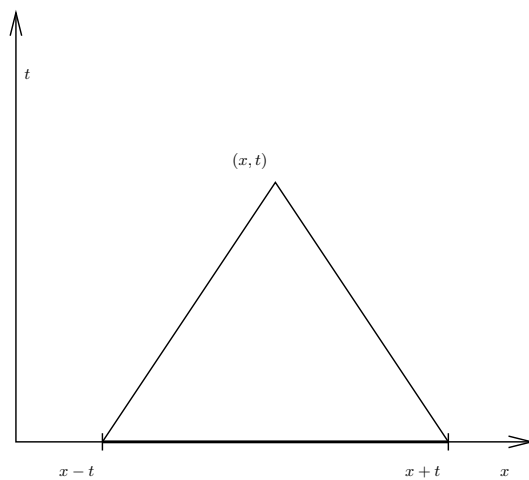


Figura 13.3: Solución general homogénea en 1+1 dimensiones.

¿Qué pasa si tenemos una fuente $f(x, t)$? Como ya tenemos la solución general (para dato de Cauchy arbitrario) de la homogénea solo necesitamos la solución de la inhomogénea con cero dato. Esto se logra integrando $f(\xi, \eta)$ primero con respecto a ξ entre $\tilde{\eta}$ y ξ y luego $\tilde{\eta}$ entre ξ y η .

$$v(\eta, \xi) = 4 \int_{\xi}^{\eta} \left(\int_{\tilde{\eta}}^{\xi} f(\tilde{\xi}, \tilde{\eta}) d\tilde{\xi} \right) d\tilde{\eta}. \quad (13.15)$$

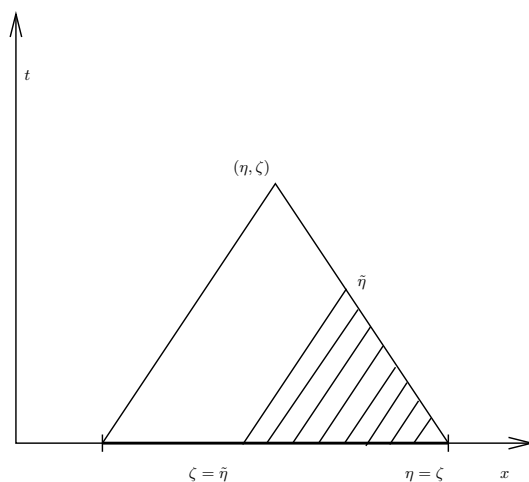


Figura 13.4: Solución general inhomogénea.

Ejercicio: Muestre que

$$v(x, t) = \int_0^t \left(\int_{x-(t-\tilde{t})}^{x+(t-\tilde{t})} f(\tilde{x}, \tilde{t}) d\tilde{x} \right) d\tilde{t}, \quad (13.16)$$

que

$$v(x, t)|_{t=0} = \frac{\partial v}{\partial t}(x, t)|_{t=0} = 0, \quad (13.17)$$

y que $v(x, t)$ satisface (13.1).

Vemos entonces que la solución buscada es,

$$u(x, t) = \frac{1}{2}(u_0(x+t) + u_0(x-t)) + \frac{1}{2} \int_{x-t}^{x+t} u_1(\tilde{x}) d\tilde{x} + \int_0^t \left(\int_{x-(t-\tilde{t})}^{x+(t-\tilde{t})} f(\tilde{x}, \tilde{t}) d\tilde{x} \right) d\tilde{t}, \quad (13.18)$$

que por construcción es única y que $u(x_0, t_0)$ depende de los valores iniciales y de f en la región cónica con vértice (x_0, t_0) dada por,

$$\begin{cases} t \leq t_0 \\ |x - x_0| \leq t_0 - t. \end{cases} \quad (13.19)$$

Esta región se llama **dominio de dependencia** del punto (x_0, t_0) , solo lo que acontece en esta región puede afectar el valor de u en dicho punto. Similarmente se define el **dominio de influencia** de un punto (x_0, t_0) como el conjunto de puntos en los cuales se puede cambiar el valor de u si se cambia el valor de u , de su derivada o de f en (x_0, t_0) . En este caso este viene dado por: $\{(x, t) | t \geq t_0, |x - x_0| \leq t - t_0\}$.

El comportamiento de las soluciones de las ecuaciones hiperbólicas es en general el mismo que el de este simple ejemplo: Dados datos de Cauchy genéricos habrá una única solución (tanto hacia el futuro como hacia el pasado). En el caso de las ecuaciones lineales esta solución se puede extender indefinidamente en ambas direcciones temporales, en el caso no-lineal las soluciones tienen solo validez en un intervalo temporal finito y es un problema físico interesante ver si las ecuaciones físicas no-lineales pueden o no ser extendidas indefinidamente y que significa físicamente la aparición de singularidades en las soluciones.³ Una cantidad de gran importancia física y matemática relacionada con la ecuación de onda es la energía de las soluciones. En dos dimensiones ésta está dada por,

$$E(u, t_0) = \frac{1}{2} \int_{t=t_0} [(\frac{\partial u}{\partial t})^2 + (\frac{\partial u}{\partial x})^2] dx. \quad (13.20)$$

Observe que la energía es positiva y su tasa de variación está dada por,

$$\begin{aligned} \frac{dE}{dt}(u, t_0) &= \int_{t=t_0} (\frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial t \partial x} \frac{\partial u}{\partial x}) dx \\ &= \int_{t=t_0} [\frac{\partial u}{\partial t} (\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2}) + \frac{\partial}{\partial x} (\frac{\partial u}{\partial t} \frac{\partial u}{\partial x})] dx \\ &= \int_{t=t_0} [\frac{\partial u}{\partial t} f] dx, \end{aligned} \quad (13.21)$$

donde en la última igualdad hemos usado (13.1) y supuesto que $\lim_{x \rightarrow \infty} \frac{\partial u}{\partial t} \frac{\partial u}{\partial x} = 0$. Si $f = 0$ luego la energía se conserva y esto nos da una prueba alternativa de la unicidad de las soluciones.

Teorema 13.1 (de Unicidad) *A lo más existe una única solución $u(x, t)$ a la ecuación de onda entre las funciones en $u(x, t) \in H^1(\mathbb{R})$, $\frac{\partial u}{\partial t}(x, t) \in H^0(\mathbb{R})$ (donde \mathbb{R} es la superficie $t = cte$) para un dado dato de Cauchy.*

³Una singularidad es un punto donde u deja de ser lo suficientemente diferenciable como para que la ecuación tenga sentido o peor aún donde u deja de tener sentido incluso como distribución.

Prueba: Supongamos existen dos soluciones u_1 y u_2 con los mismos datos de Cauchy en, digamos $t = 0$. Luego $\delta u = u_1 - u_2$ satisface la ecuación homogénea y tiene dato de Cauchy igual a cero. Por lo tanto $E(\delta u, t = 0) = 0$, pero la energía de δu es conservada y por lo tanto $E(\delta u, t - t_0) = 0 \quad \forall t_0$. Esto implica que $\|\frac{\partial \delta u}{\partial t}(x, t)|_{t=t_0}\|_{H^0(R)} = 0$ y por lo tanto $\frac{\partial \delta u}{\partial t}(x, t)|_{t=t_0} = 0$ en casi todo punto. Análogamente tenemos que $\|\frac{\partial \delta u}{\partial x}(x, t)|_{t=t_0}\|_{H^0(R)} = 0$ y por lo tanto $\frac{\partial \delta u}{\partial x}(x, t)|_{t=t_0} = 0$ o $\delta u = cte$. Pero las constantes no son de cuadrado integrable en \mathbb{R} y por lo tanto $\delta u(x, t) = u_1(x, t) - u_2(x, t) = 0$ como elemento de $H^1(\mathbb{R})$ ♠

13.3. Desigualdad de la energía para sistemas simétrico-hiperbólicos

En esta sección consideraremos un sistema simétrico-hiperbólico lineal general. Es decir un sistema de la forma:

$$M_{AB}^a \nabla_a u^B = I_A, \quad (13.22)$$

con M_{AB}^a y I_A en general dependientes de la posición, con M_{AB}^a simétrico en los índices mayúsculos y tal que exista una función τ con gradiente t_a tal que H_{AB} sea positiva definida y por lo tanto invertible.

Sea Σ_t la familia de superficies dadas por las superficies de nivel $\tau = t$ Sea Γ una región cualquiera de Σ_0 y sea Ω una región tal que $\Omega \cap \Sigma_0 = \Gamma$ sea además $\Gamma(t) = \Omega \cap \Sigma_t$

Sea u^A una solución de 13.22 y sea,

$$E(u^A, t) = \int_{\Gamma(t)} n_a M_{AB}^a u^A u^B, \quad (13.23)$$

es decir la integral de $H_{AB} u^A u^B$ sobre una región de la hipersuperficie $\tau = t = const.$, donde hemos definido H_{AB} usando $n_a = t_a/|t_a|$, es decir hemos normalizado a t_a .

Veamos la diferencia de esta cantidad entre dos superficies como muestra la figura.

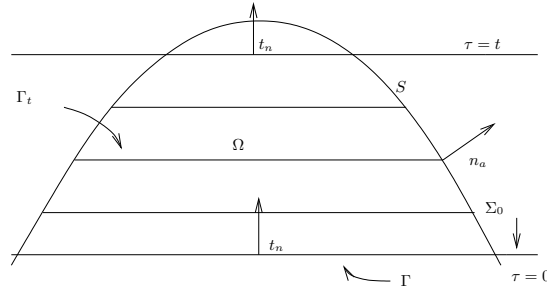


Figura 13.5: Desigualdad de la energía

Para ello usaremos el teorema de la divergencia, que nos dice que,

$$E(u^A, t) - E(u^A, 0) + E(u^A, s) = \int_{\Omega(0,t)} \nabla_a (M_{AB}^a u^A u^B), \quad (13.24)$$

donde el signo negativo del segundo término de la izquierda es debido a que en la definición de $E(u^A, 0)$ tomamos la normal entrante a la región $\Omega(0, t)$. El último término representa la integral de la energía sobre una superficie S que supondremos dada por la ecuación $\sigma = s$ donde σ es una función suave y $s \in \mathbb{R}$. Este término representa la energía que escapa de la región a través de la superficie S .

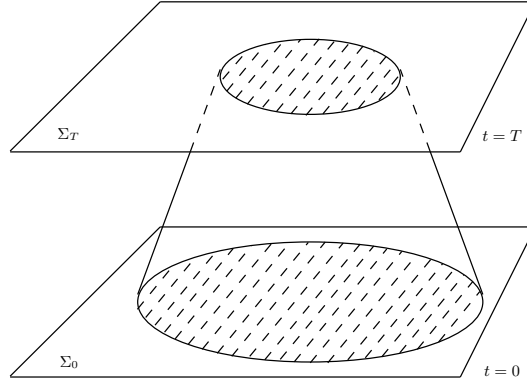


Figura 13.6: Desigualdad de la energía, vista en perspectiva.

Usando la ecuación 13.22 en el miembro derecho obtenemos,

$$\begin{aligned}
 E(u^A, t) - E(u^A, 0) + E(u^A, s) &= \int_{\Omega(0,t)} [(\nabla_a M_{AB}^a) u^A u^B + 2I_A u^A] \\
 &\leq \int_0^t \int_{\Sigma_{\tilde{t}}} \{ |(\nabla_a M_{AB}^a) u^A u^B| + 2|I_A u^A| \} d\tilde{t} \\
 &\leq \int_0^t \int_{\Sigma_{\tilde{t}}} \{ |CH_{AB} u^A u^B| \\
 &\quad + 2\sqrt{|I_A I_B H^{AB}|} \sqrt{|H_{AB} u^A u^B|} \} d\tilde{t} \\
 &\leq \int_0^t [(C+1)E(u^A, \tilde{t}) + D(\tilde{t})] d\tilde{t},
 \end{aligned} \tag{13.25}$$

donde en el primer miembro de la segunda desigualdad hemos usado la desigualdad de Schwartz (ver ejercicio) y hemos definido,

$$C^2 := \sup_{\Omega} \{ (\nabla_a M_{AB}^A) (\nabla_b M_{CD}^b) H^{AC} H^{BD} \}. \tag{13.26}$$

En la tercer desigualdad hemos usado que $2ab \leq a^2 + b^2$ y hemos definido,

$$D(t) := \int_{\Sigma_t} H^{AB} I_A I_B. \tag{13.27}$$

Ejercicio: Sea H_{AB} simétrica y positiva definida, con inversa H^{AB} . Probar:

- a.) $S^{AB} S^{CD} H_{AC} H_{BD} \geq 0$, ($= 0$ si $S^{AB} = 0$).
- b.) $|S_{AB} u^A u^B| \leq \sqrt{H^{AC} H^{BD} S_{AB} S_{CD}} H_{AB} u^A u^B$.

Haremos ahora una suposición importante que luego discutiremos en detalle:

supondremos de ahora en más que $E(u^A, s) \geq 0 \forall u^A$.

Con esta suposición podemos ignorar dicho término en la desigualdad anterior y obtener,

$$E(u^A, t) - E(u^A, 0) \leq \int_0^t [(C+1)E(u^A, \tilde{t}) + D(\tilde{t})] d\tilde{t}. \tag{13.28}$$

Diferenciando esta desigualdad integral obtendremos una cota máxima para la energía. En efecto, diferenciando esta expresión y notando que el signo de la desigualdad se mantiene obtenemos la siguiente desigualdad diferencial,

$$\frac{d}{dt}E(u^A, t) \leq (1 + C)E(u^A, t) + D(t), \quad (13.29)$$

La igualdad diferencial tiene como solución (usando el método de variación de constantes, sección 5.2),

$$Y(t) = e^{(1+C)t} [Y(0) + \int_0^t e^{-(1+C)\tilde{t}} D(\tilde{t}) d\tilde{t}]. \quad (13.30)$$

Usando ahora el lema 4.1 vemos que $E(u^A, t) \leq Y(t) \forall t \geq 0$ si $E(u^A, 0) = Y(0)$, y por lo tanto tenemos que,

$$E(u^A, t) = e^{(1+C)t} [E(u^A, 0) + \int_0^t e^{-(1+C)\tilde{t}} D(\tilde{t}) d\tilde{t}]. \quad (13.31)$$

Esta desigualdad es extremadamente importante, no solo permite inferir la unicidad de las soluciones (como veremos a continuación) sino también juega un papel fundamental para probar la existencia de soluciones y para lograr algoritmos numéricos convergentes y fiables.

13.4. Unicidad de las soluciones

Usando la desigualdad obtenida anteriormente probaremos el siguiente teorema:

Teorema 13.2 *Sea una ecuación simétrico-hiperbólica en una variedad M , Sea Σ_0 una superficie dada por la ecuación $\tau = 0$ con y tal que $M_{AB}^a \nabla_a \tau$ sea positiva definida. Sea Γ una región cualquiera de Σ_0 y sea Ω una región tal que $\Omega \cap \Sigma_0 = \Gamma$ y tal que $E(u^A, s) \geq 0 \forall u^A$. [Ver figura anterior.] Luego si u^C y \tilde{u}^C son dos soluciones que coinciden en Γ luego éstas coinciden en todo Ω .*

Prueba: Sea $\delta^A := u^A - \tilde{u}^A$. Luego δ^A satisface,

$$M_{AB}^a \nabla_a \delta^A = 0. \quad (13.32)$$

Por lo tanto tenemos una desigualdad de la energía para δ^C con $D(t) \equiv 0$ y además con $E(\delta^C, 0) = 0$ ya que las dos soluciones coinciden en Γ . Pero entonces la desigualdad nos dice que $E(\delta^C, t) = 0$ para todo t y por lo tanto $\delta^C = 0$ en todo Ω probando así el teorema ♠

13.5. Dominio de dependencia

El teorema de unicidad anterior se basó en la suposición de que

$$E(u^C, s) = \int_S n_a M_{AB}^a u^A u^B \geq 0 \quad (13.33)$$

y por lo tanto es importante determinar cuáles son las posibles regiones donde esto pasa. En particular, dada una región Γ , la mayor región Ω donde tenemos unicidad de las soluciones con idénticos datos iniciales en Γ se llama el **dominio de dependencia** de Γ , es la región que depende completamente de los datos iniciales dados en Γ , o sea dando datos iniciales en Γ podemos controlar completamente el valor de la solución en cualquier punto de su dominio de dependencia.

Veamos primero que este dominio de dependencia no es vacío. Para ello tomemos Γ compacto en Σ_0 y consideremos $H_{AB} = t_a M_{AB}^a$. Sea ahora $\sigma = \tau - \delta\xi$ con ξ una función suave en un entorno de Γ positiva en dentro de este conjunto y negativa en $\Sigma_0 - \Gamma$ (o sea que se anula en su frontera) y δ un número real que supondremos pequeño. [Ver figura 13.7.]

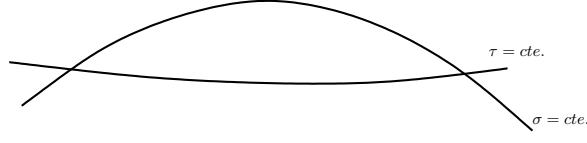


Figura 13.7: Región en forma de ampolla

En cada punto $p \in \Gamma$ H_{AB} es una métrica positiva definida y por lo tanto, como el conjunto de métricas positivas definidas es abierto en el espacio de todos los tensores simétricos, dado otro covector w_a cualquiera existirá $\varepsilon > 0$ suficientemente pequeño tal que $(t_a + \varepsilon w_a)M_{AB}^a = H_{AB} + \varepsilon w_a M^a AB$ es también positiva definida. Como Γ es compacta dado un w_a en ella habrá un ε mínimo y positivo tal que el tensor definido arriba será positivo⁴. Por el mismo argumento de continuidad habrá una región compacta alrededor de Γ y un $\varepsilon > 0$, un poco menor que el anterior tal que allí $\nabla_a \sigma M_{AB}^a$ será positiva definida. Hemos logrado así tener una región Ω entre las superficies de nivel $\tau = 0$ y $\sigma = 0$, es decir, una pequeña *ampolla* donde la integral de la energía saliente por $S = \{p \in M | \sigma(p) = 0\}$ es positiva para toda solución u^A .

¿Cuán grande podemos hacer esta ampolla, es decir cuánto más podemos inclinar la superficie S y todavía mantener positividad? Esta pregunta tiene mucho que ver con la siguiente: ¿cuánto podemos inclinar a t_a en cada punto y todavía tener positividad de $t_a M_{AB}^a$ en dicho punto?

Obsérvese primero que si $t_a M_{AB}^a$ es positivo luego $\alpha t_a M_{AB}^a$, $\alpha > 0$ también lo es, con lo que el conjunto de los t_a para los cuales tenemos positividad forman un cono. Esto también nos asegura de que no todos los co-vectores están en este cono, ya que si t_a lo está $-t_a$ no lo está.

Segundo obsérvese que si t_a y \tilde{t}_a están en el cono, [es decir cada uno de ellos da una métrica positiva definida] también lo están todos los co-vectores de la forma, $\alpha t_a + (1 - \alpha)\tilde{t}_a$, $\alpha \in [0, 1]$ ya que $(\alpha t_a M_{AB}^a + (1 - \alpha)\tilde{t}_a M_{AB}^a)u^A u^B$ es positivo si los coeficientes de α y $(1 - \alpha)$ lo son. Es decir el conjunto de los co-vectores que dan métricas positivas definidas forman un cono convexo de T_p^* , en cada punto $p \in M$. Este cono se denomina **cono característico**. [Ver figura 13.8.]

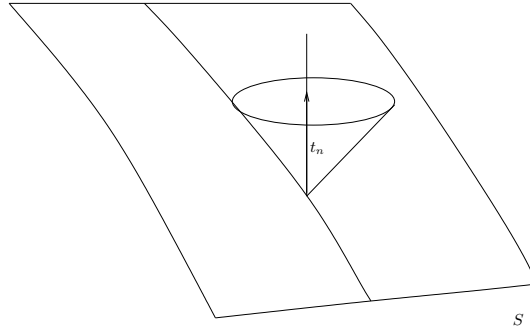


Figura 13.8: Cono característico

¿Qué pasa con los co-vectores en la frontera de dicho cono? Allí la condición de positividad definida debe fallar, es decir dado un co-vector t_a en dicha frontera habrá algún u^A tal que $t_a M_{AB}^a u^A u^B = 0$. Esto implica que allí el rango de $t_a M_{AB}^a$ deja de ser máximo.

⁴Para ver esto considere el mapa entre $(B_1 \times \Gamma) \times (B_1 \times \Gamma) \times \Gamma \rightarrow \mathbb{R}$ dado por, $w_a(p)M_{AB}^a(p)u^A(p)u^B(p)$ donde $B_1(p) = \{u^A(p) | H_{AB}(p)u^A u^B = 1\}$. Este es un mapa continuo y su dominio es compacto por lo tanto tiene un máximo, $m < \infty$. Podremos tomar entonces $0 < \varepsilon < 1/m$.

13.5.1. Construcción de una superficie característica

Ahora construiremos la superficie frontera del máximo dominio de dependencia. A partir de una región Γ dada por $\{q \in \Sigma_0 \mid \sigma_0(p) = 0\}$ construiremos una superficie S tal que su normal en cada uno de sus puntos pertenezca a la frontera del cono característico. Esta superficie vendrá dada por una función $\sigma = 0$ tal que $\sigma|_{\Sigma_0} = \sigma_0$. Para encontrar la ecuación que esta función deberá satisfacer es conveniente introducir un sistema apropiado de coordenadas, el mismo que usamos en nuestra clasificación de las ecuaciones en derivadas parciales. Una de estas coordenadas será $t = \tau$ y llamaremos a las otras x^i , también por conveniencia definiremos $\sigma_t = \frac{\partial \sigma}{\partial t}$ y $\sigma_i = \frac{\partial \sigma}{\partial x^i}$. Con estas coordenadas obtenemos que,

$$\begin{aligned} \nabla_a \sigma M_{AB}^a &= \sigma_t M_{AB}^t + \sum_i \sigma_i M_{AB}^i \\ &= \sigma_t H_{AB} + \sum_i \sigma_i M_{AB}^i. \end{aligned} \quad (13.34)$$

Multiplicando por la inversa de H_{AB} , H^{AB} , obtenemos

$$H^{CA} \nabla_a \sigma M_{AB}^a = \sigma_t \delta^C_B + \sum_i \sigma_i H^{CA} M_{AB}^i, \quad (13.35)$$

el cual es un operador, es decir un mapa lineal de un espacio vectorial en sí mismo, y por lo tanto le podemos tomar su determinante, obteniendo,

$$\det \left(\sigma_t \delta^C_B + \sum_i \sigma_i H^{CA} M_{AB}^i \right) = 0, \quad (13.36)$$

ya que el determinante de $H^{CA} \nabla_a \sigma M_{AB}^a$ se anula pues hemos supuesto que el rango de $\nabla_a \sigma M_{AB}^a$ dejaba de ser máximo. Para cada valor fijo de las derivadas espaciales σ_i esta ecuación tendrá en general n soluciones (raíces) reales, σ_t , (los autovalores del operador $\sum_i \sigma_i H^{CA} M_{AB}^i$), de todas ellas tomaremos aquella que nos dé la frontera del menor cono conteniendo a t_a , es decir la mayor raíz. Tendremos así para esta raíz una ecuación de la forma:

$$\sigma_t + H(\sigma_i, x^i, t) = 0. \quad (13.37)$$

La función H tiene una propiedad muy importante, note que si (σ_t, σ_i) es una solución también lo es $(\alpha \sigma_t, \alpha \sigma_i)$ y por lo tanto H debe ser **homogénea de primer grado**, es decir $H(\alpha \sigma_i, x^i, t) = \alpha H(\sigma_i, x^i, t)$. Estas ecuaciones, con H homogénea de primer grado se llaman ecuaciones **eikoniales** y son casos particulares de la ecuación de Hamilton-Jacobi que se estudia en mecánica.

Este tipo de ecuaciones, que en principio son ecuaciones en derivadas parciales pueden ser resueltos transformándolos a un problema equivalente en derivadas ordinarias proveniente de un Hamiltoniano. En efecto, considere el sistema de ecuaciones ordinarias Hamiltonianas:

$$\begin{aligned} \frac{dx^i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial x^i}, \end{aligned} \quad (13.38)$$

donde $H(p_i, x^i, t) := H(\sigma_i, x^i, t)|_{\sigma_i=p_i}$. Integrando este sistema con condiciones iniciales:

$$\begin{aligned} x^i(0) &= x_0^i \\ p_i(0) &= \sigma_{0i}, \end{aligned} \quad (13.39)$$

y luego restringiendo las curvas integrales obtenidas en el espacio de fase (x^i, p_j) al espacio de configuración x^i , obtenemos una serie de curvas en nuestra variedad emanando de la superficie

Γ . Llamaremos a estas curvas, **curvas características** ya que tienen la importante propiedad de que a lo largo de ellas la función σ que andamos buscando es constante! Para probar esto primero veamos que con las condiciones iniciales elegidas $p_i(t) = \sigma_i(x^i(t), t) \forall t$. Pero tomando la derivada con respecto a x^i de la ecuación 13.37 obtenemos:

$$\begin{aligned} \frac{\partial \sigma_i}{\partial t} &= \frac{\partial^2 \sigma}{\partial t \partial x^i} \\ &= \frac{\partial^2 \sigma}{\partial x^i \partial t} \\ &= - \sum_j \frac{\partial H}{\partial \sigma_j}(\sigma_k, x^k, t) \frac{\partial \sigma_i}{\partial x^j} - \frac{\partial H}{\partial x^i}(\sigma_k, x^k, t), \end{aligned} \quad (13.40)$$

y por lo tanto

$$\begin{aligned} \frac{d(p_i - \sigma_i)}{dt} &= - \frac{\partial H}{\partial x^i}(p_k, x^k, t) - \sum_j \frac{\partial \sigma_i}{\partial x^j} \frac{dx^j}{dt} - \frac{\partial \sigma_i}{\partial t} \\ &= - \frac{\partial H}{\partial x^i}(p_k, x^k, t) - \sum_j \frac{\partial \sigma_i}{\partial x^j} \frac{\partial H}{\partial p_j}(p_k, x^k, t) \\ &\quad + \sum_j \frac{\partial H}{\partial \sigma_j}(\sigma_k, x^k, t) \frac{\partial \sigma_i}{\partial x^j} + \frac{\partial H}{\partial x^i}(\sigma_k, x^k, t) \\ &= - \frac{\partial H}{\partial x^i}(p_k, x^k, t) + \frac{\partial H}{\partial x^i}(\sigma_k, x^k, t) \\ &\quad - \sum_j \frac{\partial \sigma_i}{\partial x^j} \left(\frac{\partial H}{\partial p_j}(p_k, x^k, t) - \frac{\partial H}{\partial \sigma_j}(\sigma_k, x^k, t) \right), \end{aligned} \quad (13.41)$$

donde en la segunda igualdad hemos usado las ecuaciones 13.38 y 13.40. Esta última ecuación, con las condiciones iniciales elegidas tiene solución trivial, pero el teorema de unicidad de las soluciones de ecuaciones ordinarias nos garantiza que esta es la única y por lo tanto la igualdad buscada. Por lo tanto de ahora en más no deberemos distinguir en el argumento de H si está evaluada en σ_i o en p_i .

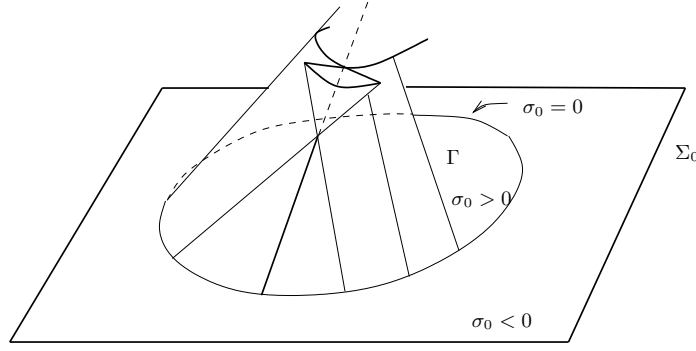
Pero entonces note que la derivada a lo largo de una curva característica de σ es,

$$\begin{aligned} \frac{d\sigma}{dt} &= \sum_i \sigma_i \frac{dx^i}{dt} + \sigma_t \\ &= \sum_i \frac{\partial H}{\partial \sigma_i} \sigma_i - H(\sigma_k, x^k, t) \\ &= 0, \end{aligned} \quad (13.42)$$

donde en la última igualdad hemos usado que por ser H homogénea de primer grado en σ_i se cumple la igualdad $H(\sigma_k, x^k, t) = \sum_i \frac{\partial H}{\partial \sigma_i} \sigma_i$.

Hemos demostrado entonces que σ será constante a lo largo de las líneas integrales de la ecuación 13.38 con condiciones iniciales dadas por 13.39. Por lo tanto conocemos S , esta será la superficie reglada por las curvas características emanando de $\partial\Gamma^5$. [Ver figura 13.9.]

⁵En general las curvas características se cruzan entre sí y por lo tanto aún dando una región Γ con frontera suave luego de un cierto tiempo la superficie reglada desarrollará singularidades y σ será multivaluada. Sin embargo estas singularidades se conocen perfectamente bien y se sabe cómo descartar regiones hasta obtener dominios de dependencia con frontera continuas.

Figura 13.9: Construcción de S y una singularidad en S

13.5.2. Dominio de dependencia, ejemplos

A continuación damos un par de ejemplos de la construcción de curvas características y determinación de los dominios de dependencia.

Ejemplo: Fluidos en una dimensión

Consideremos un fluido con densidad promedio ρ_0 , moviéndose con velocidad promedio v_0 y con una ecuación de estado para la presión en función de la densidad $p = p(\rho)$. Estamos interesados en pequeñas fluctuaciones de estas cantidades alrededor del estado de equilibrio (ρ_0, v_0) , es decir en la teoría del sonido en este fluido. En este caso $u^A = (\rho, v)$ será dichas fluctuaciones y las ecuaciones del fluido son

$$\begin{aligned} \frac{\partial \rho}{\partial t} - v_0 \frac{\partial \rho}{\partial x} - \rho_0 \frac{\partial v}{\partial x} &= 0 \\ \frac{\partial v}{\partial t} - v_0 \frac{\partial v}{\partial x} - \frac{1}{\rho_0} \frac{\partial p}{\partial x} &= 0. \end{aligned} \quad (13.43)$$

Si $c^2 := \frac{dp}{d\rho}|_{\rho_0} > 0$, (c es la velocidad del sonido en el medio) entonces el sistema es simétrico-hiperbólico con M_{AB}^a obtenido reescribiendo las ecuaciones como:

$$\begin{pmatrix} c^2 & 0 \\ 0 & \rho_0^2 \end{pmatrix} \begin{pmatrix} \rho \\ v \end{pmatrix}_t + \begin{pmatrix} -c^2 v_0 & -c^2 \rho_0 \\ -c^2 \rho_0 & -\rho_0^2 v_0 \end{pmatrix} \begin{pmatrix} \rho \\ v \end{pmatrix}_x = 0, \quad (13.44)$$

donde hemos usado que $\frac{\partial p}{\partial x} = \frac{\partial p}{\partial \rho} \frac{\partial \rho}{\partial x}$.

El determinante que debemos estudiar es entonces:

$$\det \begin{pmatrix} \sigma_t c^2 - \sigma_x c^2 v_0 & -\sigma_x c^2 \rho_0 \\ -\sigma_x c^2 \rho_0 & \sigma_t \rho_0^2 - \sigma_x \rho v \end{pmatrix}, \quad (13.45)$$

que tiene como raíces,

$$\sigma_t = (v_0 \pm c) \sigma_x. \quad (13.46)$$

Supongamos $c > v_0$ (fluido normal, subsónico), y que $\Gamma = [0, 1]$ con $\sigma_0 = x(x-1)$. En $x = 0$ σ_{0x} es negativa y entonces la mayor raíz es $v_0 - c$ y la solución es $\sigma_- = t(c - v_0) - x$. En $x = 1$ σ_{0x} es positiva y entonces la mayor raíz es $v_0 + c$ y la solución es $\sigma_- = t(v_0 + c) - (x - 1)$. [Ver figura.]

Ejemplo: Ecuación de onda en 2 + 1 dimensiones.

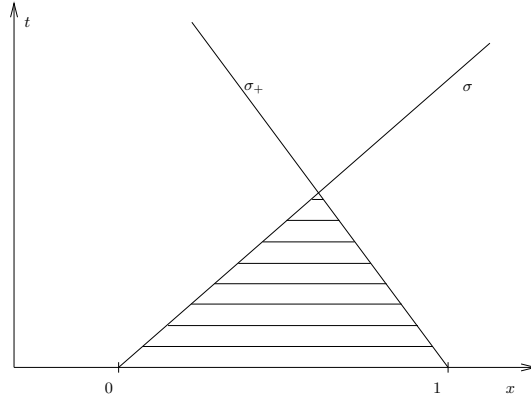


Figura 13.10: Dominio de dependencia de un fluido

La ecuación es:

$$-\frac{\partial^2 \phi}{\partial t \partial t} + \frac{\partial^2 \phi}{\partial x \partial x} + \frac{\partial^2 \phi}{\partial y \partial y} = \rho, \quad (13.47)$$

y usando $u^A = (\phi, \frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y})$ el sistema se puede escribir como,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \\ u^3 \\ u^4 \end{pmatrix}_t - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \\ u^3 \\ u^4 \end{pmatrix}_x - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \\ u^3 \\ u^4 \end{pmatrix}_y = \begin{pmatrix} -u^2 \\ -\rho \\ 0 \\ 0 \end{pmatrix}$$

Ejercicio: Pruebe que si los datos iniciales son tales que $u^2 = \frac{\partial u^1}{\partial x}$ y $u^3 = \frac{\partial u^1}{\partial y}$, luego u^1 satisface 13.47.

Las ecuaciones para σ son:

$$\det \begin{pmatrix} \sigma_t & 0 & 0 & 0 \\ 0 & \sigma_t & \sigma_x & \sigma_y \\ 0 & \sigma_x & \sigma_t & 0 \\ 0 & \sigma_y & 0 & \sigma_t \end{pmatrix} = (\sigma_t)^2 [(\sigma_t)^2 - (\sigma_x)^2 - (\sigma_y)^2], \quad (13.48)$$

o sea

$$\sigma_t = \sqrt{(\sigma_x)^2 + (\sigma_y)^2} := -H(\sigma_x, \sigma_y). \quad (13.49)$$

Las ecuaciones de Hamilton para este sistema son:

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial H}{\partial \sigma_x} = \frac{-\sigma_x}{\sqrt{(\sigma_x)^2 + (\sigma_y)^2}} \\ \frac{dy}{dt} &= \frac{\partial H}{\partial \sigma_y} = \frac{-\sigma_y}{\sqrt{(\sigma_x)^2 + (\sigma_y)^2}} \\ \frac{d\sigma_x}{dt} &= -\frac{\partial H}{\partial x} = 0 \\ \frac{d\sigma_y}{dt} &= -\frac{\partial H}{\partial y} = 0, \end{aligned} \quad (13.50)$$

y sus soluciones son:

$$\begin{aligned}x(t) &= \frac{-\sigma_{0x}}{\sqrt{(\sigma_{0x})^2 + (\sigma_{0y})^2}}t + x_0 \\y(t) &= \frac{-\sigma_{0y}}{\sqrt{(\sigma_{0x})^2 + (\sigma_{0y})^2}}t + y_0.\end{aligned}\tag{13.51}$$

14.1. Introducción

Aquí trataremos el arquetipo de ecuación parabólica, la ecuación del calor,

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= f \text{ en } \Omega, \\ u|_{t=0} &= u^0, \\ u|_L &= 0, \end{aligned} \quad (14.1)$$

donde Ω es una región cilíndrica de la forma $[0, T] \times S$ y $L = [0, T] \times \partial S$. Ver figura.

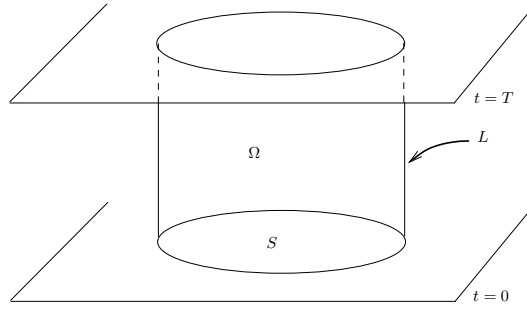


Figura 14.1: Condiciones de contorno para la ecuación del calor.

También se puede tratar el problema en que $n^a \nabla_a u|_L = 0$, o problemas mixtos, aquí solo trataremos el primero, ya que el tratamiento de los otros no reviste mayores cambios.

Usando el Teorema Espectral descomponemos u , f y u^0 en autofunciones del Laplaciano en S en $H_0^1(S)$ y obtenemos,

$$u(t, x^j) = \sum_{n=0}^{\infty} u_n(t) v_n(x^j), \quad f(t, x^j) = \sum_{n=0}^{\infty} f_n(t) v_n(x^j), \quad u^0(x^j) = \sum_{n=0}^{\infty} u_n^0 v_n(x^j), \quad (14.2)$$

y las ecuaciones anteriores nos dan el siguiente sistema infinito de ecuaciones ordinarias,

$$\begin{aligned} \dot{u}_n + \lambda_n u_n &= f_n \\ u_n|_{t=0} &= u_n^0. \end{aligned} \quad (14.3)$$

La solución de la ecuación homogénea es $u_n^H = u_n^0 e^{-\lambda_n t}$ y usando el método de variación de constantes obtenemos,

$$u_n^I = e^{-\lambda_n t} \int_0^t f_n(\tilde{t}) e^{\lambda_n \tilde{t}} d\tilde{t}. \quad (14.4)$$

La solución completa (respetando la condición inicial) es

$$u_n(t) = e^{-\lambda_n t} [u_n^0 + \int_0^t f_n(\tilde{t}) e^{\lambda_n \tilde{t}} d\tilde{t}]. \quad (14.5)$$

De manera análoga a como probamos que la solución formal para el caso hiperbólico era una solución suave se puede demostrar aquí que para $t > 0$ la solución es, en las variables espaciales dos veces más diferenciable que f y en la temporal una vez más.

Una propiedad muy importante de esta ecuación es que en general solo admite una solución para $t > 0$, esto implica que a diferencia con las ecuaciones de la mecánica o el electromagnetismo esta ecuación privilegia una dirección temporal particular. Entre otras cosas esto nos indica que esta ecuación representa o describe fenómenos que no pueden ser derivados solamente de la mecánica y que en ellos debe haber alguna clase de información de tipo termodinámico.

Para ver esto tomemos el ejemplo dado antes del Teorema Espectral,

$$u - \frac{d^2 u}{dx^2} = 0 \quad \text{en } [0, 1], \quad (14.6)$$

donde vimos que las autofunciones $v_n(x)$ con $v_n(0) = v_n(1) = 0$ son $v_n(x) = \sin(\pi n x)$ con autovalor $\lambda_n = \pi^2 n^2$.

Tomando como dato inicial $u^0(x) = \sum_{n=1}^{\infty} \frac{(-1)^{\frac{n-1}{2}}}{n^2} \sin(\pi n x)$, la cual es acotada en $[0, 1]$ ya que $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$, obtenemos, $u(t, x) = \sum_{n=1}^{\infty} \frac{(-1)^{\frac{n-1}{2}} e^{-\pi^2 n^2 t}}{n^2} \sin(\pi n x)$. Pero $u(t, 1/2) = \sum_{n=1}^{\infty} \frac{e^{-\pi^2 (2n-1)^2 t}}{(2n-1)^2}$ la cual es finita $\forall t \geq 0$ e infinita para cualquier $t < 0$ ya que el n -ésimo término tiende a infinito con n . Por otro lado dada cualquier $u^0(x) = \sum_{n=0}^{\infty} u_n^0 \sin(\pi n x)$, continua obtenemos

$$u(t, x) = \sum_{n=0}^{\infty} u_n^0 e^{-\pi^2 n^2 t} \sin(\pi n x) \quad (14.7)$$

la cual es infinitamente diferenciable para todo $t > 0$ ya que dado cualquier polinomio $P(n)$ de n $P(n) e^{-\pi^2 n^2 t} \rightarrow 0$ cuando $n \rightarrow \infty$ si $t > 0$.

14.2. Unicidad y el Teorema del Máximo

Vemos ahora que la solución obtenida en el caso general es única. Para ello supondremos que esta es C^1 con respecto al tiempo y C^2 con respecto a las coordenadas espaciales. Para llegar a esta conclusión usaremos el principio del máximo. Desarrollamos este primero para el problema de Dirichlet.

Teorema 14.1 (del Máximo) Sea $u \in C^2(S)$ y $\Delta u \geq 0$ en S , luego

$$\max_{\bar{S}} u = \max_{\partial S} u. \quad (14.8)$$

Prueba: Si $\Delta u > 0$ esto sigue simplemente del hecho de que si el máximo estuviese en $p \in S$ luego $\frac{\partial^2 u}{\partial (x^k)^2} \big|_p \leq 0 \quad \forall k = 1, \dots, n$ y por lo tanto $\Delta u \leq 0$ lo que es una contradicción. Para el caso $\Delta u \geq 0$ usamos $v = |x|^2$. Como $\Delta v > 0$ en S luego

$$\Delta(u + \varepsilon v) > 0 \quad \text{en } S \quad \forall \varepsilon > 0 \quad (14.9)$$

y así

$$\max_{\bar{S}} (u + \varepsilon v) = \max_{\partial S} (u + \varepsilon v) \quad (14.10)$$

y por lo tanto

$$\max_{\bar{S}} u + \varepsilon \min_{\bar{S}} v \leq \max_{\partial S} u + \varepsilon \max_{\partial S} v \quad (14.11)$$

y tomando $\varepsilon \rightarrow 0$ obtenemos la igualdad buscada. En el caso en que $\Delta u = 0$ luego también se cumple que

$$\min_{\bar{S}} u = \min_{\partial S} u \quad (14.12)$$

(Simplemente usando que $\min(u) = -\max(-u)$).

Este resultado nos da otra prueba de la unicidad de las soluciones al problema de Dirichlet.

Teorema 14.2 (Unicidad del Problema de Dirichlet) *El problema*

$$\begin{aligned} \Delta u &= f && \text{en } S \\ u|_{\partial S} &= u_0, \end{aligned} \quad (14.13)$$

tiene a lo sumo una única solución en $C^2(S)$.

Prueba: Sean u y \tilde{u} en $C^2(S)$ soluciones, luego $\delta = u - \tilde{u}$ satisface $\Delta \delta = 0$ y $\delta|_{\partial S} = 0$ pero entonces

$$\max_{\bar{S}} \delta = \max_{\partial S} \delta = 0 \quad (14.14)$$

y

$$\min_{\bar{S}} \delta = \min_{\partial S} \delta = 0 \quad (14.15)$$

por lo tanto concluimos que $\delta = 0$.

Ejercicio: ¿Para qué otras ecuaciones elípticas vale esta prueba de unicidad?

¿Podemos obtener un resultado similar para la ecuación del calor?

Teorema 14.3 (Unicidad para la ecuación del Calor) *Existe a lo más una única solución $u \in C^1[0, T] \times C^2(S)$ del problema,*

$$\begin{aligned} u_t - \Delta u &= f && \text{en } S \\ u|_{t=0} &= u^0, \\ u|_L &= u^1, \end{aligned} \quad (14.16)$$

u^0 y u^1 continuas.

La prueba de este teorema es una aplicación trivial del siguiente lema.

Lema 14.1 *Sea $u \in C^1[0, T] \times C^2(S)$ continua en $\bar{\Omega}$ y satisfaciendo $u_t - \Delta u \leq 0$. Luego*

$$\max_{\bar{\Omega}} u = \max_{\partial' \Omega} u, \quad (14.17)$$

donde $\partial' \Omega = S_0 \cup L$.

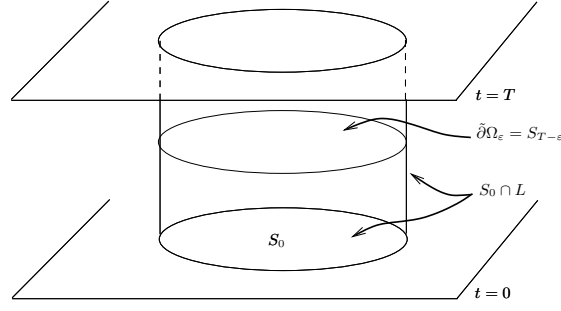


Figura 14.2: Prueba del Lema 14.1

Prueba: Veamos primero el caso $u_t - \Delta u < 0$ en Ω . Sea $0 < \varepsilon < T$ y $\Omega_\varepsilon = \{\cup_{\tilde{t} \in (0, T-\varepsilon)} S_{\tilde{t}}\}$. Como u es continua en $\bar{\Omega}_\varepsilon$ existirá $p \in \bar{\Omega}_\varepsilon$ tal que $u(p) = \max_{\Omega_\varepsilon} u$. Si $p \in \Omega_\varepsilon$ luego allí $u_t = 0$ y $\Delta u \leq 0$ lo que nos da una contradicción. Si $p \in \tilde{\partial}\Omega_\varepsilon = S_{T-\varepsilon}$ tendríamos $u_t \geq 0$ y $\Delta u \leq 0$, lo que también nos da una contradicción. Sigue entonces que $p \in \partial'\Omega_\varepsilon$ y

$$\max_{\Omega_\varepsilon} u = \max_{\partial'\Omega_\varepsilon} u \leq \max_{\partial'\Omega} u \quad (14.18)$$

tendiendo $\varepsilon \rightarrow 0$ obtenemos el resultado buscado. Para tratar el caso $u_t - \Delta u \leq 0$ en Ω consideramos $v = u - kt$, $k > 0$, luego $v_t - \Delta v = u_t - \Delta u - k < 0$ y por lo tanto

$$\max_{\bar{\Omega}} u = \max_{\bar{\Omega}} (v + kt) \leq \max_{\bar{\Omega}} v + kT = \max_{\partial'\Omega} v + kT \leq \max_{\partial'\Omega} u + kT \quad (14.19)$$

tomando $k \rightarrow 0$ obtenemos el resultado buscado.

Ejercicio: Probar que si $\frac{\partial u}{\partial t} - \Delta u \geq 0$ luego $\min_{\bar{\Omega}} u = \min_{\partial'\Omega} u$.

Bibliografía

- [1] *Mathematical Physics*, Geroch, R., The Univ. of Chicago Press .
- [2] *Topología General*, Kelley, J. L., EUDEBA .
- [3] *Ordinary Differential Equations*, Arnold, V. I., The MIT Press, 1987.
- [4] *Mathematical Methods of Classical Mechanics*, Arnold, V. I., Springer, 1980.
- [5] *Differential Equations: Geometric Theory*, Lefschetz, S, Wiley, 1962.
- [6] *General Relativity*, Wald, R., The University of Chicago Press, 1993.
- [7] *Differential Equations and their applications, an introduction to applied mathematics*, Braum, M. New York, Springer, 1983.
- [8] *Ecuaciones diferenciales ordinarias y teoría de control*. Roxin, E. Buenos Aires, EUDEBA, 1968.
- [9] *Methods of Modern Mathematical Physics. I. Functional Analysis*, Reed, M., and Simon, B., Academic Press, 1972.
- [10] *Real Analysis*, Lang, S., Addison-Wesley, 1983.
- [11] *Introduction to Linear Algebra*, Lang, S., Undergraduate Texts in Mathematics, Springer, second edition, 1986.
- [12] *Linear Algebra*, Lax, P., Pure and applied mathematics, Wiley, 1996.
- [13] *Functional Analysis*, Yosida, K., Springer-Verlag 1980.
- [14] *Introductory Relat Analysis*, A.N. Kolmogorov, and S.V. Fomin, Dover, 1970.
- [15] *Basic Linear Partial Differential Equations*, Treves, F., Academic Press, 1975.
- [16] *Partial Differential Equations*, John, F., Springer-Verlag 1982.
- [17] *Introduction to Partial Differential Equations*, Folland, G. B., Mathematical Notes, Princeton Univ. Press 1976.
- [18] *Partial Differential Equations I: Basic Theory* M. Taylor, Aplied Mathematical Sciences 115, Springer-Verlag, 1996.
- [19] *Ecuaciones de la Física Matemática*, Godunov, S.K. Moscú, MIR, 1978.
- [20] *Problemas de Ecuaciones Diferenciales Ordinarias* Kiseliiov A. , Krasnov M., and Makarenko G., Moscú, MIR, Moscú, 1968.
- [21] *Modern Differential Geometry for Physicists* Isham C., World Scientific Lecture Notes in Physics - Vol. 61 World Scientific, Second edition, 1999.