

ANÁLISIS MATEMÁTICO AVANZADO

ANÁLISIS MATEMÁTICO AVANZADO

Dr. Andrés Fraguela Collar

Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

PRÓLOGO

El texto comienza con un capítulo introductorio sobre “Teoría de conjuntos”, que resulta de gran importancia por sí mismo y es, además, de gran utilidad para el estudio del resto de los capítulos.

En el capítulo 2, partiendo de problemas de convergencia y otros problemas topológicos en espacios concretos de puntos y funciones, se llega a la definición abstracta de *espacio métrico*.

En esta nueva estructura se estudia un concepto de convergencia y de continuidad que generaliza de forma muy intuitiva a dichos conceptos en los espacios numéricos.

Mediante algunos resultados, ejemplos y problemas, se trata de llevar al lector la comprensión sobre la insuficiencia de esta abstracción para establecer una teoría general de la convergencia y continuidad, y se introduce el concepto topológico *conjunto abierto*. En este capítulo se ve que la integral de Riemann se puede definir a partir de un concepto de límite que no puede ser descrito mediante una métrica.

En los tres capítulos siguientes (3-5) se tratan subclases especiales de espacios métricos que son de particular importancia en las aplicaciones. Para el estudio de estos capítulos se cuenta con un lector motivado que conoce las nociones fundamentales de la estructura métrica y, por ello, se hacen progresos más rápidos.

En el capítulo 3 se introducen las nociones de sucesión de Cauchy y espacio métrico completo, cuya teoría es indispensable para la aplicación a los espacios funcionales, además de que permite ver con claridad el proceso de completamiento de un espacio métrico mediante el cual se obtiene el campo de los números reales a partir del conjunto de los números racionales. Precisamente, siguiendo este proceso, se obtienen los espacios completos de funciones integrables que son de tanta utilidad en las aplicaciones.

En el capítulo 4 se estudian los *espacios métricos compactos*, que son de gran importancia en el análisis matemático y en el Análisis Funcional. Los primeros epígrafes exponen las propiedades básicas de estos espacios; esta parte del capítulo requiere un estudio cuidadoso y completo. La importancia de los resultados expuestos se pone de manifiesto en el estudio de algunos teoremas fundamentales del Análisis y del Álgebra, en cuyas demostraciones se emplean esencialmente argumentos de compacidad.

En el capítulo 5 se analizan diversas nociones de conexión, las cuales se

aplican para la obtención de invariantes topológicos que permiten distinguir entre espacios métricos diferentes. Por otra parte, la conexión se utiliza también en la generalización de algunos resultados del Análisis Clásico como el teorema sobre el valor intermedio, que posee una gran importancia en las aplicaciones.

En el capítulo 6 se estudian dos casos particulares importantes de la teoría de los espacios métricos, que son los espacios normados y los espacios euclidianos. Se estudia la geometría de estos espacios y algunos nuevos conceptos que pueden ser introducidos gracias a la compatibilidad entre la estructura métrica y la estructura algebraica en dichos espacios.

Finalmente, en el capítulo 7, utilizando el concepto *operador lineal* introducido en los espacios normados, se generaliza la noción de derivada, obteniéndose así varias aplicaciones importantes, algunas de las cuales, en el caso de funciones numéricas, ya las conoce el lector de los cursos de Análisis Matemático.

Para la confección de este libro se han estudiado todos los libros de texto a nuestro alcance. Ninguno de ellos ha servido exactamente como modelo, pero se han utilizado siempre que coadyuven a lograr los objetivos didácticos. En cuanto a la estructura del libro, se ha seguido fundamentalmente la desarrollada en el libro *Topología General*, de los autores D.Hinrichsen, J.L.Fernández, A. Fragueta y A. Álvarez, el cual fue realizado con objetivos didácticos similares a los del presente libro.

El lector interesado en profundizar algún tema, hallará en las consideraciones generales, unas recomendaciones que le proporcionarán, además de algunas informaciones históricas sobre el desarrollo de los conceptos y resultados más fundamentales, una guía para el estudio posterior de temáticas afines y sus aplicaciones.

Para una información más detallada sobre la bibliografía existente en esta temática, el lector puede consultar el comentario bibliográfico expuesto en el libro *Topología General* citado anteriormente.

El libro contiene un material adecuado e importante para la impartición de cursos de pregrado y posgrado así como la superación individual de profesores de Matemáticas de la enseñanza media superior y universitaria.

Un agradecimiento especial para la Lic. Teresa Tello quien acometió la tarea de capturar el material del texto y a los estudiantes Enrique Rivera Robles y Eduardo Sánchez García quienes prepararon su versión final.

Finalmente deseo constatar mi agradecimiento a la Benemérita Universidad Autónoma de Puebla, que como alta casa de estudios y a través de su Dirección de Fomento Editorial asumió la tarea de sacar a la luz la segunda

edición corregida de este texto, cuya primera edición ya acotada fue publicada en Cuba por la Editorial Pueblo y Educación en el año 1985 por solicitud de la Dirección de Formación y Perfeccionamiento de personal Pedagógico del Ministerio de Educación de la República de Cuba.

En particular, aprovecho para expresar mi agradecimiento a la Facultad de Ciencias Físico Matemáticas y el Posgrado en Matemáticas de la Benemérita Autónoma de Puebla, a quienes me siento muy honrado de pertenecer en calidad de profesor investigador.

El autor

Puebla, Pue a 15 de junio del 2002

CONTENIDO

CONSIDERACIONES GENERALES

RELACIÓN DE SÍMBOLOS

INTRODUCCIÓN

CAPÍTULOS:

	Página
1. TEORÍA DE CONJUNTOS	35
1.1 <i>Concepto conjunto</i>	36
1.2 <i>Subconjuntos. Inclusión</i>	38
1.3 <i>Operaciones con conjuntos</i>	41
1.4 <i>Equivalencia de conjuntos. Potencia de conjuntos, Conjunto potencia +</i>	47
1.5 <i>Relación de equivalencia. Partición de un conjunto. Conjunto cociente</i>	59
1.6 <i>Relaciones funcionales. Funciones</i>	63
1.7 <i>Relaciones de orden. Conjuntos parcialmente, totalmente y bien ordenados</i>	71
1.8 <i>Axioma de elección. Teorema de Zermelo y otras proposiciones equivalentes a ellos. Inducción transfinita +</i>	77
1.9 <i>Estructuras matemáticas sobre los conjuntos +</i>	85
2. ESPACIOS MÉTRICOS	89
2.1 <i>Algunos conceptos de convergencia</i>	90
2.2 <i>Métricas. Ejemplos de espacios métricos</i>	97
2.3 <i>Subespacios. Isometrías</i>	108
2.4 <i>Seudométricas</i>	109
2.5 <i>Sucesiones convergentes en un espacio métrico</i>	118
2.6 <i>Funciones continuas</i>	124
2.7 <i>Bolas. Diámetro. Distancia entre conjuntos</i>	134
2.8 <i>Vecindades y abiertos</i>	143
2.9 <i>Puntos adherentes. Clausura de un conjunto. Conjuntos densos</i>	156
2.10 <i>Conjuntos cerrados</i>	175
2.11 <i>Punto interior, punto exterior y punto frontera</i>	178
2.12 <i>Funciones continuas, funciones abiertas y funciones cerradas</i>	183
2.13 <i>Separación de conjuntos y prolongación de funciones continuas +</i>	190
2.14 <i>Algunos conceptos de convergencia y continuidad no asociados a una métrica</i>	208

	Página
2.15 <i>Transición a una nueva teoría</i>	216
3. ESPACIOS MÉTRICOS COMPLETOS	221
3.1 <i>Sucesiones de Cauchy y funciones uniformemente continuas</i>	222
3.2 <i>Definición y propiedades generales de los espacios métricos completos</i>	227
3.3 <i>Algunos resultados importantes en los espacios métricos completos</i>	235
3.4 <i>Completamiento de un espacio métrico</i>	240
3.5 <i>Aplicaciones +</i>	242
4. COMPACIDAD EN ESPACIOS MÉTRICOS	253
4.1 <i>Espacios y subconjuntos compactos</i>	254
4.2 <i>Algunos conceptos relacionados con la noción de compacidad</i>	262
4.3 <i>Compactos y aplicaciones continuas</i>	268
4.4 <i>Compacidad en espacios producto y espacios funcionales</i>	280
4.5 <i>Espacios localmente compactos. Compactificaciones +</i>	301
5. CONEXIÓN EN ESPACIOS MÉTRICOS	319
5.1 <i>Espacios métricos conexos. Componentes conexas</i>	320
5.2 <i>Componentes conexas. Espacios localmente conexos</i>	337
5.3 <i>Espacios métricos conexos por arcos</i>	345
6. ESPACIOS NORMADOS Y EUCLIDEANOS	351
6.1 <i>Resultados auxiliares sobre espacios vectoriales</i>	352
6.2 <i>Métricas compatibles con la estructura algebraica de un espacio vectorial</i>	358
6.3 <i>Conexión, compacidad y completitud en espacios normados</i>	372
6.4 <i>Serie en espacios normados</i>	381
6.5 <i>Espacios euclidianos</i>	389
6.6 <i>Espacios de Hilbert</i>	408
6.7 <i>Operadores lineales y funcionales lineales</i>	430
7. DIFERENCIACIÓN EN ESPACIOS NORMADOS	467
7.1 <i>Diferenciación en sentido fuerte. Diferencial y derivada según Frechet</i>	468
7.2 <i>Relación entre derivación e integración abstracta</i>	477

	Página
7.3 Derivadas de orden superior y fórmula de Taylor	481
7.4 Diferenciación en sentido débil (diferencial según Gateaux)	
Relación entre las diferenciales según Frechet y Gateaux	487
7.5 Extremos locales de funcionales +	491
EJERCICIOS	505
Capítulo 1	505
Capítulo 2	507
Capítulo 3	512
Capítulo 4	515
Capítulo 5	518
Capítulo 6	519
Capítulo 7	526
BIBLIOGRAFÍA	529
ÍNDICE TEMÁTICO	531

CONSIDERACIONES GENERALES

Para el estudio de este libro se necesitaron esencialmente algunos conocimientos fundamentales del análisis, que se adquieren en los primeros dos años de la formación universitaria.

El lector necesita saber qué es una función continua (uniformemente continua, diferenciable $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$) y sería conveniente que conociera algunas nociones elementales métricas y topológicas en $\mathbb{R}^n$ (propiedades del módulo, subconjuntos abiertos y subconjuntos cerrados, sucesiones convergentes, puntos adherentes de sucesiones y puntos adherentes de conjuntos etc.), los teoremas de Heine-Borel, de Bolzano-Weierstrass y el valor medio, y necesita además, saber distinguir la convergencia puntual y la convergencia uniforme de una sucesión de funciones. Es importante que el lector estudie los problemas, ejemplos y ejercicios con el mismo interés que las definiciones, teoremas y demostraciones. Para facilitar esto, la parte teórica no ha sido separada de la parte práctica en el desarrollo del libro al contrario de como generalmente ocurre en un curso tradicional, sino el texto contiene varios ejemplos, ejercicios y problemas al final de cada epígrafe que sirven al lector de autocontrol, además de una colección de ejercicios de cada capítulo al final del libro.

Para poder trabajar productivamente con la teoría estudiada, es necesario disponer de ejemplos, así como de posibilidades de aplicación e interpretaciones intuitivas.

Insistimos en que, en el trabajo matemático, el saber formular los problemas y el saber aplicar los teoremas es tan o más importante que el saber demostrarlos. A pesar de que el libro puede servir como guía para los profesores durante la impartición de sus conferencias, este no ha sido concebido como un manual metodológico para el profesor, sino como libro de texto para el estudio individual y colectivo de los estudiantes. Esperamos que contribuya a una participación más activa de los estudiantes en las clases de esta asignatura.

En caso de que el libro sea utilizado en la impartición de cursos de posgrado o de seminarios extracurriculares para estudiantes, recomendamos intercalar en forma sistemática exposiciones de equipos sobre temas especiales. El libro contiene suficientes aplicaciones, ejercicios y material extra, idóneo para tales exposiciones.

Debe tenerse en cuenta que los temas impartidos en un curso, no pueden abarcar toda la temática expuesta en el libro por una limitación lógica de tiempo. No obstante, mediante el método propuesto, debe tratarse de trans-

mitir alguna información adicional a los participantes sobre lo que respecta a las aplicaciones y la transición a nuevas teorías.

Pensamos además que el profesor, contando con un libro de texto que puede sustituir sus conferencias como fuente de información, puede organizar el trabajo o parte de él en base a exposiciones de los alumnos, lo cual sería de gran utilidad para entrenamiento de un profesional en potencia que debe adquirir una formación no solamente científica sino también pedagógica. Podría adoptarse por ejemplo, el método siguiente:

Al comienzo del curso el profesor elabora junto con los estudiantes un plan de trabajo que determina la materia a discutir en cada reunión para las próximas semanas teniendo en cuenta el programa general de asignatura. Todos estudian individual o colectivamente las partes respectivas del libro antes de tratarlas en clase. Además, cada actividad la prepara especialmente un colectivo responsable cuya tarea es resumir los puntos principales del contenido, explicar su significación intuitiva, aclarar y elaborar preguntas, desarrollar ejercicios y buscar eventualmente ejemplos y aplicaciones adicionales en la literatura. Unos días antes de la clase, el profesor se reúne con el colectivo para discutir sobre la exposición preparada, corregir errores y aclarar dudas matemáticas y didácticas. Como todos los estudiantes deben haber analizado la materia anteriormente, pueden especificar las preguntas y dificultades individuales en la clase y participar activamente en la solución de los problemas planteados en la discusión de la materia. Es claro que este método no puede emplearse en todas las circunstancias. En particular presupone que el número de estudiantes del curso sea pequeño.

Deseamos hacer algunas aclaraciones con respecto a las ventajas que presenta este libro en relación con otros textos. En primer lugar, hemos sacrificado la reducción del texto ante la claridad de la exposición desarrollando un gran número de ejemplos, los cuales ayudan al estudiante a adaptarse a un razonamiento abstracto que no conocía anteriormente; además, el libro cuenta con una gran cantidad de ejercicios propuestos, así como problemas para futuros trabajos de curso e inclusive de diploma.

Además de estas ventajas, el libro presenta otras con respecto al texto *Elementos de la Teoría de Funciones y el Análisis Funcional* de A. N. Kolmogorov y S.V. Fomín y es que no depende de ningún concepto de Topología General, materia demasiado abstracta para ser impartida directamente en este nivel donde solo se espera que se hayan estudiado los temas del Análisis Matemático Clásico y posiblemente, la Variable Compleja. Esto amplía el campo de estudiantes para quienes puede ser interesante este texto.

En el texto de Kolmogorov se estudian las nociones de compacidad uti-

lizando los conceptos de la Topología General dados en el propio texto y lo mismo ocurre con la teoría de los espacios normados y euclidianos.

En este libro se introducen los conceptos topológicos necesarios para las Ecuaciones Diferenciales y el Análisis Funcional sin caer en conceptos abstractos para los cuales puedan no encontrarse aún maduros los estudiantes.

En nuestra exposición no es conveniente dar saltos cualitativos en la abstracción de los conceptos matemáticos hasta tanto los conocimientos básicos no hayan sido debidamente consolidados ya que se crean lagunas a veces insalvables en la formación del estudiante.

Con respecto al texto *Fundamentos del Análisis Moderno* de Dieudonné, nuestro texto presenta la ventaja de que la graduación de los ejercicios es más certera y eficaz, además de que en el libro de Dieudonné se introducen los conceptos topológicos, obviando casi totalmente la “intuición geométrica” y por ende, las representaciones gráficas. A veces abusa del método axiomático en aras de buscar una generalidad no necesaria para el nivel a que va dirigido, además de correrse el riesgo de que los estudiantes caigan en la memorización mecánica de los conceptos y teoremas. Los ejercicios que aparecen en él parten de un nivel de dominio del método axiomático y de abstracción que los estudiantes no poseen.

Los cursos básicos de Análisis Matemático (en una y varias variables), Álgebra (lineal y de estructuras) y Geometría, permiten al estudiante estar en condiciones de introducirse en algunas ramas más profundas de la matemática moderna y que exigen de él una cierta capacidad de abstracción. Una de esas ramas, quizás una de las más inmediatas a los cursos mencionados, la constituye la temática expuesta en este libro.

Después de haber cursado la materia de este texto, el lector habrá comprendido con mucha mayor profundidad la importancia y el alcance de los conocimientos adquiridos, además de que se habrá cerciorado de lo que nos puede brindar el razonamiento matemático abstracto y la generalización en Matemática. Sin embargo, hay otros temas, tan básicos como la materia estudiada en los primeros cursos y que no se deben descuidar para lograr una formación con bases sólidas. Tales son por ejemplo:

- la teoría de funciones analíticas;
- las ecuaciones diferenciales ordinarias;
- las ecuaciones en derivadas parciales;
- los problemas fundamentales de la física-matemática;
- los fundamentos de la medida e integración.

Creemos que con esta formación básica cualquier individuo está en condiciones de “elegir su rumbo” en la matemática.

Precisamente esta formación básica es la que se pretende lograr durante la estancia del alumno en las aulas universitarias. Sin embargo, ya aún desde este momento debe ir creándose el interés de trabajo e inclinación hacia determinados temas. Esta definición es útil aunque nuestro trabajo futuro no sea exactamente la investigación en matemática, sino la docencia o cualquier otro, ya que la profesión de matemático requiere un quehacer y una laboriosidad constante, la cual debe realizarse alrededor de un tema determinado. Sin la práctica constante, sin el interés en la producción científica, es imposible que logremos transmitir a nuestros educandos lo que se pretende que transmitamos. Es necesario conocer bien la profundidad e importancia de los conceptos y resultados para no transmitirlos mecánicamente. De ahí la importancia que tiene el estudiar con profundidad materias como la de este texto.

No debemos conformarnos con todo lo que hablamos anteriormente y que constituye solamente un paso inicial en nuestra formación profesional. No pretendemos tratar de hablar acerca de todas las posibles vías de continuar nuestra superación a partir de lo que hablamos anteriormente, sino solamente de algunas que son, en cierta manera, afines al propio trabajo del autor.

Una vez lograda la preparación inicial, y un tanto basándose en la temática expuesta en este texto, el estudiante se encuentra en condiciones (y esto debe ser explotado por el profesor que imparta la asignatura) de iniciar estudios más específicos sobre temas importantes como son:

- el análisis funcional;
- la teoría de aproximación;
- la teoría de optimización;
- la teoría de control;
- las ecuaciones en derivadas parciales y problemas de la física-matemática;
- la teoría de operadores;
- los métodos numéricos;

Con estas indicaciones estamos solamente proponiendo un método de promover el interés en la superación e investigación, en base a una temática propia como es la expuesta en el presente texto.

RELACIÓN DE SÍMBOLOS

Conjuntos, Espacios y Familias

$\emptyset$	Conjunto vacío
$\mathbb{N}$	Conjunto de los números naturales (incluyendo el cero)
$\mathbb{Z}$	Conjunto de los números enteros
$\mathbb{Q}$	Conjunto de los números racionales
$\mathbb{I}$	Conjunto de los números irracionales
$\mathbb{R}$	Conjunto de los números reales
$\mathbb{C}$	Conjunto de los números complejos
$\bar{\mathbb{R}}$	$\mathbb{R}$ ampliado ($\mathbb{R} \cup \{\pm\infty\}$)
$\mathbb{R}^n$	Espacio euclideo n – dimensional sobre $\mathbb{R}$
$\mathbb{C}^n$	Espacio euclideo n – dimensional sobre $\mathbb{C}$
E^*	Dual topológico de E
$l^p(\mathbb{N})$	Sucesiones tales que $\sum_{n=1}^{\infty} x_n ^p < \infty$
$l^{\infty}(\mathbb{N})$	Sucesiones acotadas
(X, d)	Espacio métrico
$C(X, Y)$	Funciones continuas de X en Y
$C_a(X, Y)$	Funciones continuas y acotadas de X en Y
$\prod_n (X_n, d_n)$	Espacio métrico producto
X/R	Espacio cociente
$[x]$	Clase del elemento x
$F(X, Y) = Y^X$	Funciones de X en Y
$[M]$	Espacio vectorial generado por M
U	Conjunto Universo
$P(X)$	Colección de los subconjuntos de X
$]a, b[$	Intervalo abierto en $\mathbb{R}$
$[a, b]$	Intervalo cerrado en $\mathbb{R}$
$[a, b[$	Intervalo cerrado abierto en $\mathbb{R}$
$]a, b]$	Intervalo abierto cerrado en $\mathbb{R}$
$B(a, r)$	Bola abierta de centro a y radio r
$B'(a, r)$	Bola cerrada de centro a y radio r
$E(a, r)$	Esfera de centro a y radio r
Δ	Diagonal de $X \times X$

C	Conjunto de Cantor
$\mathbb{R}_+$	Semieje real positivo
$\mathbb{R}_-$	Semieje real negativo
$\text{sop } f$	Soporte de f
$A + B$	Suma algebraica
$\mathbf{V}(x)$	Vecindades de x
$\mathbf{V}_A(x)$	Vecindades de x en el subespacio (A, d_A)
$\mathbf{V}(A)$	Vecindades del conjunto A
$\mathring{\mathbf{V}}(x)$	Vecindades abiertas de x
$\mathbf{F}(x)$	Sistema fundamental de vecindades de x
$L_1[a, b]$	Completamiento de $(C([a, b], \mathbb{R}), \ \cdot\ _1)$
$L_2[a, b]$	Completamiento de $(C([a, b], \mathbb{R}), \ \cdot\ _2)$
$C_{2\pi}(\mathbb{R}, X)$	Funciones continuas y de periodo 2π de $\mathbb{R}$ en X
$C([a, b], \mathbb{R})$	Funciones reales continuas sobre $[a, b]$
θ_d	Colección de abiertos en (X, d)
σ_d	Colección de cerrados en (X, d)
$P[a, b]$	Colección de las particiones finitas de $[a, b]$
$M^\perp$	Ortogonal de M en H
$(E, \ \cdot\)$	Espacio normado
$(H, \langle, \rangle)$	Espacio euclideo
$D(E, F)$	Espacio de las aplicaciones Frechet diferenciables de E en F
$L(E, F)$	Operadores lineales continuos de E en F
$M \oplus N$	Suma directa de M y N

Operaciones en Espacios y Conjuntos

$x \in X$	x es elemento de X
$x \notin X$	x no es elemento de X
$A \subset B$	Inclusión
$A \cup B$	Unión
$A \cap B$	Intersección
$A \setminus B$	Diferencia
$\cup F$	Unión de colecciones
$\cap F$	Intersección de colecciones
$\bigcup_{\alpha \in I} X_\alpha$	Unión de familias
$\bigcap_{\alpha \in I} X_\alpha$	Intersección de familias
$\prod_{\alpha \in I} X_\alpha$	Conjunto producto

$\leq$	Relación de orden
$<$	Orden estricto
$\sup . A$	Supremo de A
$\inf . A$	Infimo de A
G_f	Grafo de f

Funciones y Aplicaciones

f, g, h	Notaciones usuales para funciones
$f : X \longrightarrow Y, X \xrightarrow{f} Y$	Función de X en Y
$x \rightsquigarrow f(x)$	Función
$f \mid A$	Restricción de f a A
$f(x)$	Imagen de x por f
$f[A]$	Imagen del conjunto A por f
$f\{F\}$	Imagen de la colección F por f
$f^{-1}[A]$	Preimagen del conjunto A por f
$f^{-1}\{F\}$	Preimagen de la colección F por f
$f \circ g$	Función compuesta
$f + g$	Suma de funciones
1_X	Función identidad de X sobre X
i_A	Inyección de A en $X, A \subset X$
$\mathbb{X}_A$	Función indicadora de A
$t \rightsquigarrow \text{sign } t$	Función signo
$\langle a \rangle$	Función constante
$(f_\alpha)_{\alpha \in I}$	Familia de funciones
$(x_\alpha)_{\alpha \in I}$	Familia de puntos
(x_n)	Sucesión
$(x_{\varphi(n)})$	Subsucesión
π_i, π_x, π_α	Proyecciones
φ_R	aplicación canónica de X en X/R
P	Proyector ortogonal
$\langle, \rangle$	Producto escalar
$Df(x_0; h)$	Diferencial según Gateaux de f
$f'(x_0), f'(x_0)(h)$	Derivada y diferencial según Frechet
$d^n f(x_0)$	Derivada n – ésima de f
d, ρ	Notación usual para métricas
$\ \cdot\ , \cdot $	Notación usual para normas
$d_\infty, \ \cdot\ _\infty$	Métrica y norma de la convergencia uniforme
$d_1, \ \cdot\ _1$	Métrica y norma de la convergencia en media
$d_2, \ \cdot\ _2$	Métrica y norma de la convergencia en media cuadrática.

Operaciones Métricas y Topológicas

$\delta(A)$	Diámetro de A
$\text{int } A, \overset{\circ}{A}$	Interior de A
$\text{ext } A$	Exterior de A
$\overline{A}$	Clausura de A
A'	Conjunto derivado de A
$F_r A, \delta A$	Frontera de A
$\overset{\circ}{A}_B$	Interior de A en el subespacio (B, d_B)
$\overline{A}_B$	Clausura de A en el subespacio (B, d_B)
$(x_n) \xrightarrow{d} x$	Convergencia de sucesiones en un espacio métrico (se omite d cuando no hay confusión)

Para poder referirnos dentro del texto a proposiciones, definiciones, teoremas, etcétera, los hemos enumerado por una triada de números, de los cuales, el primero indica el capítulo, el segundo el epígrafe y el tercero, el orden que ocupa dentro del epígrafe. Los ejemplos han sido numerados de forma corrida dentro de cada capítulo.

Además, hemos utilizado los siguientes símbolos gráficos con vistas a estructurar el texto:

+ Para indicar epígrafes que pueden omitirse en una primera lectura o que no tienen que incluirse necesariamente en el desarrollo de las conferencias (aparece solamente en el índice).

* Para indicar ejercicios difíciles.

■ Para indicar el final de una demostración.

No pretendemos en este texto dedicarnos al estudio de la lógica de conjuntos ni al cálculo proposicional. No obstante, aprovechamos para informar al lector sobre la siguiente simbología que utilizaremos a veces en lugar de frases más largas del lenguaje:

$\stackrel{\sim}{\leftarrow}$	Esta definido por (declara una definición)
$:=$	Lo emplearemos cuando expresamos una definición mediante una igualdad
$\Rightarrow$	Denota la implicación lógica
$\Leftrightarrow$	Denota la equivalencia lógica
$\vee$	Denota la disyunción
$\wedge$	Denota la conjunción
$\exists x \in X$	Denota el cuantificador existencial
$\forall x \in X$	Denota el cuantificador universal

INTRODUCCIÓN

Hasta el momento, el estudiante que se enfrenta por primera vez a esta asignatura, no ha tenido necesidad de chocar con un nivel de abstracción matemática que exija de él un esfuerzo demasiado grande ni un cambio en su esquema de razonamiento que implique la formación de un verdadero hábito de pensar y razonar dentro de una teoría matemática. Desde este punto de vista, la importancia de esta asignatura radica en que trata de lograr cierto nivel de abstracción en el estudiante, que anteriormente no había necesitado, pero con ello se corre el riesgo de que el estudiante en lugar de ampliar su forma de razonamiento, al contrario, la esquematice.

Si bien el grado de abstracción que requiere la temática de este libro hemos tratado de reducirla a su mínima expresión en aras de la claridad en la exposición, no obstante el estudiante se encuentra por primera vez frente a una asignatura de carácter diferente a las vistas anteriormente y que posiblemente, quizás de manera inconsciente le produzca un efecto más profundo que el que se manifestó al pasar de la aritmética elemental y el álgebra elemental al cálculo diferencial e integral. Por todo esto resulta conveniente hacer cierta aclaración, si se quiere de carácter filosófico, sobre la importancia y como surge el proceso de abstracción en la Matemática.

Al estudiar cualquier ciencia incluso la matemática, se debe aclarar ante todo los objetos que en ella se consideran, qué corresponde a estos objetos en el mundo real en que nos desenvolvemos, cuáles principios y métodos fundamentan a esta ciencia y que lugar específico ella ocupa en el sistema del resto de las ciencias.

En lo fundamental la respuesta a estas preguntas en relación con la matemática fue dada por Federico Engels en 1877 cuando expresaba que la matemática pura tiene como objeto de estudio las formas espaciales y las relaciones cuantitativas del mundo real, por consiguiente es muy real y material. El hecho, que lo material adquiriera una forma extraordinariamente abstracta, puede solamente apagar débilmente su procedencia del mundo exterior. Pero para estar en condiciones de investigar estas formas y relaciones en forma pura, es necesario separarlas completamente de su contenido y dejar esto último a un lado como algo diferente. Al igual que el resto de las ciencias, la matemática apareció de las exigencias prácticas de los hombres; de la medición de las áreas de las parcelas de tierra y de la capacidad de las vasijas, de la estimación del tiempo y de la mecánica.

Según Engels, la matemática como ciencia apareció debido a las exigencias prácticas de los hombres y no es un producto de su raciocinio puro. Los objetos de estudio de la matemática son las formas de los objetos que existen en la realidad y sus relaciones cuantitativas; pero los conceptos de forma, cantidad y relaciones cuantitativas no son más que abstracciones; por consiguiente, los objetos de estudio de la matemática son en esencia las *abstracciones*, y el método fundamental que constituye la base de la matemática como ciencia es el *método de la abstracción*. La particularidad fundamental de la matemática consiste en que ella separa especialmente las relaciones cuantitativas y las formas espaciales, las cuales están presentes en todos los objetos y fenómenos sin excepción y hace de ellos objetos de su investigación.

Los historiadores de la matemática distinguen cuatro etapas fundamentales en su desarrollo como ciencia. La primera comenzó desde los tiempos más remotos y continuó hasta los siglos VI-VII antes de nuestra era. Los conocimientos matemáticos de esta época estaban relacionados solamente con las exigencias domésticas de la vida del hombre. La exigencia de determinar la cantidad de ganado, la cantidad de cosechas recogidas y su distribución, etcétera, dió lugar a la necesidad de contar y de generar las fracciones. La exigencia de medir las áreas de las parcelas de tierra, los volúmenes de diferentes vasijas, las piezas para la construcción etcétera, condujo a la acumulación del material práctico que sentó las bases de la más simple geometría. Sin embargo, ya en este periodo aparecen problemas en la antigua Babilonia y en China que no tienen un significado doméstico, es decir, comienza el trabajo en cuanto a la sistematización e interpretación teórica del material real hasta el momento acumulado. No obstante, la matemática de este periodo aún no es una ciencia teórica y abstracta. *Esta etapa corresponde al proceso inicial de formación de la matemática como ciencia.*

La segunda etapa corresponde al auge del desarrollo de la matemática como ciencia teórica y abstracta. Si bien en la primera etapa de su desarrollo la matemática se planteaba como pregunta fundamental *¿cómo?*, ahora en la segunda etapa la pregunta fundamental es *¿por qué?*. Durante este período tanto la historia de la Grecia antigua como del Oriente, se caracterizaron por el enriquecimiento de ciertas clases y el empobrecimiento de otras. El trabajo físico se convierte en esta época en un atributo solamente de los esclavos y artesanos, mientras que la aristocracia buscaba la tranquilidad a sus preocupaciones dedicándose a la filosofía y la ética del individuo. Esta atmósfera fué favorable para la discusión de los fundamentos de la matemática. Los griegos antiguos, sistematizando y generalizando los métodos de solución de problemas aritméticos, fundaron la Aritmética: la ciencia de los números y las

operaciones con ellos. Un alto grado de perfección lógica alcanzó también la Geometría griega, en cuya construcción fue aplicado por primera vez el método axiomático. Simultáneamente la matemática del Oriente preparaba los fundamentos del Algebra. Esta segunda etapa del desarrollo de la Matemática se denomina con frecuencia en la literatura histórico-matemática: *etapa de las magnitudes constantes*.

La tercera etapa en el desarrollo de la matemática está relacionada con la extensión del concepto *relaciones cuantitativas*. Aquí el estímulo fundamental para este desarrollo fue ocasionado por la exigencia de la técnica, de los problemas ingenieriles, de la navegación y del arte militar. La matemática de los siglos XVII al comienzo del siglo XIX, del estudio de las magnitudes constantes, pasa a investigar la dependencia entre las magnitudes variables, es decir, la representación matemática de los procesos. Este periodo del desarrollo se llama *etapa de las magnitudes variables*. En este momento la matemática ya no se restringe al estudio de los números, las magnitudes abstractas y las figuras geométricas: en ella comienzan a fijarse las ideas de continuidad, de movimiento y de cambio. En primer plano avanza el concepto *función*. Según las palabras de Engels, el punto de viraje en la matemática de este tiempo fue el concepto de magnitud variable de Descartes. Gracias a él en la matemática entró el movimiento y por ello la dialéctica.

El paso decisivo en el desarrollo de la matemática de las magnitudes variables fue la fundación del cálculo diferencial e integral, relacionada con los nombres de I. Newton (1642-1727), matemático y filósofo inglés, y G.V. Leibnitz (1646-1716), matemático y filósofo alemán, con lo cual la ciencia obtuvo un potente aparato para la investigación cuantitativa de los procesos. Gracias a esto creció extraordinariamente la aplicación de la matemática en las ciencias naturales. En este tiempo según la definición del enciclopedista francés del siglo XVIII D'Alembert, la matemática era aún una ciencia que estudiaba las propiedades de las magnitudes ya que ellas son enumerables y medibles. Pero ya en aquel entonces matemáticos geniales como R. Descartes y especialmente G.V. Leibnitz afirmaron que la matemática podía aplicarse no solamente a las formas espaciales del mundo en que vivimos al igual que a las magnitudes y sus relaciones cuantitativas, sino también a otros objetos, por ejemplo al razonamiento.

El comienzo de la cuarta etapa del desarrollo de la matemática (que continúa hasta nuestros días), está relacionado con el descubrimiento de N.I. Lobachevsky (1742-1856), destacado matemático ruso, de una geometría no euclidea y con los trabajos de E. Galois (1811-1832), matemático francés, en la teoría de grupos y cuerpos, los cuales dieron lugar a una verdadera re-

volución en la matemática. Este nuevo salto cualitativo en la matemática no fue casual; este fue condicionado por el proceso de la primera revolución industrial y por los acontecimientos políticos en Europa Occidental, es decir, la victoria de la burguesía sobre el feudalismo y el absolutismo. Y tampoco fue casual que el desarrollo exitoso de la matemática se llevara a cabo precisamente en aquellos países (primeramente Francia, y más tarde Alemania y Rusia), donde la exigencia del desarrollo tecnológico y económico social era más sensible y la ruptura ideológica con el pasado se presentaba más claramente. Precisamente en estas condiciones fue frustrada la autoridad de la hasta ese entonces bimilenaria geometría de Euclides. Aparecieron diferentes interpretaciones de la axiomática de Euclides y se fundaron nuevos sistemas de axiomas de la geometría. La generalización de la geometría se llevó a cabo también en cuanto el crecimiento del número de dimensiones del espacio. Los nuevos espacios multidimensionales hallaron rápidamente su aplicación en variados problemas de la física, la química, la economía y la sociología.

Esta nueva etapa del desarrollo también influyó cualitativamente en el álgebra. En el centro de la atención del álgebra aparecieron problemas, no solamente relacionados con la solución de ecuaciones, sino problemas de estudio de diferentes operaciones algebraicas, definidas en conjunto de naturaleza arbitraria. Si bien antes las operaciones algebraicas estaban relacionadas solamente con las cantidades, ahora las operaciones se realizan sobre objetos de diferente tipo y sus relaciones. Este enfoque abstracto de los objetos y las operaciones sobre ellos, halló su expresión más completa en la teoría de conjuntos y en el método axiomático que está íntimamente relacionado con ella. Precisamente la teoría de conjuntos brindó un nuevo método universal que rápidamente abarcó toda la matemática. En el libro de N. Bourbaki sobre Teoría de Conjuntos se dice: “Hoy en día sabemos, lógicamente hablando, que es posible deducir casi toda la matemática moderna de una fuente única: la teoría de conjuntos”. Un papel fundamental juega la teoría de conjuntos en las investigaciones relacionadas con la fundamentación lógica y filosófica de la matemática.

La cuarta etapa del desarrollo de la matemática nos muestra con claridad que el punto de vista de que la matemática es la ciencia que trata sobre las formas espaciales y las relaciones cuantitativas de las cantidades, no puede ya considerarse suficiente; esta definición de la matemática con frecuencia fue sometida a crítica, cuya fundamentación se basa en la comprensión restringida y estrecha de los conceptos de “forma”, “cantidad” y “relaciones cuantitativas”. La historia de la matemática ha demostrado de forma convincente, como poco a poco, en la medida en que se ha ido desarrollando la práctica socio-

material de la ciencia, se ha ampliado el campo de las relaciones cuantitativas y de las formas estudiadas por la matemática. El objetivo de la matemática, desde la época en que fue definido por Engels, ha enriquecido considerablemente su contenido, colocándose en un nuevo y más alto nivel de abstracción. Por ello la definición moderna de la matemática se formula aproximadamente de la manera siguiente: *los objetos de estudio de la Matemática son las formas espaciales y las relaciones cuantitativas del mundo real, así como el resto de las relaciones y formas semejantes a ellas.*

La solución general al problema sobre el surgimiento, rol y lugar de la abstracción en el proceso de conocimiento de la naturaleza fue dado por V.I. Lenin cuando expresaba que de la contemplación viva al pensamiento abstracto y de él a la práctica era el camino dialéctico para el conocimiento de la verdad, para el conocimiento de la realidad objetiva. Y además que el conocimiento es el reflejo de la naturaleza por el hombre. Pero esto no es algo simple, ni inmediato, ni un reflejo integro, sino el proceso de toda una serie de abstracciones, de formación y organización de los conceptos, leyes, etc., y la determinación y organización de los conceptos y leyes abarcan convencionalmente las leyes de la naturaleza en eterno movimiento y desarrollo. El hombre no puede abarcar, reflejar, representar la naturaleza del todo completamente ni su “espontánea integridad”, él puede tan solo aproximarse eternamente a esto, fundando abstracciones, conceptos, leyes, imágenes científicas del mundo, etcétera.

En estas palabras se encierra la esencia del materialismo dialéctico sobre las abstracciones en general y las matemáticas en particular. La abstracción es un eslabón importante y necesario y a su vez es una parte del conocimiento que refleja la naturaleza solamente en una eterna aproximación a su conocimiento total. Finalmente, la veracidad del reflejo de la naturaleza en las abstracciones la confirma la práctica.

Como resultado del proceso de abstracción aparecen conceptos, categorías y leyes que reflejan en la conciencia los aspectos esenciales de la realidad objetiva en forma general, K. Marx escribió que toda abstracción si es sensata, refleja lo general, lo fija y nos libera de una gran masa de repeticiones.

En la matemática se distinguen cuatro tipos fundamentales de abstracción aunque no pasaremos a detallarlas.

La abstracción de los objetos de la Matemática y de los métodos para su investigación determinan su carácter específico y el lugar que ocupa entre las otras ciencias.

Es claro que la matemática no es una ciencia social, y sin embargo, no pertenece a las ciencias naturales. Por una parte se diferencia de ellas por sus

objetos de estudio que son mucho más amplios que los objetos de las ciencias naturales.

Los objetos de trabajo, es decir, las relaciones espaciales y cuantitativas se encuentran también en las ciencias naturales y sociales. Por otra parte, los métodos de la matemática se diferencian radicalmente de los métodos de las ciencias naturales. El método fundamental en las ciencias naturales es la observación y el experimento. La matemática aunque parte de la observación, esto ocurre solo al inicio y aunque retorna al experimento, este se lleva a cabo una vez terminada la investigación con el objetivo de confirmar sus deducciones en la práctica y en otras ciencias. Además, a pesar de que el desarrollo de la matemática ha estado determinado en términos generales por las exigencias del desarrollo social, las investigaciones de la matemática se adelantan a estas exigencias y encuentran su aplicación en ocasiones mucho más tarde. Por ejemplo, la teoría de las secciones cónicas fue elaborada por Apolonio (Matemático de la Grecia Antigua del siglo III a.n.e), y su aplicación práctica fue hallada solamente en el siglo XVII, cuando fueron establecidas por Newton y Kepler las órbitas del movimiento de los cuerpos celestes; las geometrías no euclideanas fueron inventadas a mediados del siglo XIX y encontraron su aplicación en la física y la cosmología solamente en nuestros días; la teoría de grupos fue elaborada a inicio del siglo XX, pero prácticamente su aplicación comienza a efectuarse en la actualidad en la física de las partículas elementales. De todo esto se deduce, que la matemática no se puede adjuntar a las ciencias naturales ni por su objeto de estudio, ni por su método de trabajo. No obstante, ella está muy cerca de las ciencias naturales ya que ambas tienen un surgimiento común, poseen un importante rasgo común que es que la aplicabilidad y el desarrollo de una ha estado estimulado en gran medida por la necesidad de las otras.

En la medida en que ha ido creciendo la abstracción de los objetos matemáticos y métodos, se han ido complicando las relaciones de la matemática con la realidad objetiva, con la práctica, y por consiguiente, se plantea el problema de la autenticidad de los conocimientos matemáticos. K Marx escribió que con la práctica el hombre debe demostrar la veracidad de su pensamiento. Sin embargo, el científico soviético A.D. Alexandrof, nos indica que el problema del contenido material y de las aplicaciones prácticas de los métodos matemáticos no se puede entender de forma simple. En primer lugar, toda abstracción refleja solamente ciertas partes de la realidad. Lo general y lo abstracto existen solamente en lo particular y en lo concreto. La práctica tiene que ver siempre con objetos y fenómenos concretos y por eso en esencia ella no puede dar una confirmación absoluta de los teoremas y leyes generales. Como indicó V.I.

Lenin, no se debe olvidar, que el criterio de la práctica nunca puede en realidad confirmar o refutar completamente cualquier tipo de representación de la mente humana. Este criterio es también tan *indeterminado* para impedir que los conocimientos del hombre se conviertan en *absolutos*. En segundo lugar, los conceptos y deducciones matemáticas no son ni definitivas, ni verdades absolutamente exactas. Al contrario, ellas se desarrollan constantemente y en este desarrollo influyen todos los aspectos del desarrollo de la humanidad.

En tercer lugar, el propio carácter de comprobación en la práctica, que tienen los conocimientos matemáticos, no se mantiene invariable. Si bien en los primeros pasos del desarrollo de la humanidad la matemática estuvo estrechamente relacionada con la producción, en lo sucesivo esta relación se ha ido haciendo cada vez más completa. En la actualidad los resultados de la matemática se comprueban ante todo en las ciencias naturales y las disciplinas técnicas. En cuarto lugar no se puede exigir que cada teorema tenga una aplicación práctica inmediata. La teoría en general se verifica y se aplica como un todo y cada teorema específico se fundamenta en el sistema de aquella teoría de la cual él forma parte. Finalmente, debemos señalar que el desarrollo de la matemática no se puede reducir solamente a satisfacer las demandas de la práctica, la producción y las otras ciencias. La matemática se desarrolla independiente, al igual que el resto de las ciencias bajo la influencia de sus propias exigencias.

Este desarrollo relativamente independiente de la matemática ocurre según las leyes que determinan su objeto y métodos de trabajo.

Sin embargo, la veracidad de los resultados de este desarrollo, en fin de cuentas, aunque en ocasiones con gran retraso, se confirma siempre en la práctica socio-material de los hombres.

CAPÍTULO 1

TEORÍA DE CONJUNTOS

En este capítulo se estudian los conceptos básicos de la Teoría de conjuntos que en alguna medida serán utilizados en el resto del texto. Sin embargo, a pesar de su carácter auxiliar, la exposición se desarrolla de manera tal que pueda utilizarse para un estudio autocontenido de los fundamentos de la Teoría de Conjuntos.

Partiendo de ejemplos concretos se introduce en 1.1 el concepto primario de conjunto. En 1.2 se estudia la relación natural de inclusión entre conjuntos, lo cual nos conduce a otro concepto tan primario como el de conjunto que es la noción de pertenencia. El tema 1.3 se dedica al estudio de las operaciones con conjuntos estableciéndose un paralelo entre las operaciones del álgebra ordinaria numérica y las aquí introducidas del álgebra de conjuntos.

En 1.4 se estudia un tipo de equivalencia entre conjuntos introduciéndose el concepto de potencia que permite generalizar la noción de “cantidad de elementos de un conjunto” a los conjuntos infinitos.

En 1.5 se expone un importante método de la teoría de conjuntos para construir nuevos conjuntos a partir de uno dado, introduciendo el concepto de relación de equivalencia.

El tema 1.6 se dedica al estudio de un tipo particular de relaciones, estrechamente ligadas al desarrollo del Análisis, y que son las relaciones funcionales, también conocidas por funciones.

En 1.7 se introducen las relaciones de orden que permiten definir clases importantes de conjuntos, como son los conjuntos parcialmente ordenados, los conjuntos totalmente ordenados y los conjuntos bien ordenados.

Existen toda una serie de teoremas básicos en los fundamentos de la matemática, que tienen una gran influencia en muchas de sus ramas y que están estrechamente vinculados a las clases de conjuntos enunciadas anteriormente. Tal es, por ejemplo, el caso del lema de Zorn o axioma de elección que es estudiado en 1.8 junto con otras proposiciones equivalentes.

Finalmente, en 1.9 se discute la noción de estructura matemática sobre un conjunto con el objetivo de introducir las estructuras métricas y topológicas sobre conjuntos, lo cual constituye el objetivo de una gran parte del presente

texto.

1. 1 CONCEPTO CONJUNTO

La primera y más natural de las preguntas a que debíamos poder dar respuesta es: ¿qué es un conjunto? Sin embargo, hasta nuestros días, no es posible dar una definición rigurosa de conjunto.

Para definir cualquier concepto, es necesario contar con conceptos más simples, mediante los cuales se pueda enunciar dicha definición. Pero estos conceptos mas simples necesitan ser definidos por medio de otros aún más sencillos, quienes a su vez deben estar definidos con ayuda de otros todavía más simples. Siguiendo este proceso y suponiendo que constara de un número finito de pasos, llegaríamos a un concepto inicial, primario, el cual no se podría definir auxiliandonos de otros más sencillos; entonces, todo lo que se puede hacer es explicar, mediante una serie de ejemplos, el sentido de este concepto primario y llegar a un acuerdo de lo que vamos a entender por dicho concepto. Conjunto es precisamente este concepto primario simple; por ello, nos remitiremos a ejemplos que nos aclaren el sentido de este concepto. Por conjunto entenderemos la totalidad, reunión o colección de cualesquiera objetos, los cuales han sido agrupados teniendo en cuenta algún tipo de característica común. Por ejemplo el conjunto de todas las consonantes en el alfabeto, el conjunto de todas las escuelas primarias de la Ciudad de Puebla, el conjunto de todos los puntos de una circunferencia, el conjunto de todas las teorías sobre el surgimiento de la tierra, etc.

El fundador de la teoría de conjuntos G. Kantor escribió: “Por conjunto, yo entiendo todo aquello que puede ser pensado como una unidad, toda agrupación de determinados elementos que pueda ser relacionada formando un todo con la ayuda de una cierta ley”.

Lo esencial en el concepto *conjunto* es el acto de unir diferentes elementos en un todo. Existe una gran libertad en relación con aquellos objetos que pueden formar parte de un conjunto. Esos objetos pueden ser: números, ciudades, ideas, funciones, expresiones gramaticales, etc. en fin, cualesquiera objetos del mundo real, de nuestra intuición y de nuestro intelecto.

Lo enunciado por Kantor nos permite decir que un conjunto esta determinado completamente por sus elementos o por leyes, propiedades características, etc., de acuerdo con las cuales ocurre el acto de unión de diferentes objetos en un todo. Por eso se puede plantear que el concepto fundamental de la teoría de conjuntos es, no tanto el propio concepto *conjunto*, como la relación de pertenencia.

Con frecuencia los conjuntos se denotan por letras mayúsculas del alfabeto

latino. En particular a algunos conjuntos numéricos importantes los denotaremos a lo largo del texto por símbolos fijos:

$\mathbb{N}$: Conjunto de los números naturales (incluyendo al cero)

$\mathbb{Z}$: Conjunto de los números enteros

$\mathbb{Q}$: Conjunto de los números racionales

$\mathbb{I}$: Conjunto de los números irracionales

$\mathbb{R}$: Conjunto de los números reales

$\mathbb{C}$: Conjunto de los números complejos

Si A es un subconjunto numérico del conjunto de los números reales, denotaremos por A_+ , (A_-) al subconjunto de A formado por todos los números positivos (negativos) incluyendo al cero, contenidos en A . Por A^* designaremos al conjunto de todos los elementos de A excepto el cero, en el caso en que $0 \in A$. Los objetos que componen a un conjunto se llaman *elementos* y se representan mediante letras minúsculas del alfabeto latino. Para indicar que un conjunto A es un conjunto de elementos x , se escribe $A = \{x\}$. Las llaves indican precisamente el acto de unión de los elementos x en un todo único. Si resulta importante indicar la propiedad característica que une a los objetos en el conjunto, se utiliza la siguiente escritura: $A = \{x : \dots\}$, donde en lugar de la línea punteada se escribe la propiedad que los caracteriza.

Ejemplo 1. Se escribe: $B = \{x : x \text{ es un número real tal que } x^2 - 1 = 0\}$. Se lee: B es el conjunto de todas las raíces reales de la ecuación $x^2 - 1 = 0$.

Sería imposible tratar de estudiar todos y cada uno de los conjuntos por separado sin clasificarlos. Ante todo convendremos en diferenciar los conjuntos en conjuntos *finitos* y conjuntos *infinitos*.

Llamaremos *conjunto finito* a aquel cuya cantidad de elementos puede ser expresada por una cierta cantidad; no es importante si ese número es conocido o no, lo importante es que exista.

Ejemplo 2. Son conjuntos finitos:

el conjunto de las manos de un hombre;

el conjunto de todas las letras de esta página;

el conjunto de las letras de todo el libro;

el conjunto de las letras utilizadas en todas las páginas de la totalidad de los libros existentes hasta el momento;

el conjunto de los granos de arena en la playa de Cancún;

el conjunto de los átomos que constituyen el sistema solar.

A los conjuntos finitos pertenece el conjunto que no contiene ningún elemento, a este lo denominamos conjunto *nulo* o *vacío* y se denota por la letra griega ϕ . La necesidad de introducir el conjunto vacío se debe a que al definir un conjunto mediante cierta propiedad, no siempre se puede asegurar de antemano que contenga al menos un elemento. Por ejemplo, el conjunto de los hombres en la habitación contigua en un instante determinado puede ser vacío o no. Se puede hablar del conjunto de las raíces reales de la ecuación $x^2 - 2x + 10 = 0$, y sin embargo, esta ecuación no tiene raíces reales.

Los conjuntos *infinitos* son todos aquellos que no son finitos.

Ejemplo 3. Son conjuntos infinitos:

- el conjunto de todos los números enteros;
- el conjunto de los puntos de una circunferencia;
- el conjunto de los puntos del plano.

Al describir el concepto conjunto, planteamos que un conjunto está determinado por sus propios elementos o por sus propiedades características. Se debe señalar, además, que los conjuntos finitos pueden ser dados por una simple enumeración de sus elementos (por ejemplo, el conjunto B de todas las asignaturas de matemática que se imparten en los dos primeros años de la Licenciatura en Matemática: $B = \{\text{Análisis matemático, Álgebra, Geometría}\}$) mientras que los conjuntos infinitos se pueden representar solamente indicando las propiedades características que son comunes a todos sus elementos.

1.2 SUBCONJUNTOS. INCLUSIÓN

Comencemos por introducir algunos conceptos y notaciones fundamentales. Para indicar que x es un elemento del conjunto A , escribiremos $x \in A$ (se lee “ x pertenece a A ”; $\in$ es el símbolo de pertenencia). Por ejemplo, si x es un tractor y A es el conjunto de todos los equipos agrícolas, entonces $x \in A$. Para indicar que el elemento x no pertenece al conjunto A se escribe $x \notin A$.

Por ejemplo, si x es la luna y A el conjunto de todos los equipos agrícolas, entonces $x \notin A$.

Si los conjuntos A y B coinciden es decir si tienen los mismos elementos, entonces se escribe $A = B$.

Ejemplo 4. Si B es el conjunto de todos los alumnos en un aula, y A es el conjunto de todos los estudiantes del sexo masculino en dicha aula, entonces no se puede decir que A es igual a B . En cambio, si sabemos que todos los estudiantes del aula son varones, se tendrá que $A = B$.

Si cada elemento del conjunto B es además elemento del conjunto A , entonces se dice, que el conjunto B es una parte o un *subconjunto* del conjunto A y se escribe $B \subset A$ ($\subset$ es el símbolo de inclusión).

Ejemplo 5. Si B es el conjunto de todos los estudiantes varones en un aula y A es el conjunto de todas las personas que se encuentran en esa aula, entonces $B \subset A$. En aquellos casos, donde se presupone la posible coincidencia de los conjuntos A y B , a veces se escribe $B \subseteq A$.

Observemos que entre los símbolos $\in, =, \subset, \subseteq$ existe una diferencia sustancial. Si bien la relación $x \in A$ significa el hecho de pertenencia del elemento x al conjunto A y por ello, representa el concepto primario que corresponde a la noción de conjunto; las relaciones $A = B$, $A \subset B$ ($A \subseteq B$) expresan un concepto relativo a una parte del conjunto y son introducidas por definición a partir de la noción más simple de pertenencia. De acuerdo con la definición dada, cada conjunto A es subconjunto de sí mismo: $A \subseteq A$. Además, el conjunto vacío es subconjunto de cualquier conjunto, C . En efecto, si C es un conjunto arbitrario, entonces $\phi \subset C$, ya que en caso contrario, al conjunto ϕ pertenecería al menos un elemento que no pertenece a C , pero en el conjunto ϕ no existe un tal elemento por la propia definición de conjunto vacío.

El conjunto A y el conjunto vacío ϕ se llaman subconjuntos *impropios* de A . Todos los restantes subconjuntos de A se llaman *propios*.

Ejemplo 6. El conjunto de todos los números pares es un subconjunto propio del conjunto de todos los números enteros; el conjunto de todas las mujeres de nuestro planeta es un subconjunto propio del conjunto de todos sus habitantes.

Con frecuencia, todos los conjuntos con los que tenemos que trabajar en nuestros razonamientos son subconjuntos de un cierto conjunto U ; en tal caso, a este conjunto U se le llama conjunto universo o conjunto universal. Por ejemplo, para la aritmética, cuyo objeto de estudio son los números racionales no negativos (es decir, todas las fracciones racionales positivas y el número cero), este conjunto será el universo. Para el análisis matemático, el conjunto universal es el conjunto de todas las funciones numéricas de argumento real. Para la geometría euclidiana, su conjunto universal es el conjunto de todos los subconjuntos de puntos del espacio, ya que cualquier figura geométrica es un conjunto de puntos del espacio.

Es importante conocer las siguientes propiedades elementales de la operación de inclusión entre conjuntos.

Proposición 1.2.1. La operación de inclusión entre conjuntos satisface las propiedades siguientes:

- i) Si $A \subset B$ y $B \subset A$ entonces $A = B$ (simetría).
- ii) Para todo conjunto A , $A \subset A$ (reflexividad).
- iii) Si $A \subset B$ y $B \subset C$, entonces $A \subset C$ (transitividad)
- iv) Si $A_1 \subset A_2 \subset \dots \subset A_n \subset A_1$, entonces $A_1 = A_2 = \dots = A_n$.

La demostración se deja al lector.

Ejercicios de Autocontrol

1. Diga si es finito o infinito el conjunto de todos los átomos del sistema solar.

Fundamente su respuesta.

2. Determine los elementos que pertenecen a los conjuntos siguientes:

a) $A = \{x : x^2 - 8x + 15 = 0\}$

b) $A = \{x : x \in \mathbb{N}, -11 < x \leq 13\}$

3. Demuestre que si A es el conjunto de las raíces de la ecuación $x^2 - 7x + 6 = 0$ y $B = \{1, 6\}$, entonces $A = B$.

4. Describa el conjunto de los puntos M del plano tales que:

a) la longitud del segmento OM es igual al número r ;

b) la longitud del segmento OM es $\leq r$;

c) la longitud del segmento AM es igual a la longitud del segmento MB .

(Considere que O , A y B son puntos fijos y r es un número real positivo)

5. ¿Qué diferencia hay entre las expresiones $A \subset B$ y $a \in B$?

6. Demuestre que $\{\{1, 2\}, \{2, 3\}\} \neq \{1, 2, 3\}$

7. Diga si es verdadero o falso:

$$\{1, 2\} \in \{\{1, 2, 3\}, \{1, 3\}, 1, 2\}$$

$$\{1, 2\} \subset \{\{1, 2, 3\}, \{1, 3\}, 1, 2\}$$

8. Cite ejemplos de conjuntos A , B y C tales que:

- a) $A \in B, B \in C$ y $A \notin C$;
- b) $A \subset B, B \subset C$ y $A \subset C$;
- c) $A \subset B$ y $A \subset C$

9. Halle todos los subconjuntos de los conjuntos $\{1, 2\}$; ϕ y $\{a, b, e, d\}$.

10. Demuestre que $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$, si y solo si $a = c$ y $b = d$.

1.3 OPERACIONES CON CONJUNTOS

La teoría matemática de conjuntos parte de la idea de generar nuevos conjuntos con la ayuda de ciertas operaciones sobre una determinada familia inicial de conjuntos. El estudio de estas operaciones sobre los conjuntos es el objeto del Álgebra de conjuntos, la cual tiene mucho en común con el Álgebra ordinaria de los números, a la vez que presenta diferencias radicales con ella.

El hecho de que los métodos algebraicos puedan ser aplicados al estudio de objetos no numéricos, como son los conjuntos, nos muestra fehacientemente la generalidad de las ideas del álgebra.

Definamos las operaciones que se realizan con los conjuntos.

Unión: Se llama unión o suma de los conjuntos A y B , al conjunto C de todos los elementos contenidos al menos en uno de los conjuntos A o B .

Observación: Si un mismo elemento esta contenido simultáneamente en los conjuntos A y B , entonces en la unión, este elementos se incluye solamente una vez.

La unión C de los conjuntos A y B se designa mediante $C = A \cup B$, y los conjuntos A y B se llaman sumandos.

Ejemplo 7. Si $A = \{1, 2, 3, 4\}$ y $B = \{2, 4, 6, 8\}$, entonces $A \cup B = \{1, 2, 3, 4, 6, 8\}$; si A es el conjunto de todos los niños de una ciudad y B el conjunto de todos los habitantes del sexo masculino de esa ciudad, entonces $A \cup B$ es el conjunto de todos los habitantes del sexo masculino junto con todas las niñas de dicha ciudad.

La operación de unión de conjuntos posee algunas propiedades análogas a la operación de adición de números:

- 1) conmutatividad: $A \cup B = B \cup A$;
- 2) asociatividad: $(A \cup B) \cup C = A \cup (B \cup C)$.

Estas propiedades justifican el nombre de suma para denotar la unión de conjuntos. Precisamente, la propiedad asociativa permite definir la unión (suma) de una cantidad arbitraria de conjuntos, sin tener en cuenta la forma en que ellos se agrupan. Por ejemplo, si el número de conjuntos sumandos es finito e igual a n , su unión será el conjunto

$$C = A_1 \cup A_2 \cup \dots \cup A_n$$

o en forma más breve

$$C = \bigcup_{i=1}^n A_i.$$

Si la colección de sumandos es infinita y numerable, por ejemplo $i = 1, 2, \dots, n, \dots$, entonces

$$C = A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n \cup \dots,$$

o en forma más breve

$$C = \bigcup_{i=1}^{\infty} A_i$$

Sin embargo, la unión de conjuntos posee otras propiedades que no posee la suma de números:

$$3) \quad A \cup A = A, A \cup A \cup \dots \cup A \cup \dots = A;$$

$$4) \quad A \cup B = A, \text{ si } B \subset A.$$

Ejemplo 8. Si $A = \{1, 2, 3\}$, entonces $C = A \cup A = \{1, 2, 3\} = A$; si A es el conjunto de todos los estudiantes en un aula y B el conjunto de todos los estudiantes varones de la misma aula, entonces $C = A \cup B = A$ es el conjunto de todos los estudiantes de dicha aula.

Para ver de manera más clara las operaciones con conjuntos, representaremos a los conjuntos en forma de ciertas figuras.

En la figura 1, la parte sombreada representa la unión de los conjuntos A y B .

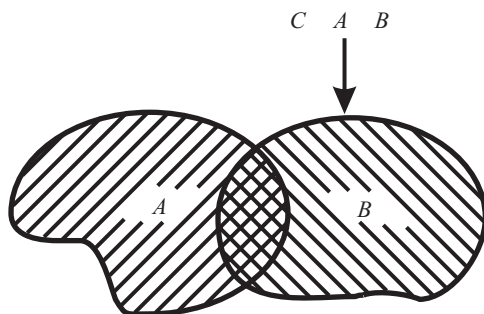


Figura 1

Intersección: Se llama intersección de los conjuntos A y B al conjunto C formado por todos los elementos que pertenecen simultáneamente a los conjuntos A y B .

De la definición se deduce que la intersección de conjuntos es la parte común de estos conjuntos. La intersección de los conjuntos A y B se denota por $A \cap B$.

En la figura 2 está sombreada la parte que representa a la intersección $C = A \cap B$.

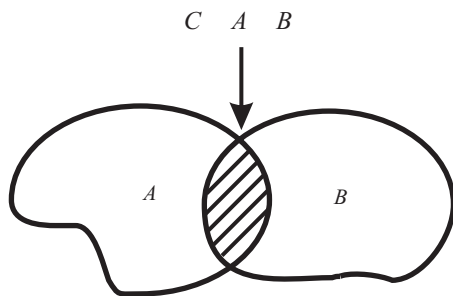


Figura 2

Ejemplo 9. Si la intersección A es el conjunto de todos los niños de una ciudad y B es el conjunto de todos los habitantes del sexo femenino de esa ciudad, entonces $A \cap B$ es el conjunto de todas las niñas de dicha ciudad:

Si $A = \{1, 3, 5, 7\}$ y $B = \{2, 4, 6, 8\}$ entonces $A \cap B = \emptyset$.

Si la intersección de conjuntos es el conjunto vacío se dice que los conjuntos son disjuntos.

No es difícil extender la definición dada de intersección a una familia arbitraria de conjuntos. La intersección de un número finito n de conjuntos A_i se denota

$$A = \bigcap_{i=1}^n A_i$$

y la intersección de una familia infinita numerable de conjuntos A_i se denota de esta forma

$$C = \bigcap_{i=1}^{\infty} A_i.$$

La operación de intersección de conjuntos posee algunas propiedades análogas a la operación de producto de números:

1) conmutatividad: $A \cap B = B \cap A$;

2) asociatividad: $(A \cap B) \cap C = A \cap (B \cap C)$;

3) distributividad con respecto a la unión: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

Las propiedades 1 y 2 de la intersección de conjuntos se deducen inmediatamente de la definición, sin embargo, la propiedad 3 no es tan evidente, y por ello, haremos su demostración analítica (no obstante el lector debe comprobar la demostración geoméricamente con ayuda de un gráfico).

Para demostrar la propiedad 3 es necesario mostrar que cada elemento del conjunto $(A \cup B) \cap C$ es también un elemento del conjunto $(A \cap C) \cup (B \cap C)$ y viceversa.

Demostración: Si $x \in (A \cup B) \cap C$, entonces $x \in A \cup B$ y $x \in C$, de donde $x \in (A \cap C) \cup (B \cap C)$.

De esta forma queda probada la inclusión

$$(A \cup B) \cap C \subset (A \cap C) \cup (B \cap C).$$

Sea ahora $x \in (A \cap C) \cup (B \cap C)$, demostraremos que $x \in (A \cup B) \cap C$.

Como x pertenece a la unión de los dos conjuntos $A \cap C$ y $B \cap C$, entonces $x \in A \cap C$ ó $x \in B \cap C$.

Si $x \in A \cap C$, entonces $x \in A$ y $x \in C$. Del hecho que $x \in A$, se deduce que $x \in A \cup B$ y como $x \in C$, entonces $x \in (A \cup B) \cap C$.

Si $x \in B \cap C$, entonces $x \in B$ y $x \in C$. Del hecho que $x \in B$, se deduce que $x \in A \cup B$ y como $x \in C$, entonces $x \in (A \cup B) \cap C$ con lo cual la propiedad queda probada. ■

Mostremos ahora otras propiedades de la intersección de conjuntos que no posee la operación de multiplicación de números:

4) $A \cap B = A$, si $A \subset B$.

Por ejemplo, si $A = \{1, 2\}$ y $B = \{1, 2, 3, 4, 5\}$ entonces $A \cap B = \{1, 2\} = A$.

5) $A \cap A = A$ y $A \cap A \cap \dots \cap A \cap \dots = A$

Por ejemplo, si $A = \{3, 4, 5\}$, entonces $A \cap A = \{3, 4, 5\} = A$.

Diferencia: Se llama diferencia de los conjuntos A y B al conjunto formado por todos los elementos del conjunto A que no pertenecen al conjunto B .

La diferencia entre los conjuntos A y B se denota $A \setminus B$.

Ejemplo 10. Si $A = \{1, 2, 3, 4, 5\}$ y $B = \{2, 4\}$, entonces $A \setminus B = \{1, 3, 5\}$.

Si B es el conjunto de todos los enteros pares, entonces $\mathbb{Z} \setminus B$ es el conjunto de todos los enteros impares; si A es el conjunto de todos los niños de una ciudad y B es el conjunto de todos los habitantes de la ciudad del sexo femenino, entonces $A \setminus B$ es el conjunto de todos los niños varones de esa ciudad.

En las figuras 3 y 4, las partes sombreadas representan la diferencia entre los conjuntos A y B .

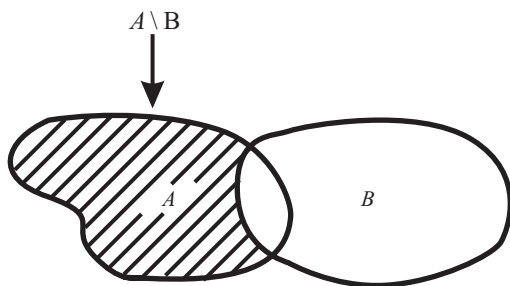


Figura 3

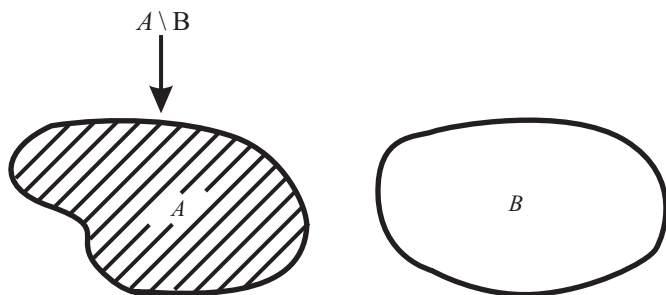


Figura 4

Observación: en el caso de la figura 4, los conjuntos son disjuntos.

Con frecuencia, en la literatura matemática, a la diferencia $A \setminus B$ se le llama complemento del conjunto B en el conjunto A y se denota por B'_A . Si todos los conjuntos en consideración se estudian como subconjuntos de un conjunto universal U , entonces por complemento del conjunto B se sobreentiende su complemento en U , y se designa por B'_U o sencillamente B' .

Producto cartesiano: Se llama producto cartesiano de los conjuntos A y B , al conjunto formado por todos los pares ordenados (a, b) donde $a \in A$ y $b \in B$. El producto cartesiano de los conjuntos A y B se denota por $A \times B$. De manera análoga se puede definir el producto cartesiano de un número finito de conjuntos $A_1 A_2, \dots, A_n$, el cual se representa por $A_1 \times A_2 \times \dots \times A_n$ y está compuesto por todos los n -uplos $(a_1, a_2, \dots, a_n)$ donde $a_i \in A_i$.

Este método de construir conjuntos de una categoría superior a partir de conjuntos más simples, se utiliza en el estudio de los campos numéricos y sus generalizaciones.

Llamaremos diagonal de $A \times A$ y la denotaremos por Δ al subconjunto de $A \times A$ definido por}

$$\Delta := \{(a; a) : a \in A\}.$$

Una propiedad importante es la siguiente:

$$A \times B = \emptyset \Leftrightarrow A = \emptyset \vee B = \emptyset$$

Resumen de Fórmulas del Álgebra de Conjuntos.

Hasta el momento hemos estudiado las operaciones con los conjuntos y algunas de sus propiedades. A continuación daremos una relación de fórmulas importantes con estas operaciones. (la propiedad 11 ya fue demostrada y las otras no presentan gran dificultad, por lo cual se dejan al lector como ejercicio de autocontrol. Recomendamos además de la demostración analítica, que se realice la representación geométrica de cada una de ellas):

- 1) $A \subset A$;
- 2) Si $A \subset B$ y $B \subset A$, entonces $A = B$;
- 3) Si $A \subset B$ y $B \subset C$, entonces $A \subset C$;
- 4) $\emptyset \subset A$;
- 5) $A \subset U$;
- 6) $A \cup B = B \cup A$;
- 7) $A \cap B = B \cap A$;
- 8) $A \cup (B \cap C) = (A \cup B) \cap C$;
- 9) $A \cap (B \cup C) = (A \cap B) \cup C$;
- 10) $A \cup A = A$;
- 11) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
- 12) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
- 13) $A \cup \emptyset = A$;
- 14) $A \cap U = A$;
- 15) $A \cup U = U$;

- 16) $A \cap A = A$;
- 17) $A \cap \emptyset = \emptyset$;
- 18) si $A \subset B$, entonces $A \cup B = B$ y $A \cap B = A$;
- 19) $A \cup A' = U$;
- 20) $A \cap A' = \emptyset$;
- 21) $\emptyset' = U$;
- 22) $U' = \emptyset$;
- 23) $(A')' = A$;
- 24) si $A \subset B$, entonces $B' \subset A'$;
- 25) $(A \cup B)' = A' \cap B'$;
- 26) $(A \cap B)' = A' \cup B'$;

Con ayuda de estas propiedades se pueden realizar operaciones con los conjuntos de forma similar a como se realizan con los números en la aritmética.

Estas fórmulas poseen la importante propiedad de “dualidad”; es decir si en alguna de ellas se cambian los símbolos $\subset$ por $\supset$, $\emptyset$ por U y $\cup$ por $\cap$ respectivamente, entonces como resultado se obtiene otra de esas fórmulas. De esto se deduce que cualquier teorema obtenido a partir de estas fórmulas, posee la propiedad de dualidad. Esta propiedad, o como se dice con frecuencia, el principio de dualidad, ayuda a simplificar considerablemente las demostraciones de algunos resultados, y es aplicable a muchas disciplinas matemáticas.

1.4 EQUIVALENCIA DE CONJUNTOS, POTENCIA DE CONJUNTOS, CONJUNTO POTENCIA

Cada disciplina matemática estudia las relaciones que pueden establecerse entre los objetos de su estudio. Sabemos que los objetos de estudio de la teoría de conjuntos son los conjuntos. Analicemos las relaciones que pueden establecerse entre ellos.

Primeramente nos ocuparemos de este análisis con respecto a los conjuntos finitos, es decir, a los conjuntos cuya cantidad de elementos se puede poner en correspondencia con un cierto número entero positivo. Para hallar este número para un conjunto finito no vacío A , es suficiente contar todos sus elementos. Si el conjunto es vacío, entonces a la cantidad de elementos de este conjunto

se le hace corresponder el número 0, que indica la ausencia de elementos en el conjunto. Pero, ¿qué significa contar? Contar los elementos de un conjunto A , significa fijar a cada elemento de dicho conjunto un número $1, 2, 3, \dots, n, \dots$, en forma sucesiva, sin repetir ni omitir ningún elemento del conjunto. Entonces los elementos del conjunto A se pueden enumerar en la forma $a_1, a_2, \dots, a_k$. El subíndice k del último elemento a_k denota la cantidad de elementos del conjunto A .

Es obvio que los elementos de un conjunto A , que contiene más de un elemento, se pueden enumerar de varias formas, pero para cualquier orden de numeración siempre fijaremos al último elemento un mismo número k , es decir, el número de elementos de cualquier conjunto finito no depende del método de numeración empleado. Esto nos da la posibilidad de considerar al número k como una característica cuantitativa de cualquier conjunto finito $A = \{a_1, a_2, \dots, a_k\}$, la cual corresponde a la cantidad de elementos de este conjunto.

Si dos conjuntos finitos A y B poseen la misma cantidad de elementos, entonces ellos poseen la misma característica cuantitativa. Tales conjuntos finitos se llaman *equinumericos*.

Si los conjuntos A y B poseen diferentes características cuantitativas $m \neq n$, no es difícil establecer una comparación entre los números m y n , es decir, las relaciones $m > n$ ó $m < n$ y así saber cual de los dos conjuntos tiene una mayor característica, o en otras palabras, en cuál conjunto hay mayor cantidad de elementos.

Sin embargo, para saber si dos conjuntos son equinumericos o no, no es necesario contar sus elementos.

Ejemplo 11. Para cerciorarse de que el número de dedos de la mano izquierda es igual al de la mano derecha (es decir, que el conjunto de dedos en una mano es equinumerico al de la otra), se puede unir cada uno de los dedos de una mano con el correspondiente en la otra, estableciendo así una correspondencia entre los dedos de ambas manos. El hecho de que cada dedo tiene su pareja indica la igualdad entre el número de dedos de las dos manos sin necesidad de contarlos.

Ejemplo 12. Supongamos que queremos saber si hay igual número de hombres que de mujeres en un local sin necesidad de contarlos. Para ello, es suficiente asociar a cada hombre su pareja del sexo opuesto. Si no queda libre ningún hombre ni ninguna mujer, entonces hay la misma cantidad de un sexo y otro.

Cuando a cada elemento a del conjunto A se le puede hacer corresponder,

mediante una cierta ley o regla, uno y solamente un elemento b del conjunto B y viceversa, se dice que entre los elementos de los conjuntos A y B se ha establecido una *correspondencia biunívoca*.

Hemos introducido el concepto *correspondencia biunívoca* teniendo en cuenta algunos ejemplos, los cuales nos muestran que una tal relación entre objetos existe efectivamente en el mundo real. Notemos, que solamente hemos establecido el hecho de la existencia de una correspondencia biunívoca entre los objetos de dos conjuntos de la realidad, y que no hemos definido este concepto a partir de las nociones primarias de *conjunto y pertenencia*. Por mucho que lo desearamos, no hubieramos podido hacerlo, ya que el concepto *correspondencia biunívoca* es tan primario como los propios conceptos de *conjunto y pertenencia*.

Definición 1.4.1. Si entre los elementos de dos conjuntos A y B se puede establecer una correspondencia biunívoca, se dice que ambos conjuntos son equivalentes y se escribe $A \sim B$.

De la definición de equivalencia de conjuntos se deducen las propiedades siguientes:

- 1) reflexividad: $A \sim A$ (cualquier conjunto es equivalente así mismo);
- 2) simetría: si $A \sim B$, entonces $B \sim A$;
- 3) transitividad: si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

Si regresamos a los conjuntos equinumerícos, podemos decir que ellos son equivalentes; se cumple también la afirmación recíproca: dos conjuntos finitos equivalentes son necesariamente equinumerícos.

Este hecho es importante desde el punto de vista lógico, ya que para el establecimiento de la equinumeridad de dos conjuntos, no es necesario conocer la noción de contar, ni el concepto *número natural* con ayuda del cual introdujimos el concepto *equinumeridad*. Pudiéramos decir que un número natural es una característica numérica común a todos los conjuntos finitos equivalentes entre sí.

Con los conceptos *conjunto, pertenencia y correspondencia biunívoca*, se puede desarrollar toda la teoría de los números naturales y después, utilizando algunas técnicas de la aritmética y el álgebra (que no expondremos aquí), construir completamente la teoría de los números enteros y racionales.

Hemos visto que para el estudio de las relaciones cuantitativas entre los conjuntos (al menos para los conjuntos finitos), se puede partir indistintamente del concepto primario de número natural o del concepto primario de correspondencia biunívoca. En nuestras deducciones posteriores, utilizaremos en calidad

de concepto primario la noción de correspondencia biunívoca en lugar de la de número natural, ya que esta última no se puede utilizar para comparar los conjuntos infinitos. En efecto, no se pueden contar todos los elementos de tales conjuntos, y por consiguiente, no se les puede hacer corresponder un cierto número natural n . Sin embargo, con ayuda del concepto *correspondencia biunívoca* los conjuntos infinitos pueden ser comparados.

Ejemplo 13. Entre los conjuntos $A = \{1, 2, 3, \dots, n, \dots\}$ y $B = \{-1, -2, -3, \dots, -n, \dots\}$ se puede establecer una correspondencia biunívoca, si consideramos relacionados los elementos con un mismo valor absoluto; luego $A \sim B$.

Ejemplo 14. Entre el conjunto de puntos del segmento $\overline{MN}$ y el conjunto de puntos del segmento $\overline{KL}$, se puede establecer una correspondencia biunívoca (fig. 5), es decir, los conjuntos de puntos de ambos segmentos son equivalentes.

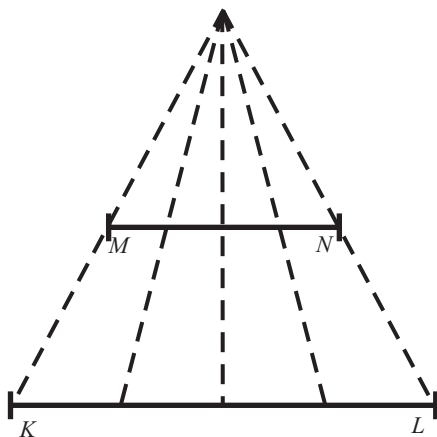


Figura 5

Si de los conjuntos finitos equivalentes decimos que son equinumerícos, para los conjuntos infinitos equivalentes diremos que son *equipotentes* o que poseen la misma *potencia*. Si bien todos los conjuntos finitos equivalentes se caracterizan por algún número natural, todos los conjuntos infinitos equivalentes se caracterizan por su *potencia*.

El concepto *número de elementos* es la noción de cantidad para los conjuntos finitos, mientras que para los conjuntos infinitos lo es la noción de *potencia*. Luego, *potencia* es la generalización del concepto *cantidad* a los conjuntos infinitos. Este concepto permite comparar los conjuntos infinitos de una manera cuantitativa.

Designemos la potencia de todos los conjuntos equivalentes a un cierto conjunto infinito A mediante el símbolo α . Si otro conjunto B , cuya potencia denotaremos por el símbolo β , es equivalente al conjunto A , entonces por analogía con los conjuntos finitos, diremos que la potencia α es igual a la potencia β y escribiremos $\alpha = \beta$. Si ambos conjuntos no son equivalentes, consideraremos que sus potencias son diferentes y pondremos $\alpha \neq \beta$. De esta forma se define la igualdad y desigualdad de potencias.

Cuando establecemos la igualdad entre las potencias de dos conjuntos, en este hecho se expresa solamente una determinada relación entre los conjuntos. La equipotencia entre conjuntos expresa una correspondencia biunívoca entre sus elementos, y no otra relación.

Para las potencias se pueden introducir los conceptos *mayor que* y *menor que*. Por analogía con los conjuntos finitos, diremos que si dos conjuntos A y B no son equivalentes y el conjunto A es equivalente a un cierto subconjunto propio del conjunto B , entonces la potencia α del conjunto A es menor que la potencia β del conjunto B y se escribe $\alpha < \beta$. De la misma forma, se dice que la potencia β es mayor que α y se escribe $\beta > \alpha$.

Este hecho de que entre las potencias se puedan establecer las relaciones de *igualdad*, *mayor que* y *menor que*, nos fundamenta el llamar también “números” a los símbolos que representan las potencias.

Estos “números” que en el dominio de lo infinito representan la transición del pensamiento matemático a una nueva etapa cualitativa, fueron denominados por G. Kantor números cardinales. Puede demostrarse el siguiente teorema de Kantor-Bernstein: si cada uno de dos conjuntos dados es equivalente a cierta parte del otro, entonces ambos conjuntos son equivalentes.

Hasta el momento hemos supuesto la existencia de conjuntos infinitos con diferentes potencias. Detengámonos ahora en la demostración de la existencia de conjuntos infinitos de diferentes potencias y más aún, mostremos que existen conjuntos con potencia tan alta como se quiera.

Teorema 1.4.2. La potencia del conjunto de todos los subconjuntos de cualquier conjunto no vacío A , es mayor que la potencia del propio conjunto A .

Demostración: Este teorema es evidente para un conjunto finito. Por ejemplo, la potencia de un conjunto A que contiene a dos elementos $A = \{a_1, a_2\}$ es igual a 2 mientras que la potencia del conjunto B , constituido por todos los subconjuntos de A ($B = \{\emptyset, \{a_1, a_2\}, \{a_1\}, \{a_2\}\}$) es igual a 4.

Demostremos el teorema para conjuntos arbitrarios. Sea A un conjunto no vacío y B el conjunto formado por todos los subconjuntos de A . Para demostrar

el teorema debemos establecer primeramente que A y B no son equivalentes y en segundo lugar que A es equivalente a una cierta parte propia de B .

Demostremos que A y B no son equivalentes utilizando el método de demostración al absurdo. Supongamos que $A \sim B$. Entonces a cada elemento $a \in A$ le corresponde un cierto elemento $b \in B$ compuesto por elementos del conjunto A . Con respecto a los elementos $a \in A$ se pueden dar dos casos: o bien el elemento a pertenece al conjunto b que le corresponde, o bien no pertenece. Siguiendo esta regla dividamos a los elementos del conjunto A en “incluidos” y “no incluidos”. Es evidente que ambas categorías son no vacías. En efecto, el elemento $a \in A$, correspondiente al propio conjunto A es “incluido”, mientras que el elemento de A correspondiente al conjunto vacío es “no incluido”.

Unamos todos los elementos no incluidos en el conjunto b_0 , el cual será un elemento del conjunto B . Como $B \sim A$, al elemento $b_0 \in B$ le corresponde un cierto elemento $a_0 \in A$. Veamos a cual categoría pertenece el elemento a_0 . Supongamos que a_0 es “incluido”. Es claro que a_0 no puede ser “incluido” en b_0 pues b_0 está constituido por todos los elementos no incluidos de A . Pero si a_0 es “no incluido”, por definición del conjunto b_0 tendría que ser $a_0 \in b_0$, lo cual no es posible por la suposición de que a_0 es “no incluido”. De esta forma hemos visto que a_0 no puede pertenecer a ninguna de las dos categorías indicadas y como no existe una tercera categoría, hemos obtenido un elemento $b_0 \in B$ al cual no le corresponde ningún elemento del conjunto A . Con esto obtenemos una contradicción y por tanto nuestra suposición, de que los conjuntos A y B son equivalentes, es falsa.

Ahora demostremos que el conjunto A es equivalente a una parte propia del conjunto B . Si en calidad de parte propia de B tomamos el conjunto B_1 de todos los subconjuntos compuestos por un solo elemento del conjunto A , entonces entre los elementos de los conjuntos A y $B_1 \subset B$ se puede establecer una correspondencia biunívoca asociando a cada elemento del conjunto A el subconjunto unitario constituido por dicho elemento. ■

Más adelante demostraremos que entre los conjuntos infinitos, el conjunto de los números naturales cumple la maravillosa propiedad de ser el que tiene la potencia de menor orden posible. Pero primeramente pasemos a ver algunas operaciones sobre las potencias.

Teniendo en cuenta que la potencia es una característica cuantitativa de los conjuntos infinitos, se extiende de forma natural el concepto de suma de potencias para dichos conjuntos.

Definición 1.4.3. Si los conjuntos A y B son disjuntos y tienen potencias α y β respectivamente, entonces la potencia γ del conjunto $C = A \cup B$ se llama *suma de las potencias* α y β y se denota por $\gamma = \alpha + \beta$

Análogamente se define la suma de cualquier número de potencias.

Para introducir el concepto producto de potencias, debemos remitirnos a la noción de producto cartesiano entre conjuntos. Si A y B son conjuntos finitos con m y n elementos respectivamente, entonces hay mn posibles pares ordenados de la forrna (a, b) con $a \in A$ y $b \in B$. Razonando de forma análoga en el caso de los conjuntos infinitos, llegamos a la definición siguiente:

Definición 1.4.4. Sean $A = \{a\}$ y $B = \{b\}$ dos conjuntos arbitrarios con potencias α y β respectivamente. Entonces la potencia del producto cartesiano $A \times B = \{(a, b)\}$ se llama producto de las potencias α y β y se denota por $\gamma = \alpha\beta$.

De forma similar se define el producto de cualquier número finito de potencias.

De la definición de suma y producto de potencias se deducen las propiedades siguientes:

suma	producto
1) conmutatividad: $\alpha + \beta = \beta + \alpha$	$\alpha\beta = \beta\alpha$;
2) asociatividad: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$,	$(\alpha\beta)\gamma = \alpha(\beta\gamma)$;
3) distributividad: $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$.	

El concepto diferencia de potencias no existe. Sería totalmente inútil tratar de introducir este concepto partiendo de la definición de complemento. En efecto, si α es la potencia del conjunto A y β la potencia del conjunto B , entonces el complemento de B en A puede, en general, ser un conjunto finito, infinito o vacío, aún en el caso cuando $\alpha = \beta$; por eso, no se puede dar un sentido determinado al símbolo $\alpha - \beta$.

Sin entrar en razonamientos detallados, indicaremos, que la potencia del conjunto de todos los subconjuntos de un conjunto dado A con potencia α , se denota por 2^α y por ello el teorema demostrado anteriormente se expresa en síntesis mediante la desigualdad $2^\alpha > \alpha$.

Demostremos una de las propiedades más importantes de los números naturales, a la cual ya hemos hecho referencia: entre todos los conjuntos infinitos, el conjunto de los números naturales tiene la menor potencia.

Para demostrar esta propiedad es necesario el teorema siguiente:

Teorema 1.4.5. Todo conjunto infinito A contiene un subconjunto B equivalente al conjunto de los números naturales y de manera tal que el conjunto $A \setminus B$ es también infinito.

Demostración: Separemos del conjunto A dos elementos a_1 y b_1 , lo cual es posible ya que en caso contrario, el conjunto A sería unitario o vacío. El conjunto restante $A_1 = A \setminus \{a_1, b_1\}$ será infinito, ya que si fuera finito también lo sería el conjunto A . Ahora del conjunto A_1 separemos otros dos elementos a_2 y b_2 . De nuevo esto es posible por ser A_1 infinito y además, el conjunto $A_2 = A_1 \setminus \{a_2, b_2\}$ será también infinito. Al repetir este razonamiento infinitamente, habremos separado dos subconjuntos infinitos $B = \{a_1, a_2, a_3, \dots, a_n, \dots\}$ y $C = \{b_1, b_2, b_3, \dots, b_n, \dots\}$. El conjunto $B \subset A$ es equivalente al conjunto de los números naturales, ya que entre ambos conjuntos se puede establecer una correspondencia biunívoca:

$$\begin{array}{cccc} a_1 & a_2 & a_3 \dots & a_n \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 1 & 2 & 3 \dots & n \dots \end{array}$$

Luego, A contiene un subconjunto B equivalente al conjunto de los números naturales y $A \setminus B$ es infinito ya que contiene al subconjunto infinito C . ■

Corolario 1. Entre todos los subconjuntos infinitos, el conjunto de los números naturales tiene la menor potencia.

Demostración: En efecto, cualquier subconjunto infinito A contiene un subconjunto B equivalente al conjunto de los números naturales. Como los conjuntos equivalentes tienen la misma potencia, si denotamos la potencia de $\mathbb{N}$ por α_0 y la del conjunto infinito A por α , entonces $\alpha_0 \leq \alpha$ ■

Definición 1.4.6. Los conjuntos equivalentes al conjunto de los números naturales se llaman *numerales*.

Veamos algunas propiedades fundamentales de los conjuntos numerables.

Teorema 1.4.7. Todo subconjunto infinito de un conjunto numerable es también numerable.

Demostración: Sea A un conjunto numerable y A_1 un subconjunto infinito de A . Si α es la potencia de A_1 y α_0 la potencia de A entonces por una parte, $\alpha \leq \alpha_0$ ya que $A_1 \subset A$, y por otra, $\alpha \geq \alpha_0$, puesto que A_1 , es un conjunto infinito y por el corolario 1 del teorema 1.4.5 su potencia no puede ser menor que la de un conjunto numerable. De las dos relaciones obtenidas, se deduce que $\alpha = \alpha_0$, es decir, A_1 es un conjunto numerable. ■

Corolario 1. Si de un conjunto numerable A se extrae un subconjunto finito

K , el conjunto $A \setminus K$ es numerable.

Demostración: En efecto, el conjunto $A \setminus K$ no puede ser finito, ya que si lo fuera también lo sería A . Por consiguiente del teorema 1.4.7 se deduce que $A \setminus K$ es numerable. ■

Teorema 1.4.8. Si un conjunto infinito A está formado por elementos a_{ij} , los cuales se diferencian entre sí por dos índices i y j que toman valores en el conjunto de los números naturales, entonces el conjunto $A = \{a_{ij}\}$ es numerable.

Demostración: Designemos la suma de los índices i y j mediante $h(h = i + j)$. Es evidente que el menor valor de h es 2 (para el elemento a_{11} ; asociemos a este elemento el número 1. Con el valor $h = 3$ habrán dos elementos a_{12} y a_{21} . Asociemos a estos dos elementos los números 2 y 3 respectivamente. Dándole a h sucesivamente los valores de todos los números naturales y numerando, para cada valor de h , todos los elementos a_{ij} tales que la suma de sus índices i y j sea igual a h , habremos enumerado todos los elementos del conjunto A utilizando todos los números naturales. Por consiguiente el conjunto A es numerable. ■

En el teorema 1.4.8 es fácil seguir el orden de numeración de los elementos del conjunto $A = \{a_{ij}\}$, si escribimos dichos elementos en la forma siguiente:

$$A = \left\{ \begin{array}{ccccccc} a_{11} & \longrightarrow & a_{12}, a_{13} & \longrightarrow & \dots, a_{1n}, \dots, \\ & \swarrow & \nearrow & & & & \\ a_{21} & & a_{22} a_{23} & & \dots, a_{2n}, \dots, \\ \downarrow & \nearrow & & & & & \\ a_{31} & & a_{32} \ a_{33} & & \dots, a_{3n}, \dots, \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ a_{n1} & & a_{n2} a_{n3} & & \dots, a_{nn}, \dots, \\ & & & & & & \end{array} \right\}$$

Pasemos a enunciar dos corolarios del teorema 1.4.8 cuyas demostraciones se dejan al lector.

Corolario 1. La unión de un número finito de conjuntos numerables es también un conjunto numerable.

Corolario 2. La unión de cualquier cantidad numerable de conjuntos nu-

merables es un conjunto numerable.

Notemos que en virtud de la correspondencia entre los elementos del conjunto $A = \{a_{ij}\}$ del teorema 1. 4.8 y el conjunto de los pares (i, j) de números naturales, este teorema puede expresarse también en la forma siguiente: el producto cartesiano de dos (y también de cualquier número finito de) conjuntos numerables es un conjunto numerable.

El lector debe comprobar la equivalencia entre ambos enunciados.

El teorema 1.4.8 con sus generalizaciones y corolarios, puede ser escrito en el lenguaje de las potencias de la manera siguiente:

$$1) \underbrace{\alpha_0 + \alpha_0 + \cdots + \alpha_0}_{n \text{ veces}} = \alpha_0;$$

$$2) \underbrace{\alpha_0 + \alpha_0 + \cdots + \alpha_0}_{\text{una cantidad numerable}} = \alpha_0;$$

$$3) \alpha_0 + n = \alpha_0, \text{ donde } n \text{ es la potencia de un conjunto finito};$$

$$4) \alpha_0 \cdot \alpha_0 = \alpha_0, \underbrace{\alpha_0 \cdot \alpha_0 \cdot \cdots \cdot \alpha_0}_{n \text{ veces}} = \alpha_0$$

Las relaciones anteriores constituyen la llamada *aritmética de los conjuntos numerables*.

No es nuestro objetivo dedicarnos a la construcción de los diferentes campos numéricos (números enteros $(\mathbb{Z})$, números racionales $(\mathbb{Q})$, números reales $(\mathbb{R})$, números complejos $(\mathbb{C})$), a partir del conjunto de los números naturales $(\mathbb{N})$; esto sería muy engorroso y alargaría considerablemente nuestra exposición. No obstante, indicaremos la esencia del método que consiste en la búsqueda de conjuntos en que sean solubles determinadas ecuaciones, las cuales representan las operaciones fundamentales de suma y producto de números. Así por ejemplo, el paso de $\mathbb{N}$ a $\mathbb{Z}$ se lleva a cabo “construyendo los números” que sean solución de la ecuación $x + n = m$ para todos n y m en $\mathbb{N}$, la cual no es siempre soluble en el conjunto $\mathbb{N}$ (en el caso $n > m$). Por otra parte el paso de $\mathbb{Z}$ a $\mathbb{Q}$ se realiza, “construyendo los números” que sean solución de la ecuación $p \cdot x = q$ para todos p y q en $\mathbb{Z}$, la cual no siempre tiene solución en $\mathbb{Z}$.

Los números racionales pueden ser interpretados como pares ordenados de números enteros y por ello, apoyándonos en el teorema 1.4.8, no es difícil demostrar que $\mathbb{Q}$ es un conjunto numerable. El paso de $\mathbb{Q}$ a $\mathbb{R}$ constituye un proceso más complejo de carácter no solamente algebraico, denominado, “completamiento” y que estudiaremos detalladamente en el capítulo 3. Desde el punto de vista cuantitativo y estructural de la teoría de conjuntos, esta

transición es fundamental, pues con ella, se pasa de los conjuntos *numerables discretos* a los *conjuntos infinitos continuos*. Pasemos a explicar estos términos en un ejemplo. Una de las propiedades del conjunto $\mathbb{Q}$ de los números racionales es su densidad, es decir, para cualesquiera números racionales $r_1 < r_2$, se puede encontrar un tercer número racional r_3 , tal que $r_1 < r_3 < r_2$, (fig. 6).



Figura 6

Geométricamente esto significa que en el segmento $\overline{r_1 r_2}$, por muy pequeño que sea, siempre se pueden encontrar infinitos puntos que corresponden a números racionales.

Este hecho nos crea la ilusión de que a cada punto A del eje numérico le corresponde un número racional, o en otras palabras, la longitud de cualquier segmento $\overline{OA}$ puede ser expresada por un número racional. Sin embargo, es conocido que esto no es así, ya que basta considerar el cuadrado de lado unidad, cuya diagonal tiene una longitud que no puede ser representada por un número racional. Este hecho nos confirma que el conjunto de los números racionales no es suficiente para medir las longitudes de segmentos arbitrarios.

Con la creación del conjunto de los números reales, los cuales nos permiten *medir la longitud de cualquier segmento*, adquirimos la idea intuitiva del concepto *continuo*. A la potencia de $\mathbb{R}$ la denotaremos por la letra c y la llamaremos *potencia del continuo*.

Veamos sin demostración la siguiente tabla que nos muestra la aritmética del continuo y cuya interpretación se deja al lector:

1) $c + n = c$, donde n es la potencia de un conjunto finito;

2) $c + \alpha_0 = c$;

3) $\underbrace{c + c + \dots + c}_{n \text{ veces}} = c$;

4) $\underbrace{c + c + \dots + c + \dots}_{\text{una cantidad numerable}} = c$

5) $\underbrace{c \cdot c \cdot \dots \cdot c}_{n \text{ veces}} = c$;

6) $\underbrace{c \cdot c \cdot \dots \cdot c \cdot \dots}_{\text{una cantidad numerable}} = c$

Del teorema 1.4.5 se concluye obviamente que $\alpha_0 < c$. Anteriormente mostramos que los conjuntos numerables tienen la menor potencia posible entre todos los conjuntos infinitos, y que no existen conjuntos con una potencia mayor que todas las demás. Esto en particular significa que se pueden encontrar conjuntos con potencia mayor que la del continuo.

Resulta natural plantearse la pregunta: ¿existen o no conjuntos infinitos cuya potencia sea a la vez mayor que la del conjunto de los números naturales y menor que la del continuo?

El fundador de la teoría de conjuntos G. Kantor no pudo encontrar tales conjuntos y enunció la hipótesis de que ellos no existían. La solución de este problema que recibió el nombre de hipótesis del continuo, por largo tiempo se resistió a los esfuerzos de los matemáticos. Solamente a principios de los años 60 de nuestro siglo, el matemático norteamericano P. Cohen logró demostrar que la hipótesis del continuo no puede ser ni demostrada ni refutada. Esto significa que para construir la teoría de conjuntos, se puede tomar en calidad de axioma lo mismo la hipótesis del continuo que su negación y como resultado obtener “diferentes” teorías de conjuntos no contradictorias. El descubrimiento de P. Cohen en la teoría de conjuntos es, en cierto sentido, análogo al descubrimiento del científico ruso N. M. Lobachevsky en la geometría, el cual mostró que existen diferentes geometrías en dependencia de que se cumpla o no en ellas el axioma de Euclides sobre el paralelismo de rectas.

Ejercicios de Autocontrol

- 1) Demuestre que la equivalencia de conjuntos posee las propiedades siguientes:
 - a) reflexividad: $A \sim A$;
 - b) simetría: si $A \sim B$, entonces $B \sim A$;
 - c) transitividad: si $A \sim B$ y $B \sim C$, entonces $A \sim C$;
 - d) si $A \setminus B \sim B \setminus A$, entonces $A \sim B$;
 - e) si $A \subset B \subset C$ y $A \sim C$, entonces $A \sim B$
- 2) Establezca una correspondencia biunívoca (al menos geoméricamente):
 - a) entre los números naturales impares y todos los números naturales;
 - b) entre los puntos de una circunferencia y los puntos del contorno de un cuadrado;

- c) entre los números naturales mayores o iguales a 1 y el conjunto de todos los números naturales (incluyendo al cero);
- d) entre los puntos de una circunferencia a la que le falta un punto y los puntos de una recta.

1.5 RELACIONES DE EQUIVALENCIA. PARTICIÓN DE UN CONJUNTO. CONJUNTO COCIENTE

Estudiaremos ahora un procedimiento de gran importancia en las aplicaciones, mediante el cual podremos formar un nuevo conjunto a partir de otro dado y que generalmente tiene propiedades mejores que el conjunto inicial con respecto a alguna estructura. adicional.

Definición 1.5.1. Llamaremos *conjunto potencia* de un conjunto X y lo denotaremos por $\mathbf{P}(X)$, al conjunto cuyos elementos son los subconjuntos de X . En particular $\emptyset \in \mathbf{P}(X)$ y $X \in \mathbf{P}(X)$.

En 1.3 hemos definido las operaciones de unión, intersección y diferencia, las cuales pueden aplicarse a un número finito de elementos de $\mathbf{P}(X)$.

Definición 1.5.2. Llamaremos *colección* a todo subconjunto de $\mathbf{P}(X)$, es decir, los elementos de una colección son parte de X .

Dada una colección arbitraria $\mathbf{F} \subset \mathbf{P}(X)$ se definen la unión e intersección de los conjuntos de $\mathbf{F}$ de manera análoga a como fueron definidas en 1.3.

$$\cup \mathbf{F} := \cup \{F : F \in \mathbf{F}\} := \{x \in X : \exists F \in \mathbf{F} \text{ tal que } x \in F\};$$

$$\cap \mathbf{F} := \cap \{F : F \in \mathbf{F}\} := \{x \in X : F \in \mathbf{F}, x \in F\},$$

es decir, la unión $\cup \mathbf{F}$ está formada por todos los elementos que pertenecen a alguno de los conjuntos de $\mathbf{F}$, mientras que la intersección $\cap \mathbf{F}$ está formada por aquellos elementos que pertenecen simultáneamente a todos los conjuntos de la colección $\mathbf{F}$.

Cuando $\mathbf{F}$ está formado por un número finito de elementos, obtenemos la unión e intersección definida en 1.3.

Para la unión e intersección de una colección $\mathbf{F} \subset \mathbf{P}(X)$ se cumplen las

siguientes propiedades, análogas a las enunciadas en 1.3:

$$(\cup \mathbf{F})' = \cap \{F' : F \in \mathbf{F}\}$$

(se lee: el complemento de la unión es la intersección de los complementos):

$$(\cap \mathbf{F})' = \cup \{F' : F \in \mathbf{F}\}$$

(se lee: el complemento de la intersección es la unión de los complementos):

$$(\cup \mathbf{F}_1) \cap (\cup \mathbf{F}_2) = \cup \{F_1 \cap F_2 : F_1 \in \mathbf{F}_1 \wedge F_2 \in \mathbf{F}_2\}$$

$$(\cap \mathbf{F}_1) \cup (\cap \mathbf{F}_2) = \cap \{F_1 \cup F_2 : F_1 \in \mathbf{F}_1 \wedge F_2 \in \mathbf{F}_2\}.$$

Definición 1.5.3. Se llama *subcolección* de una colección dada $\mathbf{F} \subset \mathbf{P}(X)$, a todo subconjunto de $\mathbf{F}$.

Definición 1.5.4. Se dice que una colección $\mathbf{F} \subset \mathbf{P}(X)$ constituye un *cubrimiento* de una parte $A \subset X$, si $A \subset \cup \mathbf{F}$. En particular $\mathbf{F}$ es un cubrimiento de X si y solo si $\cup \mathbf{F} = X$.

Definición 1.5.5. Se llama *partición* de X a un cubrimiento $\mathbf{F}$ de X tal que

$$1) \quad \forall F \in \mathbf{F}, F \neq \emptyset,$$

$$2) \quad F \in \mathbf{F} \wedge G \in \mathbf{F} \wedge (F \neq G) \Rightarrow F \cap G = \emptyset$$

La condición 2) se expresa diciendo que los elementos de $\mathbf{F}$ son disjuntos dos a dos.

Ejemplo 15. La colección formada por todos los subconjuntos de $\mathbb{N}$ de la forma $\{n, n+1\}$ con $n \geq 0$ es un cubrimiento de $\mathbb{N}$ que no constituye una partición.

En 1.3 definimos la operación producto cartesiano entre dos conjuntos X y Y de la manera siguiente:

$$X \times Y = \{(x, y) : (x \in X) \wedge (y \in Y)\}, \text{ lo cual da lugar a un nuevo conjunto.}$$

Definición 1.5.6. Dados dos conjuntos X y Y , un subconjunto R de $X \times Y$ se llama una *relación* entre X y Y . En lugar de $(x, y) \in R$ se emplea a veces la notación xRy . Cuando $X = Y$ una relación R entre X y Y se llama una *relación sobre X*

Dada una relación R sobre X y $A \subset X$, la relación $R \cap (A \times A)$ sobre A se llama *relación inducida* sobre A por R y se denota R_A .

Dadas dos relaciones R y S sobre los conjuntos X y Y respectivamente se puede definir la *relación producto* sobre $X \times Y$, que se denota $R \times S$, mediante $(x, y)R \times S(x', y') \Leftrightarrow (xRx') \wedge (ySy')$

Definición 1.5.7. Una relación R sobre X se llama *relación de equivalencia* si se cumplen las propiedades siguientes:

RE 1 $\forall x \in X, xRx$ (reflexividad)

RE 2 $(xRy) \Rightarrow (yRx)$ (simetría)

RE 3 $(xRy) \wedge (yRz) \Rightarrow (xRz)$ (transitividad)

Si xRy , se dice que x y y son equivalentes,

Dado un punto $x \in X$, al conjunto $[x] := \{y \in X : yRx\}$ se le llama *clase de equivalencia del punto x* con respecto a la relación R , y a un elemento cualquiera $y \in [x]$ se le llama un *representante* de la clase.

La colección formada por todas las clases de equivalencia correspondientes a una relación de equivalencia R , da lugar a una partición de X , es decir, que se cumple,:

$$1) \cup \{[x] : x \in X\} = X;$$

$$2) \forall x \in X, [x] \neq \emptyset$$

$$3) \forall x, y \in X, [x] \neq [y] \Rightarrow [x] \cap [y] = \emptyset$$

lo cual puede comprobar el lector sin dificultad, utilizando las propiedades de R .

Recíprocamente se tiene la proposición siguiente:

Proposición 1.5.8. Sea $\mathbf{F}$ una partición de X . Entonces la relación R definida por

$$xRy \Leftrightarrow \exists F \in \mathbf{F}, (x \in F) \wedge (y \in F),$$

es una relación de equivalencia sobre X , para la cual $\mathbf{F}$ es la colección de las clases de equivalencia.

Demostración: Probemos que R satisface las propiedades de una relación de equivalencia.

(RE 1) Sea $x \in X$, entonces, por definición, existirá un conjunto $F \in \mathbf{F}$ tal que $x \in F$, y por lo tanto, xRx .

- (RE 2) Si xRy , entonces existe $F \in \mathbf{F}$ tal que F contiene simultáneamente a x y y , luego yRx .
- (RE 3) Si xRy y yRz es porque existen conjuntos F y G en $\mathbf{F}$ tales que F contiene simultáneamente a x y y , y G contiene simultáneamente a y y z . Pero como $\mathbf{F}$ es una partición, se tendría que $F \cap G = \emptyset$ si $F \neq G$. Como hemos visto, el elemento $y \in F \cap G$ por lo tanto, $F = G$, de donde se deduce que x y z pertenecen a un mismo conjunto de $\mathbf{F}$, es decir xRz , con lo cual queda probada la proposición enunciada. ■

Definición 1.5.9. Dado, un conjunto X y una relación de equivalencia R sobre X , al conjunto X/R de todas las clases de equivalencia se llama *conjunto cociente* de X por R .

Es precisamente a este procedimiento de obtención de conjuntos cocientes al que nos referíamos al inicio del epígrafe, el cual resulta ser uno de los métodos más importantes y eficientes en la matemática para la introducción de nuevos conceptos y estructuras como veremos en otras partes del libro.

Definición 1.5.10. Dada una relación de equivalencia R sobre un conjunto X y $A \subseteq X$, se llama *saturado de A con respecto a R* y se denota $\text{sat}_R A$, al conjunto

$$\text{sat}_R A := \cup \{[x] : x \in A\}$$

$\text{sat}_R A$ es el conjunto de todos los $y \in X$ que son equivalentes a algún $x \in A$.

Ejemplo 16. Definamos sobre el conjunto de los números reales $\mathbb{R}$ una relación mediante,

$$xRy \Leftrightarrow x - y \in \mathbb{Z}$$

No es difícil comprobar que esta relación es de equivalencia, lo cual es consecuencia de que

- 1) $0 \in \mathbb{Z}$;
- 2) el opuesto de todo número entero es un número entero;
- 3) la suma de números enteros es un número entero.

Por otra parte la clase de equivalencia de cualquier número x es $[x] = \{x + k : k \in \mathbb{Z}\}$, y por lo tanto,

$[n] = \mathbb{Z}$, para todo $n \in \mathbb{Z}$.

El lector puede comprobar que $\text{sat}_R[0, 1] = \mathbb{R}$. Si sustituimos $\mathbb{Z}$ por $\mathbb{N}$ ¿por qué la relación resultante no es de equivalencia?

1.6 RELACIONES FUNCIONALES. FUNCIONES

A partir de la noción de relación, se puede introducir el concepto *función* estrictamente ligado a la expresión matemática de las *magnitudes variables*, las cuales a su vez representan de la manera más idónea los fenómenos en constante movimiento de la naturaleza que nos rodea.

Definición 1.6.1 Dados dos conjuntos X y Y , una relación $G \subset X \times Y$ la llamaremos *relación funcional* si

$$(x, y) \in G \wedge (x, y') \in G \Rightarrow y = y'.$$

Una *función o aplicación* de X en Y es una terna $(X, Y, G(f))$, donde X y Y son conjuntos y $G(f)$ es una relación funcional para la cual se cumple que

$$\forall x \in X \exists y \in Y \text{ tal que } (x, y) \in G(f).$$

Como se ve de la propia definición, cada primer elemento $x \in X$ de un par ordenado en $G(f)$ determina unívocamente su segundo elemento $y \in Y$, al cual denotaremos por $f(x)$.

La relación funcional $G(f)$ se llama *grafo* de f y se denota también Por G_f

$$G_f = \{(x, f(x)) : x \in X\}.$$

El conjunto X se llama *dominio o conjunto de partida* de f y Y se llama *conjunto de llegada* de f . Se dice que f toma valores en Y .

Tratando de conservar el razonamiento conjuntista que nos condujo a la definición del concepto *función*, diremos que dos aplicaciones $(X_1, Y_1, G(f_1))$ y $(X_2, Y_2, G(f_2))$ son iguales si y solo si $X_1 = X_2$, $Y_1 = Y_2$ y $G(f_1) = G(f_2)$

Para una aplicación f se emplea la notación $f = X \rightarrow Y$ o $X \xrightarrow{f} Y$, donde X es el dominio y Y es el conjunto de llegada de f . También, en ocasiones emplearemos la notación

$$x \rightsquigarrow f(x)$$

Ejemplo 17. La relación mediante la cual hacemos corresponder a cada número entero el propio número y su opuesto, no define una aplicación de

$\mathbb{Z}$ en $\mathbb{Z}$ ya que a cada elemento n del dominio, le hacemos corresponder dos elementos $n, -n$ y no uno solo en el conjunto de llegada.

Ejemplo 18. En la figura 7 se observa un conjunto en el espacio producto $X \times Y$, el cual representa el grafo de una función, mientras que los conjuntos de las figuras 8 y 9 no representan grafos de funciones de X en Y . (Interprete las figuras y explique nuestra afirmación.)

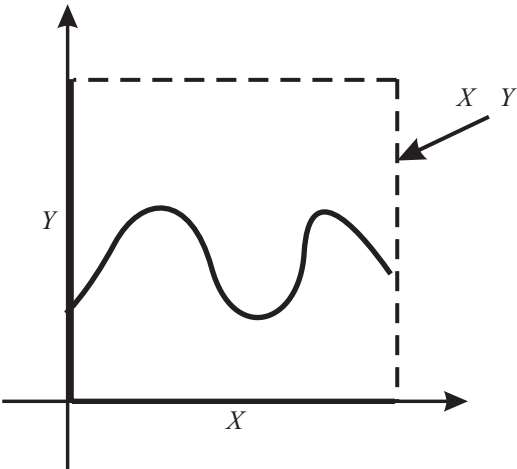


Figura 7

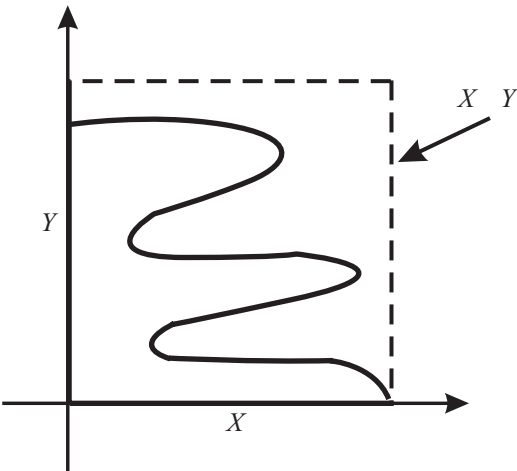


Figura 8

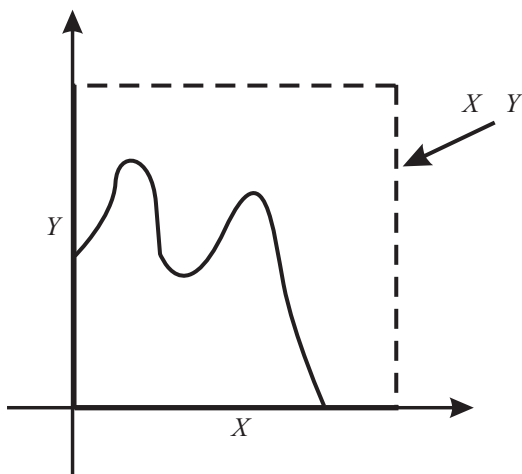


Figura 9

Definición 1.6.2. Dada una aplicación $f : X \longrightarrow Y$, $A \subset X$ llamaremos *imagen* de A por f y la denotaremos por $f[A]$, al subconjunto de Y

$$f[A] := \{f(x) : x \in A\}.$$

Cuando $A = X$ se llama a $f[X]$ *imagen* o *rango* de f .

Si $\mathbf{F}$ es una colección de subconjuntos de X , llamaremos *imagen* de $\mathbf{F}$ por f y la denotaremos por $f\{\mathbf{F}\}$, al subconjunto de Y

$$f\{\mathbf{F}\} := \{f(x) : \exists A \in \mathbf{F} \wedge x \in A\}$$

Definición 1.6.3. Decimos que una aplicación $f : X \longrightarrow Y$ es:

sobreyectiva si $\forall y \in Y, \exists x \in X$, tal que $f(x) = y$;

inyectiva si $f(x) = f(y) \Rightarrow x = y$;

biyectiva si es simultáneamente sobreyectiva e inyectiva.

Las nociones de sobreyectividad, inyectividad y biyectividad dependen no solamente de la *ley de correspondencia* f , sino también de los conjuntos de dominio y llegada, como nos muestra el ejemplo siguiente:

Ejemplo 19. Consideremos la aplicación $f : \mathbb{R} \longrightarrow \mathbb{R}$ definida mediante $f(x) = x^2$.

Evidentemente $(f, \mathbb{R}, \mathbb{R})$ no es sobreyectiva ya que el cuadrado de todo número real es siempre un número real positivo, y por lo tanto, todo número real negativo se queda fuera de la imagen de f . Por otra parte, $(f, \mathbb{R}, \mathbb{R})$ tampoco es inyectiva ya que todo número real positivo, como elemento del conjunto

imagen de f por ser el resultado de aplicar f a un número real x , y a su opuesto $-x$.

Sin embargo, el lector puede comprobar las proposiciones siguientes:

- $(f, A, \mathbb{R}_+)$ es sobreyectiva para cualquier $A \subset \mathbb{R}$ tal que $A \supset \mathbb{R}_+$, o $A \supset \mathbb{R}_-$
- $(f, \mathbb{R}_+, \mathbb{R}_+)$, y $(f, \mathbb{R}_-, \mathbb{R}_+)$ son inyectivas.
- Deduzca de a) y b) que $(f, \mathbb{R}_+, \mathbb{R}_+)$, y $(f, \mathbb{R}_-, \mathbb{R}_+)$ son biyectivas.

La no inyectividad de una aplicación puede ser representada gráficamente según nos muestra la figura 10.

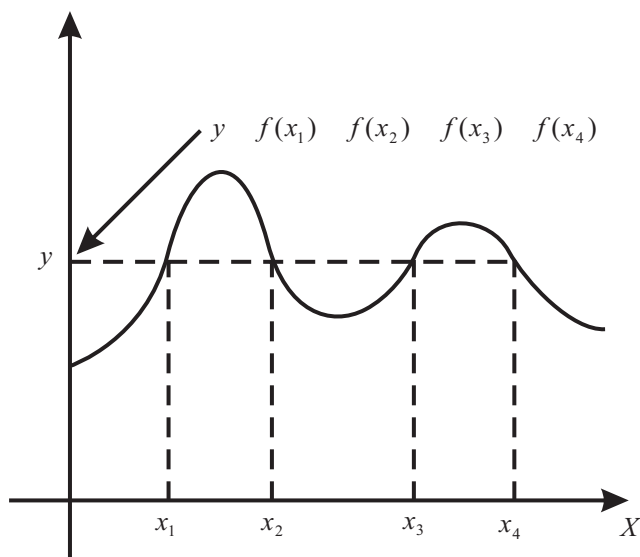


Figura 10

Observación: A partir de cualquier aplicación $f : X \longrightarrow Y$ podemos construir una aplicación sobreyectiva si sustituimos Y por el conjunto imagen $f[X]$.

Un procedimiento para construir una aplicación inyectiva o biyectiva a partir de la triada (f, X, Y) y “conservando la ley de correspondencia f ” será explicado posteriormente (véase descomposición canónica).

Observación: La importancia del procedimiento recién mencionado, consiste en que el concepto de equivalencia entre conjuntos puede ser expresado de una manera más concreta utilizando la noción de función, ya que:

$A \sim B \iff \exists f : A \longrightarrow B$ tal que (f, A, B) es biyectiva.

Definición 1.6.4. Dada una aplicación $f : X \longrightarrow Y$ y $B \subset Y$ se llama *imagen inversa* o *recíproca* de B por f y se denota por $f^{-1}[B]$ al conjunto $f^{-1}[B] := \{x \in X : f(x) \in B\}$.

En particular, si $B \cap f[X] = \emptyset$, entonces $f^{-1}[B] = \emptyset$. Si $B = \{y\}$ se denota a $f^{-1}[\{y\}]$ por $f^{-1}(y)$, o también, $[f = y]$.

La siguiente proposición recoge algunas propiedades importantes de la imagen recíproca.

Proposición 1.6.5. Sea $f : X \longrightarrow Y$, entonces:

- 1) $\forall A \subset X, A \subset f^{-1}[f[A]]$ (la igualdad ocurre si f es inyectiva);
- 2) $\forall B \subset Y, f[f^{-1}[B]] \subset B$ (la igualdad ocurre si f es sobreyectiva);
- 3) $\forall \mathbf{F} \subset \mathbf{P}(X), f[\cup\{F : F \in \mathbf{F}\}] = \cup\{f[F] : F \in \mathbf{F}\};$
- 4) $\forall \mathbf{F} \subset \mathbf{P}(X), f[\cap\{F : F \in \mathbf{F}\}] \subset \cap\{f[F] : F \in \mathbf{F}\};$
- 5) $\forall \mathbf{F} \subset \mathbf{P}(Y), f^{-1}[\cup\{G : G \in \mathbf{F}\}] = \cup\{f^{-1}[G] : G \in \mathbf{F}\};$
- 6) $\forall \mathbf{F} \subset \mathbf{P}(Y), f^{-1}[\cap\{G : G \in \mathbf{F}\}] = \cap\{f^{-1}[G] : G \in \mathbf{F}\};$
- 7) $\forall B \subset Y, f^{-1}[B'] = (f^{-1}[B])'.$

Demostración: 1) Por definición $f^{-1}[f[A]] = \{x \in X : f(x) \in f[A]\}$, por lo cual obviamente este conjunto contiene al conjunto A .

Demostraremos que si f es inyectiva entonces también se cumple la inclusión inversa. En efecto, si $x \in f^{-1}[f[A]]$, por definición tendremos que $f(x) \in f[A]$ y, por tanto, existe $y \in A$ tal que $f(x) = f(y)$. Por ser f inyectiva se tiene que $x = y$, de donde deducimos que $x \in A$, con lo cual queda probada la inclusión inversa y con ello la igualdad $A = f^{-1}[f[A]]$

3) $y \in f[\cup\{F : F \in \mathbf{F}\}] \Leftrightarrow \exists x \in \cup\{F : F \in \mathbf{F}\}$ tal que $f(x) = y \Leftrightarrow$ para algún $F \in \mathbf{F}, y \in f[F] \Leftrightarrow y \in \cup\{f[F] : F \in \mathbf{F}\}.$

7) $x \in f^{-1}[B'] \Leftrightarrow f(x) \in B' \Leftrightarrow f(x) \notin B \Leftrightarrow x \notin f^{-1}[B] \Leftrightarrow x \in (f^{-1}[B])'$

La demostración de las demás propiedades, se dejan de ejercicio al lector. ■

Definición 1.6.6. Si $f : Z \longrightarrow Y$ es inyectiva se puede definir la función *inversa* de f , la cual se denota por f^{-1} y tiene por dominio a $f[X]$, por rango

a X y su grafo viene dado por

$$G(f^{-1}) = \{(y, x) \in f[X] \times X : (x, y) \in G(f)\} .$$

Ejemplo 20. La función inversa a la función $(x^2, \mathbb{R}_+, \mathbb{R}_+)$ definida en el ejemplo 17, es la función $(\sqrt{x}, \mathbb{R}_+, \mathbb{R}_+)$, mientras que la función inversa a $(x^2, \mathbb{R}_-, \mathbb{R}_+)$, dada en el mismo ejemplo es $(-\sqrt{x}, \mathbb{R}_+, \mathbb{R}_-)$.

Definición 1.6.7. Dadas dos funciones $f : X \longrightarrow Y$ y $g : Y \longrightarrow Z$, se define la función compuesta $g \circ f : X \longrightarrow Z$, cuyo grafo viene dado por

$$G(g \circ f) = \{(x, z) \in X \times Z : \exists y \in Y \text{ tal que}$$

$$(x, y) \in G(f) \wedge (y, z) \in G(g)\}$$

Si $g \circ f$ es sobreyectiva (inyectiva), entonces necesariamente g es sobreyectiva (f es inyectiva).

Si f y g son inyectivas, entonces $g \circ f$ también lo es y se tiene

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

Veamos algunas ejemplos de aplicaciones importantes que utilizaremos en el resto del texto:

- 1) Restricción de una aplicación a un subconjunto del dominio: dada una aplicación $f : X \longrightarrow Y$ y $A \subset X$ se define la *restricción* de f a A y se denota por $f|_A : A \longrightarrow Y$, a la aplicación cuyo grafo viene dado por

$$G_f|_A := \{(x, y) \in G_f : x \in A\} := G_f \cap (A \times Y).$$

Recíprocamente, si $A \subset X$ y $g : A \longrightarrow Y$ es una aplicación dada, otra aplicación $f : X \longrightarrow Y$ tal que $f|_A = g$ se llama una *prolongación* de g a X con respecto a Y .

- 2) Aplicación constante: una aplicación $f : X \longrightarrow Y$ se llama *constante* si para algún elemento fijo $a \in Y$ se tiene

$$\forall x \in X, f(x) = a,$$

En este caso se utiliza la notación $\langle a \rangle$.

- 3) Aplicación identidad: se llama aplicación *identidad* en X y se denota por $1_X : X \longrightarrow X$ a la biyección $1(x) = x$.

Dado $A \subset X$, la restricción de la identidad $1_{X|A}$ se llama *inyección canónica* de A en X y se denota por i_A .

- 4) Función característica de un subconjunto: dados un conjunto X y $A \subset X$, se llama *función característica* del subconjunto A y se denota por $\mathcal{X}_A$ a la función $\mathcal{X}_A : X \longrightarrow \mathbb{R}$ definida por

$$\mathcal{X}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \in A' \end{cases}$$

- 5) Aplicación proyección: dados dos conjuntos X y Y , la aplicación $\pi_1 : X \times Y \longrightarrow X$ definida por $\pi_1(x, y) = x$ se llama *proyección* sobre X . Análogamente se define π_2 .

- 6) Aplicación sobreyección canónica: si $\mathbf{R}$ es una relación de equivalencia sobre X , se define la *sobreyección canónica* $\varphi_R : X \longrightarrow X/R$ mediante $\varphi_R(x) = [x]$.

Cuando la relación $\mathbf{R}$ queda clara del contexto se utiliza simplemente la notación φ .

Es claro que si $A \subset X$, $\text{sat}_R A = \varphi^{-1}[\varphi[A]]$.

Finalmente pasemos a ver la llamada *descomposición canónica* de una aplicación $f : X \longrightarrow Y$ que nos muestra el procedimiento a que nos referimos anteriormente y que consiste en construir una aplicación inyectiva a partir de la triada (f, X, Y) conservando la ley de correspondencia f .

Definición 1.6.8. Si $\mathbf{R}$ es una relación de equivalencia sobre X y $f : X \longrightarrow Y$ es una aplicación, decimos que f es compatible con $\mathbf{R}$ si $f(x) = f(y) \Rightarrow x\mathbf{R}y$. Recíprocamente, si definimos la relación $\mathbf{R}_f$ mediante $x\mathbf{R}_f y \Leftrightarrow f(x) = f(y)$, esta relación es de equivalencia sobre X y se llama relación de equivalencia asociada a f .

Es claro que f es compatible con $\mathbf{R}_f$. Si denotamos por φ la sobreyección canónica asociada a $\mathbf{R}_f$, se tiene la siguiente descomposición llamada *descomposición canónica* de $f : f = \bar{f} \circ \varphi$ donde $\bar{f} : X/\mathbf{R}_f \longrightarrow Y$ viene dada por

$$\forall [x] \in X/\mathbf{R}_f, \bar{f}([x]) = f(x).$$

Esta descomposición se expresa gráficamente diciendo que el siguiente diagrama es conmutativo (fig. 11).

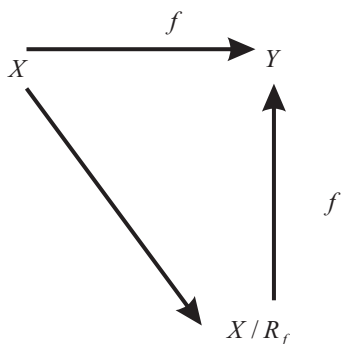


Figura 11

Observación: Lo que hacemos para construir $\bar{f}$ a partir de f , es construir una partición $\mathbf{F}$ de X agrupando en cada elemento $F \in \mathbf{F}$ a todos los elementos de X que tienen una misma imagen en Y mediante f . Si consideramos $\mathbf{F}$ como subconjunto de $P(X)$, $\mathbf{F}$ coincide precisamente con X/R_f . De esta construcción se desprende que $\bar{f}$ está bien definido si ponemos $\bar{f}([x]) = f(x)$, ya que f toma el mismo valor para cualquier representante de la clase $[x]$.

Es importante destacar que $\bar{f}$ siempre es inyectiva y que si f es sobreyectiva, entonces $\bar{f}$ es biyectiva.

Ejercicios de Autocontrol

- Sean las funciones

$$f : \mathbb{N} \rightarrow \mathbf{P}(\mathbb{N})$$

$$g : \mathbf{P}(\mathbb{N}) \longrightarrow \mathbb{N}$$

definidas mediante:

$$f(n) = \{1, 2, \dots, n\}$$

$$g(A) = \sum_{k \in A} k$$

- Analice la inyectividad y sobreyectividad de cada una de las funciones f y g .
- Calcule $g \circ f$ y diga si es sobreyectiva o inyectiva.

- Sea la función

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

definida por

$$f(x) = x^2$$

Obtenga la expresión de la descomposición canónica de f .

1.7 RELACIONES DE ORDEN. CONJUNTOS PARCIALMENTE, TOTALMENTE Y BIEN ORDENADOS

En este epígrafe expondremos varios conceptos relacionados con la idea de ordenamiento de los elementos de un conjunto, limitándonos a dar solamente las nociones elementales que necesitaremos para el posterior desarrollo del texto.

Definición 1.7.1. Una relación $\mathbf{R}$ sobre un conjunto X se llama de *preorden* si

- 1) $\forall x \in X, x\mathbf{R}x$ (reflexiva)
- 2) $(x\mathbf{R}y) \wedge (y\mathbf{R}z) \Rightarrow x\mathbf{R}z$ (transitiva)

Si además la relación $\mathbf{R}$ cumple la propiedad adicional

- 3) $(x\mathbf{R}y) \wedge (y\mathbf{R}x) \Rightarrow x = y$ (antisimétrica) entonces se dice que $\mathbf{R}$ es una relación de orden sobre X .

Al par formado por un conjunto X y una relación de preorden (de orden) se llama conjunto *preordenado* (*ordenado*).

Para las relaciones de preorden y orden se emplean indistintamente las notaciones $x < y$ o $y > x$, que se leen respectivamente *x es anterior a y* ó *y es posterior a x* y también $x \leq y$ ó $y \geq x$ que se leen respectivamente *x es menor o igual que y* ó *y es mayor o igual que x*.

Dada una relación $\leq$ sobre X , $(x, y \in X)$, el símbolo $x < y$ ó $y > x$, que se lee respectivamente *x es estrictamente menor que y* ó *y es estrictamente mayor que x*, significa

$$(x \leq y) \wedge (x \neq y)$$

La única relación que es a la vez de equivalencia y de orden es la igualdad. Con frecuencia ocurre que existen pares de elementos en un conjunto ordenado $(X, \leq)$ que no pueden ser comparables con ayuda de la relación de orden $\leq$, es decir, se encuentran elementos x y y en X para los cuales x no es menor que y ni y es menor que x ($(x \not\leq y) \wedge (y \not\leq x)$). En tales casos se dice que estamos en

presencia de un conjunto *parcialmente ordenado* y al orden correspondiente se le llama un *orden parcial*.

Ejemplo 21. En la figura 12 se representa un “árbol” ordenado mediante la relación: $x \leq y \Leftrightarrow$ se puede pasar de x a y realizando siempre un movimiento ascendente.

El lector puede comprobar que $\leq$ así definido es una relación de orden.

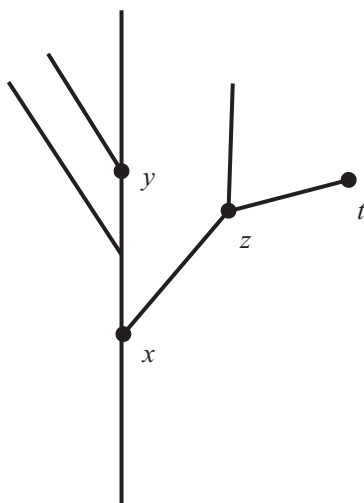


Figura 12

Es claro que $x \leq y$, $x \leq z$ y $z \leq t$, pero z y t no son comparables con y ya que para pasar de z a y es necesario “descender” primero por la “rama” que contiene a z y para pasar de y a z hay que “descender” por el tronco del árbol.

Veamos otros ejemplos de conjuntos parcialmente ordenados.

Ejemplo 22. Sea $C[a, b]$ el conjunto de todas las funciones reales continuas definidas en el segmento $[a, b]$. La relación:

$f \leq g \Leftrightarrow \forall x \in [a, b], f(x) \leq g(x)$, es un orden parcial en $C[a, b]$.

Ejemplo 23. Dado un conjunto X , se puede establecer un orden parcial en $\mathbf{P}(X)$ mediante la inclusión usual de subconjuntos de X

$$A \leq B \Leftrightarrow A \subset B.$$

Es claro que dos subconjuntos disjuntos de X no pueden ser comparables

por la relación $\leq$ ¿Cuáles otros conjuntos tampoco pueden compararse?

Ejemplo 24. El conjunto de todos los números naturales resulta parcialmente ordenado para el orden:

$$a \leq b \Leftrightarrow b \text{ es divisible por } a.$$

Demuestre que esta relación es de orden en $\mathbb{N}$.

Dentro de las relaciones de orden juegan un papel fundamental las siguientes;

Definición 1.7.2. Un orden $\leq$ sobre un conjunto X se llama *filtrante* si cumple

$$\forall x, y \in X, \exists z \in X : (x \leq z) \wedge (y \leq z).$$

Se llama *conjunto dirigido* a un conjunto ordenado con un orden filtrante. El concepto *cota* es de gran importancia en el estudio de los conjuntos ordenados y por ello introduciremos las definiciones siguientes:

Definición 1.7.3. Sea $(X, \leq)$ un conjunto ordenado y $A \subset X$. Un elemento $a \in X$ se llama *cota superior* o *mayorante* de A (notación: $A \leq a$), si

$$x \in A \Rightarrow x \leq a.$$

Un elemento $b \in X$ se llama *cota inferior* o *minorante* de A (notación: $A \geq b$), si

$$x \in A \Rightarrow b \leq x.$$

Notemos que el conjunto de las mayorantes (minorantes) de un subconjunto A puede ser vacío o por el contrario estar formado por más de un elemento, como nos muestra el ejemplo siguiente:

Ejemplo 25. Con respecto a la relación definida sobre $\mathbf{P}(X)$, en el ejemplo 23 cualquier subconjunto propio $\mathbf{F}$ de $\mathbf{P}(X)$ tiene siempre un mayorante que es $\cup \mathbf{F}$ y también todo subconjunto de X que contenga a $\cup \mathbf{F}$. Sin embargo, si eliminamos al conjunto $\emptyset$ de $\mathbf{P}(X)$, entonces un subconjunto $\mathbf{F}$ de $\mathbf{P}(X)$ tiene una cota inferior si y solo si $\cap \mathbf{F} \neq \emptyset$.

Si el conjunto de mayorantes (minorantes) de una parte $A \subset X$ no es vacío, se dice, que A es *acotado superiormente* (*inferiormente*). Si $A \subset X$ es acotado superior e inferiormente se dice que A es *acotado* como subconjunto del conjunto ordenado X .

Definición 1.7.4. Si el conjunto de las cotas superiores o mayorantes de una

parte $A \subset X$ tiene un menor elemento a , este elemento se llama *supremo* de A en X (notación: $a = \sup A$ o $a = \sup\{x : x \in A\}$). Es decir,

$$a = \sup A \Leftrightarrow a \geq A \wedge (c \geq A \Rightarrow c \geq a).$$

Si el conjunto de las cotas inferiores o minorantes de una parte $A \subset X$ tiene un mayor elemento b , este elemento se llama *ínfimo* de A en X (notación: $b = \inf A$ o $b = \inf\{x : x \in A\}$). O sea, $b = \inf A \Leftrightarrow b \leq A \wedge (d \leq A \Rightarrow d \leq b)$.

si a demás el $\sup A \in A$ ($\inf A \in A$) entonces se le llama máximo de A (mínimo de A) y se denota mediante $\max A$ ($\min A$)

Se tiene que $a = \inf A$ ($b = \sup B$) si y solo si a es el máximo (b es el mínimo) del conjunto de las cotas inferiores (superiores) de A .

Finalmente diremos que un elemento a de un conjunto ordenado X se llama *maximal* (minimal) si no existe ningún otro $x \in X$, $x \neq a$ tal que $a \leq x$ ($x \leq a$).

Observación: Notemos que tanto el elemento máximo como el mínimo de un conjunto cuando existen son únicos, no ocurriendo así con los elementos maximales y minimales, según nos muestra el ejemplo 21 (véase la figura 12), donde todos los extremos de las ramas del “árbol” son elementos maximales y, sin embargo no existe en este caso elemento máximo. ¿Existe elemento mínimo en el ejemplo mencionado?

Tanto el elemento maximo como el mínimo son comparables con el resto del conjunto, mientras que los elementos maximales y minimales solo poseen la propiedad de no permitir ningún elemento posterior ni anterior respectivamente, y sin embargo, pueden haber otros que no sean comparables con ellos (vea la figura 12).

Veamos a continuación algunas categorías importantes de conjuntos ordenados.

Definición 1.7.5. Un conjunto ordenado $(X, \leq)$ se llama *totalmente ordenado* si

$$(x \in X) \wedge (y \in X) \Rightarrow (x \leq y) \vee (y \leq x).$$

Un subconjunto A de un conjunto ordenado es una *cadena* si $(A \leq_A)$ está totalmente ordenado. Una cadena A en X se llama *maximal* si no forma parte propia de ninguna otra cadena perteneciente a X .

Un conjunto ordenado $(X, \leq)$ se llama *inductivamente ordenado* si toda cadena en X tiene una cota superior.

Un conjunto ordenado $(X, \leq)$ se dice que está *bien ordenado*, si en él cualquier subconjunto no vacío tiene un elemento mínimo.

Ejemplo 26. El conjunto $\mathbb{R}$ de los números reales con el orden usual es

totalmente ordenado. Sin embargo, $\mathbb{R}$ no es inductivamente ordenado ni bien ordenado, ya que el propio $\mathbb{R}$ representa una cadena que no tiene una cota superior en $\mathbb{R}$ ni tampoco un elemento mínimo.

Ejemplo 27. El intervalo $[0, 1]$ con el orden usual es un conjunto totalmente ordenado, y además, inductivamente ordenado, pero no está bien ordenado ya que, por ejemplo, el conjunto de todos los números racionales estrictamente positivos contenidos en $[0, 1]$ no tiene elemento mínimo.

Ejemplo 28. Los conjuntos ordenados dados en los ejemplos del 21 al 24 son todos parcialmente ordenados pero no totalmente ordenados. ¿Cuáles de ellos son además inductivamente o bien ordenados?

En un conjunto totalmente ordenado $(X, \leq)$ se pueden definir los intervalos siguientes:

$$1.]a, b[:= \{x \in X : a < x < b\} \text{ (abierto)}$$

$$2. \left. \begin{array}{l}]a, b] := \{b\} \cup]a, b[\\ [a, b[:= \{a\} \cup]a, b[\end{array} \right\} \text{ (semiabiertos)}$$

$$3. [a, b] = \{a, b\} \cup]a, b[\text{ (cerrado)}$$

$$4.]a, \rightarrow[:= \{x \in X : a < x\},$$

$$[a, \rightarrow[:= \{a\} \cup]a, \rightarrow[$$

$$5.]\leftarrow, a[:= \{x \in X : x < a\}$$

$$]\leftarrow, a] := \{a\} \cup]\leftarrow, a[$$

La figura 13 nos muestra la relación que existe entre las diferentes categorías de conjuntos ordenados que hemos definido.

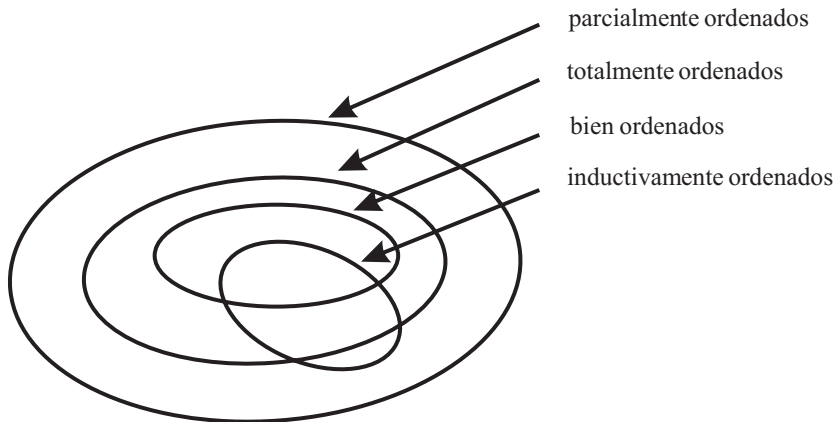


Figura 13

Finalmente introduciremos un concepto de correspondencia entre los elementos de dos conjuntos ordenados y que conserva el orden.

Definición 1.7.6. Si $(X, \leq)$ y $(Y, \leq)$ son conjuntos ordenados, entonces la función $f : (X, \leq) \rightarrow (Y, \leq)$ se llama *creciente* (*decreciente*) cuando

$$(x \in X) \wedge (y \in X) \wedge (x \leq y) \Rightarrow f(x) \leq f(y) \quad (f(x) \geq f(y))$$

La aplicación f se dice *estrictamente creciente* (*estrictamente decreciente*) si $(x < y) \Rightarrow f(x) < f(y)$ ($f(x) > f(y)$)

Se dice que f es *monótona* si ella es creciente o decreciente.

Una aplicación $f : X \rightarrow Y$ donde Y es un conjunto ordenado se llama *acotada superiormente*, *acotada inferiormente* o simplemente, *acotada*, si el conjunto $f[X]$ lo es respectivamente en Y .

En el próximo epígrafe estudiaremos las correspondencias biunívocas entre conjuntos totalmente ordenados que conservan las relaciones de orden.

Ejercicios de Autocontrol

- Compruebe la veracidad del diagrama expuesto en la figura 13, es decir:
 - todo conjunto bien ordenado es totalmente ordenado;
 - existen conjuntos inductivamente ordenados que no son ni totalmente ordenados ni bien ordenados;
 - enuncie un ejemplo de un conjunto inductivamente ordenado que es totalmente ordenado pero no bien ordenado.

2. a) Pruebe que la aplicación

$$F : (C[a, b], \leq) \longrightarrow (\mathbf{P}(\mathbb{R}), \subset)$$

(donde los órdenes respectivos vienen definidos en los ejemplos 22 y 23) $F(f) = f[a, b]$ (imagen de f) no es una aplicación monótona.

- b) Diga qué condición tienen que cumplir dos funciones f y g para que de la propiedad $f \leq g$ se tenga $F(f) \subseteq F(g)$

- c) Calcule el supremo en $(\mathbf{P}(X), \subseteq)$ de la imagen por F del subconjunto A de $C[a, b]$ definido por:

$$A = \{f \in C[a, b] : f : [a, b] \longrightarrow [0, 1]\}$$

1.8 AXIOMA DE ELECCIÓN. TEOREMA DE ZERMELO Y OTRAS PROPOSICIONES EQUIVALENCIAS A ELLOS. INDUCCIÓN TRANSFINITA

Dado un conjunto X , una *familia de partes* de X está dada por una aplicación $\alpha \rightsquigarrow X_\alpha$, cuyo dominio es un conjunto arbitrario I , llamado *conjunto de índices*, y su conjunto de llegada es $\mathbf{P}(X)$. A esta familia la denotaremos por

$$(X_\alpha)_{\alpha \in I}.$$

Llamamos *subfamilia* de $(X_\alpha)_{\alpha \in I}$ a la restricción $(X_\alpha)_{\alpha \in J}$ de $(X_\alpha)_{\alpha \in I}$ a un subconjunto J de I .

A toda familia de conjuntos $(X_\alpha)_{\alpha \in I}$ se le puede hacer corresponder una colección llamada colección asociada a dicha familia: $\{X_\alpha : \alpha \in I\}$.

Recíprocamente dada una colección $\mathbf{F}$ y tomando como conjunto de índices la propia colección, la aplicación identidad nos define una familia de conjuntos.

Para familias de conjuntos se definen, mediante su colección asociada, las operaciones de unión e intersección así como los conceptos de partición y cubrimiento. Dada una familia $(X_\alpha)_{\alpha \in I}$ emplearemos la notación:

$$\bigcup_{\alpha \in I} X_\alpha := \cup \{X_\alpha : \alpha \in I\} \text{ (unión);}$$

$$\bigcap_{\alpha \in I} X_\alpha := \cap \{X_\alpha : \alpha \in I\} \text{ (intersección).}$$

Por definición se tiene que

$$\bigcup \{X_\alpha : \alpha \in \emptyset\} = \emptyset \text{ (unión vacía);}$$

$$\bigcap \{X_\alpha : \alpha \in \emptyset\} = X \text{ (intersección vacía).}$$

En 1.3 vimos la definición de producto cartesiano para una colección finita de conjuntos. Para una familia cualquiera de conjuntos se tiene la definición siguiente:

Definición 1.8.1. Se llama *producto cartesiano de la familia* $(X_\alpha)_{\alpha \in I}$ y se denota por

$$\prod \{X_\alpha : \alpha \in I\} \text{ o } \prod_{\alpha \in I} X_\alpha,$$

al conjunto de todas las funciones $x : I \rightarrow \cup \{X_\alpha : \alpha \in I\}$ con la propiedad de que para todo α , $x(\alpha) = x_\alpha \in X_\alpha$

Los elementos de $\prod_{\alpha \in I} X_\alpha$ suelen denotarse por $(x_\alpha)_{\alpha \in I}$ y cuando no hay dudas sobre el conjunto de índices, sencillamente (x_α) . El conjunto X_α se llama el α -ésimo factor del producto y a $x_\alpha \in X_\alpha$ la α -ésima coordenada del elemento x .

Si $A_\alpha \subset X_\alpha$ para todo α , se tiene que $\prod_{\alpha \in I} A_\alpha \subset \prod_{\alpha \in I} X_\alpha$

Si la familia de conjuntos es constante, es decir, si $X_\alpha = X$ para todo $\alpha \in I$, el producto $\prod_{\alpha \in I} X_\alpha$ se denota por X^I y por definición es precisamente el conjunto de todas las funciones de I en X . En particular $X^{\mathbb{N}}$ es el conjunto de las sucesiones de elementos de X .

Definición 1.8.2. Para cada $\alpha \in I$ se llama α -ésima proyección correspondiente al producto $\prod_{\alpha \in I} X_\alpha$ a la aplicación

$$\pi_\alpha : \prod_{\alpha \in I} X_\alpha \rightarrow X_\alpha,$$

tal que para todo $x = (x_\alpha) \in \prod_{\alpha \in I} X_\alpha$

$$\pi_\alpha(x) = x_\alpha.$$

Si $A \subset X_\alpha$ entonces $\pi_\alpha^{-1}[A]$ se llama *cilindro de base* A en $\prod_{\alpha \in I} X_\alpha$.

No es difícil comprobar (se deja al lector), que las aplicaciones proyección, satisfacen las propiedades siguientes:

1) Son sobreyectivas;

2) Si $A \subset \prod_{\alpha \in I} X_\alpha$, entonces $A \subset \prod_{\alpha \in I} \pi_\alpha[A]$.

A partir de las funciones proyección π_α se puede dar una representación de las aplicaciones $f : X \rightarrow \prod_{\alpha \in I} X_\alpha$, similar a la obtenida en el análisis clásico para las funciones con valores en $\mathbb{R}^n$. En efecto, para cada $\alpha \in I$, la función compuesta:

$$f_\alpha = \pi_\alpha \circ f : X \rightarrow X_\alpha$$

se llama α -ésima coordenada de la función f . Es obvio que toda función f de X en $\prod_{\alpha \in I} X_\alpha$ puede expresarse mediante la familia de funciones $(f_\alpha)_{\alpha \in I}$ de X en X_α .

Hasta el momento hemos supuesto cierto algo que intuitivamente es trivial, pero que resulta imposible demostrar partiendo de conceptos primarios. Nos referimos a la existencia de elementos en el conjunto $\prod_{\alpha \in I} X_\alpha$ para una familia arbitraria de conjuntos no vacíos $(X_\alpha)_{\alpha \in I}$.

En efecto, puede demostrarse (aunque no lo haremos aquí), que la afirmación:

A E 1 Si $(X_\alpha)_{\alpha \in I}$ es una familia de conjuntos no vacíos, entonces $\prod_{\alpha \in I} X_\alpha \neq \emptyset$; es equivalente a la siguiente:

A E 2 Sea M un conjunto cuyos elementos son conjuntos no vacíos M_α disjuntos dos a dos. Entonces existe un conjunto $\widetilde{M}$ con la propiedad siguiente: Cada elemento m_α de M pertenece a algún conjunto M_α , y además, $\widetilde{M}$ interseca a M_α , solamente en el elemento m_α .

Sin embargo, ninguno de estos postulados puede ser demostrado partiendo de los conceptos primarios de conjunto, pertenencia y correspondencia biunívoca entre conjuntos.

La afirmación A E 1 (o bien su expresión equivalente A E 2) se postula como cierta en muchas ramas de la matemática, donde la teoría de conjuntos juega un papel fundamental, entre ellas, la Topología, y recibe el nombre de *Axioma de Elección* (A E).

En otras palabras, el axioma de elección puede ser expresado también de las formas siguientes:

A E 3 Dado un conjunto M como en A E 2, existe siempre otro conjunto $\widetilde{M}$ que se puede construir seleccionando un elemento de cada conjunto $M_\alpha \in M$.

A E 4 Si para todo $\alpha \in I$, $X_\alpha \neq \emptyset$, entonces existe una aplicación de I en $\alpha \in I \cup X_\alpha$, que asigna a cada $\alpha \in I$ un elemento $x_\alpha \in X_\alpha$.

El axioma de elección tiene muchas otras formulaciones equivalentes a algunas de las cuales nos referiremos más adelante. Este postulado fue enunciado hace más de 70 años y desde su inicio, dió lugar a múltiples investigaciones de carácter teórico acerca del lugar que debía ocupar en la estructuración lógica de la matemática moderna.

Uno de los problemas más complejos y discutidos que surgen al fundamentar la teoría de conjuntos, es la legalidad al aplicar este axioma a conjuntos arbitrarios en los razonamientos que se hacen. No tenemos posibilidad de entrar aquí en su discusión pues ello complicaría sustancialmente el contenido del libro. No obstante, observaremos, que la renuncia al axioma reduce la posibilidad de diferentes construcciones en la teoría de conjuntos.

Resulta que no se puede obviar la aplicación del axioma de elección ni tan siquiera para la demostración de algunos de los teoremas más elementales de la teoría de conjuntos y del análisis matemático clásico. Tomemos, por ejemplo, las dos siguientes definiciones de continuidad para una función real f definida sobre la recta numérica:

1. La función f se llama continua en el punto x_0 , si para cada número real positivo ϵ se puede elegir otro número positivo δ , tal que para todos los x que satisfagan la desigualdad $|x - x_0| < \delta$ se tenga $|f(x) - f(x_0)| < \epsilon$.
2. La función f se llama continua en el punto x_0 , si para toda sucesión $x_1, x_2, \dots, x_n, \dots$, que converja al punto x_0 , la sucesión $f(x_1), f(x_2), \dots, f(x_n), \dots$, converge al punto $f(x_0)$.

Es conocido que estas dos definiciones son equivalentes. Analicemos la demostración usual de esta equivalencia.

Sea f continua en el punto x_0 (según la definición 1) y sea $x_1, x_2, \dots, x_n, \dots$, una sucesión que converge a x_0 . Entonces para cada $\epsilon > 0$ se puede hallar un $\delta > 0$, tal que para todos los x contenidos en el intervalo $(x_0 - \delta, x_0 + \delta)$, se tiene $|f(x) - f(x_0)| < \epsilon$. Tomando para un ϵ dado, el correspondiente δ , asociémosle a este δ un número natural N tal que $|x_0 - x_n| < \delta$ para todos los $n \geq N$. Entonces $|f(x_0) - f(x_n)| < \epsilon$ y como esto ocurre para cualquier $\epsilon > 0$, entonces la sucesión $f(x_1), f(x_2), \dots, f(x_n), \dots$, converge a $f(x_0)$. Luego, si la función es continua según la definición 1 también lo será según la definición 2.

Notemos que la demostración de esta afirmación no se apoya en el axioma de, elección ya que la “elección” del número N se puede llevar a cabo de forma

unívoca pues basta con tomar en cada caso el primer número natural N , tal que para todos los $n > N$ se tenga $|x_0 - x_n| < \delta$.

Sea ahora f continua en el punto x_0 (según la definición 2) y, demostremos que ella también es continua en el sentido de la definición 1.

Supongamos lo contrario, entonces existe un número $\epsilon > 0$, tal que para cualquier $\delta > 0$, en el intervalo $(x_0 - \delta; x_0 + \delta)$ se tienen puntos x_δ para los cuales $|f(x_0) - f(x_\delta)| \geq \epsilon$.

Dándole al número δ el valor $\delta_n = \frac{1}{n}$ y tomando para cada δ_n un cierto x_{δ_n} , al cual representaremos para simplificar mediante x_n , obtendremos una sucesión de puntos convergente al punto x_0 y que al mismo tiempo cumple $|f(x_0) - f(x_n)| \geq \epsilon$. Con esto queda probada la equivalencia entre las definiciones 1 y 2.

Examinemos más de cerca la segunda parte de esta demostración.

La existencia de los puntos x_δ satisfaciendo simultáneamente las dos condiciones $|x_0 - x_n| < \delta$ y $|f(x_0) - f(x_\delta)| \geq \epsilon$ no significa que nosotros podemos dar una regla para la construcción efectiva de un tal punto; para poder llegar a una contradicción, es suficiente suponer que el conjunto de estos puntos no es vacío. Por eso, la suposición de que la función f no es continua en el punto x_0 en el sentido de la definición 1, significa solamente, que para algún $\epsilon > 0$ y cualquier $\delta > 0$, el conjunto M_δ de los puntos x del intervalo $[x_0 - \delta, x_0 + \delta]$, para los cuales $|f(x) - f(x_0)| \geq \epsilon$, no es un conjunto vacío. El paso de la sucesión de conjuntos no vacíos

$$M_n = M_{\delta_n}$$

a la sucesión de puntos $x_n \in M_n$, se lleva a cabo, en general, por medio de la *elección arbitraria* en cada uno de los conjuntos M_n , de un punto, al cual denotamos mediante x_n . Notemos, que los conjuntos M_n se intersectan y más aún $M_{n+1} \subset M_n$, y por ello, para poder aplicar el axioma de elección en la forma en que fue formulado en A E 2 es necesario pasar de los conjuntos M_n a los conjuntos $M_n \setminus M_{n+1}$, denotar de entre ellos a los no vacíos por $M'_1, M'_2, \dots, M'_n, \dots$, y posteriormente en base al axioma de elección, elegir los puntos x_n .

Pasaremos ahora a estudiar la relación existente entre el axioma de elección y el concepto de orden. Pero primeramente enunciaremos la definición siguiente:

Definición 1.8.3. Sean $(X, \leq)$ y $(Y, \leq)$ dos conjuntos parcialmente ordenados y $f: X \rightarrow Y$ una aplicación biyectiva.

Diremos que esta aplicación *conserva el orden* si ella es creciente (ver 1.7.6).

La aplicación f se llama isomorfismo entre los conjuntos parcialmente ordenados $(X, \leq)$ y $(Y, \leq)$, si para todos los $x, y \in X$ se cumple: $f(x) \leq f(y) \Leftrightarrow x \leq y$.

Se dice que dos conjuntos ordenados tienen el mismo *tipo ordinal* si ellos son totalmente ordenados e isomorfos.

El tipo ordinal de un conjunto bien ordenado se llama *número ordinal transfinito* o más brevemente *número transfinito*.

Cuando lo que nos interesa no es la naturaleza de los elementos del conjunto, sino solo el orden parcial que en ellos existe, se puede, sin duda, considerar como idénticos a dos conjuntos parcialmente ordenados isomorfos. Así pues, vemos que el tipo ordinal es lo común que tienen todos los conjuntos ordenados isomorfos, de la misma forma que la potencia es lo común que tienen todos los conjuntos equivalentes (considerados independientemente de cualquier relación de orden que pueda existir en ellos).

Al tipo ordinal del conjunto $\mathbb{N}$ de los números naturales, se acostumbra a denotar por ω . Es evidente que a todo tipo ordinal le corresponde una potencia, aunque la afirmación recíproca no es cierta. Por ejemplo, al tipo ordinal ω le corresponde la potencia α_0 mientras que a la potencia a_0 le corresponde una cantidad infinita, incluso no numerable de diferentes tipos ordinales.

Observación: Al igual que se define la suma y el producto de potencias, se define también la suma y el producto de tipos ordinales (vea los ejercicios de control).

Por otra parte, para dos conjuntos bien ordenados $(X, \leq)$ y $(Y, \leq)$ con números transfinitos α y β respectivamente, diremos que $\alpha = \beta$ si $(X, \leq)$ y $(Y, \leq)$ son isomorfos, $\alpha < \beta$ si $(X, \leq)$ es isomorfo a algún segmento inicial $] \leftarrow a[$ de $(Y, \leq)$ y $\alpha > \beta$ si, al contrario, $(Y, \leq)$ es isomorfo a algún segmento inicial $] \leftarrow b[$ de $(X, \leq)$. Se puede probar el siguiente resultado que enunciaremos sin demostración:

Dos números transfinitos α y β arbitrarios verifican una, y solo una, de las relaciones

$$\alpha = \beta, \alpha < \beta \text{ o } \alpha > \beta.$$

Del resultado anterior y del hecho que todo isomorfismo es además una correspondencia biunívoca, se desprende que las potencias de dos conjuntos bien ordenados son siempre comparables. Esta afirmación nos sugiere la siguiente pregunta: ¿es posible establecer de alguna manera un buen ordenamiento en un conjunto cualquiera? Si esto fuera posible, nos permitiría afirmar que *todas las potencias son comparables*, es decir, que entre dos conjuntos cualesquiera

X y Y siempre existe una aplicación $f : X \longrightarrow Y$ tal que $f : X \longrightarrow f[X]$ es una biyección (o bien, existe $g : Y \longrightarrow X$, donde $g : Y \longrightarrow f[Y]$ es una biyección). Una respuesta positiva a esta pregunta fue dada por Zermelo quién demostró el teorema siguiente:

A E 5 Teorema de Zermelo: Todo conjunto puede ser bien ordenado.

Este teorema está vinculado a toda una serie de problemas profundos y de principios en la matemática, lo cual se debe a que la demostración (que no expondremos aquí) se basa en el axioma de elección.

Más aún, puede probarse que las proposiciones A E 1 -A E 5 son equivalentes.

En efecto, si suponemos que el conjunto $X = \cup M = \cup \{M_\alpha : M_\alpha \in M\}$ está bien ordenado, donde M es el conjunto del postulado A E 2, entonces para construir el conjunto M tal que $M \cap M_\alpha = \{x_\alpha\}$, basta tomar como x_α el primer elemento en cada M_α .

A continuación enunciaremos algunas otras proposiciones, cada una de las cuales es equivalente al axioma de elección (es decir, cualquiera de ellas puede ser demostrada si se acepta el axioma de elección y viceversa, el axioma de elección puede ser demostrado si se asume como verdadera alguna de estas proposiciones), y que representan una gran ayuda, tanto en las construcciones conjuntistas en la matemática como por el método que aportan para la demostración de los resultados.

Los conceptos utilizados en el enunciado de estas proposiciones equivalentes al axioma de elección, y que daremos sin demostración, pueden ser encontrados por el lector en las definiciones 1.7.4 y 1.7.5.

A E 6 Teorema de Hausdorff: Toda cadena de un conjunto parcialmente ordenado, está contenida en alguna de las cadenas maximales de este conjunto.

El siguiente teorema es, posiblemente, la forma más conveniente en las aplicaciones de las proposiciones equivalentes al axioma de elección:

A E 7 Teorema de Zorn: En un conjunto inductivamente ordenado $(X, \leq)$, cualquier elemento de X precede a algún elemento maximal.

Los postulados equivalentes A E 1 -A E 7 resultan una potente herramienta de trabajo en la demostración de las proposiciones matemáticas, fundamentalmente en dos cuestiones.

Primero: En la generalización del método de inducción matemática (inducción transfinita).

Segundo: En la demostración de teoremas sobre la existencia de entes matemáticos con determinadas propiedades.

Pasemos a explicar estas dos cuestiones:

Como es de todos conocido, un método ampliamente extendido de demostración en la matemática es el llamado método de inducción matemática, el cual consiste en lo siguiente:

Supongamos que se tiene una proposición $P(n)$ que se enuncia para cada número natural n y que además:

- a) la proposición $P(1)$ es cierta;
- b) bajo la hipótesis de la veracidad de $P(k)$ para todo $k \leq n$, se deduce que $P(n+1)$ es cierta.

Entonces, el principio de inducción matemática nos dice que $P(n)$ es válida para todos los valores de n . En efecto, en caso contrario, podríamos encontrar el mínimo del conjunto de los n para los cuales la proposición $P(n)$ no es cierta y al cual denotamos por n_1 . Es obvio que $n_1 > 1$, es decir $n_1 - 1$ es también un número natural, y esto contradice a la condición b).

Se puede emplear un método análogo, sustituyendo la serie de los números naturales por un conjunto bien ordenado cualquiera. En este caso, decimos que estamos en presencia de la inducción transfinita. En esencia, el método de inducción transfinita consiste en lo siguiente:

Sea $(X, \leq)$ un conjunto bien ordenado y sea $P(a)$ una proposición que se enuncia para todo $a \in X$ y tal que:

- a') $P(a)$ es cierta para el primer elemento de A ;
- b') $P(a)$ es cierta si es válida $P(x)$ para todos los elementos $x < a$.

Entonces $P(a)$ es cierta para todos los elementos $a \in X$. Efectivamente, si en X existiesen elementos para los cuales $P(a)$ no se cumple, podríamos encontrar el primer elemento en el conjunto formado por ellos.

Si denotamos por a^* este primer elemento, obtendríamos una contradicción con b) ya que para todos los elementos $a < a^*$ la proposición $P(a)$ sería cierta.

Gracias al teorema de Zermelo, el principio de inducción transfinita puede ser aplicado a un conjunto cualquiera, ya que siempre puede ser bien ordenado. En este caso, el orden a que nos referimos para establecer las propiedades a') y b'), es precisamente el buen orden cuya existencia se postula en el teorema de Zermelo. Sin embargo, en la práctica generalmente conviene más, sustituir este teorema por el teorema de Zorn, donde solo se necesita la existencia

de un orden parcial en el conjunto considerado. Este resultado puede también ser aplicado en la demostración de la existencia de determinados entes matemáticos. Por ejemplo, si se quiere demostrar la existencia, dentro del conjunto X , de un elemento x con la propiedad P , basta construir en X un orden parcial $\leq$ para el cual dicho conjunto sea inductivamente ordenado. La construcción de $\leq$ debe hacerse de manera tal que si un elemento x cumple una cierta propiedad P_0 “suficientemente cercana” a P , entonces todo elemento maximal posterior a x (cuya existencia se postula en el teorema de Zorn) satisfaga la propiedad P .

Ejercicios de Autocontrol

1. Sean $(X, \leq)$ y $(Y, \leq)$ dos conjuntos totalmente ordenados y sean α y β sus tipos ordinales respectivamente.
 - a) Demuestre que la relación definida en $X \cup Y$, aceptando que cada par de elementos en X está ordenado igual que en $(X, \leq)$, que cada par de elementos en Y tiene el mismo orden que en $(Y, \leq)$ y que cualquier elemento de X precede a cualquier elemento de Y , hace de $X \cup Y$ un conjunto totalmente ordenado. (A este conjunto totalmente ordenado, que se designa por $X + Y$ se le llama suma ordenada de los conjuntos totalmente ordenados $(X, \leq)$ y $(Y, \leq)$ y su tipo ordinal se denota por $\alpha + \beta$.)
 - b) Demuestre que la relación definida en $X \times Y$ mediante: $(x_1, y_1) < (x_2, y_2)$ siempre que $y_1 < y_2$ en $(Y, \leq)$ (para cualesquiera x_1 y x_2) y $(x_1, y) < (x_2, y)$ si $x_1 < x_2$, provee a $X \times Y$ de un orden total. A este conjunto totalmente ordenado, que se designa por $X \cdot Y$ se le llama producto ordenado de los conjuntos totalmente ordenados X y Y y su tipo ordinal se denota por $\alpha \cdot \beta$.

1.9 ESTRUCTURAS MATEMÁTICAS SOBRE LOS CONJUNTOS

Hasta el momento hemos estudiado el concepto conjunto y hemos visto como este concepto está a su vez relacionado con las nociones de función, número, campo numérico y figuras geométricas, los cuales han sido objeto de estudio de los cursos de Análisis matemático, Algebra y Geometría en los primeros años de la formación matemática del lector.

Cuando nos planteamos la modelación matemática de un cierto problema, ya sea de carácter teórico o práctico, primeramente se debe determinar el

conjunto universo donde se encuentran los entes reales que tienen relación con dicho problema, y después las propiedades que nos interesa investigar tanto de los elementos como de las partes de este conjunto.

Esta segunda parte del proceso puede ser muy compleja, pues de hecho, ella determina el método matemático así como las teorías matemáticas que vamos a emplear para la solución del problema planteado. Los conceptos *método matemático* y *teoría matemática* son verdaderamente muy amplios y se salen de este contexto.

Nosotros haremos énfasis en una parte de este proceso que es fundamental desde el punto de vista estructural y conceptual en la matemática, nos referimos al concepto *estructura matemática*.

En la matemática se utiliza el concepto *estructura* para denotar toda una serie de relaciones entre los elementos o partes de un conjunto, que se introducen con el fin de obtener determinada información sobre ciertas características de la naturaleza real del problema que ha dado lugar al mencionado conjunto.

En síntesis podemos decir que el procedimiento empleado consta de tres pasos:

Primero: El problema da lugar a un conjunto.

Segundo: La estructura refleja, a partir de las relaciones entre los elementos o partes del conjunto, determinadas características del problema.

Tercero: Los métodos y teorías matemáticas aplicados a la estructura, nos proporcionan información sobre determinados aspectos de la naturaleza real del problema.

Como ya hemos visto, desde el punto de vista conjuntista, el concepto de función nos permite establecer relaciones de carácter cuantitativo entre los conjuntos a partir de la noción primaria de correspondencia biunívoca. El hecho de que mediante funciones se pueden establecer relaciones de carácter también cualitativo entre conjuntos, es el objeto de estudio de determinadas estructuras matemáticas, una de las cuales en el caso particular de los espacios métricos es el tema central de este libro: LA TOPOLOGÍA.

Una *estructura topológica* en un conjunto, es una estructura que permite introducir el concepto *convergencia* en ese conjunto.

En toda estructura matemática existe el concepto de *isomorfismo*, que son aquellos entes matemáticos que dejan invariantes las propiedades que estudia esa estructura. Por ejemplo, en el caso de la estructura topológica los isomorfismos son las aplicaciones biyectivas, tales que tanto ella como su inversa son

continuas, y una *propiedad topológica* es una propiedad invariante ante un *isomorfismo topológico*. Luego, el objeto de estudio de la topología como ciencia son los conceptos *límite y continuidad* en su forma más general y aquellas propiedades de un conjunto, provisto de una estructura topológica, que son invariantes por *aplicaciones continuas* y el paso al *límite*.

El lector ha tenido ya contacto con otras estructuras matemáticas en los conjuntos numéricos, tales como la estructura algebraica, geométrica, diferencial, de orden, etcétera.

Pasemos ahora al estudio de la estructura topológica en los espacios métricos, haciendo énfasis en los resultados inherentes y más específicos de la estructura matemática como un caso particular de estructura topológica.

CAPÍTULO 2

ESPACIOS MÉTRICOS

En este capítulo se realiza un estudio de diferentes aspectos de la topología de los espacios métricos, haciéndose énfasis en el concepto *métrica* como herramienta que permite definir nociones de convergencia y continuidad consecuentes con las que se tienen para sucesiones y funciones numéricas.

El objetivo fundamental del capítulo es motivar la generalización de los conceptos fundamentales del análisis a conjuntos no numéricos, y muy en particular a los espacios de funciones. No obstante, se hace énfasis en la insuficiencia del concepto *métrica* para describir algunas nociones importantes de *convergencia* como es, por ejemplo, la convergencia puntual de funciones, conocida con anterioridad por el lector.

Aunque no es objetivo del libro, se trata de indicar y orientar al lector hacia el estudio de una categoría superior en la cual se pueden estudiar los conceptos de convergencia y continuidad en toda su generalidad: los espacios topológicos.

En los tres primeros epígrafes se introducen a partir de conceptos conocidos de convergencia, los conceptos métrica, espacio métrico y subespacio métrico. En 2.3 se da además, un concepto de equivalencia entre espacios métricos. En el 2.4 se explica el concepto seudométrica y a la vez se brindan métodos para construir una métrica a partir de una seudométrica o de una familia de ellas. En los dos epígrafes siguientes se definen las nociones de convergencia de sucesiones y continuidad de funciones como una generalización inmediata de estos conceptos ya introducidos en el análisis clásico.

Con el objetivo de dar otra expresión equivalente a los conceptos convergencia y continuidad se introducen, en los 2.7 y 2.8, las nociones de bola, vecindad y abierto, y se estudian sus propiedades.

El 2.9 está dedicado al estudio de algunos conceptos relacionados con la convergencia tales como punto adherente, punto de acumulación, etcétera.

Los epígrafes del 2.10 al 2.12 están dedicados al estudio de otras nociones topológicas para conjuntos y funciones, así como la importancia de estas nociones topológicas en la teoría de los espacios métricos. En el 2.13 se estudia la separación de puntos y conjuntos en un espacio métrico por abiertos y por funciones continuas, así como la aplicación a la construcción de las llamadas “particiones de unidad”, que tienen una gran importancia en el proceso de

demostración de propiedades globales a partir de propiedades locales.

En el 2.14 se examinan algunos tipos de convergencia y continuidad que no pueden ser definidos por una métrica.

Finalmente, el 2.15 se dedica a explicar las insuficiencias de la estructura métrica para construir una teoría completa de convergencia y continuidad, exponiéndose las vías de transición a una teoría más general.

2.1 ALGUNOS CONCEPTOS DE CONVERGENCIA

El concepto de *convergencia* es fundamental en Análisis Matemático. Lo empleamos al hablar de sucesiones convergentes de números reales y complejos, de sucesiones de puntos de un espacio euclideo, de sucesiones de funciones, etcétera. En este último caso, incluso utilizamos diferentes tipos de convergencia: convergencia puntual, convergencia uniforme, convergencia en media, etcétera. El concepto de *convergencia* no sólo es fundamental para los razonamientos teóricos del Análisis y del Análisis Funcional, sino también en la Matemática Aplicada; basta mencionar los problemas de aproximación en Análisis Numérico y en Cálculo Operacional. En la Teoría de Optimización por ejemplo, muchas veces no es posible determinar explícitamente el valor óptimo buscado, en cuyo caso el problema consiste en hallar un algoritmo para construir una sucesión de valores que *converja* al valor optimal desconocido. Para funciones en lugar de valores reales y vectoriales existen problemas análogos: aproximación de controles optimales en Teoría de Control, aproximación de soluciones de ecuaciones diferenciales en Análisis Numérico, etcétera.

En todos los casos mencionados anteriormente, se utiliza el concepto *convergencia*, basándose en cierta noción de *distancia*; pero en cada caso las definiciones de dichas distancias son diferentes. Consideremos algunos ejemplos.

Ejemplo 1. Decimos que una sucesión de números reales (x_n) converge a un número real x si los términos x_n están tan cerca de x como se quiera cuando el subíndice n es suficientemente grande. Esta condición puede precisarse de la manera siguiente:

Para todo $\epsilon > 0$ existe un número natural $N(\epsilon)$ tal que:

$$|x - x_n| \leq \epsilon \text{ para todo } n \geq N(\epsilon)$$

t

Aquí el valor absoluto de la diferencia $x - x_n$ es una medida para la proximidad de x_n a x . A esta medida la llamamos *distancia de x_n a x* y la denotamos por $d(x, x_n)$ (fig. 14).

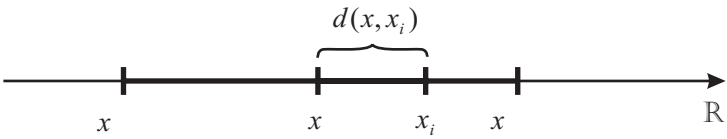


Figura 14

Evidentemente la sucesión (x_n) converge a x , si y sólo si la sucesión de las distancias $|x - x_n|$ converge a 0.

Ejemplo 2. Consideremos sucesiones (x_i) de puntos en el espacio $\mathbb{R}^n$. Para no confundir los subíndices de la sucesión con los índices de las coordenadas , denotaremos por $(x_i(1), \dots, x_i(n))$ las n coordenadas del punto $x_i (i \in \mathbb{N})$.

Decimos que la sucesión (x_i) converge al punto $x \in \mathbb{R}^n$, si las distancias

$$d_2(x, x_i) = \left(\sum_{k=1}^n (x(k) - x_i(k))^2 \right)^{1/2} \tag{1}$$

son tan pequeñas como se quiera cuando los subíndices i son suficientemente grandes. Esta definición de la distancia entre dos puntos (1) se obtiene partiendo de nuestra intuición espacial en $\mathbb{R}^3$ (fig. 15).

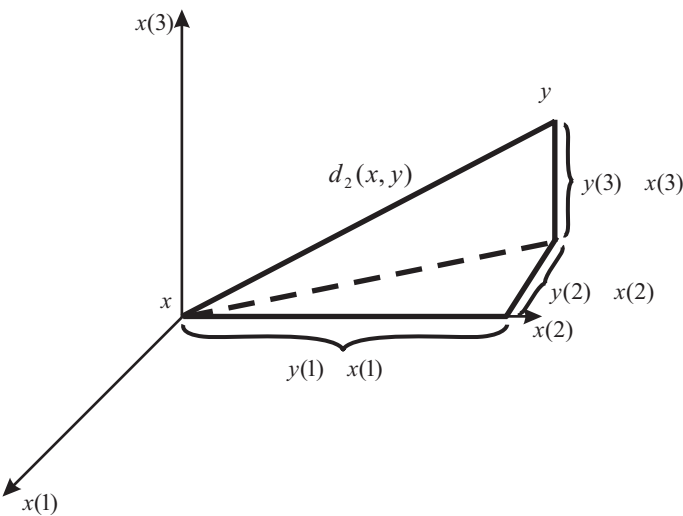


Figura 15

Ejemplo 3. Sea $A \neq \emptyset$ un conjunto cualquiera y (f_n) una sucesión de funciones $f_n : A \longrightarrow \mathbb{R}$. Se dice que (f_n) converge puntualmente a una función

$f : A \longrightarrow \mathbb{R}$, si para todo $x \in A$ la sucesión de números reales $(f_n(x))$ converge al número real $f(x)$. Consideremos por ejemplo la sucesión de las funciones reales f_n definidas sobre $A = [0, 1]$, cuya gráfica forma un triángulo isósceles de altura n , mientras que $f_n(x) = 0$ para $x \geq 1/n$ (fig. 16). Aunque esta sucesión varía de manera tan errática en la proximidad de 0, es fácil comprobar que converge puntualmente a la función constante $\langle 0 \rangle$ sobre $[0, 1]$.

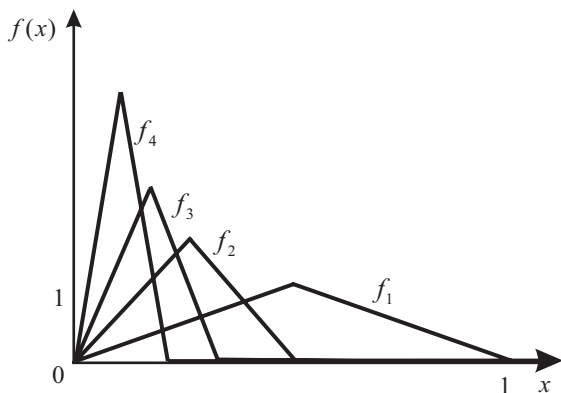


Figura 16

Para casos donde importa el comportamiento global de una función f sobre todo su espacio de partida, una aproximación puntual no sería suficiente. En este caso muchas veces se buscan aproximaciones uniformes.

Ejemplo 4. Sea $A \neq \emptyset$ un conjunto y (f_n) una sucesión de funciones $f_n : A \longrightarrow \mathbb{R}$. Se dice que f_n converge a una función $f : A \rightarrow \mathbb{R}$ uniformemente (sobre A), si para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que

$$n \geq N \Rightarrow |f_n(x) - f(x)| \leq \epsilon \text{ para todo } x \in A \quad (2)$$

Esto significa intuitivamente que a partir de un cierto índice N las gráficas de f_n se hallan dentro de la ϵ -banda de la gráfica de f (fig. 17). La condición (2) es equivalente a

$$d_\infty(f, f_n) = \sup_{x \in A} |f(x) - f_n(x)| \leq \epsilon \text{ para } n \geq N \quad (3)$$

Luego, (f_n) converge uniformemente a f , si la sucesión real $(d_\infty(f, f_n))$ converge a 0.

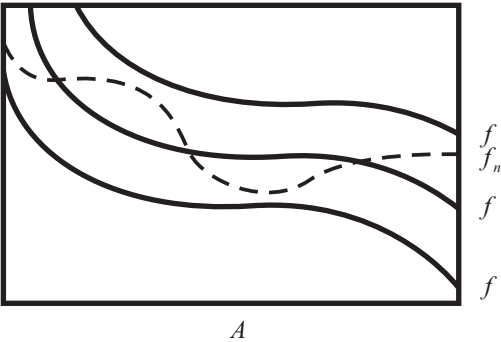


Figura 17

En la teoría de probabilidad, estadística, etcétera, los valores de una función en puntos particulares muchas veces no son importantes. Entonces no se necesitan aproximaciones que sean buenas en todos los puntos, sino sólo tales que sean buenas en media o en media cuadrática.

Ejemplo 5. Sea (f_n) una sucesión de funciones continuas definidas sobre el intervalo $[0, 1]$. La distancia entre los valores de f y de g en x es $|f(x) - g(x)|$ para todo $x \in [0, 1]$. La distancia media entre los valores de g y f :

$$d_1(f, g) = \int_0^1 |f(x) - g(x)| dx$$
 depende sólo de las dos funciones f y g , la cual corresponde al área entre los grafos de f y g (fig. 18).

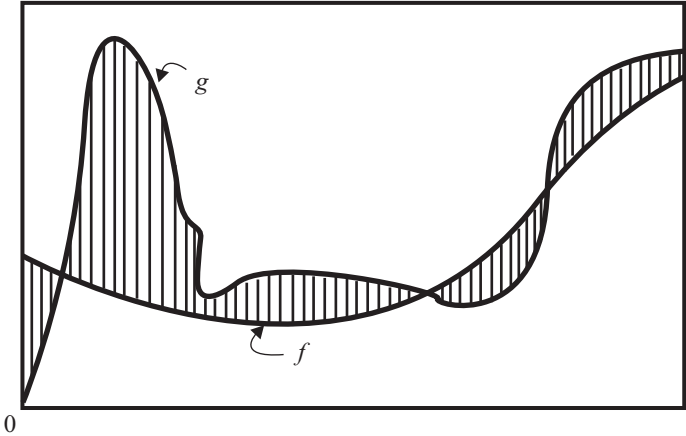


Figura 18

Se dice que una sucesión de funciones integrables (f_n) definida sobre $[0, 1]$ converge en media hacia una función f , si la sucesión $(d_1(f, f_n))$ converge a 0.

En la teoría de aproximación se utiliza también otro método de medir la proximidad entre dos funciones integrables $f, g : [0, 1] \rightarrow \mathbb{R}$. Considerando f como una función dada o buscada y g como una aproximación de f (por ejemplo, polinomial), $(f(x) - g(x))^2$ es el error cuadrático que se comete, si se sustituye f por g en el punto x , luego

$$\int_0^1 (f(x) - g(x))^2 dx \quad (4)$$

es el error cuadrático que se comete en media si se sustituye f por g . Para obtener una medida de proximidad, que es positivamente homogénea, se introduce en lugar de (4) su raíz cuadrada

$$d_2(f, g) = \left(\int_0^1 (f(x) - g(x))^2 dx \right)^{1/2} \quad (5)$$

Se dice que una sucesión de funciones integrables f_n converge en media cuadrática a f sobre $[0, 1]$, si la sucesión de los errores cuadráticos $(d_2(f, f_n))$ converge a 0.

Los conceptos de convergencia definidos en los ejemplos anteriores son aparentemente muy heterogéneos; sin embargo, para números, vectores y funciones, con respecto a todos los diversos tipos de convergencia, se plantean frecuentemente problemas similares. Por ejemplo: ¿el límite de una sucesión convergente está determinado de manera única?, ¿cada sucesión de Cauchy (de números, vectores, funciones) es convergente? ¿de cada sucesión acotada puede extraerse una subsucesión convergente? etcétera.

Parece lógico y además económico desarrollar para estos problemas comunes una teoría común, aplicable igualmente a números, vectores, funciones y otros objetos matemáticos. Para desarrollar tal teoría habría que hacer un análisis de los razonamientos usuales en el tratamiento de los diferentes tipos de convergencia, y formular la base lógica común de estos razonamientos como sistemas de axiomas de una nueva teoría. No podemos reproducir aquí el proceso histórico que terminó con la definición de los espacios métricos de Frechet (1905). Sin embargo, ya nuestros ejemplos nos dan unas sugerencias. Analicemos su estructura común.

Todos estos conceptos de convergencia quedan expresados de forma general de la manera siguiente:

una sucesión de elementos $(x_i), (f_i)$ converge a un elemento x, f , si las x_i, f_i están tan cerca de x, f como se quiera, cuando los subíndices i son suficientemente grandes. Lo esencial es, cómo se precisa en cada caso la palabra cerca.

Con excepción del ejemplo 3 (convergencia puntual) esto se hace en todos los ejemplos por una medida de proximidad:

$$d(x, y) = |y - x| \tag{6}$$

para la convergencia de números en $\mathbb{R}$;

$$d_2(x, y) = \left(\sum_{k=1}^n (x(k) - y(k))^2 \right)^{1/2} \tag{7}$$

para la convergencia de puntos en $\mathbb{R}^n$;

$$d_\infty(f, g) = \sup_{x \in A} |g(x) - f(x)| \tag{8}$$

para la convergencia uniforme;

$$d_1(f, g) = \int_0^1 |g(x) - f(x)| dx \tag{9}$$

para la convergencia en media;

$$d_2(f, g) = \left(\int_0^1 (g(x) - f(x))^2 dx \right)^{1/2} \tag{10}$$

para la convergencia en media cuadrática.

En el ejemplo 3 no tenemos una sola medida sino toda una familia que miden la proximidad de f_n a f en los distintos puntos $x \in A$

$$d_x(f, g) = |f(x) - g(x)| \quad (x \in A) \tag{11}$$

Esto parece ser una diferencia estructural tan importante que es natural excluir la convergencia puntual en una primera fase de nuestra consideración.

Regresemos a los demás ejemplos. Corresponde a nuestra intuición geométrica interpretar las dos magnitudes $d(x, y), d_2(x, y)$ en (6),(7) como distancias:

$|x - y|$ es la distancia entre los números reales x, y en la recta real;
 $\sqrt{(x(1) - y(1))^2 + (x(2) - y(2))^2 + (x(3) - y(3))^2}$
 es la distancia entre los dos puntos $x = (x(1), x(2), x(3)), y = (y(1), y(2), y(3))$
 en el espacio tridimensional.

Hasta el momento no tenemos una idea intuitiva de lo que sería una distancia entre dos funciones, estamos sólo acostumbrados a hablar de la distancia entre dos valores de funciones. Fue un gran logro en la historia de las

matemáticas cuando, en vez de fijarse en los valores de las funciones, las funciones mismas se consideraron como puntos de un nuevo espacio abstracto y las magnitudes (8),(9) y (10) se interpretaron como distancias entre estos puntos. Esta concepción es el origen del *Análisis Funcional*.

Con la interpretación de las medidas de proximidad (8)-(10) como distancias entre funciones casi tenemos ya un modelo matemático para las cuestiones de convergencia; queda sólo precisar lo que es una distancia en general. No podemos recurrir a nuestra intuición espacial, porque no tenemos una intuición geométrica de la distancia entre dos funciones, además existen diferentes conceptos de distancias como hemos visto.

Coleccionemos las propiedades más elementales de las distancias definidas. Según nuestros ejemplos, la distancia entre dos elementos de un conjunto X (de números, vectores, funciones, etcétera) es un número real positivo. Si asociamos a todo par de elementos $(x, y) \in X^2$ su distancia, obtenemos una aplicación de: $X \times X \rightarrow \mathbb{R}_+$ que se llama *métrica*. Claro, que no llamaremos métrica a cualquier función de $X \times X$ en $\mathbb{R}_+$. Exigiremos que sus valores verifiquen algunas propiedades básicas e intuitivas de las distancias. Por ejemplo, la distancia entre dos elementos x, y puede ser 0 únicamente si $x = y$. Además, la distancia de x a y debe ser igual a la distancia de y a x .

La selección de propiedades fundamentales para la caracterización axiomática de una métrica debe verificar dos criterios:

1. Las propiedades deben ser suficientemente débiles para ser comunes a todas las métricas concretas ya utilizadas.
2. Deben ser suficientemente fuertes para establecer una teoría sólida que permita el tratamiento sistemático y simultáneo de los problemas fundamentales de convergencia y de continuidad.

Fue Fréchet quien descubrió primero (en su disertación de doctorado) que es suficiente añadir la desigualdad triangular:

$$d(x, y) + d(y, z) \geq d(x, z)$$

a las dos propiedades mencionadas anteriormente, para tener una base axiomática adecuada de la teoría de los espacios métricos.

Ejercicios de Autocontrol

1. Pruebe que toda sucesión de funciones uniformemente convergente es también puntualmente convergente, pero que no se cumple el inverso.
2. Pruebe que toda sucesión (f_n) de funciones integrables $f_n : [0, 1] \rightarrow \mathbb{R}$, que converge uniformemente a $f : [0, 1] \rightarrow \mathbb{R}$, converge también en

media y en media cuadrática a f , pero que no se cumple el inverso. Pruebe además, que la convergencia puntual no implica necesariamente la convergencia en media o en media cuadrática.

3. Todo punto $x \in \mathbb{R}^n$ puede identificarse con la función $k \rightsquigarrow x(k)$ del conjunto $A = \{1, 2, \dots, n\}$, en $\mathbb{R}$, que asocia a todo índice $k \in A$ la k -ésima coordenada de x . Pruebe que para tales funciones (que tienen un espacio de partida finito), la convergencia puntual y la convergencia uniforme coinciden.

2.2 MÉTRICAS. EJEMPLOS DE ESPACIOS MÉTRICOS

Definición 2.2.1. Sea X un conjunto; una aplicación $d : X \times X \longrightarrow \mathbb{R}$, se llama métrica o distancia sobre X , si verifica los axiomas siguientes:

- (D1) $d(x, y) = d(y, x) (x, y \in X)$;
- (D2) $d(x, z) \leq d(x, y) + d(y, z) (x, y, z \in X)$
- (D3) $d(x, y) = 0 \Leftrightarrow x = y (x, y \in X)$.

La propiedad (D2) se denomina desigualdad triangular y ella corresponde a la idea intuitiva de que la menor distancia entre dos puntos en el espacio euclideo $\mathbb{R}^3$ es la longitud del segmento de recta que los une.

Para $x, y \in X$ el número real positivo $d(x, y)$ se llama entonces *distancia* entre x y y . Al par (X, d) se le llama *espacio métrico*.

Observación: Una métrica sobre X no es una función sobre X sino sobre $X^2 = X \times X$.

Ejemplo 6. Se puede probar que las expresiones (6) – (10) definen métricas sobre los espacios $\mathbb{R}; \mathbb{R}^n; B([0, 1], \mathbb{R}) = \{f \in \mathbb{R}^{[0, 1]} : f \text{ acotada}\} = \{f : [0, 1] \longrightarrow \mathbb{R} : f \text{ acotada}\}; C([0, 1], \mathbb{R}) = \{f \in \mathbb{R}^{[0, 1]} : f \text{ continua}\} = \{f : [0, 1] \longrightarrow \mathbb{R} : f \text{ continua}\}$ y $C([0, 1], \mathbb{R})$ respectivamente.

Las métricas (6) y (7) serán tratadas más adelante como casos particulares del ejemplo 7. Se deja de ejercicio al lector comprobar que las expresiones (8) y (9) definen métricas sobre los correspondientes conjuntos señalados en el ejemplo 6.

Probemos que (10) define una métrica en $(C[0, 1], \mathbb{R})$.

La propiedad (D1) resulta evidente: la propiedad (D3) en este caso, se deduce del hecho de que si la integral de una función continua positiva es nula entonces dicha función es idénticamente nula.

Para probar la desigualdad triangular en este caso (D2) comencemos por verificar la llamada desigualdad de Cauchy-Buniakovsky:

$$\left(\int_a^b x(t) y(t) dt \right)^2 \leq \int_a^b x^2(t) dt \cdot \int_a^b y^2(t) dt \quad (12)$$

Esta desigualdad puede obtenerse, por ejemplo, de la siguiente identidad que se comprueba fácilmente

$$\left(\int_a^b x(t) y(t) dt \right)^2 = \int_a^b x^2(t) dt \int_a^b y^2(t) dt$$

$$- \frac{1}{2} \int_a^b \int_a^b (x(s) y(t) - y(s) x(t))^2 ds dt.$$

Una vez obtenida la desigualdad (12) tendremos para $x, y, z \in C([a, b], \mathbb{R})$:

$$\int_a^b |x(t) - y(t)|^2 dt = \int_a^b |(x(t) - z(t) + z(t) - y(t))|^2 dt =$$

$$= \int_a^b |x(t) - z(t)|^2 dt + \int_a^b |z(t) - y(t)|^2 dt +$$

$$+ 2 \int_a^b |x(t) - z(t)| |z(t) - y(t)| dt \leq$$

$$\leq \int_a^b |x(t) - z(t)|^2 dt + \int_a^b |z(t) - y(t)|^2 dt +$$

$$+ 2 \left(\int_a^b |x(t) - z(t)|^2 dt \cdot \int_a^b |z(t) - y(t)|^2 dt \right)^{1/2} =$$

$$= \left[\left(\int_a^b |x(t) - z(t)|^2 dt \right)^{1/2} + \left(\int_a^b |z(t) - y(t)|^2 dt \right)^{1/2} \right]^2.$$

con lo cual queda probada la desigualdad triangular.

Veamos ahora otras métricas definidas en el espacio euclidiano $\mathbb{R}^n$ que según veremos posteriormente, con respecto a la convergencia de sucesiones, son *equivalentes* a la métrica dada en (7).

Ejemplo 7. La aplicación $d_p : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}_+$ definida para $p \geq 1$ mediante

$$d_p(x, y) = \left(\sum_{k=1}^n |x_k - y_k|^p \right)^{1/p}, \quad (13)$$

donde hemos denotado por x_k la k -ésima coordenada de un punto de $\mathbb{R}^n$, es una métrica sobre $\mathbb{R}^n$. En efecto, en este caso es obvio que se verifican los axiomas (D1) y (D3). Comprobemos el axioma (D2).

Sean $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, y $z = (z_1, \dots, z_n)$ tres puntos de $\mathbb{R}^n$. Consideremos $y_k - x_k = a_k$ y $z_k - y_k = b_k$. Entonces la desigualdad triangular $d_p(x, z) \leq d_p(x, y) + d_p(y, z)$ que debemos demostrar, puede representarse en la forma

$$\left(\sum_{k=1}^n |a_k + b_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^n |a_k|^p \right)^{1/p} + \left(\sum_{k=1}^n |b_k|^p \right)^{1/p} \quad (14)$$

Esta es la llamada *desigualdad* de Minkowski de orden p en $\mathbb{R}^n$. Para $p = 1$ la desigualdad de Minkowsky es obvia ya que se reduce a una propiedad conocida del módulo (el valor absoluto de la suma no sobrepasa la suma de los valores absolutos) y, por consiguiente, nos limitaremos al caso $p > 1$. Para $p < 1$ la desigualdad de Minkowsky no tiene lugar, es decir si $p < 1$, d_p no define una métrica en $\mathbb{R}^n$.

La demostración de la desigualdad (14) para $p > 1$ se basa en la llamada desigualdad de Hölder:

$$\sum_{k=1}^n |a_k b_k| \leq \left(\sum_{k=1}^n |a_k|^p \right)^{1/p} \left(\sum_{k=1}^n |b_k|^q \right)^{1/q}, \quad (15)$$

donde los números $p > 1$ y $q > 1$ cumplen la condición

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (16)$$

Observemos que la desigualdad (15) es homogénea. Ello significa que si se cumple para dos vectores cualesquiera $a = (a_1, \dots, a_n)$ y $b = (b_1, \dots, b_n)$, tambien se verifica para los vectores λa y μb donde λ y μ son números reales arbitrarios. Por eso es suficiente demostrar la propiedad (15) para el caso en que

$$\sum_{k=1}^n |a_k|^p = \sum_{k=1}^n |b_k|^q = 1. \quad (17)$$

Supongamos pues que se cumple la condición (17) y demostraremos que

$$\sum_{k=1}^n |a_k b_k| \leq 1. \quad (18)$$

Consideremos en el plano (ϵ, η) la curva dada por la ecuación $\eta = \epsilon^{p-1} (\epsilon > 0)$ o, lo que es lo mismo, por la ecuación $\epsilon = \eta^{q-1}$ (fig. 19)

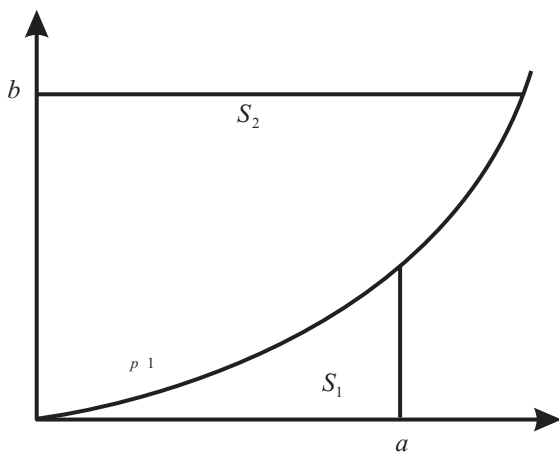


Figura 19

En la figura se ve claramente que para cualesquiera valores positivos a y b será $S_1 + S_2 \geq ab$. Calculemos las áreas S_1 y S_2

$$S_1 = \int_0^a \xi^{p-1} d\xi = \frac{a^p}{p}, \quad S_2 = \int_0^b \eta^{q-1} d\eta = \frac{b^q}{q}.$$

Por lo tanto, se verifica la siguiente desigualdad numérica

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Considerando $a = |a_k|$, $b = |b_k|$ y sumando respecto a k desde 1 hasta n , obtendremos teniendo en cuenta (16) y (17),

$$\sum_{k=1}^n |a_k b_k| \leq 1.$$

Luego, hemos demostrado la desigualdad (18), y por consiguiente, la desigualdad general (15). Cuando $p = 2$, la desigualdad de Hölder (15) se convierte en la desigualdad de Cauchy-Buniakovsky para $\mathbb{R}^n$.

Pasemos ahora a demostrar la desigualdad de Minkowsky. Para ello, consideremos la identidad

$$(|a| + |b|)^p = (|a| + |b|)^{p-1} |a| + (|a| + |b|)^{p-1} |b|.$$

Tomando en la identidad anterior $a = a_k$ y $b = b_k$ y sumando respecto a k desde 1 hasta n , obtenemos

$$\sum_{k=1}^n (|a_k| + |b_k|)^p = \sum_{k=1}^n (|a_k| + |b_k|)^{p-1} |a_k| + \sum_{k=1}^n (|a_k| + |b_k|)^{p-1} |b_k|.$$

Aplicando ahora a cada una de las sumas que aparecen a la derecha la desigualdad de Hölder y tomando en consideración que $(p-1)q = p$, encontramos

$$\sum_{k=1}^n (|a_k| + |b_k|)^p \leq \left(\sum_{k=1}^n (|a_k| + |b_k|)^p \right)^{1/q} \left(\left[\sum_{k=1}^n |a_k|^p \right]^{1/p} + \left[\sum_{k=1}^n |b_k|^p \right]^{1/p} \right)$$

Dividiendo ambos lados de esta desigualdad por

$$\left(\sum_{k=1}^n (|a_k| + |b_k|)^p \right)^{1/p},$$

obtenemos

$$\left(\sum_{k=1}^n (|a_k| + |b_k|)^p \right)^{1/p} \leq \left(\sum_{k=1}^n |a_k|^p \right)^{1/p} + \left(\sum_{k=1}^n |b_k|^p \right)^{1/p}$$

y de aquí se deduce inmediatamente la desigualdad (14), con lo cual queda comprobada la desigualdad triangular para d_p en $\mathbb{R}^n$.

La métrica d_p considerada en este ejemplo coincide para $p = 2$ con la métrica d_2 definida en (7).

El lector puede comprobar que la expresión

$$d_\infty(x, y) = \max_{1 \leq k \leq n} |x_k - y_k| \quad (19)$$

define una métrica en $\mathbb{R}^n$, que es el caso límite de las métricas $d_p(x, y)$, es decir,

$$d_\infty(x, y) = \lim_{p \rightarrow \infty} \left(\sum_{k=1}^n |x_k - y_k|^p \right)^{1/p} \quad (20)$$

Ejemplo 8. De la desigualdad $ab \leq \frac{a^p}{p} + \frac{b^q}{q} \left(\frac{1}{p} + \frac{1}{q} = 1 \right)$, establecida anteriormente, es fácil deducir la desigualdad integral de Hölder

$$\int_a^b |x(t) y(t)| dt \leq \left(\int_a^b |x(t)|^p dt \right)^{1/p} \left(\int_a^b |y(t)|^q dt \right)^{1/q} \quad (21)$$

que se verifica para cualesquiera funciones $x(t)$ y $y(t)$ siempre que las integrales que figuran a la derecha tengan sentido. A su vez, por un procedimien-

to similar al aplicado en el ejemplo 7 se obtiene la *desigualdad* integral de Minkowsky

$$\left(\int_a^b |x(t) + y(t)|^p dt \right)^{1/p} \leq \left(\int_a^b |x(t)|^p dt \right)^{1/p} + \left(\int_a^b |y(t)|^p dt \right)^{1/p},$$

de donde se concluye, sin dificultad, la desigualdad triangular para la métrica de la *media de orden p*

$$d_p(x, y) = \left(\int_a^b |x(t) - y(t)|^p dt \right)^{1/p} \quad (22)$$

definida en $C([a, b], \mathbb{R})$.

Hasta ahora, hemos obtenido los espacios métricos como un modelo para el tratamiento de cuestiones de convergencia y de aproximación. Pero como es costumbre con los modelos matemáticos, la teoría de los espacios métricos sirve para muchos otros fines. La gran variedad de posibles aplicaciones de esta teoría, ya se vislumbra en los siguientes ejemplos de espacios métricos, que no solo pertenecen al Análisis Matemático o al Análisis Funcional, sino también a disciplinas aparentemente tan ajenas como la teoría algebraica de números y la teoría de grafos.

Ejemplo 9 (métrica discreta). Sea X un conjunto cualquiera. Entonces la aplicación $d : X \times X \longrightarrow \mathbb{R}_+$ definida por

$$d(x, y) = \begin{cases} 1, & \text{si } x \neq y \\ 0, & \text{si } x = y \end{cases}$$

es una métrica. Todo conjunto puede proveerse de esta métrica trivial, que se llama discreta.

Ejemplo 10 (distancia p-ádica). Procederemos a definir sobre el conjunto $X = \mathbb{Q}$ de los números racionales la llamada distancia p -ádica, que juega un papel importante en la teoría algebraica de los números.

Sea p un número primo. Para todo $n \in \mathbb{N}$ denotaremos por $\ll v_p(n) \gg$ el exponente de p en la descomposición de n en factores primos. Entonces se cumple para todos los $n, n' \in \mathbb{N}$

$$v_p(n \cdot n') = v_p(n) + v_p(n') \quad (23)$$

Si ahora $x = \pm r/s \in \mathbb{Q}$ con $r, s \neq 0, r, s \in \mathbb{N}$ podemos definir

$$v_p(x) := v_p(r) - v_p(s). \quad (24)$$

De (23) es fácil ver que este valor no depende de la expresión particular de x como fracción y, además con esta definición, (23) sigue válida para $n, n' \in \mathbb{Q}$. Con tal convenio podemos definir la métrica p -ádica v_p por:

$$v_p(x, y) := \begin{cases} 0 & \text{si } x = y \\ p - v_p(x - y) & \text{si } x \neq y \end{cases} \quad ((x, y) \in \mathbb{Q} \times \mathbb{Q}) \tag{25}$$

Dejamos al lector verificar que v_p cumple los axiomas (D1) a (D3).

Daremos ahora un ejemplo de un espacio métrico finito. Tales espacios métricos se estudian por ejemplo en la *teoría de grafos*.

Ejemplo 11 (distancia de costos). Consideremos el problema siguiente: en una región un cierto número finito de lugares deben conectarse por carreteras, se conocen los costos de construcción de toda una serie de carreteras posibles, y se busca una red de carreteras que conecten todos los lugares y cuyos costos de construcción sean minimales. Para resolver tales problemas se toman como modelos los *digrafos evaluados* (conexos) (fig. 20).

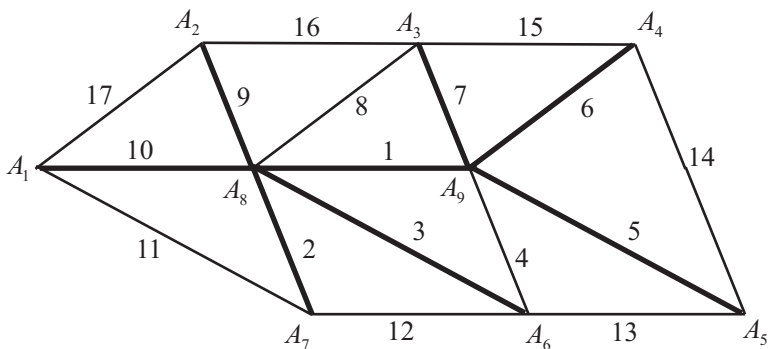


Figura 20

No daremos la definición formal de un digrafo evaluado ya que para nuestro objetivo basta el concepto intuitivo. Veremos enseguida cómo introducir una métrica sobre el conjunto de los vértices. Estos vértices pueden conectarse por al menos un camino (porque suponemos los digrafos conexos). En nuestro ejemplo tenemos entre A_1 y A_4 los caminos: (A_1, A_2, A_3, A_4) , (A_1, A_8, A_9, A_4) , $(A_1, A_2, A_8, A_3, A_9, A_4)$, etc. A cada camino $(A_{i_1}, \dots, A_{i_r})$ se le asocia su costo de construcción

$$c(A_{i_1}, A_{i_2}) + \dots + c(A_{i_{r-1}}, A_{i_r}),$$

por ejemplo al camino (A_1, A_2, A_3, A_4) se le asocian los costos $17+16+15 = 48$. Está claro que entre ellos existe (al menos) un camino cuyos costos de construcción son minimales. Evidentemente tales caminos tienen un interés particular en la economía; la teoría de los grafos ha desarrollado algoritmos para determinar explícitamente estos llamados caminos geodésicos. Si denotamos por d_{ij} el costo del camino geodésico entre A_i y A_j , es fácil comprobar que mediante la aplicación

$$d : (A_i, A_j) \rightsquigarrow \begin{cases} 0 & \text{si } i = j \\ d_{ij} & \text{si } i \neq j \end{cases} \quad (i, j = 1, \dots, 9)$$

se define una métrica sobre el conjunto de los vértices del digrafo, $d_{ij} (i \neq j)$ que se llama distancia de costos entre A_i y A_j , la matriz $(d_{ij})_{i,j=1,\dots,9}$ se llama matriz de distancias del digrafo evaluado.

En cuanto al problema planteado inicialmente puede demostrarse que la red de costos minimales que conectan todos los lugares $A_1, \dots, A_9 = A_9$ tiene $n - 1 = 8$ carreteras. Existen algoritmos para determinar explícitamente tales redes. En nuestra gráfica hemos dibujado una red de costos minimales por líneas fuertes.

En el capítulo 7 estudiaremos un tipo particular de espacios métricos de gran utilidad en las aplicaciones: los espacios normados, en los cuales la métrica es “compatible”, en un sentido que precisaremos más adelante, con la estructura algebraica del conjunto en cuestión. Entre los espacios normados estudiaremos también los llamados espacios euclidianos donde se interrelacionan la estructura métrica, la estructura algebraica y la estructura geométrica de un conjunto dado. Por el momento no nos referiremos directamente a estos casos importantes, a pesar de que como veremos en el capítulo 7, muchos de los ejemplos que hemos citado hasta el momento son casos particulares de espacios normados y euclidianos. Por el momento nos limitaremos a ver en calidad de ejemplos, otros espacios métricos cuyo conjunto subyacente posee además una estructura algebraica de espacio vectorial de dimensión infinita.

Ejemplo 12. Consideremos para $p > 1$ el conjunto $l^p(\mathbb{N})$ de todas las sucesiones de números reales (x_k) , con $\sum_{k=1}^{\infty} |x_k|^p < \infty$.

Definamos en $l^p(\mathbb{N})$ una distancia mediante

$$d(x, y) = \left(\sum_{k=1}^{\infty} |x_k - y_k|^p \right)^{1/p} \quad (26)$$

De acuerdo con la desigualdad de Minkowsky (14) tenemos para n cualquiera

$$\left(\sum_{k=1}^n |y_k - x_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p \right)^{1/p}$$

Por hipótesis, las series

$$\sum_{k=1}^{\infty} |x_k|^p \text{ y } \sum_{k=1}^{\infty} |y_k|^p$$

son convergentes, por eso, pasando al límite cuando $n \rightarrow \infty$, obtenemos

$$\left(\sum_{k=1}^{\infty} |y_k - x_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^{\infty} |y_k|^p \right)^{1/p} \quad (27)$$

lo cual demuestra que la fórmula (26), mediante la cual se define la distancia en $l^p(\mathbb{N}) = l^p$, tiene sentido para cualesquiera $x, y \in l^p$. Al mismo tiempo, la desigualdad (27) muestra que en l^p se verifica la desigualdad triangular. Los demás axiomas son evidentes.

Consideremos ahora una colección finita de espacios métricos $(E_1, d_1), \dots, (E_n, d_n)$, y definamos, a partir de las métricas d_i , diferentes métricas en el conjunto producto $\prod_{i=1}^n E_i$, a las cuales denominaremos *producto de las métricas* d_i . Para el lector resulta un ejercicio interesante, comprobar que las expresiones:

$$\rho_1(x, y) = \sum_{i=1}^n d_i(x_i, y_i); \quad (28)$$

$$\rho_2(x, y) = \left(\sum_{i=1}^n d_i(x_i, y_i)^2 \right)^{1/2}; \quad (29)$$

$$\rho_{\infty}(x, y) = \max \{d_i(x_i, y_i) : i = 1, \dots, n\} \quad (30)$$

definidas para $x = (x_1, \dots, x_n)$ y $y = (y_1, \dots, y_n) \in \prod_{i=1}^n E_i$, son métricas sobre

el conjunto producto $\prod_{i=1}^n E_i$ y que cumplen

$$\rho_{\infty} \leq \rho_2 \leq \rho_1 \leq n\rho_{\infty}. \quad (31)$$

Evidentemente el espacio euclidiano $\mathbb{R}^n$ es un caso particular de producto cartesiano con una métrica producto: la distancia euclideana sobre $\mathbb{R}^n$ se construye como la métrica producto ρ_2 de las métricas $d_i(x_i, y_i) := |x_i - y_i|$ sobre $E_i := \mathbb{R}$ $i = 1, 2, \dots, n$.

A partir de ahora en adelante, cuando nos refiramos al espacio métrico $\mathbb{R}^n$ sobreentenderemos, si no se afirma lo contrario, que está previsto de cualquiera de las *métricas comparables* ρ_1, ρ_2 o ρ_∞ , la cual seleccionaremos en cada caso en aras de la simplificación del problema que estemos tratando.

Ejemplo 13. Sea (X, d) un espacio métrico, entonces las funciones

$$d' := \frac{d}{1+d} \quad (32)$$

$$d'' := \inf \{1, d\} \quad (33)$$

son métricas sobre X (ver ejercicio 1 de 2.4). Para estas métricas se tiene

$$d' \leq d'' \leq 2d' \quad (34)$$

y ambas son acotadas por 1. Se utilizan fundamentalmente si es necesario sustituir una métrica dada d por una métrica acotada. Hemos visto como sobre el producto cartesiano de espacios métricos, pueden ser definidas diferentes métricas que son *comparables* (31) al igual que en un espacio métrico cualquiera (34). Más adelante veremos que muchos conceptos de la teoría de los espacios métricos (por ejemplo sucesión convergente, conjunto acotado, función continua y función uniformemente continua), se mantienen invariantes al pasar de una métrica a otra métrica comparable en el sentido de (31) y (34). Pero lo que es aún más importante es que los conceptos puramente topológicos, tales como sucesión convergente y función continua, pueden mantenerse invariantes aún ante métricas no comparables como pueden ser por ejemplo d y $\frac{d}{1+d}$ cuando la métrica d no está acotada. Este tipo de resultado, entre otros, nos mostrará la falta de consistencia de la estructura métrica como estructura matemática que aborde en toda su magnitud el estudio de los conceptos de convergencia y continuidad.

Ejercicios de Autocontrol

1. Demuestre que en un espacio métrico (X, d) , para todo $x, y, z \in X$ se cumple

$$|d(z, y) - d(x, y)| \leq d(x, z)$$

2. Demuestre que las expresiones (8) y (9) definidas en 2.1, representan métricas sobre los conjuntos $B(A, \mathbb{R}) = \{f : A \rightarrow \mathbb{R} \text{ acotada}\}$ y $C([0, 1], \mathbb{R})$ respectivamente.
3. Representemos en $\mathbb{R}^2$ la situación de una tribu que habita en un lugar muy boscoso con un río situado en el eje $y = 0$ con respecto al sistema de coordenadas de la (fig. 21)
 Los habitantes de la mencionada tribu para llegar al agua han hecho brechas perpendiculares al río. Si alguien desea ir de un punto (x_1, y_1) a un punto (x_2, y_2) , solo puede hacerlo por las brechas o por la orilla, debido a lo espeso del bosque.
 Dé la definición precisa de la métrica asociada a esta situación y verifique que es una métrica.
4. Demuestre la afirmación de la igualdad (20), es decir la métrica d_∞ en $\mathbb{R}^n$ es el límite de las métricas d_p .
5. Culmine los detalles en el ejemplo 8 siguiendo la idea desarrollada en el ejemplo 7.
6. Pruebe que las expresiones ρ_1, ρ_2 y ρ_∞ definidas mediante (28)-(30) son métricas en el conjunto producto $\prod_{i=1}^n E_i$ y que satisfacen las relaciones (31).

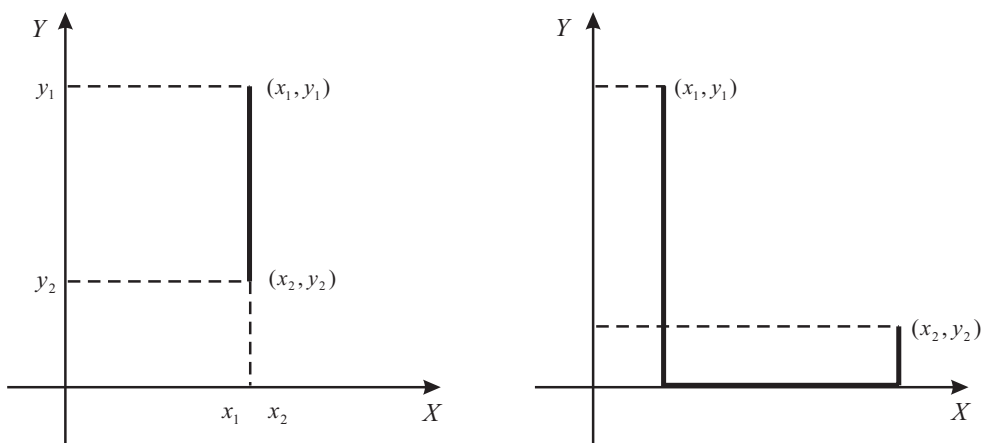


Figura 21

2.3 SUBESPACIOS. ISOMETRÍAS

Pasemos ahora a enunciar un concepto que nos proporciona un arsenal ilimitado de ejemplos de espacios métricos. Para ello notemos que si (X, d) es un espacio métrico y A un subconjunto de X entonces a todo par de puntos $x, y \in A$, como son elementos del espacio métrico (X, d) , le podemos asociar una distancia $d(x, y)$. Estas distancias $d(x, y)$, $(x, y \in A)$ verifican trivialmente los axiomas (D1)-(D3), luego “inducen” una métrica sobre A .

Definición 2.3.1. Sea (X, d) un espacio métrico y $A \subset X$; entonces la métrica

$$d_A : A \times A \longrightarrow \mathbb{R}_+$$

definida por

$$d_A(x, y) := d(x, y), (x, y \in A)$$

se llama *métrica inducida* por d sobre el subconjunto $A \subset X$ y (A, d_A) se llama *subespacio métrico* de (X, d) .

Esta definición nos provee de una gran variedad de nuevos espacios métricos partiendo de los ya definidos en los ejemplos anteriores. Por ejemplo, todo subconjunto de $\mathbb{R}^n$ nos da ahora un espacio métrico.

Después de contar con una reserva tan grande de espacios métricos, resulta natural preguntarse cuales espacios métricos podríamos considerar *isomorfos* o *equivalentes* con respecto a la estructura métrica. Es decir, de la misma forma que en el capítulo 1 definimos un concepto de equivalencia entre conjuntos abstrayéndolos de toda estructura subyacente a ellos, a partir de la noción primaria de correspondencia biunívoca, queremos ahora establecer un concepto de equivalencia entre conjuntos provistos de una estructura métrica. Como ya hemos visto, la noción de correspondencia biunívoca puede enunciarse por medio del concepto auxiliar de función de la forma siguiente:

Dos conjuntos X y Y son equivalentes (desde el punto de vista de la teoría de conjuntos) si existe una aplicación biyectiva de X en Y .

Ya que la estructura métrica se define sobre los conjuntos, el concepto de equivalencia métrica debe incluir al concepto de equivalencia en la teoría de conjuntos, a la vez que conserve las propiedades fundamentales de las métricas asociadas a los conjuntos en cuestión.

Definición 2.3.2. Sean $(X, d), (X', d')$ espacios métricos. Una biyección $f : X \rightarrow X'$ se llama una *isometría*, si para todos los $x, y \in X$ se cumple $d'(f(x), f(y)) = d(x, y)$.

Dos espacios métricos se llaman *isomorfos* o *isométricos* si existe una isometría de uno en el otro que es además sobreyectiva.

De las propiedades de la métrica se deduce que toda isometría es una aplicación inyectiva, y por tanto, dos espacios isométricos son también equivalentes en el sentido de la teoría de conjuntos.

Ejemplo 14. Cada subespacio lineal o afín de $\mathbb{R}^n$, de dimensión $m \leq n$, es isométrico al espacio $\mathbb{R}^m$, donde hemos considerado tanto en $\mathbb{R}^m$ como en $\mathbb{R}^n$ la métrica euclídeana. En efecto, consideremos el subespacio $\alpha + A$ obtenido al trasladar A mediante el vector α .

$$\alpha + A = \{\alpha + x : x \in A\}.$$

y definamos la aplicación

$$f_\alpha : x \rightsquigarrow_{\alpha+x}^A \longrightarrow \alpha + A$$

Esta aplicación es evidentemente sobreyectiva y es además una isometría, ya que

$$d_2(x + \alpha, y + \alpha) = d_2(x, y) \quad (x, y \in A)$$

donde d_2 representa la métrica euclídeana en $\mathbb{R}^n$

Ejemplo 15. Todo espacio euclídeano $\mathbb{R}^n$ es isomorfo como espacio métrico al subespacio $\{(x_i) \in l^2(\mathbb{N}) : x_{n+k} = 0, k \in \mathbb{N}^*\}$ del espacio $l^2(\mathbb{N})$ definido en el ejemplo 12.

Ejercicios de Autocontrol

1. Demuestre la afirmación del ejemplo 15
2. Pruebe que si $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es una isometría del espacio euclídeano $\mathbb{R}^n$ sobre sí mismo tal que $f(0) = 0$, entonces f es lineal.
3. Diga cuáles de las siguientes aplicaciones son isometrías sobre $\mathbb{R}^n$
 - a) las traslaciones: $x \rightsquigarrow x + b$ ($b \in \mathbb{R}^n$ fijo);
 - b) las homotecias: $x \rightsquigarrow \lambda x$ ($\lambda \in \mathbb{R}$ fijo);
 - c) las rotaciones: $x \rightsquigarrow Ax$ ($A = (\alpha_{ij})$ matriz ortogonal)
 - d) todas las aplicaciones $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ tales que $d_2(f(x), 0) = d_2(x, 0)$.

2.4 SEUDOMÉTRICAS

Sea $X \subset \mathbb{R}^A$ un conjunto de funciones reales definidas sobre un conjunto $A \neq \emptyset$. De manera natural, la proximidad de dos funciones $x, y \in X$ en un

punto $a \in A$ se mide por la distancia $|x(a) - y(a)|$ entre sus valores en el punto a . Sin embargo, la aplicación

$$f_a : (x, y) \rightsquigarrow |x(a) - y(a)|$$

no puede interpretarse como una métrica sobre X ; f_a es simétrica y cumple la desigualdad triangular, pero de $f_a(x, y) = 0$ no puede deducirse que $x = y$. Se dice que f_a es una pseudométrica sobre X .

Definición 2.4.1. Sea X un conjunto. Una aplicación $f : X^2 \rightarrow \mathbb{R}_+ := [0, \infty]$ se llama una *pseudométrica* o *semidistancia* sobre X si cumple:

$$(D1) \quad f(x, y) = f(y, x) \text{ (simetría)}$$

$$(D2) \quad f(x, y) \leq f(x, z) + f(z, y) \text{ (desigualdad triangular)}$$

$$(D3) \quad x = y \implies f(x, y) = 0$$

para todos los $x, y, z \in X$.

El par (X, f) se llama espacio *pseudométrico*.

Para ver si una pseudométrica es una métrica hay que verificar que todos sus valores son reales no negativos y se cumple que $f(x, y) \neq 0$ para todos los $x, y \in X$ con $x \neq y$.

Ejemplo 16. Sea X un conjunto y $h : X \rightarrow \mathbb{R}$ una función real sobre X . Entonces la aplicación, definida por $f(x, y) := |h(x) - h(y)|$ ($x, y \in X$) es una pseudométrica sobre X y f es una métrica, si h es inyectiva.

Partiendo de una pseudométrica o de una familia de pseudométricas dadas, y por toda una serie de operaciones bastante generales, se pueden obtener nuevas pseudométricas. Es precisamente la variedad de estas posibilidades lo que constituye la utilidad de las pseudométricas.

- 1) La suma de una familia (finita o no) de pseudométricas es una pseudométrica.
- 2) Todo límite puntual de pseudométricas es una pseudométrica.
- 3) El supremo de una familia (finita o no) de pseudométricas es una pseudométrica.
- 4) Si $\pi : X \rightarrow Y$ es una aplicación y $g : Y^2 \rightarrow \overline{\mathbb{R}}_+$ una pseudométrica, entonces se define por

$$f(x, y) := g(\pi(x), \pi(y)) \quad (x, y \in X) \quad (35)$$

una nuevaseudométrica sobre X que se llama imagen recíproca de la pseudométrica g .

- 5) Si $f : X^2 \rightarrow \overline{\mathbb{R}}_+$ es una pseudométrica y $\varphi: \overline{\mathbb{R}}_+ \rightarrow \mathbb{R}_+$ una función creciente con $\varphi(0) = 0$ y $\varphi(x + y) \leq \varphi(x) + \varphi(y)$, entonces $g := \varphi \circ f$ es otra pseudométrica sobre X .
- 6) De forma general, si $f_1, \dots, f_n : X^2 \rightarrow \overline{\mathbb{R}}_+$ son pseudométricas, $f = (f_1, \dots, f_n) : X^2 \rightarrow (\overline{\mathbb{R}}_+)^n$ y $\varphi : (\mathbb{R}_+)^n \rightarrow \overline{\mathbb{R}}_+$ una función creciente con $\varphi(0) = 0$ y $\varphi(x + y) \leq \varphi(x) + \varphi(y)$, entonces $g := \varphi \circ f$ es otra pseudométrica sobre X . Sobre el conjunto ordenado $(\overline{\mathbb{R}}_+)^n$, provisto del orden producto:

$$x \leq y \iff x_1 \leq y_1, \dots, x_n \leq y_n$$

para $x, y \in (\overline{\mathbb{R}}_+)^n$, una función real se llama creciente, si cumple: $x \leq y \Rightarrow \varphi(x) \leq \varphi(y)$ para todos los $x, y \in (\overline{\mathbb{R}}_+)^n$.

Veamos algunas aplicaciones de estas operaciones

Ejemplo 17 (convergencia uniforme sobre un subconjunto). Sea X un conjunto de aplicaciones de un conjunto $A \neq \emptyset$ en un espacio métrico (Y, d) y sea $B \subset A$. Para todo $b \in B$ la función

$$d_b : (x, y) \rightsquigarrow d_b(x, y) := d(x(b), y(b)) \quad (36)$$

es una pseudométrica sobre X .

Luego, según 3, la función

$$d_B : (x, y) \rightarrow \sup_{b \in B} d_b(x, y) \quad (37)$$

es una pseudométrica sobre X .

d_B mide la proximidad de las funciones $x, y \in X$ sobre el conjunto $B \subset A$.

Para $B = A$ obtenemos la pseudométrica $d_A = d_\infty$ de la convergencia uniforme (ejemplo 4).

Ejemplo 18 (producto finito de espacios métricos). Veamos otra forma de introducir las métricas d_2, d_1 y d_∞ ((28)-(30)) en un producto cartesiano de espacios métricos.

Mencionemos primeramente que la función

$$x \rightsquigarrow \left(\sum_{i=1}^n x_i^2 \right)^{1/2} \quad \text{sobre } (\overline{\mathbb{R}}_+)^n,$$

tiene las propiedades requeridas en 6 (por la desigualdad triangular para la norma euclidea). Sea $(X_i, d_i), i = 1, \dots, n$ una familia finita de espacios métricos; las funciones

$$(x, y) \rightsquigarrow d_i(x_i, y_i) \quad (i = 1, \dots, n)$$

son pseudométricas sobre el producto $X := \prod_{i=1}^n X_i$. Luego, según las propiedades 6.3 y 1, las aplicaciones d_2, d_∞, d_1 definidas por (28)-(30) son pseudométricas sobre X . Es evidente que son de hecho métricas, y que en efecto cumplen las desigualdades (31).

Veamos ahora como puede ser construido un espacio métrico a partir de un espacio pseudométrico.

Ejemplo 19 (espacio cociente). Sea (X, d) un espacio pseudo métrico. La relación entre elementos del conjunto X definida por:

$$x \mathbf{R} y \Leftrightarrow d(x, y) = 0$$

es una relación de equivalencia en X (probarlo). La aplicación d_R definida en $X/R \times X/R$ mediante:

$$d_R(\bar{x}, \bar{y}) = d(x, y) \quad (38)$$

donde $\bar{x}, \bar{y}$ representan las clases de los elementos x y y respectivamente, no depende de la elección del representante de la clase, y es además, una métrica en el conjunto cociente ya que

$$d_R(\bar{x}, \bar{y}) = 0 \Rightarrow d(x, y) = 0 \Rightarrow \bar{x} = \bar{y}.$$

Al espacio métrico obtenido $(X/R, d_R)$ se le llama espacio métrico cociente, el cual se obtuvo a partir de la pseudométrica d .

Sea (Y, d) un espacio métrico, X un conjunto y $f : X \longrightarrow Y$ una aplicación. Entonces se puede definir una pseudométrica ρ en X mediante:

$$\rho(x, y) = d(f(x), f(y)) \quad x, y \in X.$$

Consideremos la relación de equivalencia

$$\mathbf{R}_f : x \mathbf{R}_f y \Leftrightarrow f(x) = f(y);$$

entonces sobre el espacio $X/\mathbf{R}_f$ se puede definir la métrica cociente

$$\bar{\rho}(\bar{x}, \bar{y}) = \rho(x, y)$$

(ver la descomposición canónica de una función en el capítulo I).

Mediante el procedimiento anterior, y teniendo en cuenta que la función f realiza una *identificación* entre elementos de X , podemos encontrar ejemplos interesantes de espacios métricos cociente.

Supongamos que tenemos un espacio métrico (X, d) y veamos cómo se pueden construir nuevos espacios métricos identificando algunos de sus puntos. Se dice que la función $f : X \longrightarrow X$ realiza una identificación entre los

elementos de cada subconjunto $F \subset X$ de la familia $\mathbf{F}$, si para cada $F \in \mathbf{F}$ se tiene:

$$f(x) = f(y) \in F \text{ para todos } x, y \in F,$$

y además,

$$f(x) = x \text{ para todo } x \in X \setminus \bigcup \{F : F \in \mathbf{F}\}.$$

Con este tipo de identificación, el espacio $X/\mathbf{R}_f$ tiene una *realización* muy simple desde el punto de vista de la teoría de conjuntos: $X/\mathbf{R}_f$ se obtiene uniendo un elemento de cada $F \in \mathbf{F}$ al conjunto $X \setminus \bigcup \{F : F \in \mathbf{F}\}$.

(Analice donde se ha utilizado aquí el axioma de elección.)

La forma en que se *distribuyan* estos elementos dentro de $X/\mathbf{R}_f$ puede dar lugar a algunos espacios métricos interesantes.

Veamos algunos ejemplos de espacios métricos sumergidos en el espacio tridimensional $\mathbb{R}^3$ obtenidos por el procedimiento anterior.

Ejemplo 20. El círculo E^1 se obtiene partiendo del intervalo $[0, 1]$ por identificación de los puntos 0 y 1 (fig. 22).

Si identificamos todos los puntos de la tapa de un cilindro obtenemos un cono (fig.23).

Ejemplo 21. Partiendo del cuadrado $I^2 = [0, 1] \times [0, 1]$, obtenemos dos espacios diferentes por identificación de los puntos de dos lados opuestos (fig. 24).

Las flechas indican cuándo identificamos los lados en igual sentido o en sentido opuesto. En el caso a) obtenemos la superficie lateral de un cilindro. En el caso b) obtenemos la famosa banda de *Möbius* que tiene un único *borde*.

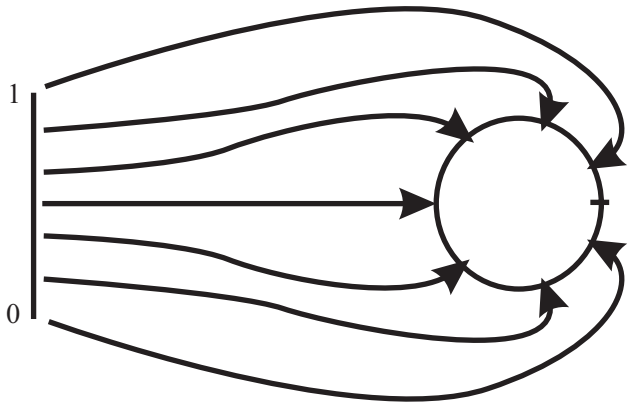


Figura 22

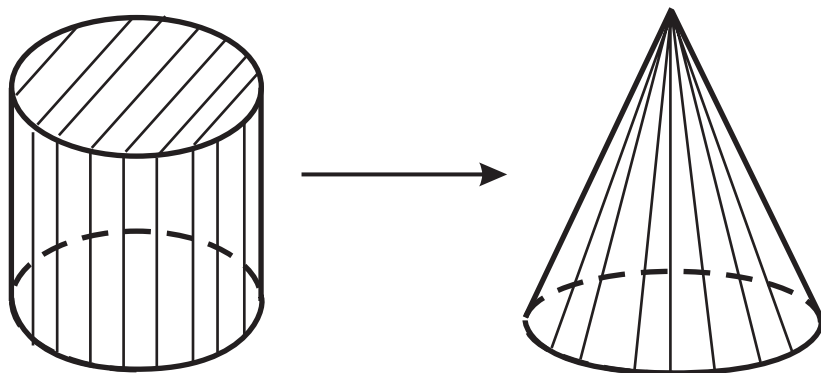


Figura 23



Cilindro



Banda de Möbius

Figura 24

Por identificación sucesiva de dos pares de lados opuestos de I^2 (en el mismo sentido) obtenemos un anillo de dimensión 2 (fig. 25).

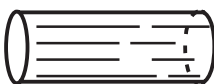
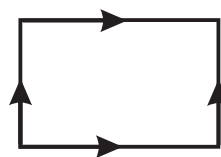


Figura 25

Trataremos ahora de identificar los lados de I^2 según el esquema que aparece en la (fig. 26).

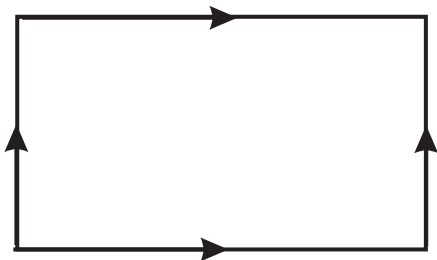


Figura 26

El resultado (la llamada botella de Klein), ya no puede representarse de manera adecuada en $\mathbb{R}^3$. Pero existe un modelo tridimensional que nos sirve para formarnos una idea, aunque tenga autointersecciones que la botella de Klein no tiene (fig. 27).

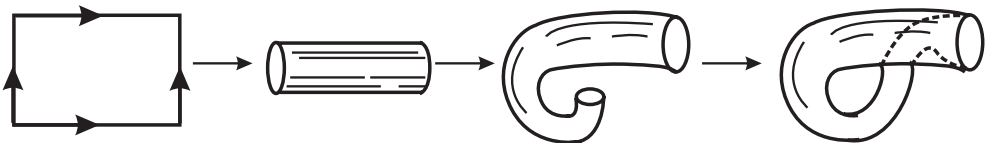


Figura 27

Una identificación de puntos tan sencilla como en el ejemplo 21, transforma una superficie elemental en un espacio relativamente complejo, de forma tal que nuestros conceptos intuitivos se confunden. Esto nos muestra por qué es necesario precisar formalmente la noción de un espacio obtenido por identificación de ciertos puntos de un espacio métrico.

Ejemplo 22 (seudométrica definida por una familia numerable de seudométricas). Sea X un conjunto y $(d_n)_{n \in \mathbb{N}}$ una familia numerable de seudométricas en X . Entonces las expresiones

$$\rho'(x,y) := \sum_{n=1}^\infty \frac{1}{2^n} \frac{d_n(x,y)}{1+d_n(x,y)} \tag{39}$$

$$\rho''(x,y) := \sum_{n=1}^\infty \frac{1}{2^n} \min \{1, d_n(x,y)\} \tag{40}$$

definen seudométricas acotadas en X .

Si las seudométricas $(d_n)_{n \in \mathbb{N}}$ constituyen además una familia separante, es decir; para todo par de puntos $x \neq y$ en X existe $n \in \mathbb{N}$ tal que $d_n(x,y) \neq 0$;

entonces las seudométricas (39) y (40) son además métricas. Estas métricas juegan un papel importante en la teoría de la convergencia en los espacios métricos que desarrollaremos en el próximo epígrafe.

Ejemplo 23 (producto numerable de espacios métricos). Sean (X_n, d_n) ($n \in \mathbb{N}$) una sucesión de espacios métricos. La sucesión de seudométricas

$$D_n(x, y) = d_n(x_n, y_n) \quad (n \in \mathbb{N})$$

donde $x = (x_n)$ y $y = (y_n)$ son elementos de $X = \prod_{n \in \mathbb{N}} Y_n$, es una familia separante.

Según el ejemplo 22 la expresión

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} \quad (41)$$

define una métrica en el producto cartesiano X .

Otras propiedades de esta métrica serán vistas en los dos epígrafes siguientes,

Pasemos ahora a estudiar las seudométricas comparables a las cuales ya hemos hecho referencia en el epígrafe 2.3.

Definición 2.4.2. Dos seudométricas d y d' sobre un conjunto K se llaman semejantes, si existen números reales $k, k' > 0$ tales que

$$d(x, y) \leq k' d'(x, y) \text{ y } d'(x, y) \leq k d(x, y) \quad (42)$$

para todos los $x, y \in X$.

Es fácil comprobar que la semejanza es una relación de equivalencia entre las seudométricas sobre un conjunto. Una seudométrica semejante a una métrica es una métrica. Según (31) las tres métricas canónicas sobre un producto finito de espacios métricos son semejantes. Esto explicará por qué pueden utilizarse equivalentemente, según convenga. Por lo mismo, las dos métricas acotadas d', d'' , canónicamente asociadas a una métrica d (ejemplo 13), son semejantes. Sin embargo, según (42), una métrica no acotada no puede ser semejante a una métrica acotada, luego una métrica no acotada d sobre un conjunto X no es semejante a la métrica $d' = \inf\{1, d\}$ ni a la métrica $d'' = d/(1 + d)$.

¿Son las métricas d_∞ (8) y d_1 (9) sobre $C([0, 1], \mathbb{R})$ semejantes?

¿Son las métricas d_1 (9) y d_2 (10) semejantes sobre $C([0, 1], \mathbb{R})$?

Regresemos a nuestro problema inicial o sea a los diferentes conceptos de convergencia para sucesiones de funciones. La diferencia estructural entre la

convergencia puntual y los demás tipos de convergencia los podemos precisar ahora de la manera siguiente:
para estudiar la convergencia puntual en un conjunto de funciones $X \subset \mathbb{R}^A$ debemos considerar toda una familia de pseudométricas $(d_a)_{a \in A}$ mientras que la convergencia uniforme, la convergencia en media y la convergencia en media cuadrada, se definen por una sola métrica, de manera tal que para su estudio, la noción de espacio métrico nos sirve como modelo abstracto.

(f_n) converge puntualmente a	$A \neq \emptyset$	(43)
$f \Leftrightarrow [d_a(f_n, f)]$ converge a 0		
para cada $a \in A$	$f_n, f : A \rightarrow \mathbb{R}$	
(f_n) converge uniformemente a	$d_a(f_n, f) := f_n(a) - f(a) $	
$f \Leftrightarrow [d_\infty(f_n, f)]$		
converge a 0	$d_\infty(f_n, f) := \sup_{a \in A} f_n(a) - f(a) $	

La teoría de los espacios métricos se puede desarrollar hasta un punto en que se puede concebir una teoría axiomática todavía más general que engloba también la convergencia puntual. Puesto que la teoría de los espacios métricos generaliza directamente los conceptos y métodos utilizados por el Análisis Clásico en su tratamiento de la convergencia y continuidad, el proceso de aprendizaje en los próximos cuatro epígrafes debe ser más rápido para el lector. De esta forma, con el contenido de esos epígrafes se persiguen dos objetivos fundamentales:

- Primero* Servir como autocontrol al lector para ver si ha entendido la estructura abstracta de los razonamientos métricos del Análisis Clásico.
- Segundo* Familiarizar al lector con conceptos topológicos a un nivel intermedio como preparación didáctica para un estudio posterior de la Topología General.

Ejercicios de Autocontrol

1. Demuestre las proposiciones 1-6 relativas a las operaciones con pseudométricas. Utilice la proposición 5 en el desarrollo del ejemplo 13.
2. Demuestre que las expresiones (39) y (40) del ejemplo 22 son pseudométricas.
3. Demuestre que la relación definida en el ejemplo 19 es de equivalencia

2.5 SUCESIONES CONVERGENTES EN UN ESPACIO MÉTRICO

Recordemos la definición de una sucesión convergente de números reales, dada en la asignatura Análisis Matemático:

Sea $(x_n)_{n=1}^{\infty}$ una sucesión de elementos de $\mathbb{R}$. Se dice que (x_n) converge al punto $x \in \mathbb{R}$, si para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que

$$|x_n - x| < \epsilon \text{ para } n \geq N$$

Sustituyendo en esta definición el conjunto $\mathbb{R}$ por X y la expresión $|x_n - x|$ por $d(x_n, x)$, podemos definir la convergencia en un espacio métrico (X, d) .

Definición 2.5.1. Sea (X, d) un espacio métrico y (x_n) una sucesión de elementos de X . Se dice que (x_n) converge a $x \in X$, si para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que

$$d(x, x_n) < \epsilon \text{ para } n \geq N. \quad (44)$$

En este caso al punto x se le denomina *límite* de la sucesión (x_n) . Se utiliza también la notación: $(x_n) \xrightarrow{d} x$ o $\lim(x_n) = x$ (también puede eliminarse el paréntesis).

Observación: La métrica d permite reducir el problema de la convergencia de la sucesión (x_n) al punto x en el espacio métrico (X, d) a un problema de convergencia de números reales. En efecto, $(x_n) \xrightarrow{d} x \iff$ la sucesión de números reales $(d(x_n, x))$ converge a 0 en $\mathbb{R}$.

Mediante la distancia d hemos subordinado el concepto abstracto de acercamiento entre dos puntos de X al ya conocido de acercamiento entre dos puntos de $\mathbb{R}$.

El enunciado $(x_n) \xrightarrow{d} x$ no depende tanto de la métrica particular d como puede parecer a simple vista. Según (42) y la definición 2.5.1, una sucesión (x_n) con $(x_n) \xrightarrow{d} x$ converge a x con respecto a toda métrica que sea semejante a d . Pero lo que es aún más sorprendente, dos métricas no semejantes como son las métricas d y $\frac{d}{1+d}$ cuando, d no es acotada, definen la misma convergencia, es decir

$$(x_n) \xrightarrow{d} x \iff (x_n) \xrightarrow{\frac{d}{1+d}} x \quad (45)$$

Esto nos muestra que en definitiva, el concepto *métrica* no parece ser el concepto primario para la construcción axiomática de una teoría general de la

convergencia.

Ejemplo 24 (sucesión convergente en un espacio métrico discreto).

En un espacio métrico discreto (véase ejemplo 9), una sucesión (x_n) converge a un punto x si, y sólo si, todos los términos x_n son iguales a x a partir de cierto número $N \in \mathbb{N}$. En efecto, tomando $\epsilon < 1$ se tendría, por la definición de límite, que existe $N \in \mathbb{N}$ tal que

$$d(x, x_n) < \epsilon < 1 \text{ para todo } n \geq N.$$

Pero por definición de la métrica discreta $d(x, x_n) < 1$ implica que $x_n = x$, y por lo tanto $x_n = x$ para todo $n \geq N$.

Ejemplo 25. Sea X el conjunto $C([0, 1], \mathbb{R})$ y consideremos los espacios métricos (X, d_∞) y (X, d_1) , donde d_∞ y d_1 están definidas en (8) y (9) respectivamente. La sucesión (f_n) de elementos de X , definida por:

$$f_n(x) = \begin{cases} 2n^3x, & \text{si } 0 \leq x \leq 1/2n^2 \\ 2n - 2n^3x & \text{si } 1/2n^2 \leq x \leq 1/n^2 \\ 0, & \text{si } 1/n^2 \leq x \leq 1 \end{cases}$$

converge a la función constante $\langle 0 \rangle$ en el espacio metrico (X, d_1) (fig. 28) ya

que $\int_0^1 |f_n(x) - \langle 0 \rangle| dx = \frac{1}{n} \longrightarrow 0$, pero la misma sucesión no converge a $\langle 0 \rangle$

en (X, d_∞) puesto que $d_\infty(f_n, \langle 0 \rangle) = n$

Sin embargo, no es difícil probar que si $f_n, f \in X$, entonces

$$(f_n)_{d_\infty} \rightarrow f \Rightarrow (f_n)_{d_1} \rightarrow f \tag{46}$$

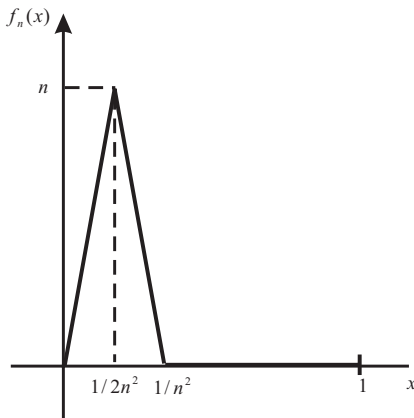


Figura 28

En el 2.8 estudiaremos con más profundidad la relación entre las métricas sobre un conjunto que tienen la propiedad de definir la misma clase de sucesiones convergentes.

Hasta el momento hemos visto como el concepto *convergienciase* define con respecto a un espacio métrico. En efecto, una sucesión puede ser convergente como sucesión de puntos de un cierto espacio métrico, y no serlo como sucesión de un cierto subespacio. Por ejemplo, la sucesión de números reales $\frac{1}{n}$ es convergente al punto 0 en el espacio métrico $[0, 1]$ con la métrica usual definida por el módulo, pero no es convergente en el subespacio $]0, 1]$ con la métrica inducida, ya que de converger sería al punto 0 que no está en este subespacio.

Denotemos por (A, d_A) un subespacio del espacio métrico (X, d) . Resulta útil para el lector distinguir cuáles de los siguientes enunciados son válidos:

- toda sucesión convergente en (A, d_A) es convergente en (X, d) ;
- una sucesión de puntos de A es convergente en (A, d_A) si y sólo si es convergente en (X, d)
- una sucesión de puntos de A es convergente en (A, d_A) si y sólo si converge a un punto de A en (X, d) .

En caso de respuesta negativa trate de encontrar un contraejemplo y demuestre las afirmaciones que considere verdaderas.

Ejemplo 26 (convergencia definida por una familia de pseudométricas). Sea $D = \{d_\alpha\}_{\alpha \in J}$ una familia separante de pseudométricas sobre el conjunto X (véase ejemplo 22). Se dice que la sucesión (x_n) converge al punto x según la familia de pseudométricas $D((x_n) \xrightarrow{D} x)$ si:

$$(x_n) \xrightarrow{d_\alpha} x \text{ para toda } d_\alpha \in D.$$

Si la familia D es numerable, entonces la convergencia según la familia D es *metrizable* en el sentido de que se puede construir una métrica ρ' , tal que:

$$(x_n) \xrightarrow{D} x \Leftrightarrow (x_n) \xrightarrow{\rho'} x.$$

En efecto, si $J = \mathbb{N}$, basta tomar ρ' como en (39)

$$\rho'(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot \frac{d_k(x, y)}{1 + d_k(x, y)}.$$

Probemos que la convergencia según ρ' implica la convergencia según D cuando D es numerable.

Sea $\epsilon > 0$, entonces si $(x_n) \xrightarrow{\rho'} x$ existe $N \in \mathbb{N}$ tal que

$$\rho'(x_n, x) < \epsilon \text{ para todo } n \geq N, \quad (47)$$

donde $\epsilon' = \frac{1}{2^k} \cdot \frac{\epsilon}{1 + \epsilon}$.

Pero de la definición de ρ' se tiene

$$\frac{1}{2^k} \cdot \frac{d_k(x_n, x)}{1 + d_k(x_n, x)} \leq \rho'(x_n, x) \text{ para todo } k \text{ y } n.$$

Luego por (47) para cada k fijo

$$\frac{d_k(x_n, x)}{1 + d_k(x_n, x)} < \frac{\epsilon}{1 + \epsilon} \text{ para todo } n \geq N. \quad (48)$$

Del hecho que la función $f : [0, \infty[\longrightarrow \mathbb{R}$ tal que

$$f(t) = \frac{t}{1 + t}$$

es estrictamente creciente, lo cual puede comprobarse verificando que la derivada $f'(t) > 0$, y de (48), se deduce que $d_k(x_n, x) < \epsilon$ para todo $n \geq N$, y por lo tanto, $(x_n) \xrightarrow{d_k} x$ para todo k .

Recíprocamente, supongamos que $(x_n) \xrightarrow{d_k} x$ para todo k y fijemos $\epsilon > 0$.

Primeramente elijamos $k_0 \in \mathbb{N}$ tal que

$$\sum_{k=k_0+1}^{\infty} \frac{1}{2^k} < \frac{\epsilon}{2}$$

lo cual es posible por ser convergente la serie

$$\sum_{k=1}^{\infty} \frac{1}{2^k}.$$

Por otra parte, para cada $k \in \{1, 2, \dots, k_0\}$ y por ser $(x_n) \xrightarrow{D} x$ se puede elegir

un número $N_k \in \mathbb{N}$ tal que $d_k(x_n, x) < \frac{\epsilon}{2}$ para todo $n \geq N_k$.

Luego, para todo $N \geq \max\{N_k : 1 \leq k \leq k_0\}$,

$$\rho'(x_n, x) = \sum_{k=1}^{k_0} \frac{1}{2^k} \frac{d_k(x_n, x)}{1 + d_k(x_n, x)} + \sum_{k=k_0+1}^{\infty} \frac{1}{2^k} \frac{d_k(x_n, x)}{1 + d_k(x_n, x)} \leq$$

$$\leq \frac{\epsilon}{2} \sum_{k=1}^{k_0} \frac{1}{2^k} + \sum_{k=k_0+1}^{\infty} \frac{1}{2^k} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

con lo cual queda probado que $(x_n) \xrightarrow{\rho'} x$.

Veremos en el próximo ejemplo que en general, la convergencia según una familia no numerable de seudométricas no es metrizable.

Ejemplo 27 (convergencia en los espacios producto). Sea

$\{(X_\alpha, d_\alpha)_{\alpha \in J}\}$ una familia de espacios métricos y sea $X = \prod_{\alpha \in J} X_\alpha$ el conjunto producto cartesiano de los X_α . Una sucesión en X la denotaremos por $(x_n) = \left(\left\{ x_n^{(\alpha)} \right\}_{\alpha \in J} \right)$, donde el supraíndice (α) denota la α -ésima coordenada del elemento n -ésimo $x_n \in X$ de la sucesión. No resulta difícil comprobar que si J es finito (producto finito), es decir, $J = \{1, 2, \dots, N\}$ entonces para el espacio métrico $X = \prod_{k=1}^N X_k$ con cualquiera de las métricas producto d_2, d_1 o d_∞ definidas en (28)-(30), se cumple la propiedad siguiente:

La sucesión $(x_n^{(1)}, x_n^{(2)}, \dots, x_n^{(N)})$ de elementos de X converge al punto $(x^{(1)}, x^{(2)}, \dots, x^{(N)})$ de X según cualquiera de las métricas producto ρ_1, ρ_2 , o ρ_∞ , si y sólo si $(x_n^{(k)}) \xrightarrow{d_k} x^{(k)}$ para todo $1 \leq k \leq N$, donde d_k es la métrica en el espacio coordenado X_k .

Esto quiere decir que la métrica producto se define para un producto finito de manera tal que la convergencia de sucesiones sea exactamente la convergencia por coordenadas, como ocurre en los espacios euclidianos $\mathbb{R}^n$, estudiados en el Análisis Matemático.

Consideremos ahora el problema de definir una métrica en $\prod_{\alpha \in J} X_\alpha$ cuando J no es un conjunto finito, de manera que siga cumpliéndose la propiedad anterior. Del resultado expuesto en el ejemplo 26 se deduce que si J es numerable, el conjunto $\prod_{\alpha \in J} X_\alpha$ puede ser provisto de una métrica d para la cual se cumpla

$$\left(\left\{ x_n^{(k)} \right\} \right) \xrightarrow{d} \left(x^{(k)} \right) \Leftrightarrow \left(x_n^{(k)} \right) \xrightarrow{d_k} x^{(k)} \text{ para todo } k \in \mathbb{N}. \quad (49)$$

En efecto, basta tomar en lugar de la métrica d , la definida por la expresión (41) en el ejemplo 23.

Puede verse además, que si J es un conjunto infinito no numerable, no es posible encontrar una métrica d sobre $\prod_{\alpha \in J} X_\alpha$ que satisfaga la propiedad de convergencia por coordenadas.

Para cerciorarse de ello basta tomar

$$J = [0, 1], X_\alpha = \mathbb{R}$$

para todo α y d_α la métrica usual definida por el módulo en $\mathbb{R}$ para todo α . En este caso el conjunto $\prod_{\alpha \in J} X_\alpha$ se identifica con el conjunto $\mathbb{R}^{[0,1]}$ de todas las funciones reales definidas sobre $[0, 1]$ y la convergencia por coordenadas no es más que la convergencia puntual de funciones de $[0, 1]$ en $\mathbb{R}$. Demostraremos

en 2.14 que no existe ninguna métrica d sobre $\mathbb{R}^{[0,1]}$ para la cual se cumpla que si f_n y $f : [0, 1] \rightarrow \mathbb{R}$, entonces:

$$f_n \text{ converge puntualmente a } f \text{ sobre } [0, 1] \Leftrightarrow (f_n) \xrightarrow{d} f.$$

Hasta el momento hemos hablado de sucesiones convergentes en espacios métricos sin referirnos a la unicidad del límite cuando este existe. En muchos problemas del análisis matemático, es fundamental el hecho de que cada sucesión converge a un solo límite. Como veremos en el próximo teorema, esto también ocurre en los espacios métricos.

Teorema 2.5.2 (unicidad del límite). Sea (X, d) un espacio métrico y (x_n) una sucesión convergente de puntos de X ; entonces el límite de (x_n) es único.

Demostración: En efecto, supongamos contrariamente que existen dos límites x y y de (x_n) ; entonces, por la desigualdad triangular:

$$d(x, y) \leq d(x, x_n) + d(x_n, y).$$

Si n tiende a infinito tendríamos que $d(x, y) = 0$, de donde $x = y$ debido al axioma (D3) de métrica. ■

Definición 2.5.3. Sea (x_n) una sucesión en el conjunto X . Se dice que la sucesión (y_n) es una *subsucesión* de la sucesión (x_n) si

$$y_n = x_{\varphi(n)}$$

donde $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ es una función creciente.

Proposición 2.5.4 (límite de una subsucesión). Si (x_n) es una sucesión en (X, d) tal que $(x_n) \xrightarrow{d} x$, entonces toda subsucesión $(x_{\varphi(n)})$ de (x_n) será convergente y

$$(x_{\varphi(n)}) \xrightarrow{d} x.$$

La demostración se deduce inmediatamente de la definición de límite.

Ejercicios de Autocontrol

1. Muestre que en un espacioseudométrico (X, d) , una sucesión convergente puede tener dos límites diferentes.
2. Demuestre que una pseudométrica d sobre X es una métrica si y sólo si las sucesiones convergentes en (X, d) tienen un límite único.
3. Sea d una métrica definida sobre el conjunto X . Demuestre que:

$$(x_n) \xrightarrow{d} x \Leftrightarrow (x_n) \xrightarrow{\min(1, d)} x$$

4. Demuestre la afirmación (46) del ejemplo 25.

5. Sea (x_n) una sucesión en el espacio métrico (X, d) . Pruebe que si (δ_n) es una sucesión de números reales positivos tal que $(\delta_n) \rightarrow 0$ y $d(x_n, a) \leq \delta_n$ entonces $(x_n) \xrightarrow{d} a$.
6. Construya una sucesión no convergente en algún espacio métrico que tenga
 - a) una subsucesión convergente;
 - b) infinitas subsucesiones convergentes.

2.6 FUNCIONES CONTINUAS

Consideremos ahora el segundo concepto fundamental del Análisis, *continuidad*; su definición para espacios métricos y espacios pseudométricos es también una generalización directa de la definición para funciones reales de una variable. En efecto, recordemos que para $f : \mathbb{R} \rightarrow \mathbb{R}$ se dice que f es continua en el punto $a \in \mathbb{R}$, si para todo número $\epsilon > 0$ existe un $\delta > 0$ (donde $\delta = \delta(\epsilon, a)$) tal que si $x \in \mathbb{R}$ satisface la relación $|x - a| < \delta$, entonces $|f(x) - f(a)| < \epsilon$.

Sustituyendo en esta definición el módulo por métricas d y d' arbitrarias, obtenemos la definición general de continuidad para funciones entre espacios métricos o espacios pseudométricos.

Definición 2.6.1 (función continua en un punto). Sean $(X, d), (Y, d')$ espacios pseudométricos. Una aplicación $f : X \rightarrow Y$ se llama continua en el punto $a \in X$ (con respecto a las métricas d y d' respectivamente), si para todo número real $\epsilon > 0$, existe un número real $\delta > 0$ ($\delta = \delta(\epsilon, a)$) tal que si $x \in X$ satisface la relación $d(x, a) < \delta$, entonces $d'(f(x), f(a)) < \epsilon$.

Ejemplo 28. Sean (X, d) un espacio métrico y $A \subset X$, $A \neq \emptyset$. Entonces para todos $x, y \in X$, $z \in A$ se tiene

$$d(x, z) \leq d(x, y) + d(y, z).$$

Tomando el ínfimo para todos los $z \in A$ y considerando

$$\begin{aligned} d(x, A) &= \inf_{z \in A} d(x, z) \quad (\text{distancia de } x \text{ a } A), \\ d(y, A) &= \inf_{z \in A} d(y, z) \quad (\text{distancia de } y \text{ a } A), \end{aligned} \tag{50}$$

tendremos $d(x, A) - d(y, A) \leq d(x, y)$, de donde intercambiando x y y se obtiene $|d(x, A) - d(y, A)| \leq d(x, y)$.

Esto nos muestra que para todo subconjunto no vacío A de un espacio métrico (X, d) , la función $f : x \rightsquigarrow d(x, A)$, de X en $\mathbb{R}$ es continua en todo $a \in X$.

Intuitivamente esto significa que la distancia $d(x, A)$ aproxima la distancia $d(a, A)$ si x está cerca de a .

En efecto, para cualquier $\epsilon > 0$ es suficiente poner $\delta = \delta(a, \epsilon) = \epsilon$ para obtener

$$d'(f(x), f(a)) = |f(x) - f(a)| = |d(x, A) - d(a, A)| < \epsilon$$

para todo $x \in X$ con $d(x, a) < \delta$.

Ejemplo 29. Sea (X, d) un espacio métrico discreto y (Y, d) un espacio métrico arbitrario, entonces toda función $f : X \longrightarrow Y$ es continua.

En efecto, para cualquier $\epsilon > 0$ basta tomar $\delta < 1$ ya que $d(x, a) < \delta < 1$ implica necesariamente $x = a$, y por lo tanto, es obvio que

$$d'(f(x), f(a)) = d(f(a), f(a)) = 0.$$

Ejemplo 30. Consideremos el espacio $C([0, 1], \mathbb{R})$ provisto de la métrica d_∞ y sea $f \in C([0, 1], \mathbb{R})$ una función fija no idénticamente nula. Entonces la aplicación

$$F : C([0, 1], \mathbb{R}) \longrightarrow \mathbb{R}$$

$$g \rightsquigarrow \int_0^1 f(x) g(x) dx$$

es continua en todo $g \in C([0, 1], \mathbb{R})$.

En efecto para todas las funciones g y h se tiene

$$|F(g) - F(h)| = \left| \int_0^1 f(x) (g(x) - h(x)) dx \right| \leq \int_0^1 |f(x)| dx d_\infty(g, h),$$

y por tanto, dado $\epsilon > 0$ basta tomar $\delta = \epsilon / \int_0^1 |f(x)| dx$ para cerciorarse de que si $d_\infty(g, h) < \delta$ entonces $|F(g) - F(h)| < \epsilon$, con lo cual queda probada la continuidad de F en g .

Ejemplo 31 (continuidad de la sobreyección canónica). Sea X un conjunto, (Y, d) un espacio métrico y $f : X \longrightarrow Y$ una función para la cual consideraremos su descomposición canónica

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & X/R_f \\ & \searrow & \downarrow \bar{f} \\ & & Y \\ f & & \end{array}$$

donde $f = \bar{f} \circ \varphi$ (ver capítulo 1).

Si sobre X y X/R_f ponemos respectivamente la pseudométrica y la métrica definidas en el ejemplo 19, entonces tanto φ como $\bar{f}$ son continuas en cada punto de sus dominios

Ejemplo 32 (aplicación no continua). La aplicación identidad de $\mathbb{R}^2$

$$i : \mathbb{R}^2 \xrightarrow{(x,y) \rightsquigarrow (x,y)} \mathbb{R}^2$$

no es continua si en el espacio de partida se pone la métrica euclídeana d_2 y en el de llegada la métrica del bosque d (ver ejercicio 3 de autocontrol de 2.2). En efecto, si $b \neq 0$, entonces i no puede ser continua en (a, b) ya que tomando $\epsilon = \frac{|b|}{2}$ no existe ningún $\delta > 0$ tal que si $d_2((x, y), (a, b)) < \delta$ se tenga que $d((x, y), (a, b)) < \epsilon$ (fig. 29).

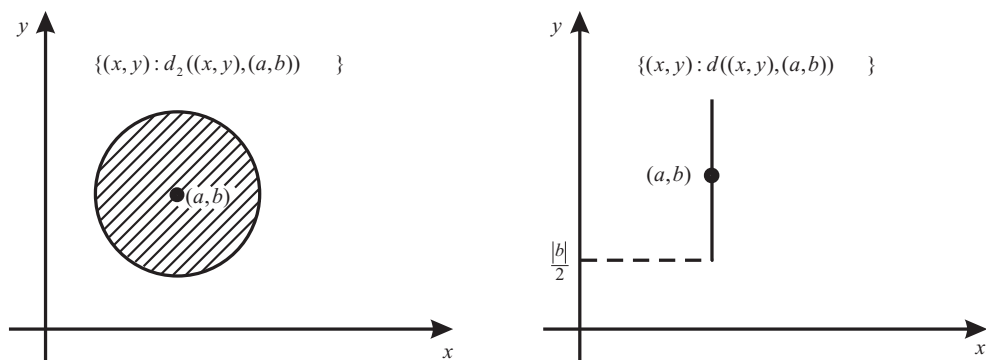


Figura 29

Veamos ahora una caracterización importante de la continuidad mediante la convergencia de sucesiones, lo cual nos destaca aún más la importancia del concepto *límite*.

Teorema 2.6.2 (caracterización de la continuidad mediante la convergencia de sucesiones). Sean (X, d) y (Y, d) espacios métricos. Una función $f : X \rightarrow Y$ es continua en el punto $a \in X$ si y sólo si para toda sucesión $(x_n) \xrightarrow{d} a$ se tiene que

$$(f(x_n)) \xrightarrow{d} f(a).$$

Demostración: Probemos primeramente que esta condición es necesaria. Sea $(x_n) \xrightarrow{d} a$, probemos que $(f(x_n)) \xrightarrow{d} f(a)$.

Dado $\epsilon > 0$, por la continuidad de f en a , existe $\delta > 0$ tal que $d(x, a) < \delta \Rightarrow d(f(x), f(a)) < \epsilon$. Pero como $(x_n) \xrightarrow{d} a$, existe $N \in \mathbb{N}$ tal que si $n \geq N$,

$d(x_n, a) < \delta$ y por tanto, para todo $n \geq N$ $d'(f(x_n), f(a)) < \epsilon$, de donde por la arbitrariedad de $\epsilon > 0$, se deduce que $f(x_n) \xrightarrow{d'} f(a)$. ■

Demostraremos la suficiencia de esta condición por el contrarecíproco, es decir, supongamos que f no es continua en el punto a , y hallemos una sucesión $(x_n) \xrightarrow{d} a$ tal que $f(x_n) \not\xrightarrow{d'} f(a)$. En efecto, si f no es continua en el punto a , existe un $\epsilon > 0$ tal que para todo $n \in \mathbb{N}$ se puede encontrar un elemento x_n con $d(x_n, a) < \frac{1}{n}$ y al mismo tiempo $d(f(x_n), f(a)) > \epsilon$. Evidentemente, la sucesión (x_n) así construida satisface los requerimientos exigidos (compruébelo) con lo cual se concluye la demostración. ■

Si se utiliza esta caracterización se obtiene la proposición siguiente:

Proposición 2.6.3 (continuidad de las aplicaciones proyección). Sean

$\{(X_n, d_n)\}_{n=1}^{\infty}$ una familia numerable de espacios métricos y $\prod_{n=1}^{\infty} X_n$ el espacio métrico producto con la métrica definida por (41) en el ejemplo 23. Entonces las aplicaciones proyección

$$\pi_k : \prod_{n=1}^{\infty} X_n \longrightarrow X_k,$$

son continuas en todo punto de $\prod_{n=1}^{\infty} X_n$.

Demostración: En efecto, según nos muestra el ejemplo 27, para la métrica d del espacio producto se tiene que:

$(\{x_n^k\}_{k=1}^{\infty}) \xrightarrow{d} \{x^k\}_{k=1}^{\infty} \Leftrightarrow (\pi_k\{x_n^k\}) = (x_n^k) \xrightarrow{d_k} x^k$ para todo $k \in \mathbb{N}$, de donde se deduce inmediatamente, al aplicar el teorema 2.6.2 que π_k es continua. ■

Veamos varias proposiciones y ejemplos que nos exponen algunas propiedades elementales de la continuidad puntual.

Ejemplo 33 (continuidad de la suma y el producto por un escalar en $\mathbb{R}$). Las aplicaciones: suma

$$\begin{aligned} S : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (x, y) &\rightsquigarrow x + y, \end{aligned} \tag{51}$$

y producto por un escalar

$$\begin{aligned} P : \mathbb{R} \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (\lambda, x) &\rightsquigarrow \lambda x \end{aligned} \tag{52}$$

donde $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ son vectores, son continuas en cada punto de su dominio.

Notemos primeramente que al hacer esta afirmación en $\mathbb{R}^n$, $\mathbb{R}^n \times \mathbb{R}^n$ y $\mathbb{R} \times \mathbb{R}^n$, implícitamente estamos sobreentendiendo que la propiedad de una función de ser continua, no varía al ser cambiada la métrica en los espacios de partida y de llegada por métricas semejantes. Esto efectivamente es así (vea ejercicio 4 de autocontrol). No es difícil demostrar que si d_2 es la métrica euclídeana en $\mathbb{R}^n$, entonces

$$D_2((x_1, y_1), (x_2, y_2)) = d_2((x_1, x_2) + d_2(y_1, y_2))$$

es una métrica equivalente a la métrica euclídeana en $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$. Si $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$, entonces

$$d_2(x + y, a + b) \leq d_2(x, a) + d_2(y, b) = D_2((x, y), (a, b)),$$

y por tanto, dado $\epsilon > 0$ basta tomar $\delta = \epsilon$ de manera que si $D_2((x, y), (a, b)) < \delta$, se tendrá $d_2(x + y, a + b) < \epsilon$, con lo cual queda probada la continuidad de la suma. La demostración de la continuidad del producto se deja de ejercicio al lector.

Teorema 2.6.4 (continuidad de la función compuesta). Sean (X, d) , (Y, d') y (Z, d'') espacios pseudométricos, y las funciones $f : X \longrightarrow Y$ y $g : Y \longrightarrow Z$, tales que f es continua en el punto $a \in X$ y g es continua en el punto $b = f(a) \in Y$.

Entonces $g \circ f : X \longrightarrow Z$ es continua en el punto $a \in X$.

Demostración: En efecto, dado $\epsilon > 0$, por la continuidad de g en el punto $b \in Y$, se puede encontrar un número $\eta > 0$ tal que si $d'(y, b) < \eta$ entonces $d''(g(y), g(b)) < \epsilon$. Además, por la continuidad de f en el punto $a \in X$, podemos hallar un $\delta > 0$ tal que si $d(x, a) < \delta$ se tenga $d'(f(x), b) < \eta$, y por tanto, $d''(g(f(x)), g(b)) = d''(g \circ f(x), g \circ f(a)) < \epsilon$, con lo cual queda probada la continuidad de $g \circ f$ en el punto a . ■

Consideremos una familia $\{X_\alpha\}_{\alpha \in I}$ de conjuntos y una función $f : X \longrightarrow \prod_{\alpha \in I} X_\alpha$. La función f queda unívocamente determinada por sus funciones coordenadas $f_\alpha : \pi_\alpha \circ f$, donde

$$f_\alpha : X \longrightarrow X_\alpha, \quad (53)$$

por lo cual podemos utilizar la notación $f = (f_\alpha)_{\alpha \in I}$.

Teorema 2.6.5 (continuidad de una función con imagen en un espacio producto). Sean (X, d) y $\{(X_n, d_n)\}_{n=1}^\infty$ una familia numerable de espacios métricos y consideremos el espacio producto $\prod_{n=1}^\infty X_n$ provisto de la métrica

producto (vea expresión (41) del ejemplo 23). Entonces $f : X \longrightarrow \prod_{n=1}^{\infty} X_n$ es continua en el punto $a \in X$ si y solo si todas las funciones coordenadas $\pi_n \circ f = f_n : X \longrightarrow X_n$ son continuas en $a \in X$.

Demostración: Utilizando la caracterización de la continuidad mediante la convergencia de sucesiones (teorema 2.6.2), tendremos que f es continua en el punto $a \in X$ si y sólo si para toda sucesión (x_n) en X :

$$(x_n) \xrightarrow{d} a \Rightarrow (f(x_n)) \xrightarrow{d'} f(a) \quad (54)$$

donde por d' hemos denotado la métrica producto en $\prod_{n=1}^{\infty} X_n$. Pero recordemos que, por la equivalencia (49) del ejemplo 27 se tiene:

$$(f(x_n)) \xrightarrow{d'} f(a) \iff f_k(x_n) \xrightarrow{d_k} f_k(a) \quad (55)$$

para todo $k \in \mathbb{N}^*$, ya que $(f(x_n)) = (\{f_k(x_n)\}_{k=1}^{\infty})$ y de (54) y (55) se deduce el resultado deseado. ■

Corolario 1 (continuidad de las aplicaciones proyección). Sea $\{(X_n, d_n)\}_{n=1}^{\infty}$ una familia numerable de espacios métricos; entonces para todo $k = 1, 2, \dots$, las aplicaciones proyección

$$\pi_k : \prod_{n=1}^{\infty} X_n \longrightarrow X_k$$

son continuas en todo elemento $x = \{x^{(n)}\}$ del espacio métrico producto $\prod_{n=1}^{\infty} X_n$.

Para la demostración basta tomar en el teorema 2.6.5, el espacio $X = \prod_{n=1}^{\infty} X_n$ y la función $f = i$ (función identidad).

Corolario 2 (suma y producto de aplicaciones continuas). Sean (X, d) un espacio pseudométrico y las funciones $f : X \longrightarrow \mathbb{R}^n$, $g : X \longrightarrow \mathbb{R}^n$ y $h : X \longrightarrow \mathbb{R}$, continuas en el punto $a \in X$; entonces las funciones

$$f + g : X \xrightarrow{x \rightsquigarrow f(x) + g(x)} \mathbb{R}^n$$

$$hf : X \xrightarrow{x \rightsquigarrow h(x)f(x)} \mathbb{R}^n$$

son también continuas en el punto $a \in X$.

Demostración: En efecto, si consideramos las aplicaciones producto:

$$(f, g) : x \xrightarrow{f(x), g(x)} \mathbb{R}^n \times \mathbb{R}^n$$

$$(h, f) : x \xrightarrow{h(x), f(x)} \mathbb{R}^n \times \mathbb{R}^n$$

no es difícil comprobar que:

$$f + g = S \circ (f, g), \quad hf = P \circ (h, g)$$

donde S y P son las aplicaciones de suma y producto por un escalar en $\mathbb{R}^n$, definidas en (51) y (52).

Aplicando el resultado del ejemplo 33 y los teoremas 2.6.4 y 2.6.5 se concluye la demostración. ■

Hasta ahora nos hemos referido a la continuidad de una función en un punto, y sin embargo, nos hemos encontrado con funciones que son continuas en todos los puntos de su dominio.

Definición 2.6.6. Una función $f : (X, d) \longrightarrow (Y, d')$, donde (X, d) y (Y, d') son espaciosseudométricos, se llama *globalmente continua* (o sencillamente, *continua*) si es continua en todo punto $a \in X$.

Notemos que por ejemplo, el teorema 2.6.4 puede enunciarse para funciones globalmente continuas diciendo que la compuesta de dos funciones continuas es también una función continua.

Proposición 2.6.7 (función continua restringida a un subespacio).

Sean (X, d) y (Y, d') espaciosseudométricos y A un subconjunto no vacío de X . Si $f : X \longrightarrow Y$ es continua, entonces la restricción

$$f_A : A \longrightarrow Y,$$

donde $f_A(x) := f(x)$ para todo $x \in A$, es también continua si sobre A se pone la métrica d_A inducida en A por la métrica d .

Demostración: Se deduce del teorema 2.6.4 si observamos que $f_A = f \circ i_A$, pues la inyección canónica $i_A : A \xrightarrow{x \mapsto x} X$ es continua. ■

Observación: Hemos visto en el corolario 1 del teorema 2.5.6, que las aplicaciones proyección definidas sobre un espacio métrico producto con una cantidad numerable de coordenadas, son continuas.

Se puede además demostrar, que la métrica producto d sobre $\prod_{n=1}^{\infty} X_n$, es la más gruesa de todas las posibles métricas a definir en el espacio producto para la cual todas las aplicaciones π_n son continuas.

La expresión: *métrica más gruesa*, significa que si para cualquier otra métri-

ca d' sobre $\prod_{n=1}^{\infty} X_n$ se cumple que todas las proyecciones π_n son continuas, entonces la identidad

$$i : \left(\prod_{n=1}^{\infty} X_n, d' \right) \longrightarrow \left(\prod_{n=1}^{\infty} X_n, d \right) \quad (56)$$

es continua.

No es difícil comprobar que la métrica producto satisface la siguiente *propiedad universal*:

Si d' es una métrica en $\prod_{n=1}^{\infty} X_n$ para la cual se cumple la equivalencia del teorema 2.6.5 con cualquier función f y cualquier espacio de partida (X, d) , entonces la aplicación i de (56) es continua.

Dicho en otras palabras, la métrica producto es la más gruesa para la cual se cumple el enunciado del teorema 2.6.5.

Ejemplo 34. Sea X un conjunto no vacío, (Y, d) un espacio métrico y

$$f_n : X \longrightarrow Y, \quad n = 1, 2, \dots$$

una sucesión *separante* de funciones de X en Y , es decir, para todo par de elementos diferentes x y y en X , se puede hallar una función f_n tal que $f_n(x) \neq f_n(y)$. No es difícil comprobar que en este caso las aplicaciones

$$d_n : X \times X \longrightarrow \mathbb{R}_+, \quad n = 1, 2, \dots$$

definidas por

$$d_n(x, y) = d(f_n(x), f_n(y)), \quad n = 1, 2, \dots$$

constituyen una sucesión separante de pseudométricas en X (vea ejemplo 22), y por lo tanto,

$$\rho(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d(f_n(x), f_n(y))}{1 + d(f_n(x), f_n(y))}$$

es una métrica en X para la cual se cumple (vea ejemplo 26):

$$(x_n) \xrightarrow{\rho} x \Leftrightarrow (x_n) \xrightarrow{d_k} x \text{ para todo } k \in \mathbb{N}^*. \quad (57)$$

De esta equivalencia se deduce que la métrica ρ es la *más gruesa* de todas las métricas sobre X para la cual se cumple la propiedad siguiente:

Todas las funciones $f_n : X \longrightarrow Y$ son continuas.

Aquí la expresión métrica más gruesa significa que si ρ' es otra métrica sobre X tal que todas las aplicaciones

$f_n : (X, \rho') \longrightarrow (Y, d)$ son continuas, entonces la identidad

$$i : (X, \rho') \longrightarrow (X, \rho) \quad (58)$$

es continua.

Probemos primeramente que las aplicaciones $f_n : (X, \rho) \longrightarrow (Y, d)$ son continuas. En efecto, por el teorema 2.6.2 de caracterización de la continuidad mediante la convergencia de sucesiones, la implicación hacia la derecha en (57) nos indica la continuidad de f_n ya que:

$$(x_n) \xrightarrow[d_k]{x} \Leftrightarrow d(f_k(x_n), f_k(x)) \xrightarrow{n \rightarrow \infty} 0$$

Supongamos ahora que $f_k : (X, \rho') \longrightarrow (Y, d)$ es continua para todo $k = 1, 2, \dots$. Por el teorema 2.6.2 esto significa que:

$$\rho'(x_n, x) \longrightarrow 0 \Rightarrow d(f_k(x_n), f_k(x)) \longrightarrow 0$$

para todo $k = 1, 2, \dots$

Pero por (57)

$$d(f_k(x_n), f_k(x)) \longrightarrow 0 \text{ para todo } k = 1, 2, \dots$$

$$\Leftrightarrow \rho(x_n, x) \longrightarrow 0,$$

y por tanto, hemos probado que

$$\rho'(x_n, x) \longrightarrow 0 \Rightarrow \rho(x_n, x) \longrightarrow 0. \quad (59)$$

Finalmente notemos que aplicando de nuevo el teorema 2.6.2, la implicación (59) nos dice que la aplicación identidad (58) es continua, lo cual queríamos probar.

Ejercicios de Autocontrol

1. Demuestre que si sobre todo el espacio $C([0, 1], \mathbb{R})$ se pone la métrica d_1 definida en (9) y $f : [0, 1] \longrightarrow \mathbb{R}$ es una función no idénticamente nula, entonces la aplicación

$$F : C([0, 1], \mathbb{R}) \longrightarrow \mathbb{R}$$

definida por

$$F(g) = \int_0^1 g(x) f(x) dx,$$

es continua.

2. Demuestre la afirmación del ejemplo 31.
3. Demuestre que toda aplicación constante $\langle a \rangle : (X, d) \longrightarrow (Y, d')$ donde $a \in Y$, es siempre continua independientemente de la elección de las métricas d y d' .

4. Pruebe que si $f : (X, d) \longrightarrow (Y, \rho)$ es continua, y las metricas d' y ρ' sobre X y Y son semejantes a d y ρ respectivamente, entonces $f : (X, d') \longrightarrow (Y, \rho')$ sigue siendo continua.
5. Demuestre que la aplicación P de producto por un escalar en $\mathbb{R}^n$ (ver (52) del ejemplo 33) es continua.
6. Analice la continuidad de las aplicaciones suma y producto en $\mathbb{R}^2$ provisto de la métrica del bosque d (ver ejercicio 3 de autocontrol de 2.2), es decir,

$$S : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$P : \mathbb{R} \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

cuando a $\mathbb{R}^2$ se provee de la metrica d y en $\mathbb{R}$ se considera la metrica usual definida por el módulo.

7. Demuestre que si $f : (X, d) \longrightarrow (Y, \rho)$ es continua, entonces $f : X \longrightarrow f[X]$ sigue siendo continua cuando sobre la imagen $f[X]$ se pone la métrica inducida por ρ .
8. Demuestre la afirmación de la observación que aparece despues de la proposición 2.6.7.
9. Sea (X, d) un espacio métrico y $C(X, \mathbb{R})$ el conjunto de las funciones reales continuas de X en $\mathbb{R}$. Pruebe que $C(X, \mathbb{R})$ es un algebra real que cumple:

$$f \in C(X, \mathbb{R}) \Rightarrow |f| \in C(X, \mathbb{R}), f, g \in C(X, \mathbb{R}) \Rightarrow \max(f, g), \min(f, g) \in C(X, \mathbb{R}).$$

Indicación: Utilice el teorema sobre la compuesta de funciones continuas y el hecho que

$$\begin{aligned} \max(f, g) &= \frac{1}{2}(f + g + |f - g|), \\ \min(f, g) &= \frac{1}{2}(f + g - |f - g|) \end{aligned}$$

10. Sean f y g dos funciones reales continuas definidas sobre el espacio metrico (X, d) si g no se anula sobre X , demuestre que el cociente

$$\begin{aligned} \frac{f}{g} : (X, d) &\longrightarrow \mathbb{R} \\ x &\rightsquigarrow \frac{f(x)}{g(x)} \end{aligned}$$

es una aplicación continua.

2.7 BOLAS. DIÁMETRO. DISTANCIA ENTRE CONJUNTOS

Tanto en la definición de convergencia de sucesiones como en la de continuidad puntual, hemos utilizado implícitamente ciertos conjuntos que permiten simplificar las expresiones analíticas de dichos conceptos, sustituyéndolas por expresiones de carácter más conjuntista.

Por ejemplo, la expresión

$$d(x_n, x) < \epsilon \text{ para } n \geq N,$$

es equivalente a

$$x_n \in B(x, \epsilon) \text{ para } n \geq N,$$

donde

$$B(x, \epsilon) := \{y \in X : d(x, y) < \epsilon\}.$$

Los conjuntos $B(x, \rho)$ ($x \in X, 0 < \rho < \infty$) juegan un papel muy importante en la teoría de los espacios métricos. Por su forma, en el caso del espacio euclideo $\mathbb{R}^3$ se acostumbra a llamarlos *bolas*.

Definición 2.7.1. Sea (X, d) un espacio sudométrico; $a \in X, 0 \leq \rho < \infty$. Entonces

$$B(a, \rho) := \{x \in X : d(a, x) < \rho\} \quad (60)$$

se llama *bola abierta* de centro a y radio ρ

$$B'(a, \rho) := \{x \in X : d(a, x) \leq \rho\}, \quad (61)$$

se llama *bola cerrada* de centro a y radio ρ

$$E(a, \rho) = \{x \in X : d(a, x) = \rho\} = B'(a, \rho) \setminus B(a, \rho). \quad (62)$$

se llama *esfera* de centro a y radio ρ .

En ocasiones se utiliza la notación $B_d(a, \rho)$, $B'_d(a, \rho)$ y $E_d(a, \rho)$ cuando se requiere hacer referencia explícita a la métrica d .

Utilizando esta definición podemos volver a escribir las definiciones de convergencia y continuidad puntual de la forma siguiente:

La sucesión (x_n) converge al punto x en el espacio métrico (X, d) , si y sólo si para todo $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que

$$x_n \in B(x, \epsilon) \text{ para todo } n \geq N. \quad (63)$$

La función $f : (X, d) \longrightarrow (Y, d')$ es continua en el punto $a \in X$ si para todo $\epsilon > 0$ existe $\rho > 0$ tal que

$$f [B_d (a, \delta)] \subset B_{d'} (f (a), \epsilon) \tag{64}$$

Por transición a la imagen recíproca obtenemos la expresión equivalente (fig. 30):

La función $f : (X, d) \longrightarrow (Y, d')$ es continua en el punto $a \in X$ si para todo $\epsilon > 0$ la imagen recíproca de la bola $B_{d'} (f (a), \epsilon)$ por f , contiene a alguna bola de centro a y cierto radio δ en (X, d) :

$$f^{-1} [B_{d'} (f (a), \epsilon)] \supset B_d (a, \delta)$$

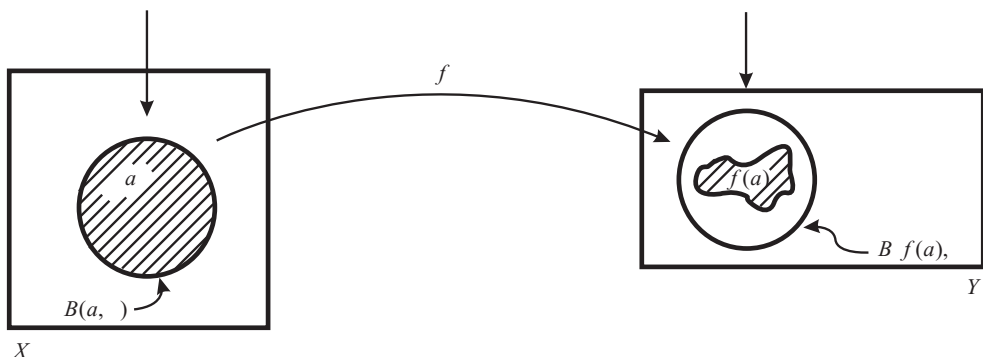


Figura 30

En el espacio euclideo $\mathbb{R}^2$ las bolas son discos y las esferas son círculos. Notemos que las bolas en un espacio pseudométrico cualquiera, no tienen en general las propiedades geométricas de las bolas euclidianas.

Ejemplo 35 (bolas en $\mathbb{R}^n$ con respecto a diversas métricas).

- 1) Las bolas en $\mathbb{R}^n$, definidas por la métrica euclidea

$$d_2 (a, x) = \left(\sum_{k=1}^n (x_k - a_k)^2 \right)^{1/2}$$

tienen la forma de una bola (fig. 31a)

- 2) Las bolas en $\mathbb{R}^n$, definidas por la métrica

$$d_\infty (a, x) = \sup \{|x_k - a_k| : k = 1, 2, \dots, n\}$$

tienen la forma de un cubo regular (fig. 31b)

3) Las bolas en $\mathbb{R}^n$, definidas por la métrica

$$d_1(a, x) = \sum_{i=1}^n |x_i - a_i|$$

tienen la forma de un rombo (fig. 31c)

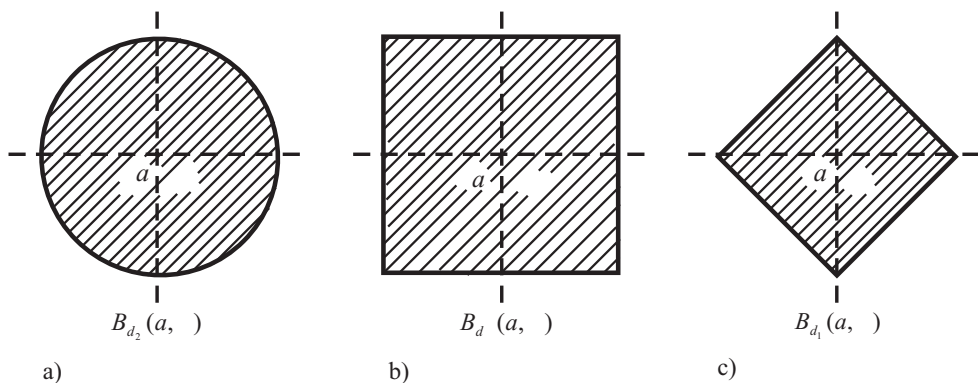


Figura 31

Según (31) tenemos $d_\infty(x, y) \leq d_2(x, y) \leq n d_\infty(x, y)$, luego:

$$B_{d_\infty}(a, 1/n) \subset B_{d_1}(a, 1) \subset B_{d_2}(a, 1) \subset B_{d_\infty}(a, 1) \text{ (Fig. 32)} \quad (65)$$

¿Qué métricas sobre $\mathbb{R}^2$, definen bolas de forma elíptica?

Ejemplo 36 (bolas de un espacio métrico discreto). Sea (X, d) es un espacio discreto (ejemplo 1.2.3). Entonces tenemos para todo $a \in X$:

$$B(a, \rho) = B'(a, \rho) = X \text{ y } E(a, \rho) = \emptyset \text{ si } \rho > 1$$

$$B(a, \rho) = B'(a, \rho) = \{a\} \text{ y } E(a, \rho) = \emptyset \text{ si } 0 < \rho < 1$$

$$B(a, \rho) = \{a\} \text{ y } B'(a, \rho) = X, E(a, \rho) = X \setminus \{a\} \text{ si } \rho = 1$$

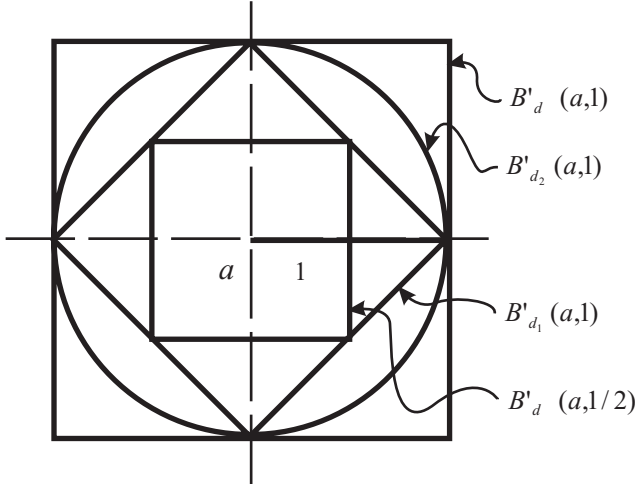


Figura 32

Ejemplo 37 (Bolas en un subespacio). Sea (X, d) un espacio métrico y $A \subset X$; entonces las bolas $B_A(a, \rho)$ de centro $a \in A$ y de radio $\rho > 0$ en el subespacio métrico (A, d_A) , son precisamente las intersecciones de las bolas del mismo centro y mismo radio en (X, d) con A :

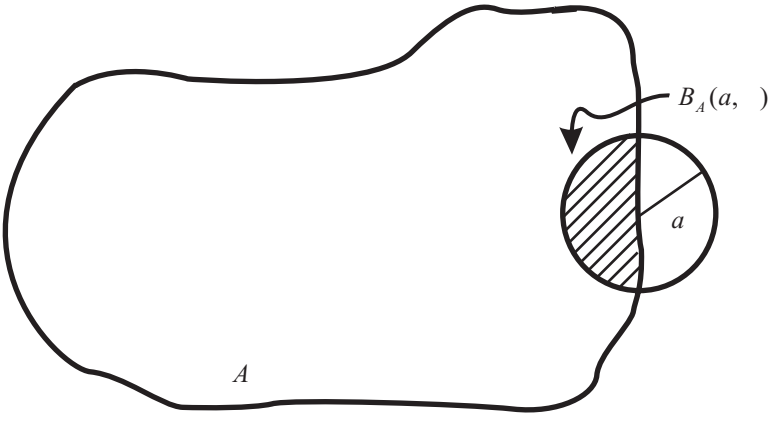
$$B_A(a, \rho) = A \cap B(a, \rho) \text{ (fig. 33).}$$


Figura 33

Ejemplo 38 (bolas en la métrica de la convergencia uniforme en $C'([0, 1], \mathbb{R})$). Consideremos la métrica

$$d_\infty(f, g) := \sup \{|f(x) - g(x)| : x \in [0, 1]\} \text{ sobre } C([0, 1], \mathbb{R}).$$

No es difícil ver que si $f \in C([0, 1], \mathbb{R})$ entonces la bola $B_\infty(f, \varepsilon)$ está formada por todas las funciones continuas cuyos grafos pasan por una franja simétrica a ambos lados del grafo de f y de ancho 2ε (fig. 34).

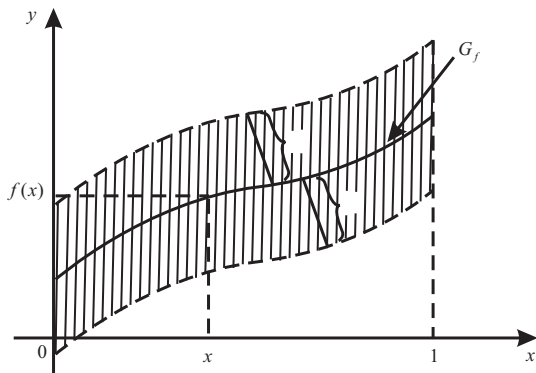


Figura 34

Ejemplo 39 (bolas en la métrica de la convergencia en media sobre $C([0, 1], \mathbb{R})$). Sea de nuevo el conjunto $C([0, 1], \mathbb{R})$, ahora provisto de la métrica

$$d_1(f, g) = \int_0^1 |f(x) - g(x)| dx$$

Contrario al caso anterior (ejemplo 38), en este caso no es posible representar geoméricamente la región barrida por los grafos de las funciones de una bola $B_{d_1}(f, \epsilon)$. En efecto, si tomamos $f = \langle 0 \rangle$, en la bola $B_{d_1}(\langle 0 \rangle, \epsilon)$ se pueden encontrar funciones que alcancen valores arbitrariamente grandes en un punto dado x , como nos muestra la figura 28. Puede demostrarse además, que para cualquier $\epsilon > 0$ se puede hallar una función en $B_{d_1}(\langle 0 \rangle, \epsilon)$ que toma valores arbitrariamente grandes.

Definición 2.7.2. Sean (X, d) un espacio pseudométrico y $A \subset X$. Entonces

$$\delta(A) := \sup \{d(x, y) : (x, y) \in A^2\} \quad (66)$$

se llama *diámetro* de A .

A se dice *acotado*, si su diámetro es finito.

Si f es una aplicación de un conjunto E en (X, d) y $B \subset E$, entonces el diámetro de $f[B] \subset X$ se llama *oscilación* de f sobre B . La función f se llama *acotada*, si $f[E] \subset X$ es acotado (fig. 35)

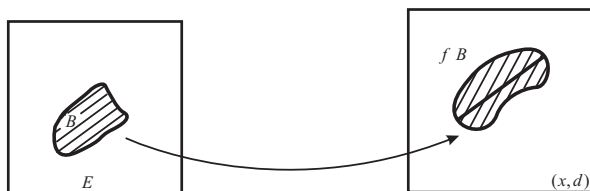


Figura 35

Observación: No necesariamente existen $x, y \in A$ tales que $\delta(A) = d(x, y)$.

El concepto *conjunto acotado*, no depende tanto de la métrica particular d como puede aparecer a primera vista. Según (42) y la definición 2.7.2, un subconjunto acotado de (X, d) es acotado con respecto a toda seudométrica d' semejante a d

Ejemplo 40. En $\mathbb{R}^n$ el diámetro de una recta o de una semirrecta es infinito.

Ejemplo 41. En $\mathbb{R}$ con su métrica usual, el diámetro de un intervalo acotado (abierto o cerrado), es igual a la diferencia de sus extremos

$$A =]a, b[\Rightarrow |b - a| \geq \delta(A) \geq \sup d(b - 1/n, a + 1/n) =$$

$$\sup_{n \in \mathbb{N}^*} |b - a - 1/2n| = |b - a|$$

Ejemplo 42. En los espacios euclidianos $\mathbb{R}^n$ el diámetro de una bola (abierta o cerrada) de radio ρ , es igual a 2ρ . Esto no se cumple en todos los espacios métricos; por ejemplo, en un espacio métrico discreto, toda bola $B(a, 1)$ tiene diámetro 0, mientras que $B'(a, 1)$ tiene diámetro 1.

¿Es necesariamente finita la oscilación de una función continua $f : \mathbb{R} \rightarrow \mathbb{R}$ sobre un subconjunto acotado $A \subset \mathbb{R}$?

Para estimar en algunos casos el diámetro de unión de dos conjuntos disjuntos A y B , hay que conocer la distancia entre ellos (fig. 36).

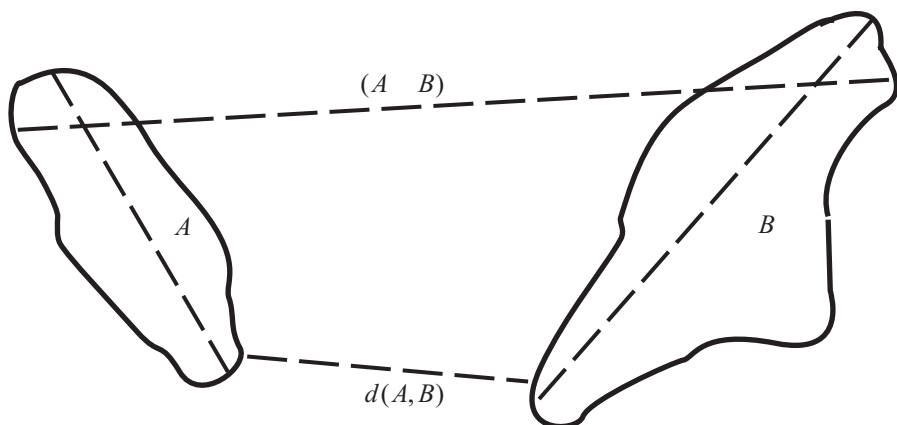


Figura 36

Definición 2.7.3 (distancia entre conjuntos). Sea (X, d) un espacio métrico y $A, B \subset X$. Entonces

$$d(A, B) = \inf\{d(x, y) : x \in A, y \in B\} \quad (67)$$

se llama *distancia* entre los conjuntos A y B .

Si $x \in X$, $d(x, B) := d(\{x\}, B) = \inf\{d(x, y) : y \in B\}$ se llama distancia entre el punto x y el conjunto B .

La aplicación $(A, B) \longrightarrow d(A, B)$ no es ni una métrica ni una seudométrica sobre $\mathbf{P}(X)$ porque no cumple la desigualdad triangular:

Si A, C son dos conjuntos de distancia $d(A, C) > 0$ y si tomamos por B un conjunto cualquiera que tiene una intersección no vacía con A y con C , entonces $d(A, B) = 0$, $d(B, C) = 0$; por tanto, $d(A, C) \not\leq d(A, B) + d(B, C)$.

Ejemplo 43. Sean A y B dos conjuntos no vacíos de un espacio métrico (X, d) . Siempre que $A \cap B \neq \emptyset$ se tendrá que $d(A, B) = 0$. Sin embargo, la relación $d(A, B) = 0$ entre A y B no implica, inversamente, que estos conjuntos tienen un punto común, como ocurre en los casos siguientes:

$$1) A := \{x \in \mathbb{R}^2 : d_2(x, 0) < 1\}$$

$$B := \{x \in \mathbb{R}^2 : d_2(x, 2) < 1\}$$

$$d(A, B) = 0 \quad (\text{fig. 37}).$$

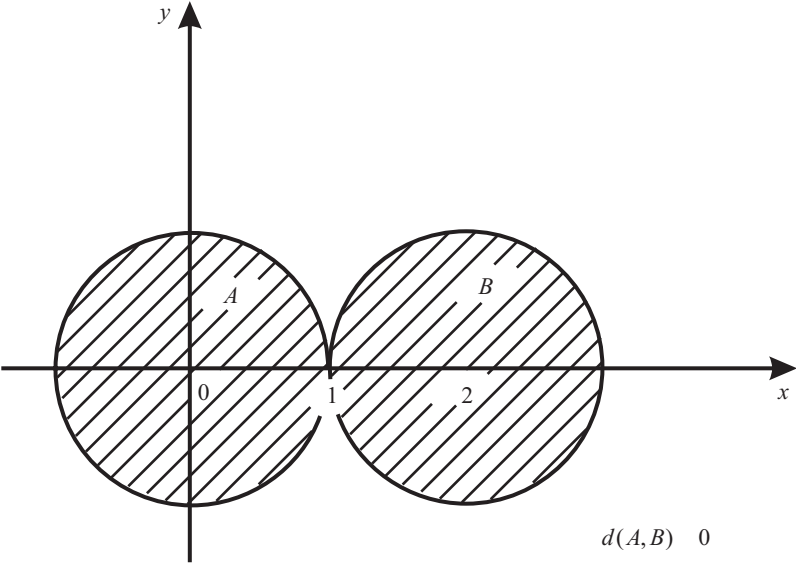


Figura 37

- 2) $A := \{(x, 1/x) : x > 0\}$ (rama de la hipérbola)
 $B := \{(x, 0) : x \in \mathbb{R}\}$ (eje OX)
 $d(A, B) = 0$ (fig. 38).

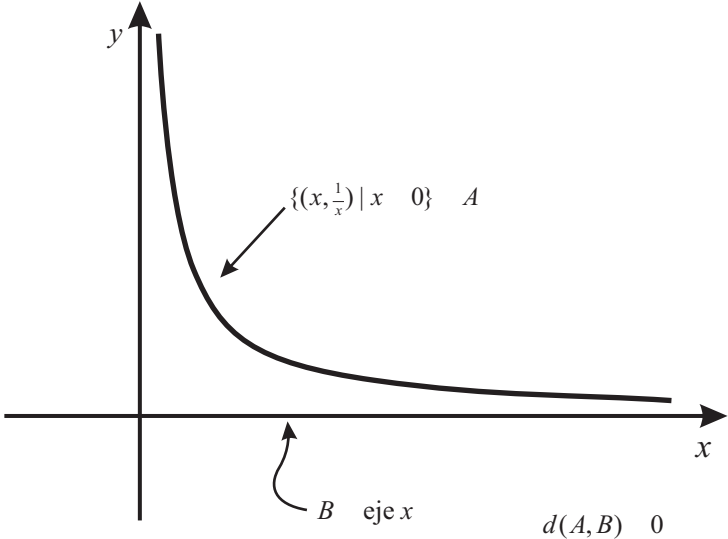


Figura 38

¿Qué diferencia fundamental existe entre los ejemplos 1 y 2 representados en las figuras 37 y 38

Ejemplo 44. La distancia entre un punto x y una recta G en el espacio euclideo $\mathbb{R}^2$ es igual a la distancia de x a su proyección ortogonal sobre (fig. 39). Veremos más adelante una generalización de este enunciado para todas las variedades lineales

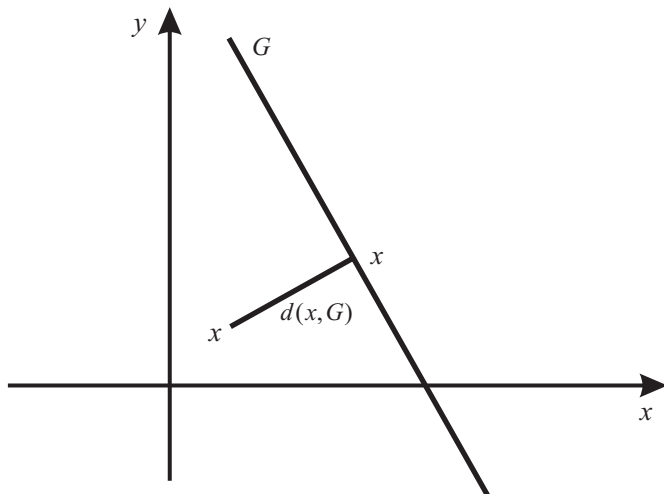


Figura 39

de $\mathbb{R}^n$, y en general, de ciertos espacios vectoriales provistos de métricas asociadas a un *producto escalar* (espacios prehilbertianos vea capítulo 6).

Proposición 2.7.4. Sea (X, d) un espacio métrico, A y B subconjuntos no vacío de X ; entonces existen sucesiones (a_n) y (b_n) de A y B respectivamente, tales que $d(A, B) = \lim (d(a_n, b_n))$.

La demostración se deduce inmediatamente de (67) y de la definición de ínfimo.

Ejemplo 45. Sean $X = \mathbb{R}^2 \setminus \{0\}$, $A = \{(x, 0) : 0 < x \leq 1\}$, $B = \{(0, x) : 0 < x \leq 1\}$. Es claro que $d(A, B) = 0$.

Si tomamos $(a_n) = ((\frac{1}{n}, 0))$ y $(b_n) = ((0, \frac{1}{n}))$, entonces $\lim (d_2(a_n, b_n)) = d(A, B) = 0$, pero las sucesiones (a_n) y (b_n) no convergen ni en A ni en B respectivamente y tampoco en X .

Ejercicios de Autocontrol

1. Describa las bolas en la métrica del bosque (ver ejercicio 3 del 2.2).
2. Demuestre que en (63) y (64) se obtienen proposiciones equivalentes si se sustituyen las bolas abiertas por cerradas.
3. Pruebe que la unión de una familia arbitraria de bolas abiertas de un mismo centro a si no es todo el espacio métrico es una bola abierta. Análogamente, la intersección de una familia cualquiera de bolas cerradas del mismo centro a es una bola cerrada o bien el conjunto unitario $\{a\}$. ¿Pueden intercambiarse *abierto* y *cerrado* en estas dos proposiciones?
4. Pruebe que en $C([0, 1], \mathbb{R})$ con la métrica d_1 , para todo $\epsilon > 0$ en la bola $B_{d_1}(\leq 0, \epsilon)$ pueden hallarse funciones que toman valores arbitrariamente grandes en una cantidad numerable de puntos del intervalo $[0, 1]$.
5. Sea (X, d) un espacio métrico. Demuestre que:
 - i) Si A es acotado, entonces $A \subset B'(a, \delta(A))$ para todo $a \in A$.
 - ii) $A \subset X$ acotado $\Rightarrow A$ está contenido en una bola.
 - iii) La unión de dos conjuntos acotados es un conjunto acotado.
 - iv) $A \cap B \neq \emptyset \Rightarrow \delta(A \cup B) \leq \delta(A) + \delta(B)$. Si A y B son no vacíos, entonces $\delta(A \cup B) \leq \delta(A) + \delta(B) + d(A, B)$.

2.8 VECINDADES Y ABIERTOS

Hasta ahora hemos estudiado el importante papel que juegan ciertos conjuntos llamados bolas en la definición de los conceptos de convergencia y continuidad en los espacios métricos. En ambos casos, las bolas pueden ser sustituidas por otros conjuntos más generales siempre que estos contengan determinadas bolas.

De esta forma llegamos a la definición de *vecindad*.

Definición 2.8.1. Sean (X, d) un espacio pseudométrico y $a \in X$. Se dice que $V \subset X$ es una *vecindad* de a o que a es un punto interior de V , si existe $\delta > 0$ tal que $B(a, \delta) \subset V$ (fig. 40).

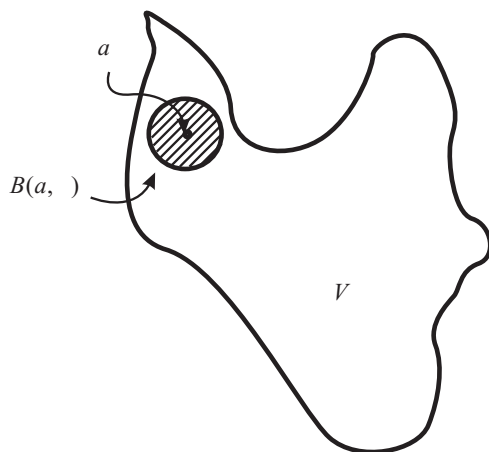


Figura 40

La colección de todas las vecindades de a en X se denota por $\mathbf{V}(a)$. Se dice que V es vecindad del subconjunto A de X si $V \in \mathbf{V}(a)$ para todo $a \in A$. La colección de todas las vecindades del subconjunto A se denota por $\mathbf{V}(A)$.

Ejemplo 46 (vecindades en un espacio métrico discreto). En un espacio métrico discreto (X, d) , cada subconjunto V que contiene un punto dado $a \in X$ como elemento, es una vecindad de este punto. En particular, el conjunto unitario $\{a\}$ es vecindad de a .

Ejemplo 47. Toda bola abierta $B(x, \rho)$ en un espacio métrico (X, d) , es vecindad de cada uno de sus elementos. En efecto, si $a \in B(x, \rho)$, entonces tenemos $d(a, x) < \rho$ y $V := B(x, \rho)$ contiene a $B(a, \delta)$ con $\delta := \rho - d(x, a) > 0$ según la desigualdad triangular (fig. 41).

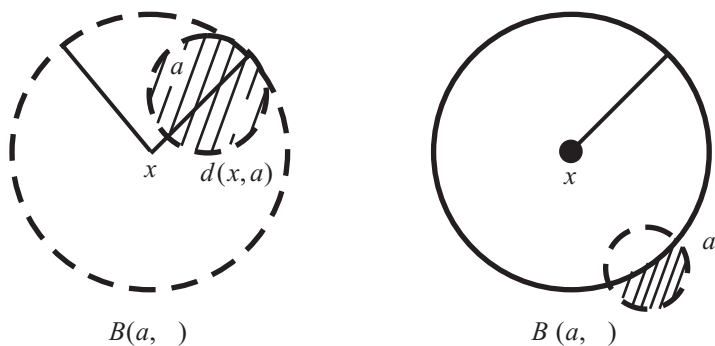


Figura 41

Para $X = \mathbb{R}^2$ esto coincide con nuestra intuición: una bola abierta en $\mathbb{R}^2$ contiene a todos sus elementos como puntos interiores. En cambio, una bola cerrada $B'(x, \rho)$ no es vecindad de sus puntos frontera $a \in E(x, \rho)$ o sea, estos puntos no son interiores a $B'(x, \rho)$.

Existen bolas cerradas en ciertos espacio métricos desiguales a $\mathbb{R}^n$ que son vecindades de cada uno de sus elementos.

El lector verificará fácilmente las siguientes propiedades elementales de las vecindades en un espacio métrico:

$$(V1) \quad V \in \mathbf{V}(a) \Rightarrow a \in V$$

$$(V2) \quad V \in \mathbf{V}(a) \wedge V' \supset V \Rightarrow V' \in \mathbf{V}(a)$$

$$(V3) \quad V, V' \in \mathbf{V}(a) \Rightarrow V \cap V' \in \mathbf{V}(a)$$

$$(V4) \quad a, b \in X \quad a \neq b \Rightarrow \exists V \in \mathbf{V}(a), W \in \mathbf{V}(b); V \cap W = \emptyset$$

Por medio del nuevo concepto podemos simplificar la definición de una función continua. Puesto que toda vecindad de $f(a)$ en Y contiene a una bola de centro $f(a)$ podemos precisar:

$$f : X \longrightarrow Y \text{ es continua en } a \in X \Leftrightarrow f^{-1}[V] \in \mathbf{V}(a) \text{ para todo } V \in \mathbf{V}(f(a)); \quad (69)$$

$$\text{o sea } f : X \longrightarrow Y \text{ es continua en } a \in X \Leftrightarrow \text{Para todo } V \in \mathbf{V}(f(a)) \text{ existe } U \in \mathbf{V}(a) \text{ con } f[U] \subset V \quad (70)$$

Como el lector podrá notar, estas caracterizaciones ya no *se refieren explícitamente* a la distancia sobre X . Nos preguntamos si podemos también caracterizar la convergencia mediante vecindades. En efecto, esto es casi trivial.

Puesto que toda vecindad de x en (X, d) contiene a una bola $B(x, \epsilon)$ e inversamente toda bola $B(x, \epsilon)$ es una vecindad de x , tenemos para todas las sucesiones (x_n) en (X, d) :

$$x_n \longrightarrow x \Leftrightarrow \text{Para toda vecindad } V \in \mathbf{V}(x), \text{ existe } N \in \mathbb{N} \text{ tal que } \{x_n : n \geq N\} \subset V; \quad (71)$$

$$x_n \longrightarrow x \Leftrightarrow \text{Para toda vecindad } V \in \mathbf{V}(x), \text{ el conjunto } \{n \in \mathbb{N} : x_n \notin V\} \text{ es finito.} \quad (72)$$

Resulta interesante que esta propiedad de $V \in \mathbf{V}(a)$ de contener a casi todos los términos de cada sucesión que converge al punto a , es característica

para las vecindades de a . (Para comprender el sentido de esta caracterización vea el ejercicio 1 de autocontrol.)

Para caracterizar la continuidad global de una función, el concepto *vecindad* no es suficiente, porque no se refiere a un punto particular.

Para un subconjunto $\overset{\circ}{A}$ de un espacio métrico (X, d) el conjunto de los puntos interiores de A , es decir, $\{a \in A : A \in \mathbf{V}(a)\}$, se denota por $\overset{\circ}{A}$ y se llama interior de A .

Ya sabemos que toda bola abierta es vecindad de cada uno de sus puntos, luego, es igual su interior. La colección de todos los subconjuntos de (X, d) con esta propiedad, juega un papel fundamental en la Topología General.

Definición 2.8.2. Sea (X, d) un espacio pseudométrico. $A \subset X$ se llama *abierto*, si A es vecindad de cada uno de sus puntos, o sea si $A = \overset{\circ}{A}$. A la familia de todos los subconjuntos abiertos en (X, d) la denotaremos por θ_d .

Ejemplo 48 (abiertos en un espacio métrico discreto). En un espacio métrico discreto todo subconjunto es abierto.

Ejemplo 49. Toda bola abierta en un espacio métrico es un abierto lo cual se deduce del ejemplo 46.

Ejemplo 50. En $\mathbb{R}^n$ todo producto de n intervalos abiertos

$$A =]a_1, b_1[\times \dots \times]a_n, b_n[, \quad a_i, b_i \in \mathbb{R}$$

es abierto. En particular, el conjunto $(\mathbb{R}_+)^n$ de los vectores estrictamente positivos es abierto.

Todo complemento de un conjunto finito en $\mathbb{R}^n$ es abierto. ¿Se cumple lo mismo para el complemento de conjuntos numerables en $\mathbb{R}^n$?

Ejemplo 51. Sea $A \subset \mathbb{R}^n$ abierto; entonces las sucesiones de $l^2(\mathbb{N})$ cuyas primeras coordenadas son recta de A , forman un subconjunto abierto de $l^2(\mathbb{N})$. El conjunto P de todas las sucesiones en $l^2(\mathbb{N})$ con términos estrictamente positivos no es abierto en $l^2(\mathbb{N})$.

¿Por qué el subconjunto

$$A = \prod_{k=1}^{\infty}]-1/k, 1/k[\subset l^2(\mathbb{N}) \text{ no es abierto en } l^2(\mathbb{N})?$$

Ejemplo 52. Consideremos el espacio pseudométrico $(\mathbb{R}^{[-1,1]}, d_{\infty})$ de todas las funciones reales y acotadas definidas sobre $[-1, 1]$; provisto de la distancia de la convergencia uniforme. Sea I un intervalo abierto en $\mathbb{R}$; entonces el

subconjunto A de todas las funciones $f : [-1, 1] \longrightarrow \mathbb{R}$ con valores en I , en general, no es abierto en $(\mathbb{R}^{[-1,1]}, d_\infty)$.

Por ejemplo, si

$$I =]-1, 1[\text{ y } f_0(x) := \begin{cases} 0 & \text{si } x = -1 \text{ ó } x = 1 \\ x & \text{si } -1 < x < 1 \end{cases}$$

entonces no existe una ϵ banda de la gráfica de f_0 que esté totalmente incluida en el área $[-1, 1] \times]-1, 1[$. Luego f_0 no es un punto interior de A , es decir, A no es abierto. La función f_0 no es discontinua por casualidad (fig. 42).

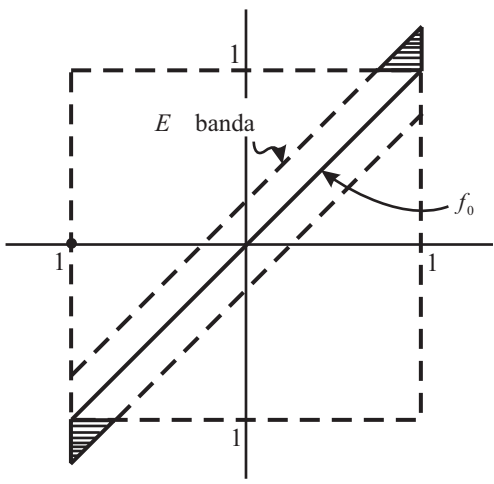


Figura 42

Si nos limitamos al espacio métrico $X = (C([-1, 1], \mathbb{R}), d_\infty)$ de todas las funciones reales continuas sobre $[-1, 1]$, el lector verificará que el subconjunto

$$A \cap X = \{f \in C([-1, 1], \mathbb{R}) : |f(x)| < 1 \text{ para } x \in [-1, 1]\}$$

es abierto en este espacio.

Realmente el concepto *abierto* es tan *fundamental* como el concepto *vecindad*, ya que las vecindades pueden caracterizarse a partir de los abiertos de la manera siguiente:

$V \subset X$ es vecindad de $x \in X \iff$ Existe un abierto $A \subset X$ con $x \in A \subset V$ (la demostración se deja al lector).

Teorema 2.8.3 (propiedades fundamentales de los abiertos). Sea (X, d) un espacio pseudométrico. Entonces se cumple:

(A1) ϕ, X son subconjuntos abiertos de (X, d) ;

(A2) si $A_1, A_2, \dots, A_n$ es una familia *finita* de conjuntos

abiertos en (X, d) , entonces $A_1 \cap A_2, \dots, \cap A_n$ es abierto;

(A3) si $(A_i)_{i \in I}$ es una familia *arbitraria* de conjuntos

abiertos, entonces $\bigcup_{i \in I} A_i$ es abierto.

Demostración:

(A1) se deduce de la definición 2.8.2;

(A2) según la propiedad (68 iii) de las vecindades.

(A3) $\bigcup_{i \in I} A_i$ es vecindad de todo $x \in A_i (i \in I)$, porque $\bigcup_{i \in I} A_i \supset A_i \in \mathbf{V}(x)$ (por la propiedad ii de las vecindades).■

Ejemplo 53. La intersección de una familia infinita de abiertos, *en general no es un conjunto abierto.*

En efecto, consideremos en $\mathbb{R}$, los conjuntos

$$I_n =]0, 1 + \frac{1}{n}[, n \in \mathbb{N}$$

Estos conjuntos son abiertos ya que

$$I_n = B\left(\frac{1}{2}\left(1 + \frac{1}{n}\right), \frac{1}{2}\left(1 + \frac{1}{n}\right)\right)$$

donde la bola se considera con respecto a la métrica usual (módulo).

Pero $A = \bigcap_{n=1}^{\infty} I_n =]0, 1]$ no es abierto ya que el punto $1 \in A$ pero $1 \notin \mathring{A}$. Esto último se debe al hecho de que no hay ninguna bola $B(1, \epsilon) =]1 - \epsilon, 1 + \epsilon[$ completamente contenida en A .

Todo subconjunto de un espacio métrico que puede ser expresado como intersección de una familia numerable de conjuntos abiertos se llama un G_δ *conjunto* o *conjunto* de tipo G_δ .

El siguiente teorema da una caracterización completa de los subconjuntos abiertos.

Teorema 2.8.4 (caracterización de los abiertos). Sea (X, d) un espacio métrico, $A \subset X$. Entonces A es abierto $\iff A$ puede representarse como unión de bolas abiertas.

Demostración: $\Leftarrow$: se deduce de (A3) y del ejemplo 47.

$\implies$: Si $A \subset X$ es abierto, entonces para cada $a \in A$ existe un radio δ_a con $B(a, \delta_a) \subset A$. Luego $A = \bigcup_{a \in A} B(a, \delta_a)$ ■

De este teorema se deducen los cuatro resultados siguientes, los cuales son muy importantes.

Teorema 2.8.5 (caracterización de los conjuntos abiertos en $\mathbb{R}$). Z un subconjunto A de $\mathbb{R}$ es abierto si y sólo si puede expresarse como la unión de una familia, a lo más numerable, de intervalos abiertos disjuntos.

Demostración: Del hecho de que un intervalo $]a, b[$ no es más que la bola de centro $\frac{a+b}{2}$ y radio $\frac{b-a}{2}$, se deduce la suficiencia de la condición aplicando el teorema 2.8.4. Recíprocamente si $A \subset \mathbb{R}$ es abierto, entonces por 2.8.4, A se puede expresar como la unión de una cierta familia de intervalos (bolas) abiertos $\{I_\alpha\}_{\alpha \in I}$.

Para cada $a \in A$ definimos

$$J_a = \{\alpha : a \in I_\alpha\}$$

y pongamos

$$I_a = \bigcup \{I_\alpha : \alpha \in J_a\}$$

Entonces cada I_a es un intervalo abierto, y para cada dos puntos cualesquiera a y b en A , se tiene una de las dos condiciones excluyentes:

- 1) $I_a = I_b$;
- 2) $I_a \cap I_b = \emptyset$.

Definamos una relación de equivalencia en A mediante:

$$a \mathbf{R} b \iff I_a = I_b.$$

Entonces, se pueden definir por cada $a \in A$ los intervalos

$$I_{[a]} = I_a$$

y obtener una familia de intervalos abiertos disjuntos

$$\mathbf{F} = \{I_{[a]} : [a] \in A/\mathbf{R}\},$$

tal que

$$A = \bigcup \{I_{[a]} : [a] \in A/\mathbf{R}\}.$$

Probemos que la familia $\mathbf{F}$ es a lo más numerable. En efecto, si para cada $n \in \mathbb{N}$, la familia

$$\mathbf{F}_n = \{I_{[a]} \cap]-n, n[: [a] \in A/\mathbf{R}\}$$

es numerable, entonces $\mathbf{F}$ también es numerable. En caso contrario, existiría un número $N \in \mathbb{N}$ tal que $\mathbf{F}_N$ no es numerable y esto sería una contradicción ya que si denotamos por $l_{[a]}$ la longitud de $I_{[a]} \cap]-N, N[$ entonces por un lado la serie

$$\sum \{l_{[a]} : [a] \in A/\mathbf{R}\}$$

es divergente ya que tiene una cantidad más que numerable de términos positivos, y por otro lado, debe converger pues la suma está acotada por la longitud del intervalo $] -N, N[$ es decir, por $2N$. ¡Contradicción! ■

Teorema 2.8.6 (abiertos en subespacios). Sean (X, d) un espacio métrico y $A \subset X$; entonces los abiertos en (A, d_A) son los subconjuntos de A obtenidos al intersectar los abiertos en X con A . (fig. 43)

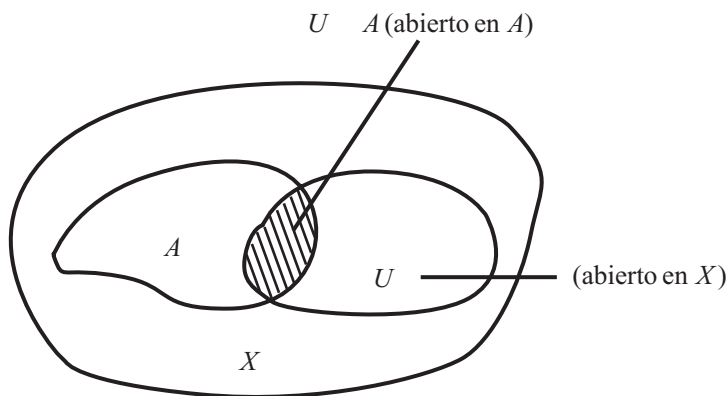


Figura 43

Demostración: En efecto, si U_A es abierto en A , entonces según la caracterización anterior, se puede expresar

$$U_A = \bigcup_{a \in A} B_{d_A}(a, \epsilon_a),$$

y por tanto,

$$U_A = \bigcup_{a \in A} (B_d(a, \epsilon_a) \cap A) = (\bigcup_{a \in A} B_d(a, \epsilon_a)) \cap A = U \cap A,$$

donde U es obviamente un abierto en A . ■

No es difícil generalizar el resultado del ejemplo 48 a un producto finito de espacios métricos cualesquiera. En efecto, si $(X_k, d_k); k = 1, 2, \dots, n$ son espacios métricos y U_k es abierto en X_k entonces $\prod_{k=1}^n U_k$ es abierto en $\prod_{k=1}^n X_k$ con respecto a la métrica producto (el lector puede verificarlo). Sin embargo, esto no ocurre para un producto numerable según nos muestra el teorema siguiente:

Teorema 2.8.7 (abiertos en los espacios producto). Sea $(X_n, d_n), n = 1, 2, \dots$, una familia numerable de espacios métricos y sea U_n abierto en X_n

para cada n . $\prod_{n=1}^{\infty} U_n$ es abierto en $\prod_{n=1}^{\infty} X_n$ con respecto a la métrica producto (ver (41)), si y sólo si existe $n_0 \in \mathbb{N}$ tal que

$$U_n = X_n, n \geq n_0.$$

Demostración: Denotemos por d la métrica producto en $\prod_{n=1}^{\infty} X_n$. Si $\prod_{n=1}^{\infty} U_n$ es abierto, entonces necesariamente contiene una bola $B_d((x_n), \epsilon)$, es decir, si $y_n \in X_n$ para cada n, y

$$\sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} < \epsilon, \quad (73)$$

entonces $(y_n) \in \prod_{n=1}^{\infty} U_n$, o sea, $y_n \in U_n, n = 1, 2, \dots$

Pero de (73) se deduce que

$$\frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} < 2^n \epsilon, n = 1, 2, \dots \quad (74)$$

Tomemos $n_0 \in \mathbb{N}$ tal que $2^n \epsilon > 1$ para $n \geq n_0$. Entonces como para todo $y_n \in X_n (n \geq n_0)$ se cumple (74), se tiene que $X_n = U_n$ si $n \geq n_0$.

Recíprocamente supongamos que:

$$U_n = X_n, n \geq n_0$$

y sea $\prod_{n=1}^{\infty} U_n = \prod_{k=1}^{n_0-1} U_k \times \prod_{k=n_0}^{\infty} X_k$.

Por ser U_k un abierto, para cada k , existe un número δ_k tal que

$$B_{d_k}(x_k, \delta_k) \subset U_k.$$

Si tomamos

$$\delta = \min \{\delta_k : 1 \leq k \leq n_0\}$$

no resulta difícil comprobar que

$$B_d\left((x_n), \frac{1}{2^{n_0}} \frac{\delta}{1+\delta}\right) \subset \prod_{n=1}^{\infty} U_n,$$

y por tanto, $\prod_{n=1}^{\infty} U_n$ es abierto. ■

Teorema 2.8.8 (caracterización de funciones continuas). Sean (X, d) y (Y, d') espaciosseudométricos; entonces una aplicación $f : X \longrightarrow Y$ es continua, si y sólo si la imagen recíproca de cada abierto en Y es abierto en X .

Demostración: La suficiencia de la condición es evidente. En efecto, para cada $a \in X$, la bola $B_d(f(a), \epsilon)$ es un abierto en Y , y si la imagen recíproca $f^{-1}[B_d(f(a), \epsilon)]$ es un conjunto abierto en X , esto implica en particular que existe $\delta > 0$ tal que

$$B_d(a, \delta) \subset f^{-1}[B_d(f(a), \epsilon)]$$

de donde se concluye la continuidad de f .

Recíprocamente sean f continua y A abierto en Y . Por el teorema 2.8.4, para cada punto $a \in A$ existe un número $\epsilon_a > 0$ tal que

$$A = \bigcup_{a \in A} B_{d'}(a, \epsilon_a),$$

y por tanto,

$$f^{-1}[A] = \bigcup_{a \in A} f^{-1}[B_{d'}(a, \epsilon_a)] =$$

$$= \bigcup \{f^{-1}[B_{d'}(a, \epsilon_a)] : a \in f[f^{-1}[A]]\}$$

Por ser f continua, para todo $a \in f[f^{-1}[A]]$, existe $b \in f^{-1}[A]$ y $\delta_a > 0$ tales que:

$$f^{-1}[B_{d'}(a, \epsilon_a)] \supset B_d(b, \delta_a),$$

$$f(b) = a.$$

Pero entonces

$$f^{-1}[A] \supset \{B_d(b, \delta_a) : b \in f^{-1}[A]\}$$

y como la inclusión en sentido contrario es evidente, se tendrá que

$$f^{-1}[A] = \bigcup \{B_d(b, \delta_a) : b \in f^{-1}[A]\}$$

de donde se concluye, utilizando el teorema 2.8.4, que $f^{-1}[A]$ es abierto. ■

Definición 2.8.9 (base, segundo axioma de numerabilidad: AN2).

Sea (X, d) espacio pseudométrico y $\mathbf{F}$ una familia formada por conjuntos abiertos. La familia $\mathbf{F}$ se llama una *base topológica* en (X, d) , si cualquier abierto en (X, d) se puede poner como unión de elementos de $\mathbf{F}$. Se dice que (X, d) satisface el *segundo axioma de numerabilidad* o que es un *AN2-espacio*, si tiene una base topológica numerable.

Observación: Del teorema 2.8.4 se deduce que la familia de todas las bolas abiertas forman una base topológica en cualquier espacio métrico. Hemos utilizado la terminología de segundo axioma de numerabilidad, sin haber definido un primer axioma (AN1). La definición de AN1 espacio existe en la topología general y además todo espacio métrico es AN1. Por este motivo, no la hemos introducido, aunque el lector puede hallarla en la literatura existente sobre topología general.

Definición 2.8.10 (homeomorfismos, espacios métricos topológicamente equivalentes).

Una aplicación f de un espacio métrico (X, d) en otro

espacio métrico (Y, d') se llama *homeomorfismo*, si f es una biyección tal que f y f^{-1} son continuas.

Dos espacios métricos (X, d) y (Y, d') se dice *homeomorfos* o *topológicamente equivalentes*, si existe un homeomorfismo de X sobre Y .

Como podemos ver del teorema 2.8.8, un homeomorfismo entre los espacios métricos (X, d) y (Y, d') establece una correspondencia biunívoca no solo entre los puntos de X y Y , sino también entre la familia de todos los conjuntos abiertos en X y la correspondiente familia en Y .

Ejemplo 54. Toda homotecia $x \rightsquigarrow \lambda x$ ($\lambda \neq 0$) de $\mathbb{R}^n$ sobre $\mathbb{R}^n$ es un homeomorfismo.

Más aún, toda aplicación lineal sobreyectiva f de $\mathbb{R}^n$ sobre $\mathbb{R}^n$ es un homeomorfismo. En efecto, del hecho que $f(x) = f(x^{(1)}, x^{(2)}, \dots, x^{(n)}) = y = (y^{(1)}, \dots, y^{(n)})$ se expresa en la forma equivalente

$$a_{11}x^{(1)} + \dots + a_{1n}x^{(n)} = y^{(1)}$$

$$a_{21}x^{(1)} + \dots + a_{2n}x^{(n)} = y^{(2)}$$

.....

$$a_{n1}x^{(1)} + \dots + a_{nn}x^{(n)} = y^{(n)},$$

donde las a_{ij} son números reales, y de la continuidad de las operaciones de suma y producto por un escalar, se deduce que si $(x_k) \longrightarrow x$ en $\mathbb{R}^n$ entonces $(f(x_k) \longrightarrow f(x))$ en $\mathbb{R}^n$ (recordemos que la convergencia en $\mathbb{R}^n$ es equivalente a la convergencia por coordenadas.)

Si f es además sobreyectiva, entonces existe la inversa f^{-1} que también es lineal, y por lo tanto, continua.

Ejemplo 55. La aplicación $f : x \rightsquigarrow x/(1 + |x|)$ es un homeomorfismo de $\mathbb{R}$ sobre $] -1, 1[$ que además conserva el orden. No es difícil comprobar que f es continua y además estrictamente creciente y sobreyectiva, por lo cual existe

$$f^{-1} :] -1, 1[\longrightarrow \mathbb{R}$$

cuya expresión explícita es

$$f^{-1}(y) = \frac{y}{1 - |y|}$$

que también es continua (demuéstrelo).

Ejemplo 56 (proyección estereográfica). Consideremos la circunferencia en $\mathbb{R}^2$ de centro $(0, 1/2)$ y de radio $1/2$, a la cual hemos quitado el polo ubicado en $(0, 1)$. El conjunto S obtenido es un espacio métrico con respecto a la métrica inducida por la distancia euclideana en $\mathbb{R}^2$. Definamos una aplicación f de S en $\mathbb{R}$, haciéndole corresponder a todo $x \in S$ el punto en que intersecta a $\mathbb{R}$ la recta que une x con $(0, 1)$ (fig. 44.)

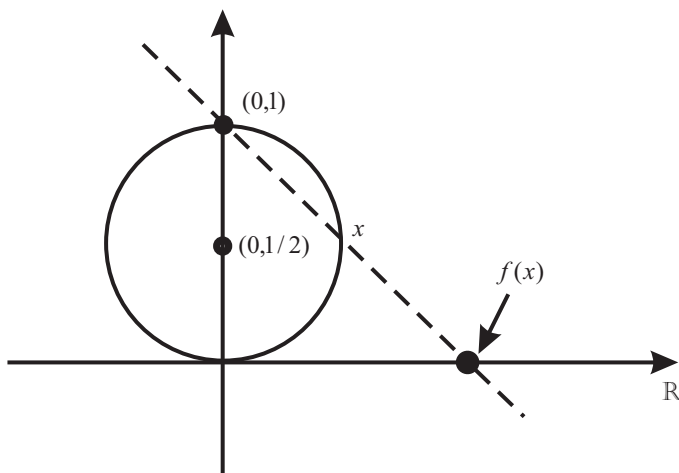


Figura 44

Se deja al lector la demostración de que f define un homeomorfismo de S en $\mathbb{R}$. Recomendamos primeramente hallar una expresión explícita para f .

Pregunta: ¿Como podría definirse la proyección estereográfica de la esfera unitaria de $\mathbb{R}^3$ en el plano $\mathbb{R}^2$? ¿Esta proyección será también un homeomorfismo?

Observación: Es conveniente señalar que hay aplicaciones que son continuas y biyectivas, pero que no son necesariamente bicontinuas (o sea, su inversa no es continua). Por ejemplo, la aplicación idéntica.

$$1_{\mathbb{R}} : x \rightsquigarrow x$$

de la recta real, provista de la métrica discreta sobre $\mathbb{R}$ provisto de la métrica usual es continua, mientras que su inversa evidentemente no lo es.

No resulta difícil probar que la inversa de un homeomorfismo y la compuesta de dos homeomorfismos son homeomorfismos, y por lo tanto, la homeomorfía es una relación de equivalencia entre los espacios métricos (ver ejercicio 5 de autocontrol).

Al final de 2.2 hicimos algunas afirmaciones que nos hacen suponer que diferentes métricas sobre un conjunto pueden determinar la misma colección de abiertos. Esto da lugar a la definición siguiente:

Definición 2.8.11. Dos (seudo) métricas d y d' sobre un conjunto X se llaman *topológicamente equivalentes* si ambas definen los mismos conjuntos abiertos en X , es decir, si se cumple para todo $A \subset X$:

A es abierto en $(X, d) \Leftrightarrow A$ es abierto (X, d') .

Proposición 2.8.12. Sea X un conjunto y d, d' dos métricas sobre X ; entonces las afirmaciones siguientes son equivalentes:

1) d y d' son métricas topológicamente equivalentes en X ;

2) La aplicación idéntica

$1_X : (X, d) \rightsquigarrow (X, d')$ es un homeomorfismo;

3) toda bola $B_d(a, \rho)$ contiene una bola $B_{d'}(a, \rho')$ e inversamente. (Es evidente que podemos limitarnos aquí a bolas con radios ρ y $\rho' \leq \epsilon$, donde $\epsilon > 0$ es arbitrario.)

La demostración se deja al lector.

Ejemplo 57. Las seudométricas semejantes son también topológicamente equivalentes. En efecto, sean d y d' dos seudométricas semejantes sobre un conjunto X y $B_d(a, \rho)$ una bola cualquiera en (X, d) . Existe $k' > 0$ con $d \leq k'd'$; luego, es suficiente tomar $\rho' := \rho/k'$ para obtener $B_{d'}(a, \rho') \subset B_d(a, \rho)$. Puesto que la semejanza es una relación simétrica, el inverso se obtiene análogamente.

De esta forma concluimos que d y d' son topológicamente equivalentes.

El ejemplo siguiente muestra que no todas las métricas topológicamente equivalentes son semejantes.

Ejemplo 58. La distancia euclídeana d sobre $\mathbb{R}^n$ no es semejante a la distancia $d' = \inf \{1, d\}$ sobre $\mathbb{R}^n$. Sin embargo, ambas distancias son topológicamente equivalentes, puesto que para $\rho \leq 1$ las bolas abiertas de radio ρ con respecto a ambas métricas coinciden.

Este ejemplo nos dice que a toda seudométrica le corresponde una seudométrica acotada topológicamente equivalente.

Ejemplo 59. Si en $\mathbb{N}$, consideramos la métrica inducida $d_{\mathbb{N}}$ por la métrica usual d en $\mathbb{R}$, entonces $d_{\mathbb{N}}$ es topológicamente equivalente a la métrica discreta sobre $\mathbb{N}$. En efecto, como los abiertos en $(\mathbb{N}, d_{\mathbb{N}})$ se expresan mediante la intersección de abiertos de $\mathbb{R}$ con $\mathbb{N}$, cada conjunto unitario

$$\{n\} =]n - 1/2, n + 1/2[\cap \mathbb{N}$$

es abierto en $(\mathbb{N}, d_{\mathbb{N}})$ de donde se deduce que los conjuntos abiertos definidos por $d_{\mathbb{N}}$ y por la métrica discreta coinciden. Para finalizar veamos un resultado que nos pone de manifiesto el hecho de que la noción de convergencia en un

espacio métrico no depende de la métrica de dicho espacio, sino más bien de la familia de conjuntos abiertos determinados por esa métrica.

Teorema 2.8.13. Dos métricas d y d' sobre un conjunto X son topológicamente equivalentes si y sólo si definen la misma convergencia.

La demostración se deduce inmediatamente de la definición de convergencia y de la proposición 2.8.12.

Ejercicios de Autocontrol

- Sean (X, d) un espacio métrico, $a \in X$ y $V \subset X$. Pruebe que $V \notin \mathbf{V}(a)$ si y sólo si existe una sucesión (x_n) , con $x_n \in X \setminus V$ para todo $n \in \mathbb{N}$, que converge a a .
- Sean (X, d) un espacio pseudométrico y $A \subset X$. Demuestre que las vecindades en (A, d_A) están dadas por la intersección de las vecindades en (X, d) con A .
- Pruebe la equivalencia siguiente:
 $V \subset X$ es vecindad de $x \in X \Leftrightarrow$ existe un abierto $A \subset X$ con $x \in A \subset V$.
- Demuestre que toda función creciente y continua de $[a, b]$ en $\mathbb{R}$ define un homeomorfismo de $[a, b]$ sobre $[f(a), f(b)]$.
- Sean $f : (X, d) \rightarrow (Y, d')$ y $g : (Y, d') \rightarrow (Z, d'')$ homeomorfismos. Demuestre que $f^{-1} : (Y, d') \rightarrow (X, d)$ y $g \circ f : (X, d) \rightarrow (Z, d'')$ son también homeomorfismos y deduzca de esto que la relación: $(X, d) \mathbf{R} (Y, d') \Leftrightarrow (X, d)$ es homeomorfo a (Y, d') , es una relación de equivalencia en la categoría de los espacios métricos.
- Sean φ y φ' dos funciones crecientes de $\mathbb{R}_+$ en $\mathbb{R}_+ = [0, \infty]$ continuas en 0 y tales que $\varphi(0) = \varphi'(0) = 0$. Pruebe que si d y d' son dos métricas sobre un conjunto X tales que $d' \leq \varphi \circ d$ y $d \leq \varphi' \circ d'$ sobre X , entonces d y d' son topológicamente equivalentes.
- Pruebe que la métrica del bosque sobre $\mathbb{R}$ (ejercicio 3 de autocontrol de 2.2) no es topológicamente equivalente a la métrica euclídeana sobre $\mathbb{R}$.

2.9 PUNTOS ADHERENTES. CLAUSURA DE UN CONJUNTO. CONJUNTOS DENSOS

Los epígrafes anteriores establecen y desarrollan las definiciones fundamentales de convergencia y continuidad sobre la base de los conceptos *vecindad*

y *abierto*. En este epígrafe y en los dos siguientes, desarrollaremos algunas nociones elementales para analizar los subconjuntos de un espacio métrico en general.

En análisis elemental un punto a se llama adherente a un subconjunto $A \subset \mathbb{R}^n$, si puede aproximarse tanto como se quiera por puntos del conjunto A , o sea, si a es límite de una sucesión convergente (x_k) de elementos de A . Intuitivamente, un punto x es adherente de un conjunto C si no puede separarse de C por una vecindad. Formalizaremos este concepto para subconjuntos de un espacio métrico cualquiera, y veremos más adelante que un punto x es adherente de C si y sólo si es un punto límite de una sucesión de elementos de C .

Ejemplo 60. Los extremos $a, b \in \mathbb{R}$ son puntos adherentes del intervalo abierto $]a, b[$ ($a < b$), puesto que son límites de las sucesiones convergentes $(a + 1/n)$ (resp. $(b - 1/n)$), cuyos elementos pertenecen a partir de cierto n al intervalo $]a, b[$.

El conjunto de todos los puntos adherentes de $]a, b[$ en $\mathbb{R}$, provisto de su topología usual, es el intervalo cerrado $[a, b]$.

Se demuestra que $a \in \mathbb{R}^n$ es punto adherente de un conjunto $A \subset \mathbb{R}^n$ si y sólo si todo entorno V de a intersecta a A (fig. 45).

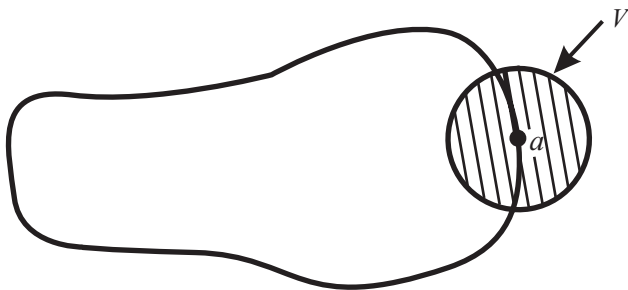


Figura 45

Esta caracterización se generaliza inmediatamente a espacios métricos cualesquiera.

Definición 2.9.1. Sea A un subconjunto de un espacio seudométrico (X, d) . El punto $a \in X$ se llama *punto adherente* de A si toda vecindad $V \in \mathbf{V}(a)$ tiene una intersección no vacía con A :

$$V \in \mathbf{V}(a) \implies V \cap A \neq \emptyset \quad (75)$$

El conjunto de todos los puntos adherentes de A se llama *adherencia* (*clausura*) de A y se denota por $\overline{A}$. Un punto $a \in X$ no es *adherente* de A si puede ser separado de A por una vecindad:

$$\exists V \in \mathbf{V}(a) \quad (V \cap A = \emptyset) \quad (76)$$

Del hecho que las vecindades de un punto no varían al pasar de una métrica a otra topológicamente equivalente, deducimos que la noción de clausura no depende intrínsecamente de la métrica sino de la estructura topológica por ella determinada.

Veamos la siguiente caracterización de los puntos adherentes:

Teorema 2.9.2 (caracterización de los puntos adherentes). Sea (X, d) un espacioseudométrico y $A \subset X$. Entonces son equivalentes:

$$\text{i) } a \in \overline{A}; \quad (77)$$

ii) dada una sucesión de números positivos $(\delta_n) \rightarrow 0$ se tiene

$$B(a, \delta_n) \cap A \neq \emptyset; \quad (78)$$

$$\text{iii) existe una sucesión } (x_n) \text{ en } A \text{ con } (x_n) \rightarrow a; \quad (79)$$

$$\text{iv) } d(a, A) = 0 \quad (80)$$

Demostración: i) $\Rightarrow$ ii) ya que para cada δ_n , $B(a, \delta_n)$ es una vecindad de a .
ii) $\Rightarrow$ iii) Sea la sucesión (δ_n) tal que se cumple (78) y elijamos para cada n , un elemento $x_n \in B(a, \delta_n) \cap A$.

Es obvio que (x_n) es una sucesión de elementos de A con $(x_n) \rightarrow a$, ya que $d(a, x_n) < \delta_n$.

iii) $\Rightarrow$ iv) Sea (x_n) una sucesión de elementos de A que satisface (79), entonces para cada n

$$d(a, A) = \inf \{d(a, x) : x \in A\} \leq d(a, x_n).$$

Si n tiende a ∞ se obtiene (80).

iv) $\Rightarrow$ i) Probemos la proposición equivalente no i) $\Rightarrow$ no iv) En efecto, si $a \notin \overline{A}$ entonces existe $V \in \mathbf{V}(a)$ tal que $V \cap A = \emptyset$.

Pero existe $\delta > 0$, tal que $B(a, \delta) \subset V$, y por lo tanto,

$$B(a, \delta) \cap A = \emptyset,$$

Es decir, todo punto de A está fuera de la bola $B(a, \delta)$, luego $d(a, x) \geq \delta \quad \forall x \in A$. De aquí se concluye que

$$d(a, A) = \inf \{d(a, x) : x \in A\} \geq \delta > 0$$

de donde obtenemos la negación de iv).■

Ejemplo 61. Consideremos el espacio $\mathbb{R}^n$ provisto de su topología usual. Sea $B(z, \rho)$ la bola abierta de centro z y de radio ρ con respecto a la métrica euclídeana. Entonces todo punto a de la esfera $E(z, \rho)$ es adherente de A . En efecto, la función

$$\varphi : \alpha \rightsquigarrow z + \alpha(a - z)$$

de $[0, 1]$ en $\mathbb{R}^n$ es continua y aplica el intervalo $[0, 1[$ sobre el segmento $[z, a[\subset B(z, \rho)$. Puesto que $\varphi(1) = a$, existe para todo $V \in \mathbf{V}(a)$ un $\alpha_V \in [0, 1[$ tal que $\varphi(\alpha) \in V$ para todo $\alpha \in [\alpha_V, 1[$.

Luego

$$(\alpha_V) \in V \cap [z, a[\subset V \cap B(z, \rho), \text{ o sea, } V \cap A \neq \emptyset.$$

Concluimos que en el espacio $\mathbb{R}^n$ la adherencia de una bola abierta es la bola cerrada correspondiente, o sea

$$\overline{B(a, \rho)} = B'(a, \rho) \text{ en } \mathbb{R}^n \quad (81)$$

Observación: Esto no se cumple en espacios métricos cualesquiera, como muestran las bolas $B(a, 1) = \{a\}$ en un espacio métrico discreto X compuesto por más de un punto, que son iguales a su adherencia y diferentes de las bolas cerradas $B'(a, 1) = X$.

Ejemplo 62. Consideremos en el espacio métrico $\mathbb{R}^2$ el grafo de la aplicación

$$f : x \rightsquigarrow \text{sen } (1/x) \text{ sobre }]0, 1/\pi]:$$

$$A = \{(x, \text{sen } (1/x)) : 0 < x \leq 1/\pi\}.$$

Entonces, todo punto $(0, y)$ sobre el eje OY tal que $-1 \leq y \leq 1$ es un punto adherente de A (fig. 46).

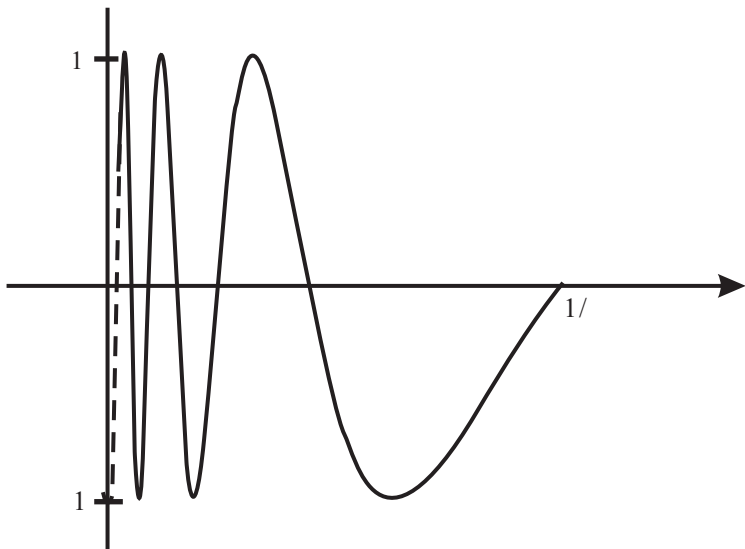


Figura 46

Ejemplo 63. En un espacio discreto (X, d) un punto $a \in X$ es adherente a un subconjunto $A \subset X$ si y sólo si $a \in A$. Para ver esto basta aplicar la definición de punto adherente a la vecindad $\{a\}$ de a :

$$A \subset X \Rightarrow A = \overline{A} \text{ en todo espacio métrico discreto } (X, d).$$

Proposición 2.9.3 (puntos adherentes en un subespacio). Sean (X, d) un espacioseudométrico y $B \subset A \subset X$; entonces, si denotamos por $\overline{B}_A$ la clausura de B en el subespacio métrico (A, d_A) , tendremos

$$\overline{B}_A = \overline{B} \cap A. \quad (82)$$

Demostración: En efecto, se tiene la siguiente cadena de equivalencias:

$$a \in \overline{B} \cap A \Leftrightarrow a \in \overline{B} \wedge a \in A \Leftrightarrow$$

$$\Leftrightarrow \forall V \in \mathbf{V}(a) (V \cap B \neq \emptyset) \wedge a \in A \Leftrightarrow$$

$$\Leftrightarrow \forall V \in \mathbf{V}(a) (V \cap B \cap A \neq \emptyset) \Leftrightarrow$$

$$\Leftrightarrow \forall V_A \in \mathbf{V}_A(a) (V_A \cap B \neq \emptyset) \Leftrightarrow$$

$$\Leftrightarrow a \in \overline{B}_A \text{ (Notación: } V_A = V \cap A \text{)}$$

Notemos que en la penúltima equivalencia hemos utilizado el hecho de que las vecindades V_A de a en A son de la forma

$$V_A = V \cap A; V \in \mathbf{V}_A(a). \blacksquare$$

Ejemplo 64. Como vemos de la proposición 2.9.3 $\overline{B} \supset \overline{B}_A$ y en general la inclusión puede ser estricta. Por ejemplo, si tomamos $B =]0, 1]$, $A =]0, \infty[$, $X = \mathbb{R}$, entonces $0 \in \overline{B}$ y $0 \notin \overline{B}_A$.

Consideremos un conjunto finito $A = \{x_i \mid i = 1, \dots, m\}$ en el espacio $\mathbb{R}^2$. Todo punto en el plano que no pertenece a A tiene una vecindad que no contiene a ningún punto de A ; luego los únicos puntos adherentes de A son los x_i . Además, todo punto x_i posee una vecindad V_i que excluye a todos los demás puntos de A (fig. 47).

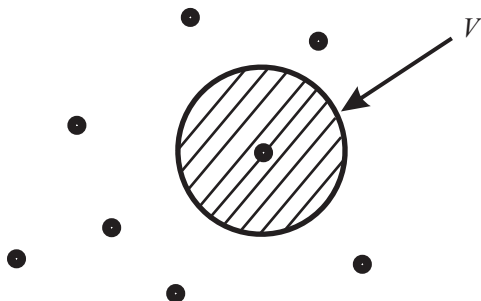


Figura 47

Tales puntos se llaman *aislados* en A . Un conjunto finito A en $\mathbb{R}^2$ posee solamente puntos aislados, es decir, los puntos de A “no se acumulan” en ningún punto de $\mathbb{R}^2$.

Definición 2.9.4. Sea (X, d) un espacio sudométrico y $A \subset X$. Un punto $x \in X$ se llama *punto de acumulación* o *punto límite* de A , si toda vecindad $V \in \mathbf{V}(x)$ contiene un punto de A distinto de x :

$$V \in \mathbf{V}(x) \Rightarrow A \cap (V \setminus \{x\}) = V \cap (A \setminus \{x\}) \neq \emptyset \quad (83)$$

y se llama *punto aislado* si existe $V \in \mathbf{V}(x)$ tal que

$$V \cap A = \{x\} \quad (84)$$

El conjunto de los puntos de acumulación de A se llama *conjunto derivado* de A y se denota A' .

Luego, distinguimos dos clases de puntos de adherencia: *los puntos de acumulación* y *los puntos aislados*.

Los puntos aislados de A pertenecen necesariamente a A , mientras que los puntos de acumulación pueden pertenecer o no a A :

$$x \in A' \iff x \in \overline{A \setminus \{x\}}. \quad (85)$$

Todo punto adherente que no pertenece a A es necesariamente un punto de acumulación. $\overline{A}$ se obtiene uniendo los puntos de acumulación al conjunto A :

$$\overline{A} = A \cup A'. \quad (86)$$

Ejemplo 64. (conjunto de puntos aislados que posee un punto de acumulación) Consideremos el conjunto $A = \{\frac{1}{n} : n \in \mathbb{N}^*\}$ en $\mathbb{R}$. Obviamente A contiene solamente puntos aislados. Sin embargo, no es difícil probar que $0 \in A'$.

Ejemplo 65. Sea (x_n) una sucesión convergente a x en un espacio métrico (X, d) y posee un número infinito de términos diferentes, entonces x es el único punto de acumulación del conjunto de los términos $\{x_n : n \in \mathbb{N}\}$.

Teorema 2.9.5 (Caracterización de los puntos de acumulación). En un espacio métrico (X, d) un punto $x \in X$ es un punto de acumulación de $A \subset X$ si y sólo si existe una sucesión (x_n) en A de términos diferentes entre sí que converge a x .

Demostración: Probemos por inducción que se puede obtener una sucesión (x_n) de elementos de A , cuyos términos son diferentes entre sí y a su vez diferentes a x , y tales que $d(x_n, x) < \frac{1}{n}$. En efecto, tomemos $x_1 \neq x$ tal que $x_1 \in B(x, 1)$. A partir de x_1 seleccionemos $x_2 \neq x$ en la bola $B(x, \min\{d(x_1, x), 1/2\})$.

Teniendo ya elegido x_n , sabemos que podemos encontrar $x_{n+1} \neq x$ con $x_{n+1} \in B(x, \min\{d(x_n, x), 1/(n+1)\})$.

Evidentemente $x_{n+1} \notin \{x_1, x_2, \dots, x_n\}$ ya que $d(x_{n+1}, x) < d(x_n, x) < \dots < d(x_1, x)$.

Por la hipótesis de inducción, podemos entonces construir una sucesión (x_n) cuyos términos son todos diferentes y tal que para todo n $d(x_n, x) < 1/n$ y por tanto, $(x_n) \rightarrow x$. ■

En el análisis clásico, además del concepto *límite de una sucesión*, se utiliza el concepto *punto adherente o punto de acumulación de una sucesión*, el cual difiere de los ya definidos para un conjunto.

Definición 2.9.6. Sea (X, d) un espacio pseudométrico y (x_k) una sucesión

de elementos de X . El punto $x \in X$ se llama *punto adherente* o *punto de acumulación* de la sucesión (x_k) si:

$$\forall \epsilon > 0, \forall n \in \mathbb{N}, \exists k \in \mathbb{N} \text{ tal que } k \geq n \text{ y } d(x_k, x) < \epsilon \quad (87)$$

No es difícil verificar que la propiedad de un punto x de ser punto de acumulación de una sucesión (x_n) , es equivalente a una de las proposiciones siguientes:

$$\text{i) } \forall \epsilon > 0, \{n : x_n \in B(x, \epsilon)\} \text{ es infinito.} \quad (88)$$

Observación: Notemos que si x fuera $\lim(x_n)$ entonces dentro de cada bola $B(x, \epsilon)$ no habrían solo un número infinito de términos de la sucesión (x_n) sino que estarían todos los términos a partir de un cierto $n_0 = n_0(\epsilon)$.

$$\text{ii) Sea } S_n(x_k) = \{x_k : k \geq n\} \text{ la } n\text{-ésima sección final de la sucesión } (x_k). \\ \text{Entonces}$$

$$\forall \epsilon > 0 \wedge \forall n \in \mathbb{N}, S_n(x_k) \cap B(x, \epsilon) \neq \emptyset \quad (89)$$

$$\text{iii) } \forall V \in \mathbf{V}(x) \wedge \forall n \in \mathbb{N}, V \cap S_n(x_k) \neq \emptyset \quad (90)$$

Observación: esta última caracterización nos relaciona los conceptos de puntos adherentes a un conjunto y a una sucesión ya que, x es un punto de acumulación de la sucesión (x_n) si y sólo si x es un punto adherente de todas sus secciones finales $S_n(x_k)$.

La demostración de iii se deja al lector.

Ejemplo 66. La sucesión (x_n) definida en $\mathbb{R}$ mediante

$$x_n = \begin{cases} 1 & \text{si } n \text{ es par} \\ -1 & \text{si } n \text{ es impar} \end{cases}$$

tiene dos puntos de acumulación: 1 y -1 . Sin embargo, el conjunto de sus términos $A = \{x_n : n \in \mathbb{N}\} = \{1, -1\}$ no tiene ningún punto de acumulación.

Ejemplo 67. Por ser $\{x_n : n \in \mathbb{N}\} = S_1(x_k)$ y por (90) se tiene que todo punto de acumulación de una sucesión (x_n) en un espacio métrico (X, d) , es punto adherente del conjunto de todos sus términos. Sin embargo, el inverso no se cumple ya que una sucesión que contenga al menos un elemento que

corresponda a un número infinito de términos de la sucesión y que sea aislado, será un punto de acumulación de la sucesión.

Ejemplo 68. Una sucesión puede tener infinitos puntos de acumulación. Por ejemplo, en $\mathbb{R}$ la sucesión constituida por todos los números racionales del intervalo $[0, 1]$ (ordenados de la forma indicada en el capítulo 1) es tal, que todos sus puntos son de acumulación.

Pregunta: ¿Cuál es el conjunto de todos los puntos de acumulación de esta sucesión?

Ejemplo 69 (límite superior y límite inferior de una sucesión de números reales). Hemos visto cómo una sucesión puede tener más de un punto de acumulación. Por otra parte, si la sucesión converge, puede probarse (ver ejercicios de autocontrol) que el límite es su único punto de acumulación.

Sea (x_n) una sucesión de números reales y sea

$$A = \{ x : x \text{ es un punto de acumulación de } (x_n) \}.$$

Entonces la estructura de orden en $\mathbb{R}$, permite definir el límite superior y el límite inferior de (x_n) mediante:

$$\limsup (x_n) = \sup A, \quad (91)$$

$$\liminf (x_n) = \inf A. \quad (92)$$

La relación entre la estructura topológica de $\mathbb{R}$ y su estructura de orden, se evidencia en un hecho bien conocido del Análisis Clásico: una sucesión (x_n) de números reales converge si y sólo si

$$\limsup (x_n) = \liminf (x_n) = x, \text{ y en este caso } \lim (x_n) = x$$

Pasemos ahora a introducir el concepto de *conjunto denso*, el cual resulta de gran utilidad en los razonamientos del Análisis. Ya hemos visto que los puntos de $\mathbb{Q}$ son densos en la recta en el sentido intuitivo de que no existe un segmento (no reducido a un punto), que no contenga a un número racional. Esto significa topológicamente que $\overline{\mathbb{Q}} = \mathbb{R}$. Los conjuntos con esta propiedad juegan un papel importante en la Topología General, así como en el Análisis Matemático y el Análisis Funcional.

Definición 2.9.7. Un subconjunto A de un espacio pseudométrico (X, d) se llama *denso* en X , si $\overline{A} = X$.

Es claro que las caracterizaciones generales de la clausura se aplican en este caso. Por ejemplo:

$$A \subset X \text{ es denso en } X \Leftrightarrow A \cap V \neq \emptyset \quad (93)$$

para toda vecindad V de cualquier punto $x \in X$.

$$A \subset X \text{ es denso en } X \iff A \cap B(x, \epsilon) \neq \emptyset \quad (93i)$$

para todo $x \in X$ y todo $\epsilon \geq 0$.

$$A \subset X \text{ es denso en } X \iff \text{para todo } x \in X \text{ se puede hallar una sucesión } (x_n) \text{ en } A \text{ tal que } (x_n) \longrightarrow x. \quad (93ii)$$

$$A \subset X \text{ es denso en } X \iff \text{para todo } x \in X, d(x, A) = 0. \quad (93iii)$$

Ejemplo 70. En un espacio métrico discreto (X, d) el único subconjunto denso en X es el propio X .

Ejemplo 71. El conjunto $\mathbb{Q}^n$ de los puntos con coordenadas racionales es denso en $\mathbb{R}^n$. En efecto, del Análisis clásico sabemos que todo elemento de $\mathbb{R}$ puede ser aproximado por una sucesión de números racionales. Luego, si

$$x = (x^{(1)}, x^{(2)}, \dots, x^{(n)}) \in \mathbb{R}^n$$

y construimos la sucesión

$$(r_k) = \left(r_k^{(1)}, r_k^{(2)}, \dots, r_k^{(n)} \right), r_k^{(j)} \in \mathbb{Q},$$

donde para cada $j = 1, 2, \dots, n$

$$\left(r_k^{(j)} \right) \longrightarrow x^{(j)} \text{ en } \mathbb{R},$$

entonces

$$(r_k) \longrightarrow x \text{ en } \mathbb{R}^n,$$

y por lo tanto, $x \in \overline{\mathbb{Q}^n}$.

Este ejemplo 71 nos muestra que $\mathbb{R}^n$ tiene un conjunto denso numerable.

Definición 2.9.8. Un espacio métrico es *separable*, si tiene algún subconjunto denso numerable

Ejemplo 72. Un espacio métrico discreto (X, d) es separable si y sólo si X es numerable.

Ejemplo 73. El espacio $l^2(\mathbb{N})$ es separable. Un subconjunto numerable denso de $l^2(\mathbb{N})$ se construye de la manera siguiente.

Sea el conjunto de todas las sucesiones racionales cuyos términos se anulan a partir del índice n :

$$D_n := \{(x_k) \in l^2(\mathbb{N}) : \forall k \in \mathbb{N} (x_k \in \mathbb{Q} \wedge k > n \implies x_k = 0)\}.$$

D_n es equipotente a $\mathbb{Q}^n$, luego es numerable.

Veamos que el conjunto numerable $D := \bigcup_{n=1}^{\infty} D_n$ es denso en $l^2(\mathbb{N})$. En

efecto, sea $Z = (z_k) \in l^2(\mathbb{N})$ arbitrario y sea $\epsilon > 0$ dado. Puesto que $\sum_{k=1}^{\infty} z_k^2$ es

convergente, existe $N \in \mathbb{N}$ tal que $\sum_{k=N+1}^{\infty} z_k^2 < \frac{\epsilon}{2}$.

Para todo $p = 1, \dots, N$ podemos hallar un número racional r_p tal que

$$|r_p - z_p|^2 < \epsilon/2N$$

Si definimos $R := (r_1, r_2, \dots, r_n, 0, 0, \dots) \in D$, entonces

$$d_2(Z, R)^2 \leq \sum_{k=1}^n |z_k - r_k|^2 + \sum_{k=N+1}^n z_k^2 \leq$$

$$\leq N\epsilon/2N + \epsilon/2 = \epsilon.$$

Luego, D es denso en $l^2(\mathbb{N})$.

Ejemplo 74. (espacio métrico no separable) Denotaremos por l^∞ el espacio vectorial de todas las sucesiones reales acotadas, provisto de la métrica:

$$d_\infty((x_n), (y_n)) := \sup\{|x_n - y_n| : n \in \mathbb{N}\}.$$

Probemos que ningún subconjunto numerable de l^∞ puede ser denso en l^∞ .

En efecto, si:

$$A = \left\{ \left\{ x_n^{(k)} \right\}_{n=1}^{\infty} : 1 \leq k < \infty \right\},$$

donde

$$\left\{ x_n^{(k)} \right\}_{n=1}^{\infty} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}, \dots) \in l^\infty,$$

entonces el elemento $\{x_n\}_{n=1}^{\infty}$, definido por:

$$x_n = x_n^{(n)} + 1$$

satisface para todo $k = 1, 2, \dots$

$$d_\infty(\{x_n\}, \{x_n^{(k)}\}) = \max \left\{ |x_p - x_p^{(k)}|, 1 \leq p < \infty \right\} \geq$$

$$\geq |x_n - x_n^{(n)}| = 1,$$

y por lo tanto, el elemento $\{x_n\}_{n=1}^{\infty}$ no puede ser aproximado por ninguna sucesión de elementos de A .

Como acabamos de ver existen espacios métricos no separables; un ejemplo muy simple de este hecho lo brinda también cualquier espacio métrico discreto no numerable.

No resulta difícil probar que todo $AN2$ -espacio (definición 2.8.9) es separable. En efecto, sea $B = \{B_n : n \in \mathbb{N}\}$ una base numerable en (X, d) . Escogemos de todo $B_n \neq \emptyset$ un punto arbitrario $x_n \in B_n$. El conjunto $D := \{x_n : n \in \mathbb{N}, B_n \neq \emptyset\}$ es un conjunto numerable denso en (X, d) pues todo abierto no vacío U en (X, d) contiene a un $B_n \neq \emptyset$, luego $x_n \in U \cap D \neq \emptyset$.

En este caso se cumple también el inverso según nos muestra la proposición siguiente.

Proposición 2.9.9. Un espacioseudométrico cumple $AN2$ si y sólo si es separable.

Demostración: Solo tenemos que probar que todo espacioseudométrico separable cumple $AN2$.

Sea $\{x_n : n \in \mathbb{N}\}$ un subconjunto denso de (X, d) . Probaremos que la colección numerable de bolas abiertas

$B := \{x_n, 1/m) : n \in \mathbb{N}, m \in \mathbb{N}^*\}$ forma una base del espacio métrico (X, d) .

Sea $U \in \mathcal{X}_d$ y $x \in U$.

Hay que demostrar que existe un $B \in \mathbf{B}$ tal que $x \in B \subset U$.

Ahora bien, existe $m \in \mathbb{N}^*$ tal que $B(x, \frac{1}{m}) \subset U$ y existe $n \in \mathbb{N}$ tal que $x_n \in B(x, 1/2m)$. Concluimos que $x \in B(x_n, 1/2m) \subset B(x, 1/m) \subset U$. ■

Para finalizar este epígrafe, expondremos un resultado importante que se obtiene al aplicar los conceptos topológicos a los espacios métricos de funciones.

Veamos primeramente la definición siguiente:

Definición 2.9.10. Una función $f : \mathbb{R} \longrightarrow X$ se llama *periódica*, si existe un número $T \in \mathbb{R}$ tal que para todo $x \in \mathbb{R}$:

$$f(x + T) = f(x) \quad (94)$$

Al número T se le llama *período* de f

Notemos que si f es periódica, de período T , también tendrá periodo kT para todo $k \in \mathbb{Z}$

Ejemplo 75 (polinomios trigonométricos). Un tipo importante de funciones periódicas de $\mathbb{R}$ en $\mathbb{R}$ lo constituyen los *polinomios trigonométricos*.

Un polinomio trigonométrico es una combinación lineal de las funciones

$$\{1, \text{sen } Wnx, \cos Wnx\}, n = 1, 2, \dots$$

donde W es un número real fijo no nulo, es decir:

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n \{a_k \cos Wkx + b_k \operatorname{sen} Wkx\}.$$

El número $T = \frac{2\pi}{W}$ es el periodo del polinomio trigonométrico.

Ejemplo 76 (polinomios algebraicos). Un *polinomio algebraico* es una función de $\mathbb{R}$ en $\mathbb{R}$ definida mediante una combinación lineal de las funciones:

$$\{1, x^k\} \quad k = 1, 2, \dots, n \text{ es decir} \quad P_n(x) = \sum_{k=1}^n a_k x^k.$$

No resulta difícil demostrar que un polinomio algebraico no es una función periódica salvo si es constante.

Teorema 2.9.11 (primer teorema de Weierstrass, sobre la aproximación uniforme de cualquier función continua periódica por polinomios trigonométricos). Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función continua periódica de periodo T . Entonces existe una sucesión $(\sigma_n(x))$ de polinomios trigonométricos de periodo T que converge uniformemente a $f(x)$ en todo $\mathbb{R}$:

$$d_\infty(\sigma_n, f) = \sup_{x \in \mathbb{R}} |\sigma_n(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0 \quad (95)$$

Demostración: Notemos en primer lugar que en vista de la periodicidad de σ_n y f basta demostrar la convergencia uniforme en cualquier intervalo de longitud T . Para simplificar los cálculos, tomaremos además $W = 1$, ya que si para cualquier función continua f de período 2π se prueba la existencia de una sucesión de polinomios trigonométricos (σ_n) de período 2π que satisfaga (95), entonces se cumplirá la afirmación del teorema para cualquier T . En efecto, si g es continua de periodo T la función $f(x) = g\left(\frac{T}{2\pi}x\right)$ seguiría siendo continua pero con período 2π . Si la sucesión $(\sigma_n(x))$ de polinomios trigonométricos de período 2π converge uniformemente a $f(x)$ en $\mathbb{R}$, entonces la sucesión $(\sigma_n(\frac{2\pi}{T}x))$ converge uniformemente a $g(x)$. No es difícil verificar que $(\sigma_n(\frac{2\pi}{T}x))$ es un polinomio trigonométrico de periodo T .

Supongamos entonces que $f : [-\pi, \pi] \rightarrow \mathbb{R}$ es una función continua de periodo 2π , y construyamos una sucesión $(\sigma_n(x))$ de polinomios trigonométricos de período 2π que converja uniformemente a $f(x)$ en el intervalo $[-\pi, \pi]$. Trazaremos el esquema de la demostración y dejaremos al lector encargado de culminar algunos detalles técnicos.

Consideremos el polinomio trigonométrico

$$S_k(x) = \frac{a_0}{2} + \sum_{j=1}^k a_j \cos jx + b_j \operatorname{sen} jx \quad (96)$$

donde los coeficientes a_j los hemos tomado de la forma

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx \quad (97)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx, \quad (98)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx dx, \quad (99)$$

y se denominan *coeficientes de Fourier* de la función f . Al polinomio trigonométrico $S_k(x)$ se le llama *suma parcial de orden k de la serie de Fourier* de f y al polinomio trigonométrico

$$\sigma_k(x) = \frac{S_0(x) + S_1(x) + \dots + S_{n-1}(x)}{n} \quad (100)$$

obtenido como la media aritmética de los polinomios $S_k(x)$, lo llamaremos *suma de Fejer* de f de orden n .

Obviamente (σ_n) es una sucesión de polinomios trigonométricos de periodo 2π .

Probemos que $(\sigma_n) \xrightarrow{d_\infty} f$.

Utilicemos la siguiente representación integral para las sumas parciales de la serie de Fourier de f :

$$S_k(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \frac{\operatorname{sen} \frac{2k+1}{2}(t-x)}{2 \operatorname{sen} \frac{t-x}{2}} dt, \quad (101)$$

la cual puede ser verificada por el lector.

Introduciendo la expresión (101) en la igualdad (100), obtenemos la siguiente expresión para $\sigma_n(x)$:

$$\sigma_n(x) = \frac{1}{2n\pi} \int_{-\pi}^{\pi} \left\{ \sum_{k=0}^{n-1} \frac{\operatorname{sen} \frac{2k+1}{2}(t-x)}{2 \operatorname{sen} \frac{t-x}{2}} \right\} f(t) dt, \quad (102)$$

que mediante la fórmula

$$\sum_{k=0}^{n-1} \operatorname{sen} (2k+1)u = \frac{\operatorname{sen}^2 nu}{\operatorname{sen} u}, \quad (103)$$

puede ser reducida a la forma

$$\sigma_n(x) = \frac{1}{2n\pi} \int_{-\pi}^{\pi} \left(\frac{\operatorname{sen} n \frac{t-x}{2}}{\operatorname{sen} \frac{t-x}{2}} \right)^2 f(t) dt. \quad (104)$$

La expresión

$$\phi_n(z) = \frac{1}{2n\pi} \left(\frac{\operatorname{sen} n \frac{z}{2}}{\operatorname{sen} \frac{z}{2}} \right)^2 \quad (105)$$

es el llamado núcleo de Fejer. Tomando $t - x = z$ y teniendo en cuenta que el valor de la integral de una función periódica es siempre el mismo en un segmento de longitud igual al período, podemos escribir (104) en la forma

$$\sigma_n(x) = \int_{-\pi}^{\pi} f(x+z) \phi_n(z) dz. \quad (106)$$

Debemos demostrar que para $n \rightarrow \infty$ esta expresión converge uniformemente hacia $f(x)$. Señalemos primeramente las propiedades siguientes del núcleo de Fejer:

$$1) \quad \phi_n(z) \geq 0.$$

$$2) \quad \int_{-\pi}^{\pi} \phi_n(z) dz = 1,$$

$$3) \quad \text{para todo } \delta > 0 \text{ fijo y } n \rightarrow \infty \text{ tenemos}$$

$$\int_{-\pi}^{-\delta} \phi_n(z) dz = \int_{\delta}^{\pi} \phi_n(z) dz = \eta_n(\delta) \rightarrow 0$$

La primera de estas propiedades es evidente, la segunda se obtiene directamente de la igualdad (106) si tomamos $f(x) \equiv 1$ y observamos que para esta función $\sigma_n(x) = 1$ para todo n ; la tercera propiedad se desprende inmediatamente de que para $\delta < z \leq \pi$ se tiene $\operatorname{sen} \frac{z}{2} \geq \frac{2\delta}{\pi}$ y consecuentemente

$$\left(\frac{\operatorname{sen} n \frac{z}{2}}{\operatorname{sen} \frac{z}{2}} \right) \leq \left(\frac{\pi}{2\delta} \right)^2.$$

Tomando en consideración las propiedades del núcleo de Fejer no es difícil demostrar el teorema 2.9.11. Como f es una función continua y periódica, es acotada y uniformemente continua en toda la recta (estos resultados se obtienen en los cursos básicos de Análisis Matemático y serán generalizados en el capítulo 4).

En otras palabras, existe una constante M tal que para todo x

$$|f(x)| \leq M \quad (107)$$

y, además, para $\epsilon > 0$ existe un $\delta > 0$ tal que

$$|f(x'') - f(x')| < \frac{\epsilon}{2} \quad (108)$$

siempre que

$$|x'' - x'| < \delta$$

Para demostrar el teorema debemos estimar la diferencia

$$f(x) - \sigma_n(x) = \int_{-\pi}^{\pi} [f(x) - f(x+z)] \phi_n(z) dz,$$

que puede ser representada como suma de las tres integrales siguientes:

$$J_1 = \int_{-\pi}^{-\delta} [f(x) - f(x+z)] \phi_n(z) dz,$$

$$J_2 = \int_{-\delta}^{\delta} [f(x) - f(x+z)] \phi_n(z) dz,$$

$$J_3 = \int_{\delta}^{\pi} [f(x) - f(x+z)] \phi_n(z) dz.$$

De (107) y (108) se obtienen directamente los estimados siguientes:

$$|J_1| \leq 2M\eta_n(\delta)$$

$$|J_3| \leq 2M\eta_n(\delta)$$

$$|J_2| \leq \frac{\epsilon}{2} \int_{-\delta}^{\delta} \phi_n(z) dz < \frac{\epsilon}{2}$$

Escojamos ahora n_0 tan grande que para $n \geq n_0$ y un δ dado se cumpla la desigualdad

$$2M\eta_n(\delta) < \frac{\epsilon}{4}.$$

Entonces,

$$|f(x) - \sigma_n(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon$$

y de aquí, debido a la arbitrariedad de ϵ se desprende la afirmación del teorema. ■

Teorema 2.9.12 (segundo teorema de Weierstrass sobre la aproximación uniforme de funciones $f : [a, b] \longrightarrow \mathbb{R}$, por polinomios algebraicos). Sean a y b números reales con $a < b$ y $f : [a, b] \longrightarrow \mathbb{R}$ continua.

Entonces existe una sucesión $(Q_m(x))$ de polinomios algebraicos tal que

$$d_\infty(Q_m, f) = \sup \{|Q_m(x) - f(x)| : a \leq x \leq b\} \xrightarrow{n \rightarrow \infty} 0 \quad (109)$$

Demostración: Se deduce inmediatamente del primer teorema de Weierstrass. En efecto, si $f : [a, b] \rightarrow \mathbb{R}$ y hacemos el cambio de variable $x = \frac{t(b-a)}{\pi} + a$, obtenemos una función $\varphi(t)$ de la variable t definida en el segmento $[0, \pi]$. Prolonguemosla primero al segmento $[-\pi, 0]$ haciendo $\varphi(-t) = \varphi(t)$, y después, por periodicidad, a toda la recta real. Construyamos ahora un polinomio trigonométrico T_n que verifique la condición $|T_n(t) - \varphi(t)| < \frac{\epsilon}{2}$ para todo t .

Ahora bien, este polinomio trigonométrico puede ser desarrollado en una serie de Taylor que converge uniformemente en cualquier intervalo finito. Sea P_m la suma parcial de la serie de Taylor para T_n tal que $|T_n(t) - P_m(t)| < \frac{\epsilon}{2}$ para $0 \leq t \leq \pi$.

Efectuando en $P_m(t)$ la sustitución $t = \frac{x-a}{b-a}$ obtenemos un polinomio $Q_m(x)$ que satisface, evidentemente, la condición

$$|f(x) - Q_m(x)| < \epsilon \text{ para } a \leq x \leq b. \blacksquare$$

Utilizando la afirmación (93ii) podemos dar las siguientes formulaciones equivalentes de los teoremas 2.9.11 y 2.9.12

El conjunto de los polinomios trigonométricos de periodo T es denso en $(C_T(\mathbb{R}, \mathbb{R}), d_\infty)$, donde

$$C_T(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ es continua y tiene periodo } T\}$$

El conjunto de los polinomios algebraicos es denso en $(C([a, b], \mathbb{R}), d_\infty)$.

De los resultados anteriores se tiene el corolario siguiente:

Corolario 1. El espacio métrico $(C([a, b], \mathbb{R}))$ con cualquiera de las métricas d_∞, d_1, d_2 (ver (8)-(10)) es separable.

Demostración: Partiendo del hecho que:

$$d_2(f, g) \leq \sqrt{b-a} d_\infty(f, g),$$

$$d_1(f, g) \leq (b-a) d_\infty(f, g),$$

y de (93iii) basta probar que $(C([a, b], \mathbb{R}), d_\infty)$ tiene un subconjunto denso numerable, el cual también será denso con respecto a d_1 y d_2 .

El conjunto $\mathbf{P}_\mathbb{Q}$ de todos los polinomios algebraicos cuyos coeficientes son números racionales es un conjunto numerable ¿por qué?.

Mostremos que este conjunto es denso en $(C([a, b], \mathbb{R}), d_\infty)$.

En efecto, si $f : [a, b] \rightarrow \mathbb{R}$ y $\epsilon > 0$ por el teorema 2.9.12, podemos encontrar un polinomio algebraico

$$Q_m(x) = \sum_{k=0}^{m-1} a_k x^k, a_k \in \mathbb{R}$$

tal que

$$d_\infty(Q_m, f) < \frac{\epsilon}{2} \quad (110)$$

Sea $v_k \in \mathbb{Q}$ para cada $k = 0, 1, \dots, m-1$, que cumple

$$|a_k - v_k| < \frac{\epsilon}{2mc^k}, \quad (111)$$

donde $c = \max\{|a|, |b|\}$.

Consideremos

$$\tilde{Q}_m(x) = \sum_{k=0}^{m-1} r_k x^k \in \mathbf{P}_{\mathbb{Q}}$$

De (111) se deduce que

$$d_\infty(Q_m, \tilde{Q}_m) < \frac{\epsilon}{2} \quad (112)$$

De (110) y (112) tendremos

$$d_\infty(f, \tilde{Q}_m) < \epsilon.$$

Luego, $\mathbf{P}_{\mathbb{Q}}$ es denso en $(C([a, b], \mathbb{R}), d_\infty)$ ■

En los teoremas 2.9.11 y 2.9.12 hemos dado la demostración expuesta en el libro *Elementos de la teoría de funciones y del análisis funcional* de A.N. Kolmogorov y S.V Fomin.

El primer teorema de Weierstrass con la demostración *constructiva* expuesta, es también llamado teorema de Fejer. Este tipo de demostración se llama constructiva pues no solamente hemos demostrado la existencia de las sucesiones (σ_n) y (Q_n) en ambos teoremas, sino además, hemos dado su expresión explícita. Notemos que esto no quiere decir de ninguna forma que la única sucesión de polinomios que satisface cada teorema sea σ_n o Q_m .

Por ejemplo si $f : [0, 1] \rightarrow \mathbb{R}$ y $Q_m(x) \xrightarrow{d_\infty} f$, entonces también

$$(Q_m(x) + x^m(1-x)^m) \xrightarrow{d_\infty} f, \text{ ya que } (x^m(1-x)^m) \xrightarrow{d_\infty} \langle 0 \rangle \text{ en } [0, 1].$$

Observemos que en la demostración del teorema 2.9.11 se han empleado solamente las propiedades 1), 2) y 3) del núcleo de Fejer; esto permite obtener diferentes generalizaciones del teorema 2.9.11 dentro de la llamada *Teoría de Aproximación*. En esta teoría se estudian, entre otras cosas, las condiciones

bajo las cuales una cierta familia P de funciones simples, es capaz de aproximar, mediante algún concepto de convergencia, a las funciones de una clase más amplia $\mathbf{F} \supset \mathbf{P}$. La teoría de aproximación es una de las herramientas más útiles del Análisis Funcional y de la aplicación de los métodos numéricos a la solución de problemas en la Matemática.

Ejercicios de Autocontrol

1. Sea (x_n) una sucesión en el espacio pseudométrico (X, d) , en la cual se repiten a lo más un número finito de elementos. Pruebe que un punto $x \in X$ es de acumulación para la sucesión (x_n) si y sólo si es punto adherente del conjunto de sus elementos $\{x_n : n \in \mathbb{N}\}$.
2. Demuestre que si $(x_{\phi(n)})$, es una subsucesión de (x_n) , entonces todo punto de acumulación de $(x_{\phi(n)})$ también lo es de (x_n) .
3. Demuestre que un punto x es punto de acumulación de la sucesión (x_n) si y sólo si existe una subsucesión $(x_{\phi(n)})$ de (x_n) con $(x_{\phi(n)}) \xrightarrow{d} x$.
4. Demuestre que un punto x es adherente a una sucesión (x_k) si y sólo si x es adherente a toda sección final:

$$x \in \overline{S_n(x_k)} = \overline{\{x_k : k \geq n\}}$$
 para todo $n \in \mathbb{N}$.
5. Sea (X, d) un espacio pseudométrico y $B \subset A \subset X$. Halle la relación existente entre los puntos de acumulación de B en (A, d_A) y en (X, d) .
6. Demuestre que todo punto aislado de $\overline{A}$ en (X, d) es también punto aislado de A .
7. Pruebe que una sucesión convergente tiene un único punto de acumulación, el cual es precisamente el límite de la sucesión. Halle una sucesión en un espacio métrico con un único punto de acumulación y que no sea convergente.
8. Demuestre que si $B \subset A \subset X$ entonces B es denso en (A, d_A) si y sólo si $\overline{B} \supset A$, donde $\overline{B}$ es la clausura de B en (X, d) .
9. Construya un conjunto $A \subset \mathbb{R}$ tal que $A \neq A'$ y $(A')' = \{0\}$.
10. Sean $A_1, \dots, A_n$ subconjuntos de un espacio métrico (X, d) .
 Pruebe que

$$(A_1 \cup \dots \cup A_n)' = A_1' \cup \dots \cup A_n'.$$

11. Sea (X, d) un espacio pseudométrico. Demuestre que una familia F de conjuntos abiertos en (X, d) es una base topológica, si y sólo si para todo abierto A y cada $x \in A$ existe $B \in F$ con $x \in B \subset A$.
12. Demuestre utilizando el principio de inducción y la expresión polar para un número complejo, que se cumplen las igualdades (101)-(104) del teorema 2.9.11.

2.10 CONJUNTOS CERRADOS

Para problemas de convergencia y de continuidad, los conjuntos que contienen a todos sus puntos límites tienen una importancia fundamental

Definición 2.10.1. Sea (X, d) un espacio pseudométrico, un subconjunto $A \subset X$ se llama *cerrado* si $A = \overline{A}$.

El conjunto de todos los cerrados se denota por τ_d .

Para comprobar que un subconjunto A de (X, d) es cerrado hay que verificar que contiene a todos sus puntos de acumulación.

Luego, un subconjunto A de un espacio métrico es cerrado si contiene a todos los límites de sucesiones convergentes (x_n) con $x_n \in A$ para todo $n \in \mathbb{N}$.

Ejemplo 77. Toda bola cerrada $B'(a, \rho)$ en un espacio métrico (X, d) es cerrada, pues, si una sucesión (x_n) de puntos $x_n \in B'(a, \rho)$ converge a $x \in X$, se tiene que $d(x, a) = \lim_n (d(x_n, a)) \leq \rho$, por la continuidad de $z \rightsquigarrow d(z, a)$.

Ejemplo 78. Sea (x_n) una sucesión convergente en un espacio métrico (X, d) . Entonces el conjunto $\{x_n : n \in \mathbb{N}\} \cup \{\lim x_n\}$ es cerrado en (X, d) .

Proposición 2.10.2. La adherencia $\overline{A}$ de un subconjunto A de (X, d) es cerrado, es decir, $\overline{A} = \overline{\overline{A}}$. Con más precisión $\overline{A}$ es el menor cerrado que contiene a A .

Demostración: Evidentemente $\overline{A} \subset \overline{\overline{A}}$. Probemos la inclusión en sentido contrario.

$$\text{Sea } x \in \overline{\overline{A}}, \text{ entonces para todo } \epsilon > 0, B(x, \epsilon) \cap \overline{A} \neq \emptyset. \quad (113)$$

Si suponemos que para algún $\delta > 0$ se tiene $B(x, \delta) \cap A = \emptyset$, entonces se tendría que $B(x, \frac{\delta}{2}) \cap \overline{A} = \emptyset$, lo cual está en contradicción con (113) ya que $d(y, A) \geq \frac{\delta}{2}$ para todo $y \in B(x, \delta/2)$ y sabemos que $y \in \overline{A}$ si y sólo

si $d(y, A) = 0$. Por otra parte, si B es un cerrado que contiene a A , entonces todo punto adherente de A lo será también de B , y por lo tanto, $\overline{A} \subset \overline{B} = B$. Luego, $\overline{A}$ es el menor conjunto cerrado que contiene a A . ■

El siguiente teorema nos muestra una gran variedad de nuevos ejemplos y un nuevo criterio para probar que un conjunto dado es cerrado.

Teorema 2.10.3. Sea (X, d) un espacio pseudométrico. Entonces $A \subset X$ es cerrado, si y sólo si $X \setminus A$ es abierto.

Demostración: Si A es cerrado, entonces ningún punto $x \in X \setminus A$ es adherente de A , luego existe una vecindad $V \in \mathbf{V}(x)$ tal que $A \cap V = \emptyset$ o sea $V \subset X \setminus A$. Esto nos muestra que $X \setminus A$ es abierto.

Consideremos ahora que $X \setminus A$ abierto, entonces ningún punto $x \in X \setminus A$ puede ser límite de una sucesión de puntos de A , puesto que la vecindad $X \setminus A \in \mathbf{V}(x)$ tiene intersección vacía con A . Por lo tanto, A contiene a todos sus puntos adherentes. ■

Observación: los conceptos conjunto abierto y conjunto cerrado no son conceptos que se excluyen mutuamente. Existen conjuntos que son a la vez abiertos y cerrados, por ejemplo el $\emptyset$ y X en todo espacio métrico (X, d) .

En un espacio métrico discreto todo subconjunto es a la vez abierto y cerrado. Estos dos conceptos tampoco son exhaustivos, existen subconjuntos $A \subset (X, d)$ que no son ni abiertos ni cerrados, como muestran los intervalos semiabiertos en $\mathbb{R}$.

El teorema anterior establece una dualidad entre los abiertos y los cerrados de un espacio pseudométrico. Por transición al complemento se deduce de todo enunciado sobre conjuntos abiertos un enunciado dual sobre los conjuntos cerrados.

Por ejemplo, partiendo de los axiomas (A1)-(A3) se obtienen las siguientes propiedades fundamentales de los cerrados.

Corolario 1 (propiedades fundamentales de los cerrados). En todo espacio métrico (X, d) se cumple para la colección τ_d de los cerrados:

$$(C1) \quad \emptyset \in \tau_d \text{ y } X \in \tau_d.$$

$$(C2) \quad C_1, C_2 \in \tau_d \implies C_1 \cup C_2 \in \tau_d.$$

$$(C3) \quad \forall \alpha \in I, (C_\alpha \in \tau_d) \implies \bigcap_{\alpha \in I} C_\alpha \in \tau_d.$$

Demostración:

(C1) $\emptyset \in \tau_d$ pues $X \setminus \emptyset = X \in \theta_d$ por (A1).

$X \in \tau_d$ pues $X \setminus X = \emptyset \in \theta_d$ por (A1).

(C2) $X \setminus C_1, X \setminus C_2$ son abiertos, luego

$X \setminus (C_1 \cup C_2) = (X \setminus C_1) \cap (X \setminus C_2)$ es abierto

y por lo tanto $C_1 \cup C_2$ es cerrado.

(C3) Como para todo $\alpha \in I$, $X \setminus C_\alpha$ es abierto, por (A3) se tiene que

$$\bigcup_{\alpha \in I} (X \setminus C_\alpha) = X \setminus (\bigcap_{\alpha \in I} C_\alpha) \in \theta_d, \text{ y por tanto } \bigcap_{\alpha \in I} C_\alpha \in \tau_d. \blacksquare$$

Partiendo de las propiedades (C1)-(C3) de los cerrados, pueden deducirse las propiedades (01)-(03) de los abiertos.

La unión de una familia cualquiera de subconjuntos cerrados en un espacio métrico no es necesariamente un conjunto cerrado. Todo subconjunto de un espacio métrico que pueda representarse como la unión de una familia numerable de subconjuntos cerrados se llam F_σ -conjunto o *conjunto* de tipo F_σ .

Ejemplo 79. De la propiedad (C3) y la proposición 2.10.2 se deduce que para todo $A \subset (X, d)$

$$\overline{A} = \bigcap \{C \in \tau_d : C \supset A\} \quad (114)$$

Teorema 2.10.4. Sean (X, d) y (Y, d') espaciosseudométricos y $C(X, Y)$ el espacio de todas las aplicaciones continuas de (X, d) en (Y, d') . Entonces $C(X, Y)$ es cerrado en el espacioseudométrico (Y^X, d_∞) , donde $d_\infty(f, g) = \sup_{x \in X} d'(f(x), g(x))$.

Es decir, si una sucesión (f_n) de funciones continuas $f_n : X \longrightarrow Y$ converge uniformemente a una función $f : X \longrightarrow Y$, entonces f es continua.

Demostración: Probemos la continuidad de f en un punto arbitrario $x \in X$. Para cada $y \in X$ y todo $n \in \mathbb{N}$, se tiene

$$\begin{aligned} d'(f(x), f(y)) &\leq d'(f(x), f_n(x)) + d'(f_n(x), f_n(y)) + \\ &+ d'(f_n(y), f(y)). \end{aligned}$$

Dado $\epsilon > 0$ y por la convergencia uniforme de f_n a f , se puede hallar un número $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$, $d_\infty(f_n, f) < \epsilon/3$, y por lo tanto,

$$\text{si } n \geq n_0 d'(f(x), f_n(x)) < \epsilon/3 \text{ para todo } x \in X. \quad (115)$$

Seleccionemos $n \geq n_0$ fijo. Por la continuidad de f_n en el punto x , existe $\delta > 0$ tal que

$$d(x, y) < \delta \Rightarrow d'(f_n(x), f_n(y)) < \epsilon/3. \quad (116)$$

Teniendo en cuenta (115) y (116) en (114) obtendremos que si $d(x, y) < \delta$ entonces

$$d'(f(x), f(y)) \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

y de la arbitrariedad de $\epsilon > 0$ se deduce la continuidad de f . ■

Ejercicios de Autocontrol

1. Sean (X, d) y (Y, d') espacios métricos, $A \subset X$ cerrado y $f : A \longrightarrow Y$ una aplicación continua. Pruebe que el grafo de f , $G_f = \{(x, f(x)) : x \in A\}$ es cerrado en el espacio métrico producto $X \times Y$ provisto de una cualquiera de las métricas canónicas. Tomando $X = Y$ y f como la función identidad concluir que la diagonal en $X \times Y$ es un conjunto cerrado.
2. Exprese cómo se caracterizan los cerrados en un subespacio métrico a partir de los cerrados en todo el espacio.
3. Sean (X, d) y (Y, d') espacios métricos, $A \subset X$ y $B \subset Y$ cerrados en X y Y respectivamente. Pruebe que $A \times B$ es cerrado en el espacio métrico producto $X \times Y$.

2.11 PUNTO INTERIOR, PUNTO EXTERIOR Y PUNTO FRONTERA

Consideremos un subconjunto arbitrario A de un espacioseudométrico (X, d) (fig. 48).

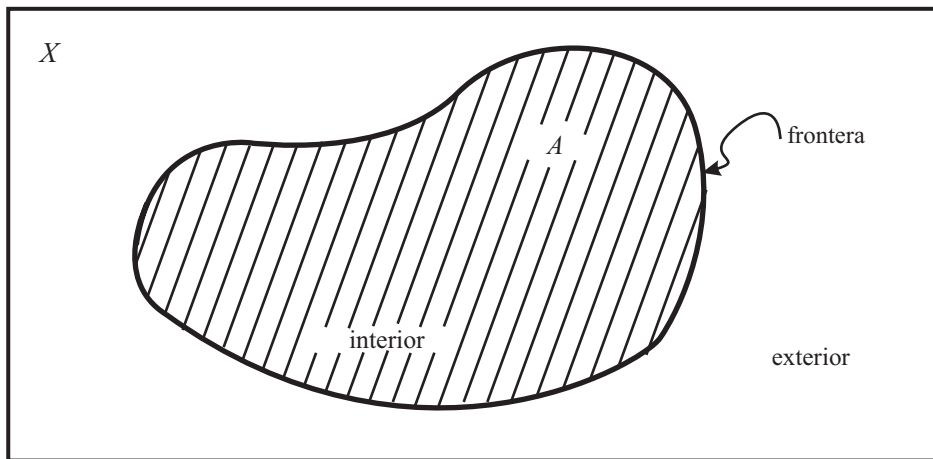


Figura 48

Intuitivamente distinguimos los puntos interiores de A , los puntos exteriores de A y los puntos frontera de A que no son interiores ni exteriores. Formalicemos estos conceptos intuitivos

Definición 2.11.1. Sea A un subconjunto de un espacio métrico (X, d) .

- i) decimos que $x \in X$ es *punto interior* de A si existe una vecindad $V \in \mathbf{V}(x)$ tal que $V \subset A$;
- ii) decimos que $x \in X$ es *punto exterior* de A si existe una vecindad $V \in \mathbf{V}(x)$ tal que $V \subset X \setminus A$;
- iii) decimos que $x \in X$ es *punto frontera* de A si no es exterior ni interior de A , o sea, si cumple:

$$V \in \mathbf{V}(x) \implies V \cap A \neq \emptyset \quad (117)$$

$$V \in \mathbf{V}(x) \implies V \cap (X \setminus A) \neq \emptyset \quad (118)$$

El conjunto de los puntos interiores de A se llama *interior* de A y se denota por " $\overset{\circ}{A}$ " o " $\text{int}A$."

El conjunto de los puntos exteriores de A se llama *exterior* de A y se denota por " $\text{ext}A$."

El conjunto de los puntos frontera se llama *frontera* de A y se denota por " ∂A " o " $\text{fr}A$ ".

Utilizando la definición de vecindad es evidente que:

$$x \in \overset{\circ}{A} \iff A \in \mathbf{V}(x); \quad (119)$$

$$x \in \text{ext}A \iff (X \setminus A) \in \mathbf{V}(x) \iff x \in \text{int}(X \setminus A). \quad (120)$$

Resulta que

$$\text{int}A = \text{ext}(X \setminus A) \quad \text{y} \quad \text{ext}A = \text{int}(X \setminus A). \quad (121)$$

Utilizando la definición de clausura obtenemos, comparando (117) y (118) con (75)

$$x \in \partial A \iff x \in \overline{A} \wedge x \in \overline{(X \setminus A)}, \quad (122)$$

$$\partial A = \partial(X \setminus A) \quad (123)$$

Según la definición 2.11.1., un punto $x \in X$ pertenece necesariamente a uno de estos tres conjuntos: $\text{int}A$, $\text{ext}A$ o $\text{fr}A$. De esto se obtiene la proposición siguiente

Proposición 2.11.2. Sea A un subconjunto de un espacioseudométrico (X, d) ; entonces la colección $\text{int}A$, $\text{ext}A$, ∂A forma una partición de X , en el sentido de que estos tres conjuntos son mutuamente disjuntos y recubren a X :

$$\text{int}A \cup \text{ext}A \cup \partial A = X \quad (124)$$

Un punto adherente de $A \subset (X, d)$ no puede pertenecer al exterior de A , porque cada una de sus vecindades intersecta a A , luego $\overline{A} \subset X \setminus \text{ext}A$. Inversamente, es trivial que todo punto interior y todo punto frontera de A son adherentes de A . Podemos concluir según la proposición anterior que

$$\overline{A} = X \setminus \text{ext}A \quad (125)$$

Corolario 1. Sea (X, d) un espacioseudométrico y A un subconjunto de X . Se tienen las siguientes igualdades:

$$X \setminus \text{int}A = \text{adh}(X \setminus A), \text{ o sea, } \text{int}A = X \setminus \overline{(X \setminus A)} \quad (126)$$

$$X \setminus \text{adh}A = \text{ext}A \text{ o sea, } \overline{A} = X \setminus \overline{\overset{\circ}{(X \setminus A)}}. \quad (127)$$

En la proposición 2.10.2 se demostró que la clausura $\overline{A}$ es el menor cerrado que contiene A . Como caracterización dual se tiene la proposición siguiente:

Proposición 2.11.3 (caracterización del interior). Sea A subconjunto de un espacio pseudométrico (X, d) . Entonces A es el mayor abierto de θ_d contenido en A .

Demostración: Ya sabemos que A es abierto. Puesto que el operador $\text{int} : A \rightsquigarrow \overset{\circ}{A}$ es creciente según la definición 2.11.1 tenemos para todo abierto U

$$U \subset A \implies U = \overset{\circ}{U} \subset \overset{\circ}{A}. \blacksquare$$

Partiendo de las propiedades fundamentales (A1)-(A3) de los abiertos y ((C1)-(C3)) de los cerrados, habríamos podido introducir por definición el interior de un conjunto A (y la adherencia de un conjunto A) como el mayor abierto contenido en A , (como el menor cerrado que contiene a A), pues estos axiomas implican que $\cup\{U : U \in \theta_d \wedge U \subset A\}$ es abierto y $\cap\{C : C \in \tau_d \wedge C \supset A\}$ es cerrado. A continuación, de la definición de interior y de adherencia, se pueden obtener las definiciones 2.9.1 y 2.11.1.

Este procedimiento, que permite destacar todavía más la dualidad conceptual entre el interior y la adherencia, es más elegante pero también muy formal y poco intuitivo. En principio no se entiende por qué es necesario introducir el cerrado (la adherencia) si sus propiedades son reflejos fieles de las propiedades ya estudiados de los abiertos, (del interior). Esto parece ser una duplicidad superflua; pero esto es solamente una consecuencia de una exposición artificial, que sacrifica el sentido intuitivo a una elegancia que, sin embargo, puede traernos confusión. El matemático no tiene interés en los cerrados por la dualidad con los abiertos, sino por la propiedad de ser cerrados con respecto a la operación de límite, y descubre sólo *a posteriori* que los cerrados son los complementos de los abiertos.

Ejemplo 80. Sea A una bola cerrada $B'(a, \rho)$ de centro a y de radio ρ en un espacio $\mathbb{R}^n$; entonces el interior de A es la bola abierta correspondiente $B(a, \rho)$; el $\text{ext}A$ es el conjunto $\{y \in \mathbb{R}^n : d(a, y) > \rho\}$ y la frontera de A es la esfera $E(a, \rho)$. (Repetimos que esto no es el caso en un espacio métrico cualquiera.)

No siempre tenemos una situación tan cerca de nuestra intuición como en el caso de una bola. Existen, por ejemplo, conjuntos sin ningún punto exterior y a la vez sin ningún punto interior, o sea, cuya frontera es todo el espacio.

Ejemplo 81. Sean (X, d) la recta real $\mathbb{R}$ con su métrica usual, $A = \mathbb{Q}$. Ningún

intervalo abierto no vacío está contenido enteramente en $\mathbb{Q}$, luego $\overset{\circ}{\mathbb{Q}} = \emptyset$. Tampoco existe un intervalo abierto no vacío enteramente incluido en $\mathbb{R} \setminus \mathbb{Q}$. Por lo tanto, $\text{ext } \mathbb{Q} = \emptyset$. Esto implica $\partial \mathbb{Q} = \mathbb{R}$ según (124).

La asociación del interior o de la adherencia a todo subconjunto A de (X, d) nos define dos operadores: $\text{int} : A \rightsquigarrow \overset{\circ}{A}$ y $\text{adh} : A \rightsquigarrow \overline{A}$ sobre el conjunto potencia $\mathbf{P}(X)$. Deduiremos a continuación, de forma dual unas propiedades fundamentales de estos operadores.

Proposición 2.11.4. Sean (X, d) un espacio pseudométrico, $A, B \subset X$; entonces:

$$\text{adh } \emptyset = \emptyset \quad (128)$$

$$A \subset \text{adh } A \quad (129)$$

$$\text{adh } (A \cup B) = \text{adh } A \cup \text{adh } B \quad (130)$$

$$\text{adh } (\text{adh } A) = \text{adh } A \quad (131)$$

$$\text{lint } X = X \quad (132)$$

$$\text{int } A \subset A \quad (133)$$

$$\text{lint } (A \cap B) = \text{int } A \cap \text{int } B \quad (134)$$

$$\text{int}(\text{int } A) = \text{int } A \quad (135)$$

Demostración: Es suficiente probar las propiedades fundamentales del interior, puesto que las propiedades correspondientes de la clausura se obtienen por transición al complemento según el corolario 1 de la proposición 2.11.2.

De la proposición 2.11.3 se implica inmediatamente las propiedades (132), (133) y (135).

Probemos que $\text{int } (A \cap B) = \text{int } A \cap \text{int } B$.

La inclusión $\subset$ es evidente. Sea inversamente x punto interior de A y de B . Entonces $\overset{\circ}{A} \cap \overset{\circ}{B}$ es una vecindad abierta de x contenida en $A \cap B$, luego $x \in \text{int}(A \cap B)$. ■

La monotonía de los operadores int y adh :

$$A \subset B \implies \text{int } A \subset \text{int } B \quad (136)$$

$$A \subset B \implies \text{adh } A \subset \text{adh } B \quad (137)$$

parece ser también una propiedad fundamental de ellos, pero no es lógicamente independiente de las propiedades enumeradas en 2.11.3. Por ejemplo, (136) se deduce de (134), puesto que:

$$A \subset B \implies A \cap B = A \implies \overset{\circ}{A} \cap \overset{\circ}{B} = \overset{\circ}{A} \implies \overset{\circ}{A} \subset \overset{\circ}{B}.$$

Ejercicios de Autocontrol

1. Dé un ejemplo de un espacio métrico (X, d) tal que

$$A \subset X \implies \partial A = \emptyset$$

2. Pruebe las siguientes caracterizaciones:

$$i) A \text{ es abierto} \iff A \cap \partial A = \emptyset$$

$$ii) A \text{ es cerrado} \iff \partial A \subset A \iff A' \subset A$$

$$iii) A \text{ es abierto y cerrado} \iff \partial A = \emptyset$$

3. Demuestre las siguientes inclusiones en un espacio métrico cualquiera:

$$\overset{\circ}{A} \subset \overset{\circ}{\overset{\circ}{A}} \subset \overset{\circ}{A} \subset \overline{\overline{\overline{A}}}$$

$$\overset{\circ}{A} \subset \overline{\overline{\overline{A}}} \subset \overline{\overline{\overline{A}}} \subset \overline{\overline{\overline{A}}}$$

Muestre con contraejemplos que en general no se tiene igualdad entre estos conjuntos.

2.12 FUNCIONES CONTINUAS, FUNCIONES ABIERTAS Y FUNCIONES CERRADAS

En 2.10 hemos visto como las propiedades topológicas en un espacio métrico pueden expresarse mediante los conjuntos cerrados y el operador clausura en lugar de los abiertos; esto se pone de manifiesto en el siguiente teorema que nos muestra una caracterización de las funciones continuas por medio de los cerrados y del operador clausura.

Teorema 2.12.1. (caracterización de las funciones continuas). Una aplicación $f : (X, d) \rightarrow (Y, d')$ es continua si y sólo si cumple una de las condiciones equivalentes siguientes:

$$i) U \in \theta_{d'} \implies f^{-1}[U] \in \theta_d$$

$$ii) C \in \tau_{d'} \implies f^{-1}[C] \in \tau_d$$

$$iii) f[\overline{A}] \subset \overline{f[A]}$$

Demostración: i) es equivalente a la continuidad de f según el teorema 2.8.8.

i) $\implies$ ii): se tiene por dualidad

$$C \in \tau_{d'} \implies X \setminus C \in \theta_{d'} \implies f^{-1}[Y \setminus C] = X \setminus f^{-1}[C] \in \theta_d \implies$$

$$f^{-1}[C] \in \tau_d.$$

ii) $\implies$ iii): $f^{-1}[\overline{f[A]}]$ es cerrado según ii) y contiene a A , luego contiene a $\overline{A}$, también por lo tanto, $f[\overline{A}] \subset \overline{f[A]}$.

iii) $\implies$ i): Sea $U \in \theta_{d'}$, y supongamos por el absurdo que $A := f^{-1}[U] \notin \theta_d$. Entonces existe $a \in A$ que es punto adherente de $X \setminus A$. Según iii), $f(a)$ es punto adherente de $f[X \setminus A] = f[X \setminus f^{-1}[U]] = f[f^{-1}[Y \setminus U]] \subset Y \setminus U$. Luego $f(a)$ sería también punto adherente de $Y \setminus U$, en contradicción con $U \in \theta_{d'}$. En fin, si se cumple iii) entonces $f^{-1}[U] \in \theta_d$. ■

Corolario 1. Sean f y g aplicaciones continuas del espacio pseudométrico (X, d) en $\mathbb{R}^n$; entonces el conjunto de coincidencia $\{x \in X : f(x) = g(x)\}$, es cerrado en (X, d) .

Demostración: Basta notar que el conjunto de puntos de coincidencia de f y g es la preimagen por la función continua $f - g$ del conjunto unitario $\{0\}$ el cual es cerrado en $\mathbb{R}^n$. ■

Corolario 2 (prolongación de igualdad). Sean f y g aplicaciones continuas del espacio métrico (X, d) en $\mathbb{R}^n$ y sea $A \subset X$ denso en (X, d) . Si $f(x) = g(x)$ para todo $x \in A$, entonces $f = g$ sobre X .

Demostración: Del corolario 1 del teorema 2.12.1 sabemos que el conjunto de los puntos de coincidencia de f y g es cerrado y como contiene al conjunto A denso en X , tiene que se necesariamente todo X .

Ejemplo 82. Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ es continua tal que $f(x + y) = f(x) + f(y)$ para todos $x, y \in \mathbb{R}$, entonces $f(x) = ax$ para algún $a \in \mathbb{R}$ fijo y todo $x \in \mathbb{R}$. En efecto, de la propiedad aditiva de f se deduce que si ponemos $f(1) = a$, entonces

$$f(n) = nf(1) = an, \text{ para todo } n \in \mathbb{R}^*.$$

Por otra parte, $f(0) = f(0 + 0) = f(0) + f(0) = 2f(0)$, de donde $f(0) = 0$.

De aquí obtenemos $0 = f(0) = f(n + (-n)) = f(n) + f(-n)$, y por lo tanto, $f(-n) = -f(n)$.

Luego, $f(n) = an$, para todo $n \in \mathbb{Z}$.

$$\text{Además, } a = f(1) = f\left(\frac{q}{q}\right) = f\left(\underbrace{\frac{1}{q} + \dots + \frac{1}{q}}_{q \text{ veces}}\right) = qf\left(\frac{1}{q}\right),$$

es decir, $f\left(\frac{1}{q}\right) = a \cdot \frac{1}{q}$, para todo $q \in \mathbb{N}^*$.

Todo número racional positivo r se pone de la forma $r = \frac{p}{q}$ con $p, q \in \mathbb{N}^*$.

Entonces,

$$f\left(\frac{p}{q}\right) = f\left(\underbrace{\frac{1}{q} + \dots + \frac{1}{q}}_{p \text{ veces}}\right) = pf\left(\frac{1}{q}\right) = a\frac{p}{q}.$$

Por otra parte,

$$0 = f(0) = f(r + (-r)) = f(r) + f(-r),$$

de donde

$$f(-r) = -f(r).$$

Con esto hemos demostrado que $f(r) = ar$ para todo $r \in \mathbb{Q}$.

Como $\mathbb{Q}$ es un subconjunto denso de $\mathbb{R}$ y las aplicaciones $f(x)$ y ax son continuas de $\mathbb{R}$ en $\mathbb{R}$, se tendrá que $f(x) = ax$, para todo $x \in \mathbb{R}$.

Observación: En el análisis clásico se definen, mediante ciertas propiedades de aditividad, las funciones elementales no racionales tales como la función exponencial y la función logaritmo. En la teoría de la medida se demuestra la existencia de funciones $f : \mathbb{R} \rightarrow \mathbb{R}$ que satisfacen $f(x + y) = f(x) + f(y)$ y que no son continuas, por lo tanto, $f(x) \neq ax$.

Los corolarios 1 y 2 del teorema 2.12.1 pueden ser generalizados de la forma siguiente:

Corolario 3. Sean (X, d) un espacio pseudométrico y (Y, d') un espacio métrico, y sea $f : (X, d) \rightarrow (Y, d')$ continua. Entonces:

- i) el conjunto de los puntos de coincidencia $\{x \in X : f(x) = g(x)\}$, es cerrado en (X, d) :
- ii) si $A \subset X$ es denso en (X, d) y $f(x) = g(x)$ para todo $x \in A$, tendremos $f \equiv g$ sobre X .

Demostración: Notemos que no se puede utilizar la misma demostración que en los corolarios anteriores ya que no tiene que haber una operación de resta en (Y, d') que permita definir la función $f - g$.

- i) Del ejercicio de control 1 de 2.10 se tiene que la diagonal Δ en $Y \times Y$ es cerrada con respecto a la métrica producto.

Si definimos la función producto

$$(f, g) : \underset{x \rightsquigarrow (f(x), g(x))}{X} \longrightarrow Y \times Y$$

tendremos que (f, g) es continua y no es difícil ver que el conjunto de puntos de coincidencia de f y g es la preimagen por (f, g) de la diagonal Δ de $Y \times Y$.

ii) Se obtiene de i) al igual que el corolario 2 se obtiene del corolario 1. ■

El teorema 2.12.1 es uno de los instrumentos más eficientes para comprobar que un conjunto dado $C \subset (X, d)$ es cerrado.

Ejemplo 83. Sea $P(x_1, \dots, x_n)$ una función polinomial sobre $\mathbb{R}^n$. El conjunto de los ceros de esta función $\{x \in \mathbb{R}^n : P(x_1, \dots, x_n) = 0\} = P^{-1}[\{0\}]$ es una superficie cerrada de $\mathbb{R}^n$ ya que $P : \mathbb{R} \longrightarrow \mathbb{R}$ es continua, luego la preimagen del conjunto cerrado $\{0\} \subset \mathbb{R}$ es cerrado. Los subconjuntos de $\mathbb{R}^n$ definidos de esta manera por polinomios en n variables se llaman *variedades algebraicas* (curvas algebraicas si $n = 2$). Por ejemplo, la hipérbola $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 - 1 = 0\}$ en $\mathbb{R}^2$.

Si bien la preimagen de un conjunto abierto o cerrado por una aplicación continua sigue siendo abierto o cerrado respectivamente, no se tiene el mismo resultado para la imagen directa.

La imagen directa de un conjunto abierto o cerrado por una función continua no es necesariamente abierto o cerrado.

Ejemplo 84. Consideremos sobre $\mathbb{R}$ la métrica discreta ρ y la métrica usual d ; entonces, la aplicación idéntica $\mathbb{I}_{\mathbb{R}} : (\mathbb{R}, \rho) \rightarrow (\mathbb{R}, d)$ es continua, pero la imagen del abierto $\{0\} \in \theta_{\rho}$ no es abierta en $(\mathbb{R}, d)$.

Definición 2.12.2. Sean (X, d) (Y, d') espacios pseudométricos. Una aplicación $f : X \rightarrow Y$ se llama *abierta* si

$$(\forall U \in \theta_d) f[U] \in \theta_{d'} \quad (138)$$

y se llama *cerrada* si

$$(\forall C \in \tau_d) f[C] \in \tau_{d'} \quad (139)$$

Es evidente que la compuesta de dos aplicaciones abiertas (cerradas) es abierta (cerrada).

Una función biyectiva $f : (X, d) \longrightarrow (Y, d')$ es abierta (cerrada) si y sólo si su inversa es continua, puesto que las imágenes directas por f son imágenes recíprocas por f^{-1} . Luego para f biyectiva se tiene:

$$f \text{ es abierta} \iff f^{-1} \text{ es continua} \iff f \text{ es cerrada} \quad (140)$$

$$\boxed{\text{Una biyección continua } f : (X, d) \longrightarrow (Y, d') \text{ es un homeomorfismo si y sólo si es abierta o cerrada.}} \quad (141)$$

Para demostrar que una aplicación dada es abierta no es necesario comprobar que todo abierto tiene una imagen abierta; es suficiente comprobarlo para los abiertos de una base topológica

Proposición 2.12.3. Sean (X, d) , (Y, d') espacios sudométricos y $\mathbf{B}$ una base topologica en (X, d) ; entonces una aplicación $f : X \longrightarrow Y$ es abierta si y sólo si cumple:

$$B \in \mathbf{B} \implies f[B] \in \theta_{d'}. \quad (142)$$

Demostración: es una consecuencia inmediata de la fórmula

$$f[\bigcup_{\alpha \in I} A_\alpha] = \bigcup_{\alpha \in I} f[A_\alpha]. \blacksquare$$

Puesto que la igualdad correspondiente para intersecciones no es válida, no tenemos un criterio análogo para las aplicaciones cerradas.

Otro criterio útil para aplicaciones abiertas es el siguiente criterio local, cuya demostración se deja al lector.

Proposición 2.12.4. Una función $f : (X, d) \rightarrow (Y, d')$ es abierta si y sólo si aplica toda vecindad $V \in \mathbf{V}(x)$ de cualquier punto $x \in X$ en una vecindad $f[V]$ de $f(x)$ en (Y, d') .

(La demostración se deja de ejercicio al lector)

Ejemplo 85. La aplicación *módulo* de $\mathbb{R}^n$ en $\mathbb{R}_+$ definida en cada $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, mediante $|x| = \left(\sum_{k=1}^n x_k^2 \right)^{1/2}$ es continua y abierta. La continuidad del módulo se obtiene inmediatamente.

Para demostrar que es abierta, basta notar que la imagen por la aplicación módulo de cualquier bola $B(x, \epsilon)$ en $\mathbb{R}^n$ es el intervalo $] |x| - \epsilon, |x| + \epsilon[\cap [0, \infty[$ que es abierto en $\mathbb{R}_+$.

Mientras que las propiedades de ser abierta y de ser cerrada son equivalentes para biyecciones (140) esto no se cumple para funciones cualesquiera, pues no tenemos, en general, que $f[X \setminus A] = Y \setminus f[A]$.

Ejemplo 86. Denotaremos por " π_1 " la proyección vertical $(x_1, x_2) \rightsquigarrow x_1$ de $\mathbb{R}^2$ en $\mathbb{R}$. π_1 transforma toda vecindad de un punto x en el plano, en una

vecindad de su proyección sobre el eje x ; por tanto es abierta. Sin embargo, π_1 no es cerrada; en efecto, si consideramos la hipérbola $H := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 = 1\}$, H es cerrado en $\mathbb{R}^2$ (4.5.2), pero $\pi_1[H] = \mathbb{R}^*$ no es cerrado en $\mathbb{R}$. (fig. 49).

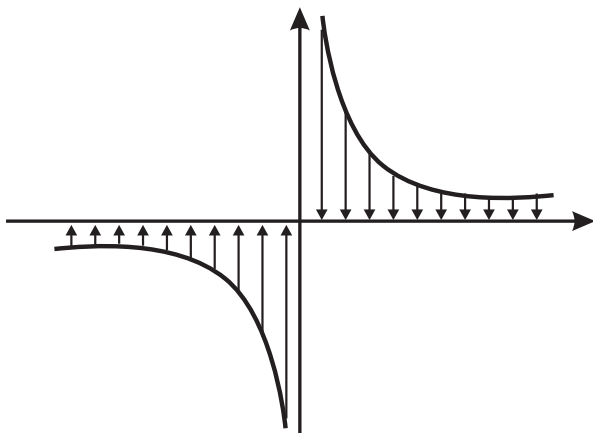


Figura 49

Inversamente una aplicación cerrada no es necesariamente abierta.

Ejemplo 87. Consideremos la función continua $f : x \mapsto \sin x$ de $[0, 2\pi]$ en $\mathbb{R}$. Del análisis elemental es conocido que f posee una inversa continua sobre los intervalos $[0, \pi/2]$, $[\pi/2, 3\pi/2]$ y $[3\pi/2, 2\pi]$. Luego, su restricción a cada uno de estos intervalos es cerrada. Si $C \subset [0, 2\pi]$ es cerrado, entonces sus intersecciones C_1, C_2, C_3 con cada intervalo, son cerrados (fig. 50).

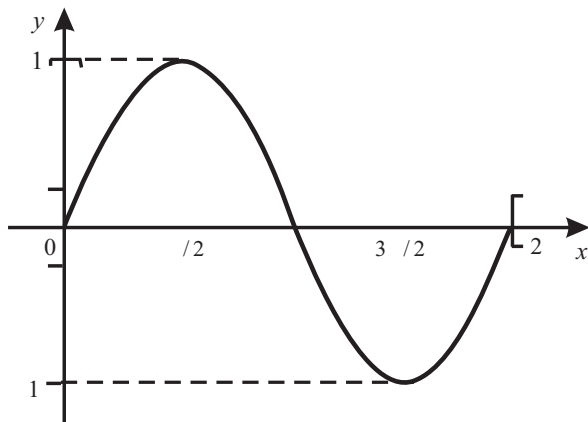


Figura 50

$f[C] = f[C_1] \cup f[C_2] \cup f[C_3]$ es cerrado para todo cerrado $C \subset [0, 2\pi]$, o sea, la aplicación f es cerrada. En cambio, f no es abierta. Por ejemplo, $A :=]0, 2\pi[$ es abierto en $[0, 2\pi]$, pero $f[A] = [-1, 1]$ no es abierto en $\mathbb{R}$.

Veamos ahora algunas formas en que se manifiesta el llamado principio del máximo para aplicaciones abiertas.

Teorema 2.12.5 (principio del máximo para aplicaciones abiertas).

Sean (X, d) un espacioseudométrico, $A \subset X$ y $f : (X, d) \longrightarrow \mathbb{R}^n$ abierta, entonces:

- i) Si A es abierto, la aplicación $x \rightsquigarrow |f(x)|$ no alcanza su supremo en A ;
- ii) Si A es abierto y f no se anula, entonces $x \rightsquigarrow |f(x)|$ no alcanza su ínfimo sobre A ;
- iii) Si A es arbitrario y $x \rightsquigarrow |f(x)|$ alcanza su máximo sobre $\overline{A}$ en un punto $a \in \overline{A}$, entonces $a \in \partial \overline{A}$.

Demostración: Del ejemplo 8.3 se deduce que $x \rightsquigarrow |f(x)|$ es una aplicación abierta de (X, d) en $\mathbb{R}_+$, luego la imagen del abierto $A \subset X$ es un abierto en $\mathbb{R}_+$ y por ello $\sup |f| [A] \notin |f| [A]$ de donde concluimos i).

Para demostrar ii) basta notar que si f no se anula entonces $|f| [A]$ es un abierto en $\mathbb{R}$, y por tanto, no contiene a su ínfimo.

La demostración de iii) se deduce de i) y del hecho que $\overline{A} = A \cup \partial A$. ■

Observación: Los principios del máximo y del mínimo para las funciones armónicas en el caso de $\mathbb{R}^2$ (partes real e imaginaria de funciones diferenciables de variable compleja), se deducen precisamente del hecho de que una función analítica no constante es una aplicación abierta.

Ejercicios de Autocontrol

1. Sean f y g funciones reales continuas sobre el espacioseudométrico (X, d) .

¿Cuáles de los siguiente conjuntos son cerrados (abiertos)?

$$\{x \in X : f(x) \leq g(x)\}$$

$$\{x \in X : f(x) \neq g(x)\}$$

$$\{x \in X : f(x) < g(x)\}.$$

2. Pruebe utillizando el teorema 2.12.1 que toda bola cerrada en un espacio métrico es un conjunto cerrado.

3. Sea $f : (X, d) \rightarrow (Y, d')$ una aplicación. Demuestre que f es abierta si y sólo si para todo $A \subset X$:

$$f[\text{int } A] \subset \text{int } f[A],$$

y que f es cerrada si y sólo si para todo $A \subset X$

$$f[\overline{A}] \supset \overline{f[A]},$$

(Observe el sentido de las inclusiones y compare con el teorema 2.12.1.)

2.13 SEPARACIÓN DE CONJUNTOS Y PROLONGACIÓN DE FUNCIONES CONTINUAS

En este epígrafe nos dedicaremos a estudiar la *separación* de puntos y conjuntos en un espacio métrico, mediante abiertos y la relación de este problema con la *prolongación* de funciones continuas definidas en un subespacio a todo el espacio. Veremos bajo qué condiciones es posible realizar la *separación* y la *prolongación* en un espacio métrico. Finalmente, estudiaremos como se relaciona todo lo anterior con la existencia de las llamadas *descomposiciones de unidad*, las cuales resultan de gran importancia en la demostración de *propiedades globales* (es decir, que se manifiestan en todo el espacio) a partir de propiedades locales (es decir, que se manifiestan en ciertas vecindades).

Para comenzar veamos cómo se relaciona la propiedad de separación de puntos con las nociones topológicas en un espacio métrico.

Resulta natural plantearnos como definición de separación entre dos puntos en un espacioseudométrico, la definición siguiente:

los puntos x y y del espacioseudométrico (X, d) están separados si la distancia entre ellos es estrictamente positiva

$$d(x, y) > 0$$

Es evidente que el punto $x \in X$ está separado de cualquier otro $y \in Y (y \neq x)$, si y sólo si el conjunto unitario $\{x\}$ es cerrado (demuéstrello). Luego, no es difícil ver que las siguientes proposiciones son equivalente en un espacioseudométrico (X, d) :

- i) d es una métrica
- ii) todo conjunto unitario de (X, d) es cerrado:
- iii) todo par de puntos diferentes en (X, d) pueden ser separados por vecindades disjuntas (fig. 51).

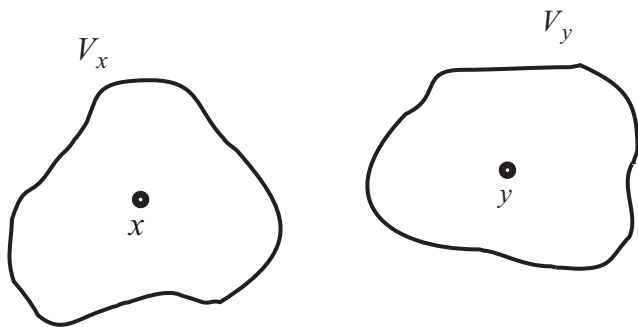


Figura 51

La propiedad iii) enunciada anteriormente es llamada *propiedad de Hausdorff* y es precisamente la propiedad topológica que asegura la unicidad del límite en los espacios métricos (teorema 2.5.2). En efecto, si se cumple la propiedad iii), x y y son elementos diferentes en X y V_x , V_y son vecindades disjuntas de x y y respectivamente, entonces ninguna sucesión (x_n) podría converger simultáneamente a x y y , ya que a partir de un cierto $n_0 \in \mathbb{N}$ todos los elementos x_n de la sucesión tendrían que estar simultáneamente en V_x y V_y y esto es imposible pues $V_x \cap V_y = \emptyset$.

El lector puede verificar que en esencia, la propiedad de Hausdorff, es la que permite demostrar el corolario 3 del teorema 2.12.1 cuya segunda parte puede ser enunciada de la forma equivalente siguiente:

Si una aplicación continua definida sobre un subespacio denso A del espacio pseudométrico (X, d) y con valores en el espacio métrico (Y, d') puede prolongarse continuamente a todo el espacio, entonces su *prolongación* está determinada de manera única.

Veamos una forma de determinar esta prolongación.

Sea $f : (A, d_A) \rightarrow (Y, d')$ continua y supongamos que existe una prolongación $\bar{f}$ de f a todo X , es decir,

$$\bar{f} : (X, d) \rightarrow (Y, d'),$$

$$\bar{f}|_A = f.$$

Consideremos el valor de $\bar{f}$ en un punto arbitrario $z \in X$. Puesto que A es denso en (X, d) existe una sucesión (x_n) en A tal que $(x_n) \xrightarrow{d} x$. Para la prolongación continua $\bar{f} : (X, d) \rightarrow (Y, d')$ tendremos necesariamente:

$$\bar{f}(z) = \lim f(x_n)$$

Como este límite no depende de la elección de la sucesión (x_n) en A siempre que $(x_n) \xrightarrow{d} x$ pondremos

$$\overline{f}(z) = \lim_{\substack{x \rightarrow z \\ x \in A}} f(x). \quad (143)$$

Luego el valor de la prolongación $\overline{f}$ en $z \in X$ puede calcularse por medio de los valores de f en los puntos de A que son *cercanos* a z .

Veremos ahora que, inversamente, mediante (143) se define una prolongación continua de f si el espacio de llegada es métrico. La demostración de este hecho se basa en la siguiente propiedad de separación en los espacios métricos cuya demostración se deja al lector.

Regularidad de los espacios métricos: En un espacio métrico cualquier conjunto cerrado puede ser separado de todo punto que no pertenece a él, por medio de vecindades abiertas disjuntas (fig. 52).

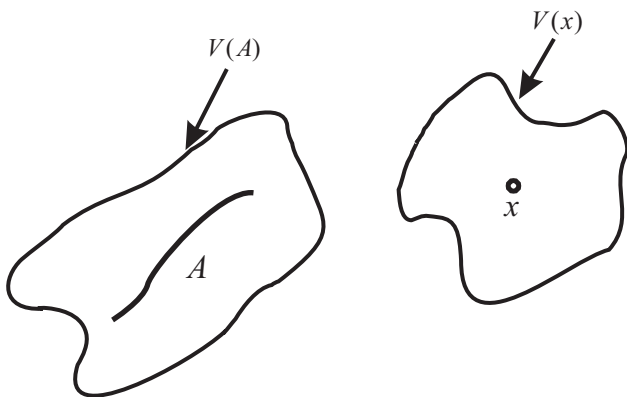


Figura 52

Teorema 2.13.1 (prolongación de funciones continuas a la clausura).

Sea A un subconjunto denso de un espacio pseudométrico (X, d) y f una aplicación continua de (A, d_A) en un espacio métrico (Y, d') . Entonces existe una única prolongación continua $\overline{f}$ de f a (X, d) si y sólo si para todo $z \in X$ existe

$$\lim_{\substack{x \rightarrow z \\ x \in A}} f(x) \text{ en } (Y, d').$$

Demostración: El razonamiento que nos condujo a la fórmula (143) muestra que la condición es necesaria.

Supongamos ahora que existe $\lim_{\substack{x \rightarrow z \\ x \in A}} f(x)$ para todo $z \in X$. Tomemos $\overline{f}(z) =$

$\lim_{\substack{x \rightarrow z \\ x \in A}} f(x)$ y probemos que $\overline{f}$ es continua.

Sea W una vecindad de $\overline{f}(z)$ en (Y, d') . Por la propiedad de regularidad de (Y, d') existe una vecindad cerrada V de $\overline{f}(z)$ tal que $V \subset W$ (por qué).

(En calidad de V se puede tomar una bola cerrada.) Según la suposición de continuidad de f , hay una vecindad abierta

$$U \in V(z) \text{ tal que } f[U \cap A] \subset V$$

(ver los teoremas 2.12.1 y 2.8.6). Es fácil observar que $\overline{U \cap A} = \overline{U}$ (demuéstrelo), luego

$$\overline{f[U]} \subset \overline{f[\overline{U}]} \subset \overline{f[\overline{U \cap A}]} \subset V.$$

Esto demuestra la continuidad de $\overline{f}$ en z . ■

Ejemplo 88 (integral de Cauchy). El teorema 2.13.1 nos muestra una condición necesaria y suficiente para que una función continua f pueda ser extendida por continuidad a la clausura. En el capítulo 6 demostraremos que toda aplicación lineal continua definida sobre un subespacio denso del espacio de partida, satisface esta propiedad, y por tanto, puede ser extendida a una aplicación lineal continua en todo el espacio. Veremos ahora cómo puede ser utilizado este resultado en el caso particular de la construcción de la integral de Cauchy.

Consideremos sobre un intervalo $[a, b] \subset \mathbb{R}$ el espacio vectorial E de todas las *funciones escalonadas*:

$$h : [a, b] \longrightarrow \mathbb{R}, h[[a, b]] = \{h_1, \dots, h_n\}$$

es finito y $h^{-1}[\{h_i\}]$ es para cada $i = 1, \dots, n$ una unión finita de subintervalos de $[a, b]$ (fig. 53)

A estas funciones se les asigna de manera natural una integral

$$\int_a^b h(x)dx = \sum_{i=1}^n h_i(x_i - x_{i-1}) \text{ con } \{h_i\} = h[[x_{i-1}, x_i]].$$

El problema de la construcción de la integral de Cauchy, la integral de Riemann o la integral de Lebesgue, consiste en generalizar esta integral *natural* a una clase más amplia de funciones. En el caso particular de la integral de Lebesgue que abarca los dos anteriores, los conjuntos $h^{-1}[\{h_i\}]$ pueden tener una estructura más compleja (conjuntos medibles) y la posibilidad de establecer un concepto de *longitud* para tales conjuntos se estudia en la llamada *Teoría de la Medida*.

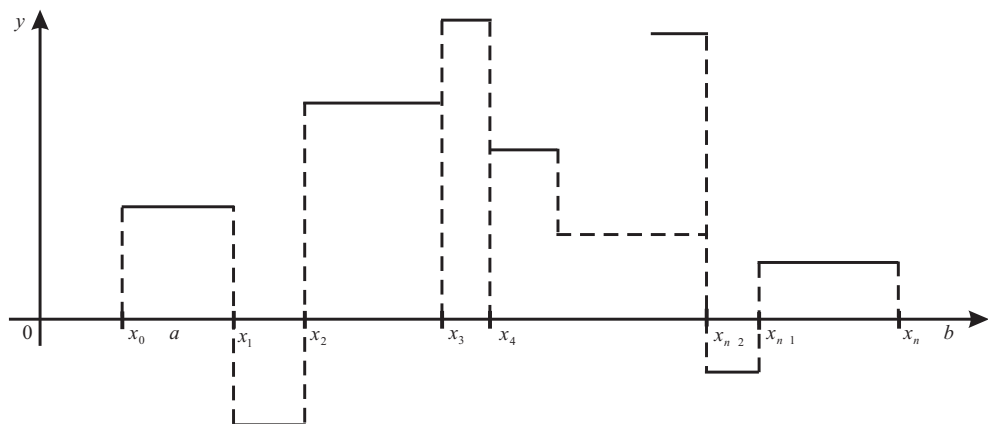


Figura 53

Consideremos aquí sólo la integral de Cauchy, cuya construcción es elemental. Sea F el espacio vectorial de todas las funciones *regladas* sobre $[a, b]$, o sea, el espacio de todas las funciones que pueden aproximarse uniformemente sobre $[a, b]$ por funciones escalonadas. El conjunto F , provisto de la métrica de la convergencia uniforme, es un espacio métrico que contiene a E como subespacio vectorial denso. No es difícil verificar que $C[a, b]$ es también un subespacio vectorial de F .

La aplicación $I : h \rightsquigarrow I(h) = \int_a^b h(x)dx$ es una forma lineal sobre E y cumple

$$|I(h) - I(g)| \leq (b - a)d_\infty(h, g) \quad (144)$$

para todo $h, g \in E$.

Para probar que $\lim_{\substack{h \rightarrow f \\ h \in E}} I(h)$ existe para todo $f \in F$, es suficiente demostrar la existencia de $\lim_{n \rightarrow \infty} I(h_n)$ para toda sucesión (h_n) de funciones escalonadas que converja a f en (F, d_∞) (teorema 2.13.1).

Pero $I(h_n)$ es una sucesión de Cauchy en $\mathbb{R}$ lo cual se comprueba tomando en (144) $h = h_n, g = h_m$. Luego, existe $\lim_{n \rightarrow \infty} I(h_n)$.

Concluimos que I posee una prolongación continua $\bar{I} : F \rightarrow \mathbb{R}$. Para comprobar que $\bar{I}$ es lineal tomemos dos funciones $f, g \in F$ y dos sucesiones (f_n) (g_n) de elementos de E que converjan uniformemente sobre $[a, b]$ a f y g respectivamente.

Entonces

$$\bar{I}(f + g) = \lim_{n \rightarrow \infty} I(f_n + g_n) = \lim_{n \rightarrow \infty} (I(f_n) + I(g_n)) =$$

$$\lim_{n \rightarrow \infty} I(f_n) + \lim_{n \rightarrow \infty} I(g_n) = \bar{I}(f) + \bar{I}(g).$$

De forma similar se demuestra que $\bar{I}(\lambda f) = \lambda \bar{I}(f)$ para todo $f \in F$ y todo $\lambda \in \mathbb{R}$.

Luego existe una prolongación lineal continua $\bar{I}$ de I a F , cuyos valores denotaremos también por $\int_a^b f(x)dx$ ($f \in F$). $\bar{I}(f)$ se llama *integral de Cauchy* de la función reglada $f : [a, b] \rightarrow \mathbb{R}$.

Como dijimos anteriormente, es fácil probar que las funciones continuas $f : [a, b] \rightarrow \mathbb{R}$ pertenecen a F , y además, que la integral de Cauchy y la integral de Riemann (definidas a través de las sumas superiores y sumas inferiores), coinciden sobre $\mathbf{C}([a, b], \mathbb{R})$.

Presentaremos a continuación algunos resultados sobre la propiedad de separación de conjuntos cerrados disjuntos en un espacio métrico y su relación con la separación por funciones continuas y el problema de prolongación de funciones continuas definidas sobre subespacios cerrados.

Comenzaremos por las definiciones siguientes:

Definición 2.13.2. (propiedad de normalidad). Se dice que el espacio métrico (X, d) satisface la propiedad de normalidad ((X, d) es normal) si todo par de conjuntos cerrados disjuntos en (X, d) pueden ser separados por vecindades abiertas disjuntas (fig. 54):

$$(C_1 \text{ y } C_2 \text{ cerrados}) \wedge (C_1 \cap C_2 = \emptyset) \Rightarrow$$

$$(\exists V_1 \in \mathbf{V}(C_1), V_2 \in \mathbf{V}(C_2)) \wedge V_1 \cap V_2 = \emptyset$$

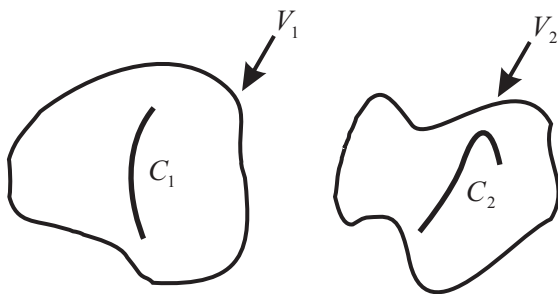


Figura 54

Veamos otra formulación equivalente de esta propiedad.

Proposición 2.13.3. El espacio métrico (X, d) cumple la propiedad de normalidad si y sólo si para todo C cerrado en (X, d) y U vecindad abierta de C , existe una vecindad V de C , tal que

$$C \subset V \subset \overline{V} \subset U \quad (145)$$

Demostración: Veamos primeramente que si (X, d) es normal, entonces se cumple (145). En efecto, si C es un cerrado en (X, d) y U es un abierto tal que $C \subset U$, entonces C y $X \setminus U$ son cerrados disjuntos. Según la propiedad de normalidad existen vecindades disjuntas $V \in \mathbf{V}(C)$ y $W \in \mathbf{V}(X \setminus U)$ que podemos suponer abiertas. Resulta que $C \subset V$ y $V \subset X \setminus W$, luego $\overline{V} \subset X \setminus W \subset U$, puesto que $X \setminus W$ es cerrado.

Probemos ahora la implicación inversa. Si C_1 y C_2 son dos cerrados disjuntos en (X, d) , entonces tenemos para $U := X \setminus C_2$, que $C_1 \subset U \in \theta_d$.

Según (145) existe una vecindad V de C_1 tal que $C \subset V \subset \overline{V} \subset U$. Concluimos que V y $X \setminus \overline{V} \supset C_2$ son vecindades disjuntas de C_1 y C_2 respectivamente. ■

Definición 2.13.4. Sea f una función real sobre un conjunto X . Se dice que f separa dos subconjuntos A y B de X , si

$$\sup f[A] < \inf f[B] \text{ o } \sup f[B] < \inf f[A]. \quad (146)$$

Intercambiando eventualmente A y B , una aplicación $f : X \longrightarrow \mathbb{R}$ separa dos subconjuntos A, B de X si y sólo si existen números reales α, β tales que se cumple:

$$x \in A \wedge y \in B \Rightarrow f(x) \leq \alpha < \beta \leq f(y). \quad (147)$$

Esto significa intuitivamente que A y B son separados por dos líneas (superficies) de nivel de f (fig. 55).

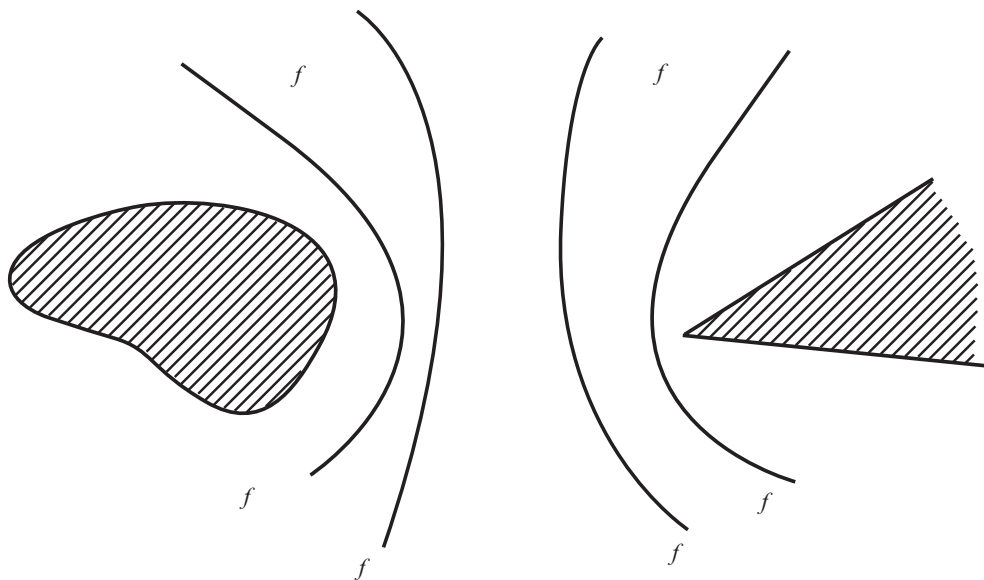


Figura 55

Nos interesaremos especialmente por la pregunta, ¿cuándo dos conjuntos cerrados disjuntos de un espacio topológico pueden separarse por una función continua? Para analizar esta cuestión, será útil considerar un tipo particular de funciones:

Definición 2.13.5. Sean A, B dos subconjuntos disjuntos de un espacio métrico (X, d) . Una función real continua f sobre (X, d) se llama *función de Urysohn* para los conjuntos A, B , si cumple:

$$f[A] = \{0\} \wedge f[B] = \{1\} \wedge f[X] \subset [0, 1]. \quad (148)$$

Está claro que toda función de Urysohn para A y B separa estos dos subconjuntos de (X, d) , pero se cumple también el inverso: toda función continua que separa A y B puede transformarse en una función de Urysohn para A y B . En efecto, si una función continua $f : (X, d) \rightarrow \mathbb{R}$ cumple (147), entonces la función continua $\tilde{f} = \max(\alpha, \min(\beta, f)) : x \mapsto \max(\alpha, \min(\beta, f(x)))$ toma el valor α sobre A y el valor constante β sobre B . Luego la función $\bar{f} = (\tilde{f} - \alpha)/(\beta - \alpha)$ es una función de Urysohn para A y B en (X, d) .

Si dos conjuntos A y B en un espacio métrico pueden separarse por una función continua, entonces pueden también separarse por una función de Urysohn para A y B .

La figura 55 nos muestra que dos conjuntos A y B en (X, d) , que pueden separarse por una función continua, también pueden separarse mediante vecindades.

Proposición 2.13.6. Sean A y B dos subconjuntos de un espacio métrico (X, d) . Si existe una función real continua sobre (X, d) que separa A y B , entonces existen dos vecindades $U_A \in \mathbf{V}(A)$ y $U_B \in \mathbf{V}(B)$ que son disjuntas.

Demostración: Supongamos que una función continua $f : (X, d) \longrightarrow \mathbb{R}$ cumple (147). Si $\gamma := \frac{1}{2}(\alpha + \beta)$ entonces $U_A := f^{-1}[-\infty, \gamma]$ y $U_B := f^{-1}[\gamma, \infty]$ son vecindades abiertas disjuntas de A y B respectivamente. ■

Según esta proposición, un espacio (X, d) es normal, si para dos cerrados disjuntos cualesquiera A y B en (X, d) existe siempre una función de Urysohn. Este criterio suficiente es a veces útil para demostrar la normalidad de un espacio. Como ejemplo, probaremos por este método que cada espacio métrico es normal.

Ejemplo 89. Sea (X, d) un espacio métrico. Si A y B son dos cerrados disjuntos en (X, d) que están separados por una distancia estrictamente positiva $\epsilon = d(A, B) > 0$, entonces es inmediato como pueden separarse mediante vecindades. $V(A, \epsilon/2) = \{x \in X : d(x, A) < \epsilon/2\}$ y $V(B, \epsilon/2) = \{x \in X : d(x, B) < \epsilon/2\}$ forman dos vecindades disjuntas de A y B respectivamente. Sin embargo, sabemos que dos cerrados disjuntos en un espacio métrico no tienen necesariamente una distancia estrictamente positiva, basta considerar $A = \{(x, 0) : x \in \mathbb{R}\}$ y $B = \{(x, 1/x) : x > 0\}$ en $\mathbb{R}^2$ (fig. 56)

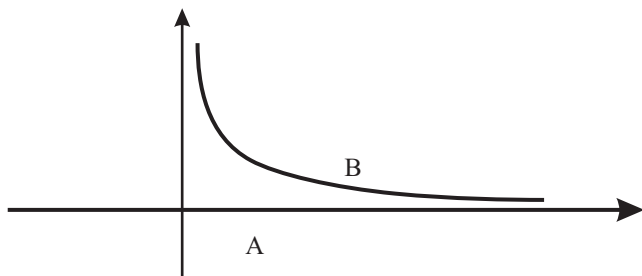


Figura 56

En este caso parece natural separar los cerrados A y B por una función continua sobre $\mathbb{R}^2$ que tenga una curva de nivel que se quede estrictamente por encima de A y por debajo de B .

Teorema 2.13.7 (Teorema de Urysohn). Todo espacio métrico satisface la propiedad de normalidad. Más exactamente, todo par de conjuntos cerrados disjuntos en un espacio métrico pueden ser separados por una función de Urysohn.

Demostración: Sean A y B dos cerrados disjuntos en el espacio métrico (X, d) . Entonces la función $f : (X, d) \longrightarrow \mathbb{R}$ definida por

$$f : x \rightsquigarrow \frac{d(x, A)}{d(x, A) + d(x, B)} \quad (149)$$

es una función real continua por ser cociente de dos funciones reales continuas (ver ejercicio 10 de autocontrol 2.6) cuyo denominador no se anula ¿por qué?. No resulta difícil verificar que f es una función de Urysohn para el par de cerrados A, B de donde se deduce la normalidad de (X, d) . ■

Según la proposición anterior en un espacio métrico la separación por funciones es equivalente a la separación por vecindades. Más exactamente, se puede sustituir en la propiedad de normalidad la separación por vecindades por la separación por funciones continuas.

Este problema tiene una gran importancia por que relaciona la separación por abiertos con la existencia de funciones reales continuas sobre el espacio de base con ciertas propiedades. Por lo tanto, es de suponer que sobre un espacio métrico podemos obtener algunos nuevos resultados referentes al prolongamiento continuo de funciones reales.

Veamos primeramente que el teorema de Urysohn puede interpretarse como un teorema de prolongación. Sean A y B dos subconjuntos cerrados disjuntos de un espacio métrico (X, d) . Definamos una función $f : A \cup B \rightarrow [0, 1]$ para la cual $f(x) = 0$ si $x \in A$ y $f(x) = 1$ si $x \in B$.

Esta función sobre el subespacio $A \cup B$ es continua, puesto que A y B son a la vez abiertos y cerrados en $A \cup B$. Las funciones de Urysohn son precisamente las prolongaciones continuas $\bar{f} : X \rightarrow [0, 1]$ de f . El teorema de Urysohn nos asegura que existen tales prolongaciones

$A \cup B$ es cerrado en (X, d) , f es una función muy especial sobre el cerrado $A \cup B$. Luego, el teorema de Urysohn es un teorema de prolongación para una clase muy especial de funciones sobre un cerrado en un espacio métrico. No obstante veremos que el mismo puede ser generalizado a la clase de todas las funciones reales continuas sobre un cerrado. Más aún, demostraremos el teorema siguiente:

Teorema 2.13.8 (Teorema de Tietze). La propiedad de normalidad de

los espacios métricos es equivalente a la siguiente:

Si $C \subset (X, d)$ es cerrado, entonces existe para cada aplicación continua f de C en un intervalo $[a, b]$ de $\mathbb{R}$, una prolongación continua $\tilde{f} : X \rightarrow [a, b]$ de f a X .

Demostración: Según el teorema de Urysohn veamos que esta propiedad implica la normalidad

Para probar la implicación inversa podemos suponer sin restricción de la generalidad que $[a, b] = [-1, 1]$.

Construiremos la prolongación continua de $f : C \rightarrow [-1, 1]$ por aproximación sucesiva de f sobre C por restricciones de funciones continuas $g_n : X \rightarrow [-1, 1]$, basandonos en el siguiente:

Lema. Sea (X, d) un espacio métrico $C \subset X$ cerrado y $f : C \rightarrow [-1, 1]$ una función continua. Entonces existe una función continua $g : X \rightarrow \mathbb{R}$ tal que:

- i) $-1/3 \leq g(x) \leq 1/3$ para $x \in X$
- ii) $|f(x) - g(x)| \leq 2/3$ para $x \in C$.

Demostración: Sean $A = f^{-1}[[1/3, 1]]$ y $B = f^{-1}[-1, -1/3]]$ dos cerrados disjuntos en C , luego lo serán en (X, d) ya que C es cerrado. Por el teorema de Urysohn existe una función $g : X \rightarrow \mathbb{R}$ tal que $g[A] = \{1/3\}$, $g[B] = \{-1/3\}$ y $-1/3 \leq g(x) \leq 1/3$ para $x \in X$, lo cual prueba i). Como $-1 \leq f(x) \leq 1$ para $x \in C$ se obtiene sucesivamente, considerando los tres casos posibles $x \in A$, $x \in B$, $x \notin A \cup B$, que $|f(x) - g(x)| \leq 2/3$ o sea ii). ■

Continuemos ahora la demostración del teorema 2.13.8

Construiremos por inducción las funciones continuas $g_n : (X, d) \rightarrow \mathbb{R}$ tales que:

- a) $-1 + (2/3)^n \leq g_n(x) \leq 1 - (2/3)^n$ para $x \in X$;
- b) $|f(x) - g_n(x)| \leq (2/3)^n$ para $x \in C$;
- c) $|g_{n+1}(x) - g_n(x)| \leq 1/3(2/3)^n$ para $x \in X$.

Para $n = 0$ definamos $g_0(x) \equiv 0$ sobre X . Supongamos ya definidas $g_0, \dots, g_n$ que cumplen las condiciones a), b), c). Pongamos:

$$f_n(x) = (3/2)^n(f(x) - g_n(x))(x \in C).$$

La función $f_n : C \rightarrow [-1, 1]$ es continua. De acuerdo con el lema anterior, existe una función continua $g : X \rightarrow \mathbb{R}$ tal que $-1/3 \leq g(x) \leq 1/3$ sobre X y $|f_n(x) - g(x)| \leq 2/3$ sobre C .

Para la función continua $g_{n+1} : X \longrightarrow \mathbb{R}$ definida por:

$$g_{n+1}(x) = g_n(x) + (2/3)^n g(x)$$

se cumplen las tres propiedades a), b), c):

- a) $|g_{n+1}(x)| \leq 1 - \left(\frac{2}{3}\right)^n + \left(\frac{2}{3}\right)^n \frac{1}{3} = 1 - \left(\frac{2}{3}\right)^{n+1};$
 b) $|f(x) - g_{n+1}(x)| = |f(x) - g_n(x) - \left(\frac{2}{3}\right)^n g(x)| =$
 $= \left(\frac{2}{3}\right)^n |f_n(x) - g(x)| \leq \left(\frac{2}{3}\right)^{n+1};$
 c) $|g_{n+1}(x) - g_n(x)| = \left(\frac{2}{3}\right)^n |g(x)| \leq \frac{1}{3} \left(\frac{2}{3}\right)^n.$

La sucesión de funciones (g_n) así definida es uniformemente convergente sobre X , porque

$$|g_n(x) - g_m(x)| \leq$$

$$\leq |g_{m+1}(x) - g_m(x)| + |g_{m+2}(x) - g_{m+1}(x)| + \dots + \leq$$

$$\leq |g_n(x) - g_{n-1}(x)| \leq 1/3(2/3)^m \sum_{i=0}^{n-m-1} (2/3)^i \leq (2/3)^m$$

para $n > m$.

Por lo tanto $\bar{f} : x \rightsquigarrow \lim_{n \rightarrow \infty} g_n(x)$, como límite uniforme de (g_n) , es continua sobre (X, d) . Según a) y b) $\bar{f}$ coincide con f sobre C y aplica X en $[-1, 1]$. ■

Sea (X, d) un espacio métrico, si f es una función real continua sobre un cerrado C tal que

$$|f(x)| \leq \alpha \text{ para todo } x \in C; \alpha > 0 \quad (150)$$

entonces existe, según el teorema 2.13.8 una prolongación continua $\bar{f}$ de f a (X, d) tal que

$$|\bar{f}(x)| \leq \alpha \text{ para todo } x \in X. \quad (150i)$$

Un resultado análogo a (150i) puede deducirse para la desigualdad estricta.

Corolario 1. Sea (X, d) un espacio métrico y $C \subset X$ cerrado. Si $f : C \longrightarrow \mathbb{R}$ es continua tal que $|f(x)| < \alpha$ para todo $x \in C$, entonces existe una prolongación continua

$\bar{f} : X \longrightarrow \mathbb{R}$ de f tal que $|\bar{f}(x)| < \alpha$ para todo $x \in X$.

Demostración: Existe una prolongación continua g de f a X tal que

$$\sup_{x \in X} |g(x)| \leq \alpha.$$

$B = \{x \in X : |g(x)| = \alpha\}$ es cerrado en (X, d) y disjunto de C . Sea h una función de Urysohn para el par $B, C : h[B] = \{0\}$ y $h[C] = \{1\}$. Tomando $f = h \cdot g$ se tiene la prolongación pedida, ya que $\bar{f}(x) = g(x) = f(x)$ sobre C y

$$|\bar{f}(x)| = \begin{cases} 0 & , \text{ si } x \in B \\ |h(x)g(x)| \leq |g(x)| < 1 & , \text{ si } x \notin B. \blacksquare \end{cases}$$

Hasta ahora hemos considerado solamente la prolongación de funciones continuas acotadas; pero es fácil ver que los resultados obtenidos se aplican también a funciones no acotadas.

Sea h el homeomorfismo $x \rightsquigarrow \begin{cases} x/(1+|x|) & \text{ si } x \in \mathbb{R} \\ \pm 1 & \text{ si } x = \pm\infty \end{cases}$

de $\overline{\mathbb{R}}$ sobre $[-1, 1]$. Si $f : \overline{\mathbb{R}} \longrightarrow \overline{\mathbb{R}}$ es una función continua, entonces $h \circ f : C \longrightarrow [-1, 1]$ es una función continua y acotada. Si g es una prolongación continua de $h \circ f$, entonces $\bar{f} = h^{-1} \circ g$ es evidentemente una prolongación continua de $h^{-1} \circ h \circ f = f$. Además, si f es real o sea si $h \circ f[C] \subset]-1, 1[$, entonces existe una prolongación g de $h \circ f$ tal que $g[C] \subset]-1, 1[$ o sea $\bar{f}[X] \subset \mathbb{R}$.

De esta forma, del teorema 2.13.8 y del corolario 1 se obtiene el corolario siguiente:

Corolario 2. Sea (X, d) un espacio métrico y $C \subset X$ cerrado; entonces

- i) Cada función numérica continua $f : C \longrightarrow \overline{\mathbb{R}}$ tiene una prolongación numérica continua $\bar{f}$ de X en $\overline{\mathbb{R}}$.
- ii) Cada función real continua $f : C \longrightarrow \mathbb{R}$ tiene una prolongación real continua $\bar{f}$ de X en $\mathbb{R}$ tal que

$$\sup \{|\bar{f}(x)| : x \in X\} = \sup \{|f(x)| : x \in C\}.$$

Uno de los problemas fundamentales de la Topología métrica, es dar respuesta a la pregunta de cuando una aplicación continua sobre un subespacio de un espacio métrico puede prolongarse continuamente a todo el espacio, manteniendo el mismo espacio de llegada

El corolario 2 nos da una respuesta satisfactoria para las funciones reales sobre un subconjunto cerrado en un espacio métrico. Hay que destacar que para espacios de llegada diferente de $\mathbb{R}$, $\overline{\mathbb{R}}$ o de intervalos reales $[a, b]$ el problema puede complicarse mucho más.

Ejemplo 90. Sea X el disco unitario $B^2 \subset \mathbb{R}^2$, $C \subset X$ el círculo unitario E^1 y Y el mismo círculo E^1 . Entonces puede probarse que la aplicación idéntica de E^1 sobre E^1 , no puede extenderse continuamente a todo el disco B^2 manteniendo E^1 como espacio de llegada. Este resultado no es fácil de obtener; es equivalente al teorema del punto fijo de Brouwer para dimensión 2. La razón intuitiva es que toda aplicación del disco B^2 sobre el círculo E^1 que coincide sobre E^1 con la aplicación idéntica, rompe necesariamente el disco, luego no es continua. Para una demostración rigurosa necesitaríamos las técnicas de la topología algebraica.

El ejemplo 88 plantea el problema de para cuáles espacios de llegada en vez de $[a, b]$ se cumple el enunciado del teorema de Tietze.

Dejamos al lector probar que en este teorema el espacio de llegada $[a, b]$ puede sustituirse por cualquier potencia de $\mathbb{R}$ o de $I = [0, 1]$

En el capítulo 6 estudiaremos espacios vectoriales provistos de una métrica *compatible* con su estructura algebraica en un sentido que precisaremos de forma muy natural. Al problema de separar dos subconjuntos cerrados disjuntos en un espacio métrico, por una función real continua, le corresponde en tales espacios vectoriales al problema lineal siguiente:

separar dos conjuntos cerrados convexos disjuntos por un funcional lineal continuo.

Este problema juega un papel fundamental en diversas teorías de Optimización y en Análisis Funcional en general. Sin embargo, veremos en el Capítulo 6 que aunque los teoremas correspondientes en el *caso lineal* se establecen principalmente gracias a razonamientos topológicos, los métodos empleados son completamente diferentes a los que se utilizan para la demostración del teorema de Tietze.

Existe una clase muy importante de espacios topológicos que tienen localmente la estructura de un espacio euclideo $\mathbb{R}^n$, o sea en dichos espacios cada punto tiene una vecindad que es homeomorfa a $\mathbb{R}^n$ (las llamadas variedades topológicas). Sobre tales espacios no es difícil definir lo que son las funciones diferenciables, puesto que la diferenciabilidad es, como la continuidad, una noción local. En cambio, la integral de una función depende del comportamiento de la función sobre todo el espacio, la integral es una noción global. Para introducir conceptos globales sobre espacios localmente euclideos, se necesitan técnicas de vincular nociones y propiedades globales con nociones y propiedades locales.

Estudiaremos a continuación uno de estos métodos estrechamente ligado con la propiedad de normalidad: *la partición de unidad*.

Empecemos por analizar los cubrimientos abiertos de un espacio métrico.

La propiedad de normalidad puede interpretarse como una propiedad que asegura que en todo cubrimiento (U_1, U_2) de X uno de sus elementos, por ejemplo U_1 , puede sustituirse por un abierto más pequeño W_1 tal que

$$\overline{W_1} \subset U_1 \quad (151)$$

De manera que el par (W_1, U_1) continua siendo un cubrimiento. En efecto, si $U_1 \cup U_2 = X$ y $U_1, U_2 \in \theta_d$ entonces $C = X \setminus U_2 \subset U_1$ es cerrado, luego existe, según (145) un abierto V tal que $C \subset V \subset \overline{V} \subset U_1$. Basta tomar $W_1 = V$, para obtener un cubrimiento abierto (W_1, U_2) de (X, d) tal que $\overline{W_1} \subset U_1$.

Aplicando este método sucesivamente a todos los abiertos de un cubrimiento abierto finito $(U_1, \dots, U_n)$ se obtiene un cubrimiento abierto “más fino” $(W_1, \dots, W_n)$ tal que $\overline{W_i} \subset U_i (i = 1, \dots, n)$.

Definición 2.13.9. Sea $(U_\alpha)_{\alpha \in I}$ un cubrimiento de un espacio métrico (X, d) . Se dice que $(U_\alpha)_{\alpha \in I}$ es *puntualmente finito*, si cada $x \in X$ está contenido solamente en un número finito de los conjuntos U_α :

$$\forall x \in X, \{\alpha \in I : x \in U_\alpha\} \text{ es finito} \quad (152)$$

El cubrimiento $(U_\alpha)_{\alpha \in I}$ es *localmente finito*, si para cada $x \in X$ existe una vecindad $v \in \mathbf{V}(x)$ que intersecta solamente a un número finito de los conjuntos U_α :

$$\forall x \in X \exists V \in \mathbf{V}(x) : \{\alpha \in I : V \cap U_\alpha \neq \emptyset\} \text{ es finito} \quad (153)$$

Definición 2.13.10. Sea $(U_\alpha)_{\alpha \in I}$ un cubrimiento de un espacio métrico (X, d) . Un cubrimiento $(W_\beta)_{\beta \in J}$ de X se dice que es más *fino* o que es un *refinamiento* de $(U_\alpha)_{\alpha \in I}$ si

$$\forall \beta \in J \exists \alpha \in I : W_\beta \subset U_\alpha. \quad (154)$$

$(W_\alpha)_{\alpha \in I}$ se llama *encogimiento* de $(U_\alpha)_{\alpha \in I}$ si

$$\overline{W_\alpha} \subset U_\alpha \text{ para todo } \alpha \in I. \quad (154i)$$

Se plantea el problema de cuáles cubrimientos abiertos poseen un encogimiento.

Teorema 2.13.11. Todo cubrimiento abierto puntualmente finito de un espacio métrico posee un encogimiento.

Demostración: Sea $(U_\alpha)_{\alpha \in I}$ un cubrimiento abierto puntualmente finito del espacio métrico (X, d) .

Sobre el conjunto G de todos los cubrimientos abiertos de (X, d) de la forma

$$M_J = (W_\beta)_{\beta \in J} \cup (U_\alpha)_{\alpha \in I \setminus J} (\overline{W_\beta} \subset U_\beta \text{ para } \beta \in J \subset I) \quad (155)$$

introducimos una relación de orden $<$ por:

$$M_J < M_{J'} \Leftrightarrow J \subset J' \wedge W_\beta = W'_\beta \text{ para } \beta \in J \quad (156)$$

Probemos que $(G, <)$ es un conjunto inductivamente ordenado. Sea $(M_{J_s})_{s \in S}$ una familia totalmente ordenada de cubrimientos en (G) . Debemos hallar una cota superior de esta familia. Sean

$$J = \bigcup_{s \in S} J_s \text{ y } M_J = (W_\beta)_{\beta \in J} \cup (U_\alpha)_{\alpha \in I \setminus J} \text{ con}$$

$$W_\beta = W_{\beta_s}^s \text{ para } \beta \in J_s$$

Esta colección de abiertos está bien definida, puesto que $W_{\beta_s}^s = W_{\beta_s'}^{s'}$ si $\beta \in J_s$, y $\beta \in J_{s'}$. Evidentemente M_J es una cota superior de $(M_{J_s})_{s \in S}$ en $(G, <)$ si logramos demostrar que es un cubrimiento de X .

Sea $x \in X$, por hipótesis tenemos $x \in U_\alpha$ sólo para un número finito de índices $\alpha \in \{\alpha_1, \dots, \alpha_n\}$. Si uno de estos subíndices α_i pertenece a $I \setminus J$, entonces se cumple $x \in U_{\alpha_i} \subset \bigcup M_J$. En cambio, si $\alpha_i \in J$ para $i = 1, \dots, n$, entonces existe J_{s_0} ($s_0 \in S$) tal que $\alpha_i \in J_{s_0}$ para $i = 1, \dots, n$ por ser $(J_s)_{s \in S}$ totalmente ordenado. Pero $M_{J_{s_0}}$ es un cubrimiento de X , luego existe un $\beta \in J_{s_0}$ con $x \in W_\beta$. Por definición, $J_{s_0} \subset J$ y $x \in W_\beta \subset \bigcup M_J$. Por lo tanto M_J recubre X .

Por el lema de Zorn existe un recubrimiento maximal M_{J_0} en $(G, <)$.

Supongamos por el absurdo que $J_0 \neq I$. Tomamos un $\alpha \in I \setminus J_0$; según el razonamiento anterior, el abierto U_α en M_{J_0} puede sustituirse por un abierto $W_{\alpha'}$ tal que $\overline{W_{\alpha'}} \subset U_\alpha$.

El recubrimiento $M_{J_0} \cup \{W_{\alpha'}\}$ así definido es posterior a M_{J_0} en $(G, <)$ lo cual implica una contradicción. Por tanto, concluimos que todo cubrimiento maximal en $(G, <)$ es un encogimiento de $(U_\alpha)_{\alpha \in I}$. ■

Notemos que el teorema anterior formula una propiedad equivalente a la normalidad.

Sea $(U_\alpha)_{\alpha \in I}$ un cubrimiento abierto localmente finito de un espacio métrico (X, d) y sea $(W_\alpha)_{\alpha \in I}$ un encogimiento de dicho cubrimiento. Según el teorema de Urysohn existe para todo $\alpha \in I$ una función continua $f_\alpha : X \rightarrow [0, 1]$ que es igual a 1 sobre $\overline{W_\alpha}$ y se anula fuera de U_α . Estas funciones nos servirán para descomponer toda función real continua sobre X en una familia de funciones que sólo localmente son diferentes de 0.

Definición 2.13.12. Sea (X, d) un espacio métrico y $f : X \rightarrow \mathbb{R}$ una aplicación. Entonces el conjunto cerrado

$$\text{sop}(f) := \overline{\{x \in X : f(x) \neq 0\}} \quad (157)$$

se llama *soporte* de f .

$\text{Sop } f$ es el cerrado más pequeño fuera del cual f se anula.

Los soportes de las funciones de Urysohn f_α construidas anteriormente no están necesariamente contenidos en U_α . En el lema siguiente se muestra que existen funciones de Urysohn para el par $\overline{W_\alpha}, X \setminus U_\alpha$ con esta propiedad:

Lema 2.13.13. Sea (X, d) un espacio métrico, $C \subset X$ cerrado y $A \in \mathbf{V}(C)$. Entonces existe una función continua $f : X \rightarrow [0, 1]$ tal que $f[C] = \{1\}$ y $\text{sop } f \subset A$.

La demostración se deja al lector (utilice (145)).

Si $(f_\alpha)_{\alpha \in I}$ es una familia de funciones reales continuas sobre (X, d) tal que la familia de sus soportes $(\text{sop } f_\alpha)_{\alpha \in I}$ es puntualmente finita, entonces la función suma

$$f = \sum_{\alpha \in I} f_\alpha : x \rightsquigarrow \sum_{\alpha \in I} f_\alpha(x) \quad (x \in X) \quad (158)$$

está bien definida, puesto que para todo $x \in X$ sólo un número finito de valores $f_\alpha(x)$ ($\alpha \in I$) es diferente de 0. Se plantea el problema de cuándo f será continua.

Lema 2.13.14. Si $(f_\alpha)_{\alpha \in I}$ es una familia de funciones reales continuas sobre (X, d) tal que la familia de sus soportes $(\text{sop } f_\alpha)_{\alpha \in I}$ es localmente finita, entonces la función suma $f = \sum_{\alpha \in I} f_\alpha : X \rightarrow \mathbb{R}$ es continua.

Demostración: Para todo $x \in X$ existe una vecindad abierta U_x de x tal que $U_x \cap \text{sop } f_\alpha \neq \emptyset$ sólo para un número finito de índices $\alpha \in I$. Pero $f|_{U_x}$ es suma finita de funciones continuas, luego es continua. No es difícil demostrar que la continuidad de f sobre cada U_x implica la continuidad de f en todo X . ■

Definición 2.13.15. Sea (X, d) un espacio métrico. Una familia $(\varphi_\alpha)_{\alpha \in I}$ de funciones continuas $\varphi_\alpha : X \rightarrow [0, 1]$ se llama *partición de unidad* sobre (X, d) si cumple:

- i) La familia $(\text{sop } \varphi_\alpha)_{\alpha \in I}$ es localmente finita;
- ii) $\sum_{\alpha \in I} \varphi_\alpha(x) = 1$ para todo $x \in X$.

Definición 2.13.16. Sea (U_α) un cubrimiento de (X, d) ; entonces una partición de unidad $(\varphi_\alpha)_{\alpha \in I}$ se llama *subordinada al cubrimiento* $(U_\alpha)_{\alpha \in I}$ si para todo $\alpha \in I$ se tiene $\text{sop } \varphi_\alpha \subset U_\alpha$.

La importancia del teorema 2.13.11 consiste principalmente en su corolario

Corolario 1. En un espacio métrico (X, d) , todo cubrimiento abierto, localmente finito, posee una partición subordinada de unidad.

Demostración: Sea $(U_\alpha)_{\alpha \in I}$ un cubrimiento abierto localmente finito de (X, d) . Existe un encogimiento $(W_\alpha)_{\alpha \in I}$ de $(U_\alpha)_{\alpha \in I}$. Sabemos además por el lema 2.13.13 que existen funciones continuas $f_\alpha : X \longrightarrow [0, 1]$ tales que $f_\alpha[\overline{W_\alpha}] = \{1\}$ y $\text{sop } f_\alpha \subset U_\alpha$. Como $(U_\alpha)_{\alpha \in I}$ es localmente finita también lo será $(\text{sop } f_\alpha)_{\alpha \in I}$, y por esta razón, la función suma $f = \sum_{\alpha \in I} f_\alpha$ es continua (2.13.14). Para $x \in X$ se tiene que $f(x) \geq 1$ ya que x tiene que estar en algún $\overline{W_\alpha}$.

Definamos ahora $\varphi_\alpha = f_\alpha/f$. La función $\varphi_\alpha : X \longrightarrow [0, 1]$ es continua y cumple $\text{sop } \varphi_\alpha = \text{sop } f_\alpha \subset U_\alpha$ ($\alpha \in I$). Por definición, tenemos $\sum_{\alpha \in I} \varphi_\alpha(x) = (1/f(x)) \cdot \sum_{\alpha \in I} f_\alpha(x) = 1$ para todo $x \in X$. Luego $(\varphi_\alpha)_{\alpha \in I}$ es una partición de unidad subordinada a $(U_\alpha)_{\alpha \in I}$. ■

Ejemplo 91. Sea (X, d) un espacio topológico y sean A y B dos cerrados disjuntos de (X, d) . Entonces $(X \setminus A, X \setminus B)$ es un cubrimiento abierto localmente finito de (X, d) . Supongamos que (φ_1, φ_2) es una partición de unidad subordinada al cubrimiento $(X \setminus A, X \setminus B)$. Entonces $\text{sop } \varphi_1 \subset (X \setminus A)$ y $\text{sop } \varphi_2 \subset (X \setminus B)$, o sea $\varphi_1[A] = \varphi_2[B] = 0$. Como $\varphi_1(x) + \varphi_2(x) = 1$ para todo $x \in X$ concluimos que $\varphi_1[A] = \varphi_2[B] = 1$. Vemos que φ_1 es una función de Urysohn para el par (A, B) y φ_2 una función de Urysohn para el par (B, A) . Esto nos muestra que el enunciado del corolario anterior para todo cubrimiento de dos abiertos (U_1, U_2) es una propiedad equivalente a la normalidad en los espacios métricos.

Observación: Como ya hemos planteado anteriormente sobre ciertos espacios métricos está definida una estructura diferenciable. Sobre tales espacios se necesitan particiones de unidad $(\varphi_\alpha)_{\alpha \in I}$ cuyas funciones no sean solamente continuas sino diferenciables. En los ejercicios que aparecen al final del libro se da una indicación de cómo construir tales particiones de unidad diferenciables

sobre $\mathbb{R}^n$.

Las particiones de unidad pueden aplicarse para obtener soluciones globales partiendo de soluciones locales. Tales problemas de *sintetización* o *pegamento* se plantean particularmente en la teoría de la llamadas *variedades diferenciales* y *variedades de Riemann*, las cuáles constituyen el objeto de la *Topología Diferencial* y la *Geometría Diferencial*. Estas variedades forman probablemente, después de los espacios funcionales, la segunda clase de espacios métricos (y en general de espacios topológicos) concretos de mayor importancia tanto en la teoría como en la práctica.

Para una mayor información al respecto remitimos al lector a los libros de M. Spivak (Calculus on Manifolds), S. Lang (Diferenciable Manifolds), y D. Hinrichsen y otros (Topología general). Vea la bibliografía al final del libro.

Ejercicios de Autocontrol

1. Sean (X, d) un espacio métrico y A y B subconjuntos cerrados de (X, d) :
 - a) Pruebe que existe una función de Urysohn para el par (A, B) tal que $A = f^{-1}[\{0\}]$ si y solo si A es un G_δ -conjunto.
 - b) Pruebe que si A y B son G_δ -conjuntos entonces existe una función de Urysohn para el par (A, B) tal que $A = f^{-1}[\{0\}]$ y $B = f^{-1}[\{1\}]$.
2. Demuestre el teorema de Tietze en los tres casos siguientes:
 - a) cuando el espacio de llegada es $[0, 1]^n$
 - b) cuando el espacio de llegada es $\mathbb{R}^n$;
 - c) cuando el espacio de llegada es $\overline{\mathbb{R}}^n$.

2.14 ALGUNOS CONCEPTOS DE CONVERGENCIA Y CONTINUIDAD NO ASOCIADOS A UNA MÉTRICA

No es nuestro objetivo ser exhaustivo en este epígrafe. Solamente mencionaremos algunos de los innumerables ejemplos que pueden ser comprensibles al lector y que muestra la ineficiencia del concepto métrica para el desarrollo de una teoría unificada de la convergencia.

Con esto no estamos negando la importancia de los espacios métricos cuyo campo de acción para el estudio de la convergencia y continuidad es suficientemente amplio y abarca los aspectos fundamentales que satisfacen nuestro

objetivo inmediato, ante todo, como concepto unificador de los temas estudiados en las asignaturas de Análisis Matemático sobre los conjuntos numéricos.

Nuestro objetivo final con este epígrafe y el próximo es dejar una “puerta abierta” al lector interesado en profundizar en estos temas y sus aplicaciones, fundamentalmente, al Análisis Funcional.

A continuación desarrollaremos previamente los siguientes temas, en los cuales se utilizarán ciertos conceptos de convergencia que no pueden ser definidos por una métrica:

- a) límite superior y límite inferior de sucesiones de números reales.
- b) semicontinuidad superior y semicontinuidad inferior de funciones reales
- c) convergencia de redes en espacios métricos (integral de Riemann)
- d) convergencia en espacios producto y convergencia puntual de funciones
- e) convergencia en espacios de funciones continuas
- f) algunos conceptos de convergencia relacionados con el desarrollo de las ecuaciones diferenciales.

a) Límite superior y límite inferior de sucesiones de números reales.

En el ejemplo 68 hemos definido estos conceptos mediante (91) y (92). No puede existir una métrica d sobre $\mathbb{R}$ para la cual se tenga:

$$\limsup (x_n) = x \Leftrightarrow (x_n) \xrightarrow{d} x$$

En efecto, cualquier subsucesión $(x_{\varphi(n)})$ de la sucesión convergente (x_n) tendría que cumplir $(x_{\varphi(n)}) \rightarrow x$ y sin embargo, no necesariamente se cumple que $\limsup x_{\varphi(n)} = x$. Basta considerar, en calidad de ejemplo, la sucesión (x_n) definida por:

$$x_{2n} = 1, x_{2n+1} = -1, n = 1, 2, \dots,$$

cuyo límite superior es 1, mientras que para la subsucesión (x_{2n+1}) el límite superior es -1 .

Teniendo en cuenta la noción de vecindad en un espacio métrico (definición 2.8.1), podríamos definir una *vecindad generalizada* V de $x \in \mathbb{R}$ como cualquier subconjunto de $\mathbb{R}$ que contenga a un intervalo $]a, b]$ tal que $a < x \leq b$.

Con esta definición no es difícil probar la siguiente afirmación:

$$\limsup (x_n) = x \Leftrightarrow \text{para cada vecindad generalizada } V \text{ de } x \text{ se puede encontrar un } N \in \mathbb{N} \text{ tal que } (x_n) \in V \text{ si } n \geq N. \quad (159)$$

Un razonamiento análogo puede hacerse para el límite inferior sustituyendo los intervalos $]a, b]$ por $[a, b[$.

b) Semicontinuidad superior y semicontinuidad inferior de funciones reales.

En Análisis se dice que una aplicación $f : \mathbb{R} \longrightarrow \mathbb{R}$ es semicontinua inferiormente (superiormente) si y solo para todo $\alpha \in \mathbb{R}$ y todo $\epsilon > 0$ existe $\delta > 0$ tal que:

$$x \in \mathbb{R} \wedge |x - \alpha| < \delta \Rightarrow f(x) > f(\alpha) - \epsilon \text{ (fig. 57a)} \quad (160)$$

$$x \in \mathbb{R} \wedge |x - \alpha| < \delta \Rightarrow f(x) < f(\alpha) + \epsilon \text{ (fig. 57b)} \quad (161)$$

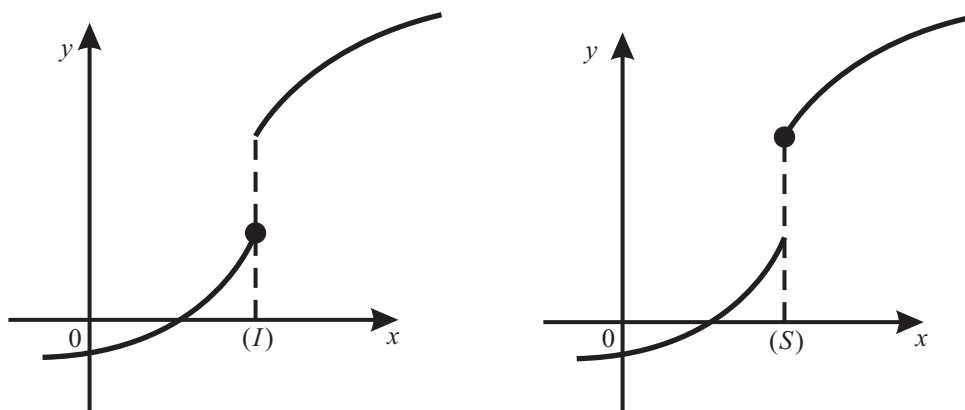


Figura 57

Trataremos de exponer esta definición en la forma de la afirmación (69)

Si $f(a)$ es estrictamente mayor que cualquier número b (que puede escribirse en la forma $b = f(a) - \epsilon$ con $\epsilon > 0$, entonces por (160) existe todo un intervalo de centro $a : \{x \in \mathbb{R} : |x - a| < \delta\}$, sobre el cual f es estrictamente mayor que b , es decir, que para todo $b \in \mathbb{R}$ la preimagen $f^{-1}(]b, \infty[)$, o sea, el conjunto de todos los $a \in \mathbb{R}$ con $f(a) > b$, es abierto en $\mathbb{R}$, pues es vecindad de todos sus puntos.

Si ahora llamamos *vecindad generalizada* de x en el espacio de llegada $\mathbb{R}$, a todo subconjunto V de $\mathbb{R}$ que contenga a un intervalo $]b, \infty[$ con $x > b$, entonces podemos enunciar la afirmación siguiente:

$$\text{La función } f : \mathbb{R} \longrightarrow \mathbb{R} \text{ es semicontinua inferiormente en el punto } a \text{ si } f^{-1}[V] \text{ es vecindad de } a \text{ con respecto a la métrica euclidea para toda vecindad generalizada } V \text{ de } f(a). \quad (162)$$

En general podemos decir que una función real f definida sobre un espacio métrico (X, d) es semicontinua inferiormente en el punto $a \in X$ si $f^{-1}[V]$ es vecindad de a en (X, d) para toda vecindad generalizada V de $f(a)$ en $\mathbb{R}$.

Un enunciado similar se obtiene para las funciones semicontinuas superiormente, sustituyendo los intervalos $]b, \infty[$, por los intervalos $]-\infty, b[$.

De acuerdo a (69), la semicontinuidad inferior correspondería a un concepto de continuidad en espacios métricos si pudiera construirse una métrica d en $\mathbb{R}$ para la cual las vecindades fueran precisamente las *vecindades generalizadas* definidas anteriormente. Obviamente, en un espacio métrico en el cual las vecindades sean precisamente estas vecindades generalizadas, los abiertos serían únicamente los intervalos del tipo $]b, \infty[$, y por lo tanto, los cerrados serían los intervalos $]-\infty, b[$.

Pero en este caso la bola abierta $B_d(a, \epsilon)$ no podría contener la bola cerrada $B'_d(a, \epsilon/2)$, lo cual es una contradicción.

c) Convergencia de redes en espacios métricos (integral de Riemann)

Una red en un conjunto X es una aplicación $\varphi : a \rightsquigarrow x_a$ de un conjunto dirigido $(I, \leq)$ en X (ver capítulo 1). Comunmente se denota como la familia $(x_\alpha)_{\alpha \in I}$.

Puesto que en un conjunto dirigido la expresión “apartir de un elemento $\alpha \in I$ ” tiene un sentido evidente, la noción de sucesión convergente puede generalizarse análogamente a redes cualesquiera en un espacio métrico general.

En efecto, si $(x_\alpha)_{\alpha \in I}$ es una red en el espacio métrico (X, d) , se dice que $(x_\alpha)_{\alpha \in I}$ converge a $x \in X$, si para toda $V \in \mathbf{V}(x)$ existe un $\alpha_\gamma \in I$ tal que:

$$x_\alpha \in V \text{ para todo } \alpha \in I \text{ con } \alpha_\gamma < \alpha$$

y se utiliza la misma notación que para una sucesión convergente $((x_\alpha) \xrightarrow[d]{} x)$.

Se puede tener un teorema sobre la caracterización de la continuidad mediante la convergencia de redes, análogo al teorema 2.6.2, sustituyendo simplemente las sucesiones convergentes por redes convergentes. Luego, en un espacio métrico la convergencia de redes no introduce ningún concepto esencialmente nuevo a diferencia de como ocurre en los espacios topológicos generales donde no puede ser sustituida por la de sucesiones.

Sin embargo, existen algunos conceptos que se definen mediante el límite de cierta red que no pueden ser definidos utilizando sucesiones. La integral de Riemann nos proporciona un ejemplo importante.

Sea $[a, b]$ un intervalo cerrado y acotado en $\mathbb{R}$ y sea (P_n) una sucesión de particiones de $[a, b]$ tal que los diámetros de los subintervalos P_n convergen a 0 (si $n \rightarrow +\infty$). Si escogemos en cada intervalo $[x_{i-1}^{(n)}, x_i^{(n)}]$ de cada partición $P_n = (x_0^{(n)}, \dots, x_{k(n)}^{(n)})$ ($n \in \mathbb{N}$) un número real $\epsilon_i^{(n)}$, entonces la integral de una función continua $f : [a, b] \rightarrow \mathbb{R}$ es precisamente el límite de la sucesión convergente (S_n) de las sumas

$$S_n := \sum_{i=1}^{k(n)} f(\epsilon_i^{(n)})(x_i^{(n)} - x_{i-1}^{(n)}).$$

No es difícil comprobar que esta definición es precisamente la integral de Cauchy para funciones continuas introducida en el ejemplo 86 y que no depende de la elección de la sucesión de particiones (P_n) siempre que los diámetros de los subintervalos de P_n convergan a 0. Sin embargo, no existe una caracterización general de la integral de Riemann por medio del límite de una sucesión, en este caso necesitamos auxiliarnos de las redes.

En efecto, el conjunto $P[a, b]$ de todas las particiones: $P = \{x_0, x_1, \dots, x_n\}$ ($x_0 = a, x_n = b$) del intervalo $[a, b]$ provisto del orden inducido por la inclusión entre conjuntos $\subset$, es un conjunto dirigido.

A toda partición P se le asocia la *suma inferior*

$$\underline{S}(f, P) := \sum_{i=1}^n \underline{f}_i(x_i - x_{i-1})$$

$$\text{con } \underline{f}_i := \inf \{f(x) : x_{i-1} < x < x_i\}$$

y la *suma superior*

$$\overline{S}(f, P) := \sum_{i=1}^n \overline{f}_i(x_i - x_{i-1}),$$

$$\text{con } \overline{f}_i := \sup \{f(x) : x_{i-1} < x < x_i\}.$$

Es conocido que las redes $\{\overline{S}(f, P) : P \in P[a, b]\}$ y $\{\underline{S}(f, P) : P \in P[a, b]\}$ son convergentes en el espacio pseudométrico $\overline{\mathbb{R}}$. Sus límites se denotan respectivamente por:

$$\overline{\int} f(x) dx \text{ (integral superior).}$$

$$\underline{\int} f(x) dx \text{ (integral inferior).}$$

Se dice que f es integrable en sentido de Riemann sobre $[a, b]$ si sus integrales superior e inferior coinciden, y en este caso, a ese valor común se le llama integral de Riemann de f en $[a, b]$ y se denota por

$$(R) \int_a^b f(x) dx.$$

d) Convergencia en espacios producto y convergencia puntual de funciones

Sabemos que el conjunto producto $\prod_{i \in I} X_i$ de los espacios métricos $\{(X_i, d_i)\}_{i \in I}$ puede ser provisto de una métrica para la cual la convergencia de sucesiones sea precisamente la convergencia por coordenadas, cuando el conjunto I es numerable (vea ejemplo 27).

Tomando $X_i = X$ para todo $i \in I$, el enunciado anterior nos afirma que la convergencia puntual para las funciones de I en X , puede ser obtenida mediante una métrica, cuando el conjunto I es numerable.

Veamos ahora que esto no ocurre cuando I no es numerable y para ello, consideremos el caso particular en que $X = \mathbb{R}$,

$$I = [0, 1], \mathbb{R}^{[0,1]} = \{f : [0, 1] \longrightarrow \mathbb{R}\}.$$

Notemos primeramente que para una métrica d sobre $\mathbb{R}^{[0,1]}$ para la cual la convergencia coincida con la convergencia puntual de funciones, se tendría que las aplicaciones proyección

$$\pi_x : (\mathbb{R}^{[0,1]}, d) \longrightarrow \mathbb{R}, x \in [0, 1]$$

son todas continuas. Luego los subconjuntos

$$\pi_x^{-1} :]a, b[, a < b; b \in \mathbb{R},$$

serían abiertos en $\mathbb{R}^{[0,1]}$ al igual que sus intersecciones finitas las cuales pueden ser expresadas en la forma

$$U = I_1 \times I_2 \times \dots \times I_n \times \mathbb{R}^{[0,1] \setminus \{x_1, \dots, x_n\}} \quad (163)$$

donde los intervalos I_k son intervalos abiertos en $\mathbb{R}$.

Teniendo en cuenta que $\mathbb{R}^{[0,1]} = \prod_{x \in I} \mathbb{R}_x$ veamos que en $(\mathbb{R}^{[0,1]}, d)$ cualquier abierto tendría que contener a alguno del tipo (163).

En efecto, si esto no fuera así habría un abierto V en $(\mathbb{R}^{[0,1]}, d)$ con la propiedad siguiente:

Existe una sucesión (x_n) en $[0, 1]$ y una sucesión (y_n) en $\mathbb{R}$ tal que $y_n \in \mathbb{R}_{x_n}$ y $y_n \notin \pi_{x_n}[V]$.

Sea ahora $f = (f(x))_{x \in [0,1]} \in V$ y consideremos una función $g : [0, 1] \longrightarrow \mathbb{R}$ tal que $g(x_n) = ny_n$ para todo $n \in \mathbb{N}$.

Obviamente la sucesión $(\bar{f}_n) = (\bar{f} + \frac{1}{n}g)$ converge en $\mathbb{R}^{[0,1]}$ puntualmente a f , y por lo tanto deberá existir un número natural N , tal que $f_n \in V$ para todo $n \geq N$, lo cual contradice la suposición de que $\frac{1}{n}g(x_n) = ny_n \notin \pi_{x_n}[V]$.

Con esto probamos que todo abierto en $(\mathbb{R}^{[0,1]}, d)$ debería contener a algún abierto de la forma (161). En particular cualquier bola $B_d(f, \epsilon)$ debería contener un conjunto U del tipo (161) tal que $U \ni f$; Entonces se tendría por una parte que

$$\{f\} = \bigcap_{n \in \mathbb{N}^*} B_d(f, \frac{1}{n}) \quad (164)$$

ya que por unicidad del límite puntual, d debería ser necesariamente una métrica y no una pseudométrica. Por otro lado

$$\bigcap_{n \in \mathbb{N}^*} B_d(f, \frac{1}{n}) \supset \bigcap_{n \in \mathbb{N}^*} U_n \quad (165)$$

donde U_n es de la forma (161) y como $f \in U_n$ para todo n , entonces $\bigcap_{n \in \mathbb{N}^*} U_n$ es un subconjunto infinito de $\mathbb{R}^{[0,1]}$.

De la contradicción entre (164) y (165) se deduce que la convergencia puntual en $\mathbb{R}^{[0,1]}$ no puede ser obtenida mediante una métrica.

e) *Convergencia en espacios de funciones continuas*

De acuerdo con nuestra motivación inicial para el estudio de los espacios métricos, en el transcurso de este capítulo hemos introducido espacios de funciones continuas provistos de diferentes métricas que corresponden a determinados tipos de convergencia usuales en el Análisis Funcional.

Como ya sabemos, la idea de considerar las funciones como puntos de un espacio métrico constituyó el punto de partida para el Análisis Funcional, así como el paso intermedio esencial para la definición abstracta de los espacios métricos y más tarde de los espacios topológicos. Esta idea surgió como resultado de un largo proceso de abstracción y aún a principios del siglo XIX solamente era conocida la noción de una función arbitraria. Los diferentes conceptos de convergencia (puntual y uniforme) no se distinguieron con precisión antes de los años 1840-50. El estudio formal de la convergencia uniforme (encabezado por la escuela italiana: Dini, Arzelá, Ascoli) comenzó en el último tercio del siglo pasado y no fue hasta finales de siglo que se utilizaron sistemáticamente métricas y topologías sobre conjunto de funciones.

Los conceptos *convergencia puntual* y *convergencia uniforme* junto con otros conceptos importantes como son la *convergencia compacta* y la *convergencia acotada*, pueden ser estudiados desde un punto de vista unificado, introduciendo la noción de σ -*convergencia* o convergencia uniforme sobre una familia σ de subconjuntos de un conjunto dado X (ver D. Hinrichsen y otros, Topología General, capítulo 11).

Constituye un hecho interesante ver que esta noción unificadora de convergencia para funciones y muy en particular para las funciones continuas, puede ser estudiada de forma general en el marco de los espacios topológicos y no de los espacios métricos ya que generalmente la σ -convergencia no puede ser descrita por una métrica.

f) *Algunos conceptos de convergencia relacionadas con el desarrollo de las ecuaciones diferenciales*

El impetuoso desarrollo de las ecuaciones diferenciales y sus aplicaciones, ha dado lugar a la introducción de espacios de funciones en los cuales la convergencia está definida de manera tal que los operadores de diferenciación sean continuos, o al menos, satisfagan determinadas propiedades de regularidad. Espacios de este tipo son, por ejemplo, los llamados espacios de funciones de

prueba que aparecen en la teoría de distribuciones (entes matemáticos que pueden ser “derivados” infinitamente en cierto sentido generalizado).

Los diferentes conceptos de convergencia definidos tanto en los espacios de funciones de prueba como en su espacio dual (el espacio de distribuciones), o para los operadores diferenciales definidos sobre estos espacios, no pueden generalmente ser definidos por una métrica.

Ejercicios de Autocontrol

1. Sea X_A la función característica de un subconjunto A de un espacio métrico (X, d) :

$$X_A(x) = \begin{cases} 1, & \text{si } x \in A \\ 0, & \text{si } x \notin A \end{cases}$$

Pruebe que X_A es semicontinua inferiormente si y solo si A es abierto.

2. Sea (X, d) un espacio métrico y sean $f_a : X \longrightarrow \mathbb{R}$ ($a \in I$) funciones semicontinuas inferiormente (superiormente)

- a) Pruebe que $\sup_{\alpha \in I} f_\alpha (\inf_{\alpha \in I} f_\alpha)$ es semicontinua inferiormente (superiormente).

En particular la envoltura superior (inferior) de funciones numéricas continuas es semicontinua inferiormente (superiormente)

- b) Demuestre que $\inf_{\alpha \in I} f_\alpha (\sup_{\alpha \in I} f_\alpha)$ es semicontinua inferiormente (superiormente), si I es finito.

- c) Dé un ejemplo de que esto no se cumple en general si I no es finito.

3. Demuestre que una aplicación $f : (X, d) \longrightarrow \mathbb{R}$ es continua (para la métrica usual sobre $\mathbb{R}$) si y solo si es superior e inferiormente continua.
4. Demuestre la afirmación (159)
5. Demuestre que una aplicación $f : (X, d) \longrightarrow (Y, d')$ es continua en el punto $a \in X$ si y solo sí, para toda red $(x_\alpha)_{\alpha \in I} \xrightarrow{d} a$ en (X, d) se tiene que $(f(x_\alpha))_{\alpha \in I} \xrightarrow{d'} f(a)$ en (Y, d')
6. Demuestre que la integral superior y la integral inferior existen para una función acotada cualquiera $f : [a, b] \longrightarrow \mathbb{R}$ (ver punto d).

2.15 TRANSICIÓN A UNA NUEVA TEORÍA

Los resultados del epígrafe anterior nos muestran el papel central de los abiertos para la topología de los espacios métricos. Si conocemos la colección de todos los abiertos de un espacio métrico (X, d) , entonces podemos determinar:

- cuáles conjuntos $V \subset X$ son vecindades de un punto $x \in X$;
- cuáles sucesiones (x_n) en X convergen a un punto $x \in X$;
- cuáles aplicaciones sobre X son continuas;
- cuáles aplicaciones sobre X son o no continuas en cierto punto $a \in X$.

Luego, para tratar problemas de convergencia y de continuidad sobre X , podemos prescindir de la métrica d , si conocemos los abiertos de X . Además esto parece más adecuado, porque continuidad y convergencia dependen esencialmente de la colección de los abiertos pero no de la métrica específica que la define: Una función f continua sobre un espacio métrico (X, d) sigue siendo continua sobre (X, d') para todas las métricas d' topológicamente equivalentes a d según el teorema 1.6.20. Esto nos sugiere axiomatizar el concepto de *conjunto abierto* sin la utilización de una distancia. Para una tal axiomatización deberíamos partir de propiedades básicas de los abiertos en un espacio métrico, y considerar conjuntos que tienen asociadas colecciones de subconjuntos que cumplen estas propiedades básicas. Con esto desaparecería para nosotros la hasta ahora imperiosa necesidad de una pseudométrica para definir las nociones de continuidad y convergencia.

Además de que una *axiomatización de los abiertos* es más adecuada a la estructura de los conceptos de continuidad y convergencia que la *axiomatización de las distancias* podemos esperar ventajas esenciales de tal procedimiento:

Primero: Como hemos visto, la teoría de los espacios métricos no comprende la convergencia puntual. Esta insuficiencia de la teoría de los espacios métricos es otro índice, de que la axiomatización del concepto de distancia todavía no nos proporciona la teoría axiomática definitiva de la continuidad y de la convergencia. Escogiendo propiedades de los abiertos que son suficientemente generales podemos construir una teoría que nos servirá para un análisis de la convergencia puntual así como de toda convergencia definida por una *familia de pseudométricas*.

Segundo: En ciertos conjuntos resulta más natural y más sencillo definir directamente los abiertos que definirlos mediante el método de introducir una distancia, a priori innecesaria. Un ejemplo claro de esto es $\mathbb{R}$ donde podemos definir directamente los abiertos como uniones arbitrarias de intervalos abiertos. Además podemos esperar que una axiomatización de los abiertos basada sobre la teoría de conjuntos en vez de la teoría compleja de los números reales

(métrica), nos permita hacer las demostraciones más claras y evitar la *ep-silóntica*.

Tercero: Otra ventaja sistemática tiene algo que ver con la construcción estructural de la matemática. La teoría de los espacios métricos presupone una teoría de los números reales. Pero $\mathbb{R}$ tiene una estructura muy compleja, que es una combinación de tres estructuras elementales: estructura algebraica, estructura de orden y estructura topológica. Dentro de la jerarquía de las estructuras que constituyen la matemática moderna no parece muy lógico confeccionar una teoría de una estructura fundamental como la topológica, presuponiendo en esta teoría una estructura tan concreta como la de los números reales. Con una axiomatización de los abiertos, la Topología General se fundamenta únicamente en la teoría de conjuntos, base común de todas las teorías estructurales.

El problema de axiomatizar el concepto de *subconjunto abierto*, es el problema de seleccionar propiedades básicas sencillas, suficientemente generales y suficientemente sustanciales, para servir como axiomas de una teoría general de la convergencia y continuidad.

El sistema de solo tres axiomas que constituyen la base de la Topología General, es uno de los logros más importante de los primeras décadas del siglo XX

En la Topología General existe un concepto de homeomorfismo, que permite establecer una relación de equivalencia entre los espacios topológicos.

Para probar que dos espacios topológicos son homeomorfos, hay que mostrar que existe un homeomorfismo entre ellos.

En general, es más difícil demostrar que no existe un tal homeomorfismo. En capítulos posteriores veremos criterios que nos permitirán asegurar que ciertos espacios métricos no son topológicamente homeomorfos.

Para dar una idea de lo que significa la homeomorfía topológica, agrupamos en las siguientes líneas (sin demostración), algunas figuras y letras que son equivalentes como espacios topológicos (fig. 58)

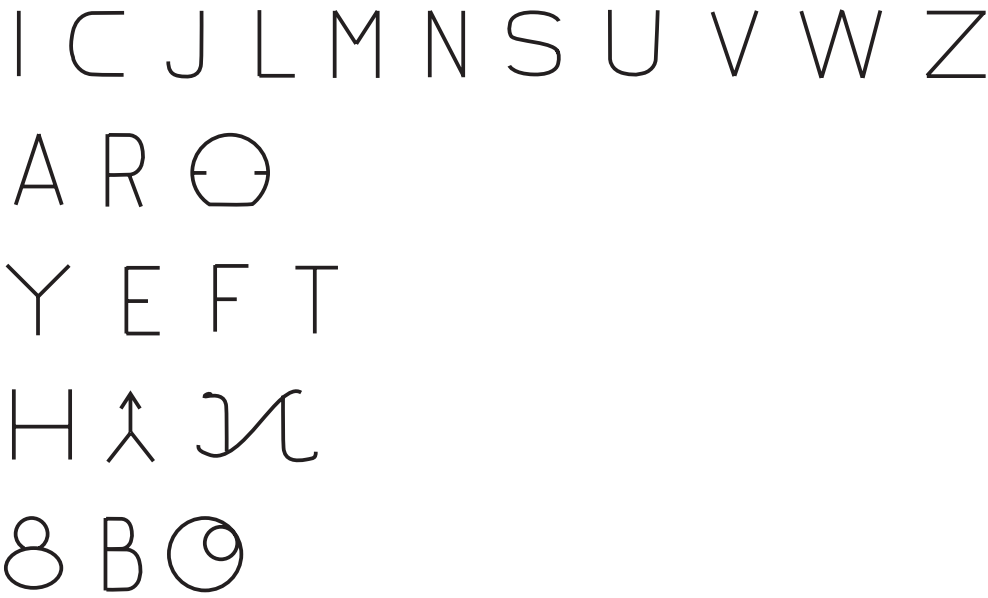


Figura 58

Intuitivamente podemos ver si curvas en el plano $\mathbb{R}^2$ o superficies en el espacio $\mathbb{R}^3$ son homeomorfas, suponiendo que son de goma.

Si se pueden transformar una en otra por una serie de extensiones y compresiones sin romperlas, ni pegarlas, son homeomorfas.

Se ha definido a un topólogo como un matemático que no encuentra la diferencia entre una rosquilla y una taza de café.

El siguiente dibujo muestra la diferencia estados en la deformación de una rosquilla en una taza (fig. 59)



Figura 59

Cada una de las geometrías clásicas tiene su concepto de configuraciones equivalentes. Para un geómetra euclideano, no es bueno el concepto de equivalencia topológica. Para él, una rosquilla y una taza de café son configuraciones completamente diferentes. En geometría euclideana se dice que dos configuraciones son equivalentes si son congruentes, o sea, si existe un movimiento rígido

que lleve una sobre otra. Sabemos que una congruencia en $\mathbb{R}^3$ está definida por una isometría de $\mathbb{R}^3$ en sí mismo, luego a un geómetra le interesan, más que las propiedades topológicas de los subconjuntos de $\mathbb{R}^3$, aquellas que son específicamente métricas.

Las propiedades de los espacios topológicos que pueden definirse por medio de los abiertos son invariantes con respecto a homeomorfismos. Si se cumplen para un espacio topológico, se cumplen también para todos los espacios homeomorfos. Tales propiedades que pueden considerarse como propiedades de las clases de espacios topológicamente equivalentes, se llaman *propiedades topológicas*.

Las propiedades topológicas de los objetos matemáticos forman el tema de la Topología General.

Algunos autores dan la siguiente definición breve y abstracta de la topología como disciplina matemática: *La tarea de la topología es la investigación de la categoría de los espacios topológicos*. Eso nos parece una buena definición del objeto general de la topología pero el objeto general no determina la tarea de una disciplina. Según nuestra opinión, la categoría de los espacios topológicos no tiene un interés en sí mismo, es decir, los topólogos no deben encerrarse con esta categoría para investigarla con un interés estructuralista.

Los espacios topológicos son una abstracción artificial por la cual se aísla un aspecto fundamental de las estructuras concretas de la realidad, para hacer una teoría coherente y eficiente de este aspecto, y producir conocimientos instrumentales para el análisis concreto. En la aplicación (que es la máxima justificación final para el desarrollo de tal teoría abstracta), los razonamientos topológicos vuelven a combinarse con razonamientos algebraicos, estadísticos, etcétera, para adaptarse a la estructura compleja de los problemas reales (tecnológicos, económicos, etcétera). Luego, es necesaria una vinculación estrecha entre la topología general y las demás disciplinas matemáticas.

Uno de los objetivos fundamentales de los siguientes capítulos es el estudio de los invariantes topológicos en los espacios métricos (compacidad, conexidad, etcétera). No expondremos una teoría autosuficiente de la categoría de los espacios métricos, sino estudiaremos los conceptos y resultados básicos que se utilizan en el Análisis, Análisis Funcional y otras ramas de la matemática cercana a la aplicación.

Sin embargo, continuaremos en el siguiente capítulo con el tema de la completitud en los espacios métricos el cual no puede formularse en un espacio topológico cualquiera y cuyos resultados son una herramienta muy potente en el desarrollo del Análisis y sus aplicaciones.

CAPÍTULO 3

ESPACIOS MÉTRICOS COMPLETOS

La veracidad de muchos teoremas importantes del Análisis, depende en gran medida, de la completitud de los espacios en que se desarrolla un problema determinado. Precisamente esto explica la insuficiencia de los números racionales y de la integral de Riemann (si hablamos solamente de los ejemplos más conocidos) así como el éxito que se alcanza al ser sustituidos por los números reales y la integral de Lebesgue respectivamente (véase 3.5).

En este capítulo nos dedicaremos al estudio de la teoría de los espacios métricos completos, procediendo en dos direcciones. Por una parte, veremos cómo una métrica permite expresar de forma intuitiva no solamente el concepto *cercanía* entre puntos y conjuntos, si no además el concepto *cercanía uniforme*, gracias al cual podemos introducir las nociones de *continuidad uniforme* y *sucesión de Cauchy* en un espacio métrico. No obstante, trataremos de indicar porqué estas nociones son propias de una estructura superior a la estructura métrica aunque tampoco son nociones topológicas ya que no son invariantes ante métricas topológicamente equivalentes.

Este capítulo tiene por tanto una tendencia hacia la generalización sistemática de ciertos conceptos, y resultados básicos que aparecen en la teoría de los espacios métricos pero que no corresponden propiamente a la estructura métrica.

Por otra parte expondremos algunos resultados más específicos sobre los espacios métricos completos y algunos teoremas importantes en estos espacios, con vista a la aplicación en otras ramas de la matemática. La mayoría de estos resultados podrían generalizarse a otros espacios topológicos más generales, pero esta temática se aparta del objetivo del texto y además, con nuestra exposición llegaremos directamente a los resultados en su forma más útil para las aplicaciones.

Presentaremos los conceptos y resultados más fundamentales que generalizan los ya conocidos por el lector para sucesiones de Cauchy y funciones uniformemente continuas en espacios numéricos, y que además, pueden ser aplicados al estudio de los espacios funcionales.

En 3.1 se presentan los resultados básicos sobre las sucesiones de Cauchy y las aplicaciones uniformemente continuas, explicandose sus propiedades fundamentales y demostrandose mediante ejemplos que estos conceptos no per-

manecen invariantes al cambiar la métrica del espacio por otra topológicamente equivalente. Motivado por los resultados obtenidos en este epígrafe, se definen en 3.2 los espacios métricos completos y se estudian algunas de sus propiedades.

Basándose en estos resultados generales se deducen una serie de teoremas importantes sobre los espacios métricos completos (3.3), y el completamiento de un espacio métrico (3.4). Estos resultados son básicos para las teorías de los espacios de Hilbert y espacios de Banach que veremos en el capítulo 6. Para ilustrar su utilidad en diversas ramas de la matemática, el último epígrafe (3.5) expondrá algunas de las aplicaciones en Análisis, Análisis Numérico. Teoría de Ecuaciones Diferenciales e Integrales y en Teoría de Variable Compleja.

3.1 SUCESIONES DE CAUCHY Y FUNCIONES UNIFORMEMENTE CONTINUAS

Del análisis conocemos dos conceptos fundamentales aparentemente topológicos y que aún no hemos generalizado para los espacios métricos, estos conceptos son *sucesión de Cauchy* y *continuidad uniforme*. Ambas nociones tienen una generalización obvia a los espacios métricos.

Definición 3.1.1 (sucesión fundamental o de Cauchy). Una sucesión (x_n) en un espacioseudométrico (X, d) se llama de *Cauchy*, si para todo $\epsilon > 0$ existe un número $N \in \mathbb{N}$ tal que

$$m, n \geq N \Rightarrow d(x_m, x_n) < \epsilon \quad (1)$$

Definición 3.1.2 (aplicación uniforme continua). Una aplicación f de un espacioseudométrico (X, d) en otro espacioseudométrico (Y, d') se llama *uniformemente continua* si para todo $\epsilon > 0$ existe $\delta > 0$ tal que para todo $x, y \in X$:

$$d(x, y) < \delta \Rightarrow d'(f(x), f(y)) < \epsilon \quad (2)$$

Observación: Notemos que en la definición anterior a diferencia de la definición de continuidad, el número δ depende de ϵ pero no del punto $x \in X$.

Veamos los siguientes ejemplos y propiedades elementales de los conceptos introducidos

Ejemplo 1. Sean (X, d) un espacioseudométrico y $A \subset X$. Hemos probado en el capítulo anterior que $|d(x, A) - d(y, A)| \leq d(x, y)$, por lo cual la aplicación $x \rightsquigarrow d(x, A)$ es uniformemente continua.

Proposición 3.1.3. Sean (X, d) y (Y, d') espacios pseudométricos. Toda aplicación uniformemente continua $f : (X, d) \longrightarrow (Y, d')$ es además continua.

La demostración es obvia por lo cual se deja al lector.

Ejemplo 2 (función continua que no es uniformemente continua). Consideremos la función continua

$$\begin{aligned}]0, 1] &\longrightarrow R \\ x &\rightsquigarrow \frac{1}{x} \end{aligned}$$

Para probar que ella no es uniforme continua, debemos encontrar un número $\epsilon > 0$ tal que para cualquier $\delta > 0$ existan $x, y \in]0, 1]$ con $|x - y| < \delta$ y

$$\left| \frac{1}{x} - \frac{1}{y} \right| > \epsilon.$$

Tomemos $\epsilon = 1$, $x = \frac{1}{n}$, $y = \frac{1}{2n}$;

entonces, para cualquier δ se puede tomar n suficientemente grande para que $x, y \in]0, 1]$ y $|x - y| = \frac{1}{2n} < \delta$. Pero por otra parte,

$$\left| \frac{1}{x} - \frac{1}{y} \right| = n > 1 = \epsilon.$$

Observación: En el siguiente capítulo estudiaremos los llamados espacios métricos compactos para los cuales se cumple que toda función continua definida sobre ellos es también uniformemente continua.

La formulación de las nociones de sucesión de Cauchy y continuidad uniforme en los espacios métricos, presupone que se tiene un instrumento (en este caso, la métrica) que permite expresar la noción intuitiva de cercanía uniforme entre pares de puntos cualesquiera.

Las vecindades $V \in \mathbf{V}(a)$ y $V \in \mathbf{V}(A)$ en un espacio métrico, nos determinan la cercanía de un punto x cualquiera al punto prefijado a o al conjunto prefijado A . Aunque no lo veremos en este libro, podemos señalar que en un espacio topológico general no se pueden comparar las vecindades de diferentes puntos, ya que no se tiene un criterio para determinar el conjunto de todos los pares de puntos (x, y) que tienen un cierto orden de cercanía. A pesar de que gracias a las propiedades de la métrica esta comparación puede realizarse en un espacio métrico, mostraremos mediante un ejemplo que la propiedad de Cauchy no es una propiedad topológica de las sucesiones en un espacio métrico, es decir, no es invariante al cambiar la métrica por otra topológicamente equivalente, sino que depende en cierta medida de una estructura adicional

que introduce la métrica.

Ejemplo 3. Consideremos sobre $\mathbb{R}$ las dos distancias

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

$$d'(x, y) = \left| \frac{x}{1+x} - \frac{y}{1+y} \right|.$$

No es difícil comprobar que d y d' son topológicamente equivalentes, o sea definen los mismos abiertos.

La sucesión (n) es una sucesión de Cauchy en $(\mathbb{R}, d')$, pero claramente no es una sucesión de Cauchy en $(\mathbb{R}, d)$ (demuéstrelo).

Notemos que para demostrar que una sucesión (x_n) no es de Cauchy en (X, d) , debe poder hallarse un número $\epsilon > 0$, tal que para cualquier $N \in \mathbb{N}$ se tengan elementos x_n y x_m con $n, m > N$ y que cumplan: $d(x_n, x_m) > \epsilon$.

Observación: No resulta difícil verificar mediante un ejemplo, que la noción de continuidad uniforme tampoco es una propiedad topológica. Es decir, si $f : (X, d_1) \longrightarrow (Y, d_2)$ es uniformemente continua, no tiene por qué continuar siéndolo al cambiar d_1 o d_2 por una métrica topológicamente equivalente (construya un ejemplo). Sin embargo, tanto la propiedad de Cauchy como la continuidad uniforme permanecen invariantes al sustituir una métrica por otra semejante.

Tal parece que al pasar por la abstracción de los espacios métricos o por la abstracción topológica, se ha saltado una estructura intermedia; esta es precisamente la llamada *estructura uniforme de un conjunto* y es a partir de ella, que podemos obtener una axiomatización conjuntista del concepto de cercanía uniforme.

El orden de cercanía entre dos puntos $x_0, y_0 \in X$ puede ser definida mediante una pseudométrica d por un número real:

$$d(x_0, y_0) < \epsilon$$

La misma relación puede expresarse utilizando el lenguaje de la teoría de conjuntos:

$$(x_0, y_0) \in U_\epsilon := \{(x, y) \in X \times X : d(x, y) < \epsilon\} \quad (3)$$

y diremos que el par de puntos $x_0, y_0 \in X$ están cercanos en un orden ϵ si $(x_0, y_0) \in U_\epsilon$.

Si conocemos la colección:

$$B = \{U_\epsilon : \epsilon > 0\} \quad (4)$$

podemos olvidar que sus elementos se definieron por medio de una distancia y tomar los propios conjuntos $U \in B$ para indicar el orden de cercanía entre dos puntos planteando que:

$$x, y \text{ son vecinos de orden } U \Rightarrow (x, y) \in U \quad (5)$$

Nosotros no analizaremos aquí las propiedades básicas de la colección B y mucho menos nos detendremos en la axiomatización y estudio de los llamados *espacios uniformes*. El lector interesado, puede consultar el capítulo 10 del libro Topología general de D. Hinrichsen y otros.

El objetivo de nuestra explicación ha sido solamente transmitir al lector una inquietud que le estimule a profundizar en esta temática, propia de una estructura superior a la métrica, a pesar de que nosotros la desarrollaremos en los espacios métricos donde obtendremos una cantidad considerable de resultados importantes. Una vez aclarado esto, continuaremos con el estudio de los conceptos *sucesión de Cauchy y aplicación uniformemente continua* en los espaciosseudométricos.

Proposición 3.1.4 (propiedades de las sucesiones de Cauchy). (X, d) un espacio pseudométrico; entonces:

- i) toda sucesión convergente de (X, d) es de Cauchy;
- ii) si la sucesión de Cauchy (x_n) en (X, d) tiene una subsucesión convergente $(x_{\phi(n)})$ $((x_{\phi(n)}) \rightarrow x)$, entonces la propia sucesión (x_n) es convergente, y además, $(x_n) \rightarrow x$.

Demostración: i) Sea (x_n) una sucesión convergente en $(X, d) : (x_n) \rightarrow x$, entonces:

$$d(x_n, x_m) \leq d(x_n, x) + d(x_m, x)$$

Para cada $\epsilon > 0$ se puede encontrar $n_0, m_0 \in \mathbb{N}$, tales que:

$$n \geq n_0 \Rightarrow d(x_n, x) < \epsilon/2, \quad m \geq m_0 \Rightarrow d(x_m, x) < \epsilon/2,$$

y por tanto,

$n \geq N = \max(n_0, m_0) \Rightarrow d(x_n, x_m) < \epsilon$, con lo cual queda probado que la sucesión (x_n) es de Cauchy.

ii) Sea $(x_{\phi(n)} \rightarrow x)$. Entonces: $d(x_n, x) \leq d(x_n, x_{\phi(n)}) + d(x_{\phi(n)}, x)$.

Sea $\epsilon > 0$. Por ser $(x_{\phi(n)})$ convergente, se puede encontrar $n_0 \in \mathbb{N}$ tal que: $n \geq n_0 \Rightarrow d(x_{\phi(n)}, x) < \epsilon/2$.

Por ser (x_n) de Cauchy existe $n_1 \in \mathbb{N}$, tal que:

$$\varphi(n) \geq n \geq n_1 \Rightarrow d(x_n, x_{\phi(n)}) < \epsilon/2,$$

luego $n \geq N = \max(n_0, n_1) \Rightarrow d(x_n, x) < \epsilon$ y de la arbitrariedad de ϵ se deduce la convergencia de (x_n) a x . ■

Observación: Como nos muestra la proposición anterior, toda sucesión de Cauchy converge a uno de sus puntos adherentes y por ello, en un espacio métrico donde el límite es único, una sucesión de Cauchy puede tener a lo sumo un punto adherente que será necesariamente su límite.

Ejemplo 4 (sucesión de Cauchy no convergente). La proposición anterior nos sugiere la existencia de sucesiones de Cauchy en un espacio métrico que no son convergentes. En efecto, consideraremos el intervalo $]0, 1]$ con la métrica usual de $\mathbb{R}$ (módulo). La sucesión $(\frac{1}{n})$ es una sucesión de Cauchy en $]0, 1]$ (compruébelo) pero no es convergente, ya que su límite sería el punto cero que no pertenece al espacio métrico en consideración.

Ejercicios de Autocontrol

1. Demuestre que los conceptos sucesión de *Cauchy* y *función uniformemente continua* permanecen invariantes al ser sustituidas las métricas correspondientes por otras semejantes.
2. Sean (X, d) un espacio métrico y B la familia de subconjuntos de $X \times X$ definida en (4). Dados $U, V \subset X \times X$, definamos

$$V \circ V := \{(x, y) \in X \times X : \exists z \in X : (x, z) \in V\}$$

$$U^{-1} := \{(x, y) \in X \times X : (y, x) \in U\}$$

Pruebe que se cumple

- i) $\Delta \subset U$ para todo $U \in B$ (Δ : diagonal de $X \times X$)
 - ii) Para todo $U \in B$ existe un $V \in B$ tal que $V \subset U^{-1}$
 - iii) Para todo $U \in B$ existe un $V \in B$ tal que $V \circ V \subset U$
3. Demuestre que toda función uniformemente continua $f : (X, d) \longrightarrow (Y, d')$ transforma sucesiones de Cauchy en X , en sucesiones de Cauchy en Y . Construya una función continua que no satisfaga la propiedad anterior.
 4. Demuestre que ninguna función polinomial de grado $n > 1$ $p : x \rightsquigarrow p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ de $\mathbb{R}$ en $\mathbb{R}$ es uniformemente continua. Más aún, pruebe que una función derivable es uniformemente continua si y solo si su derivada $f' : \mathbb{R} \longrightarrow \mathbb{R}$ es acotada.

5. Para cada par A y B de subconjuntos cerrados disjuntos de $\mathbb{R}$ se define la función

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}$$

Encuentre dos subconjuntos A y B con las propiedades antes señaladas para los cuales f no sea uniformemente continua.

3.2 DEFINICIÓN Y PROPIEDADES GENERALES DE LOS ESPACIOS MÉTRICOS COMPLETOS

Para evitar complicaciones técnicas que no aportan nada conceptual, a partir de ahora y hasta el final de este capítulo consideraremos que los espacios en cuestión son métricos y no pseudométricos.

Motivados por el ejemplo 4 introduciremos los llamados espacios métricos completos, los cuales han resultado de gran utilidad a las aplicaciones en diferentes ramas de la matemática.

Definición 3.2.1. Un espacio métrico (X, d) es *completo*, si toda sucesión de Cauchy en (X, d) es convergente en (X, d) .

La mayoría de los espacio métricos completos utilizados en la práctica, poseen generalmente una estructura vectorial. De especial importancia son la clase de los espacios normados completos llamados *Espacios de Banach* y los espacios prehilbertianos completos, llamados *Espacios de Hilbert*, los cuales estudiaremos en el capítulo 6.

Ejemplo 5. El llamado criterio de Cauchy introducido en el Análisis Matemático para asegurar la convergencia de sucesiones en $\mathbb{R}^n$, no es más que la confirmación de la completitud del espacio $\mathbb{R}^n$.

Ejemplo 6. El espacio $l_2(\mathbb{N})$ de las sucesiones (x_n) de números reales tales que $\sum_{n=0}^{\infty} |x_n|^2 < \infty$ provisto de la métrica

$$d((x_n), (y_n)) = \left(\sum_{n=0}^{\infty} |x_n - y_n|^2 \right)^{1/2}$$

es completo. En efecto, sea $(x^{(m)}) = (x_n^{(m)})$ una sucesión de Cauchy en $l_2(\mathbb{N})$. Esto significa que para cada $\epsilon > 0$ existe un m_0 tal que:

$$p, q \geq m_0 \Rightarrow \sum_{n=0}^{\infty} \left| x_n^{(p)} - x_n^{(q)} \right|^2 \leq \epsilon$$

Para cada n fijo esto implica en primer lugar que la sucesión $(x_n^{(m)})$ es de Cauchy en $\mathbb{R}$, y por tanto, converge hacia un límite x_n .

En segundo lugar, dejando $p \geq m_0$ fijo y haciendo tender q a ∞ en $\sum_{n=0}^N |x_n^{(p)} - x_n^{(q)}|^2$ ($N \in \mathbb{N}$) obtenemos por la continuidad del módulo que $\sum_{n=0}^N |x_n^{(p)} - x_n|^2 \leq \epsilon$ para todo $N \in \mathbb{N}$, luego $\sum_{n=0}^{\infty} |x_n^{(p)} - x_n|^2 \leq \epsilon$ para $p \geq m_0$.

Esto demuestra que la sucesión $(x_n^{(p)} - x_n)$ pertenece a $l_2(\mathbb{N})$, por tanto, $x = (x_n)$ pertenece a $l_2(\mathbb{N})$ y se tiene que $x^{(m)} \rightarrow x$ en $l_2(\mathbb{N})$.

Ejemplo 7. Dado un conjunto A y un espacio métrico (X, d) , el espacio

$$B(A, X) := \{f : A \rightarrow X \text{ acotada}\} \quad (6)$$

provisto de la métrica de la convergencia uniforme.

$$d_{\infty}(f, g) = \sup_{x \in A} d(f(x), g(x))$$

es completo si (X, d) es completo. Supongamos que (X, d) es completo y sea (f_n) una sucesión de Cauchy en $B(A, X)$. Como la convergencia uniforme implica la convergencia puntual, es claro que para cada $x \in A$ la sucesión $(f_n(x))$ será de Cauchy en (X, d) , y por lo tanto, converge a un elemento de X que denotaremos por $f(x)$. De esta forma hemos definido la función

$$f : A \longrightarrow X$$

tal que $f_n(x) \xrightarrow{d} f(x)$ para cada x , es decir, f_n converge a f puntualmente. Demostraremos que f_n converge a f uniformemente.

En efecto, por ser (f_n) de Cauchy en $B(A, X)$, dado $\epsilon > 0$ existe n_0 , tal que si $m, n \geq n_0$ entonces

$$d(f_m(x), f_n(x)) \leq \epsilon \quad \forall x \in A \quad (7)$$

Dejando n fijo y haciendo tender $m \rightarrow \infty$ en (7) se obtiene para $n \geq n_0$:

$$d(f(x), f_n(x)) \leq \epsilon \quad \forall x \in A \quad (8)$$

Concluimos que (f_n) converge a f uniformemente sobre A y que $f \in B(A, X)$.

Sean (X, d) y (Y, d') espacios métricos. Denotaremos mediante $C(X, Y)$ al espacio de las funciones continuas y acotadas de X y Y , provisto de la métrica

de la convergencia uniforme.

Teorema 3.2.2 (completitud de $C_a(X, Y)$). Si el espacio métrico (Y, d') es completo, entonces $C_a(X, Y)$ también es completo

Demostración: Sea (f_n) una sucesión de Cauchy en $C_a(X, Y)$. La sucesión (f_n) es, por lo tanto, de Cauchy en $B(X, Y)$, y aplicando el resultado del ejemplo 7 convergerá uniformemente sobre X a una función $f \in B(X, Y)$. En el capítulo anterior hemos demostrado que el límite uniforme de una sucesión de funciones continuas es también una función continua, y por lo tanto, $f \in C_a(X, Y)$. Luego $C_a(X, Y)$ es completo. ■

Veamos ejemplos importantes de espacios métricos no completos.

Ejemplo 8. Denotemos por $C_1[0, 1]$ ($C_2[0, 1]$) al espacio $C([0, 1], \mathbb{R})$ provisto de la métrica:

$$d_1(f, g) = \int_0^1 |f(x), g(x)| dx.$$

$$\left(d_2(f, g) = \left(\int_0^1 |f(x), g(x)| dx \right)^{1/2} \right).$$

Demostremos que ninguno de estos espacio métricos es completo. Comencemos por $C_1[0, 1]$. Para cumplir nuestro objetivo debemos hallar una sucesión de Cauchy en $C_1[0, 1]$ que no sea convergente en $C_1[0, 1]$. Definamos para $n \geq 2$ (fig. 60):

$$f_n(x) = \begin{cases} 1, & \text{si } x \in [0, \frac{1}{2} - \frac{1}{n}] \\ -\frac{n}{2}x + \frac{n}{4} + \frac{1}{2}, & \text{si } x \in [\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n}] \\ 0, & \text{si } x \in [\frac{1}{2} + \frac{1}{n}, 1] \end{cases}$$

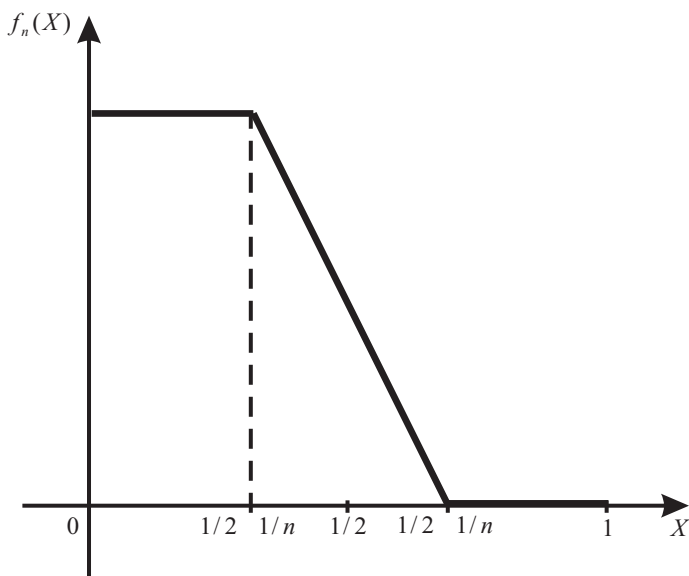


Figura 60

$$\begin{aligned}
 I &= 2 \left[\int_{1/2-1/n}^{1/2-1/n} (1 - f_m(x)) dx + \right. \\
 &\quad \left. + \int_{1/2-1/n}^{1/2} (f_n(x) - f_m(x)) dx \right] = \\
 &= 2 \left[\int_{1/2-1/n}^{1/2-1/n} \left(mx - \frac{m}{4} + \frac{1}{2} \right) dx + \right. \\
 &\quad \left. + \int_{1/2-1/n}^{1/2} \left(\frac{m-n}{2}x + \frac{n-m}{4} \right) dx = \right] \\
 &= \frac{1}{2} \left(\frac{1}{m} - \frac{1}{n} \right)_{n,m \rightarrow \infty} 0.
 \end{aligned}$$

luego la sucesión (f_n) es de Cauchy en $C_1[0, 1]$. Veamos que (f_n) no es convergente en $C_1[0, 1]$. En efecto, supongamos por el contrario que existe $f \in C_1[0, 1]$ tal que $(f_n) \xrightarrow{d_1} f$. Si consideramos las restricciones $f_n[0, 1/2]$, $f_n[1/2, 1]$, $f[0, 1/2]$ y $f[1/2, 1]$, no es difícil ver que la convergencia de f_n a f en $C_1[0, 1]$ se deduce:

$$(f_n [0, 1/2]) \longrightarrow f [0, 1/2] \text{ en } C_1 [0, 1/2],$$

$$(f_n [1/2, 1]) \longrightarrow f [1/2, 1] \text{ en } C_1 [1/2, 1].$$

Pero por otra parte, podemos comprobar inmediatamente que

$$(f_n [0, 1/2]) \longrightarrow \langle 1 \rangle \text{ en } C_1 [0, 1/2],$$

$$(f_n [1/2, 1]) \longrightarrow \langle 0 \rangle \text{ en } C_1 [1/2, 1].$$

Por unicidad del límite en $C_1 [0, 1/2]$ y $C_1 [1/2, 1]$ se tendría que

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

lo cual es una contradicción con la suposición de que f es continua sobre $[0, 1]$. Hemos demostrado que $C_1 [0, 1]$ no es completo.

Utilizando ahora la desigualdad de Cauchy-Buniakowski obtenida en el capítulo 2, podemos también demostrar la no completitud de $C_2 [0, 1]$. En efecto, no es difícil comprobar que la sucesión (f_n) definida anteriormente es también de Cauchy en $C_2 [0, 1]$ (demuéstrela). Aplicando la desigualdad de Cauchy-Buniakowski a las funciones $\langle 1 \rangle$ y $|f_n(x) - g(x)|$, donde g es una función continua arbitraria sobre $[0, 1]$ obtenemos:

$$\int_0^1 |f_n(x) - g(x)| dx \leq \left(\int_0^1 \langle 1 \rangle^2 dx \right)^{1/2} \left(\int_0^1 |f_n(x) - g(x)|^2 dx \right)^{1/2}$$

o sea,

$$d_1(f_n, g) \leq (b - a)^{1/2} d_2(f_n, g).$$

Supongamos que $f_n \xrightarrow{d_2} g$, entonces de la desigualdad anterior se tendría que $f_n \xrightarrow{d_1} g$ y por unicidad del límite $f \equiv g$, lo cual está en contradicción con la suposición de que g es continua.

Veamos algunas propiedades de los espacios métricos completos:

Proposición 3.2.3 (subespacios completos). Todo subespacio cerrado de un espacio completo es completo. Recíprocamente todo subespacio completo de un espacio métrico es cerrado.

Demostración: Verifiquemos inicialmente la primera afirmación de la proposición.

En efecto, sea A cerrado en el espacio métrico completo (X, d) y (x_n) una sucesión de Cauchy en (A, d_A) . La sucesión (x_n) será también de Cauchy en (X, d) y por la completitud de (X, d) se tiene que $(x_n) \xrightarrow{d} x$.

Por ser A cerrado $x \in A$, luego (A, d_A) es completo.

Supongamos recíprocamente que (A, d_A) es completo en el espacio métrico (X, d) y sea (x_n) una en A convergente en X a un punto $x : (x_n) \xrightarrow{d} x$. Por ser la sucesión (x_n) convergente será una sucesión de Cauchy (3.1.4) y de la completitud de (A, d_A) se deduce que su límite $x \in A$, con lo cual queda probado que A es cerrado. ■

Proposición 3.2.4 (producto numerable de espacios métricos completos). El producto de una familia finita o numerable de espacios métricos completos, provisto de la métrica producto, es un espacio métrico completo.

Demostración: sea $((X_n, d_n))_{n=1}^{\infty}$ una familia numerable de espacios métricos; sabemos que la métrica producto $\prod_{n=1}^{\infty} X_n$ se define por

$$d((x_n), (y_n)) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}$$

Sea $(x^{(m)}) = (x_n^{(m)}) = ((x_1^{(m)}), \dots, (x_n^{(m)}), \dots$

una sucesión de Cauchy en $\left(\prod_{n=1}^{\infty} X_n, d\right)$.

Esto significa que para cada $\epsilon > 0$ existe un m_0 tal que:

$$p, q \geq m_0 \Rightarrow \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n^{(p)}, x_n^{(q)})}{1 + d_n(x_n^{(p)}, x_n^{(p)})} \leq \epsilon$$

Para cada n fijo, esto implica en primer lugar que la sucesión $(x_n^{(m)})$ es de Cauchy en X_n , y por lo tanto, converge hacia un límite $x_n \in X_n$.

Dejando $p \geq m_0$ fijo y haciendo tender q a ∞ en

$$\sum_{n=1}^n \frac{1}{2^n} \frac{d_n(x_n^{(p)}, x_n^{(q)})}{1 + d_n(x_n^{(p)}, x_n^{(p)})}$$

obtenemos por la continuidad de la métrica que

$$\sum_{n=1}^N \frac{1}{2^n} \frac{d_n(x_n^{(p)}, x_n)}{1 + d_n(x_n^{(p)}, x_n)} \leq \epsilon \text{ para todo } N \in \mathbb{N}$$

Luego, $d((x_n^{(p)}), (x_n)) \leq \epsilon$ para $p \geq m_0$.

Esto demuestra que la sucesión $((x_n^{(p)}))$ converge en $\left(\prod_{n=1}^{\infty} X_n, d\right)$, con lo cual queda probada la completitud del producto. (Compare esta demostración con el ejemplo 6).■

Proposición 3.2.5 (espacios completos isométricos). Dos espacios métricos isométricos son simultáneamente completos o no.

Demostración: Sea $f : (X, d) \longrightarrow (Y, d')$ una isometría, es decir f es sobreyectiva y $d(x, y) = d'(f(x), f(y))$, para todo $x, y \in X$, y supongamos además que (X, d) es completo. Sean (y_n) una sucesión de Cauchy en (Y, d') y (x_n) la sucesión en (X, d) tal que

$$f(x_n) = y_n.$$

Por ser f una isometría, la propiedad de Cauchy es satisfecha también por (x_n) , y por lo tanto, existe $x = \lim(x_n)$ en (X, d) .

Por la continuidad de f existirá también $y = \lim(f(x_n)) = \lim(y_n)$, luego (Y, d') es completo.

Para probar que la completitud de (Y, d') implica la completitud de (X, d) basta notar que

$$f^{-1} : (Y, d') \longrightarrow (X, d)$$

también es una isometría. ■

Como nos muestra la proposición 3.2.5, la relación de isometría entre espacios métricos conserva la completitud pero esto no ocurre para la relación de homeomorfía; la completitud no es una propiedad topológica. Veamos el ejemplo siguiente:

Ejemplo 9. El espacio $\mathbb{R}$ con la métrica

$$d(x, y) = \left| \frac{|x|}{1 + |x|} - \frac{|y|}{1 + |y|} \right|$$

es topológicamente equivalente a $\mathbb{R}$ con la métrica usual, pero no es completo. En efecto, la sucesión (n) es de Cauchy en $(\mathbb{R}, d)$ y no converge a ningún punto de $\mathbb{R}$.

Este ejemplo no se puede confundir con el ejemplo 3. El ejemplo 3 muestra que una sucesión puede ser de Cauchy para una métrica y no serlo para otra métrica topológicamente equivalente, mientras que en el ejemplo 9 vemos que, aunque en un espacio métrico toda sucesión de Cauchy sea convergente, al sustituir su métrica por otra topológicamente equivalente puede haber sucesiones de Cauchy que no sean convergentes. Estas sucesiones de Cauchy no

convergentes con la nueva métrica, no pueden ser de Cauchy para la métrica inicial (¿por qué?).

Para finalizar el epígrafe analicemos el problema de la prolongación de funciones uniformemente continuas

Proposición 3.2.6 (prolongación de funciones uniformemente continuas). Sea A un subespacio de un espacio métrico (X, d) y (Y, d') un espacio métrico completo. Entonces toda aplicación uniformemente continua $f : (A, d_A) \longrightarrow (Y, d')$ tiene una única prolongación uniformemente continua $\bar{f} : \bar{A} \longrightarrow Y$.

Demostración: Por la propiedad de regularidad de los espacios métricos, sabemos que existe una única prolongación continua $\bar{f} : \bar{A} \longrightarrow Y$; solo queda probar que $\bar{f}$ es además uniformemente continua. En efecto, por la continuidad uniforme de f , para cada número $\epsilon > 0$ se puede hallar un número $\delta > 0$ tal que:

$$d(x, y) < 2\delta \Rightarrow d'(f(x), f(y)) < \epsilon/2.$$

Por las propiedades de la clausura y la continuidad de $\bar{f}$, para cada par $\bar{x}, \bar{y} \in \bar{A}$ podemos encontrar $x, y \in A$ tal que

$$d(x, \bar{x}) < \delta/2,$$

$$d(y, \bar{y}) < \delta/2,$$

y además,

$$d(f(x), \bar{f}(\bar{x})) < \epsilon/4,$$

$$d(f(y), \bar{f}(\bar{y})) < \epsilon/4.$$

Si $d(\bar{x}, \bar{y}) < \delta$, tendremos

$$d(x, y) \leq d(x, \bar{x}) + d(\bar{x}, \bar{y}) + d(y, \bar{y}) < \delta/2 + \delta + \delta/2 = 2\delta,$$

y por lo tanto

$$d(\bar{f}(\bar{x}), \bar{f}(\bar{y})) \leq d(\bar{f}, (\bar{x}), f(x)) + d(f(x), f(y)) + d(f(y),$$

$$d(f(\bar{y})) < \epsilon/4 + \epsilon/2 + \epsilon/4 = \epsilon,$$

con lo cual queda probada la continuidad uniforme de $\bar{f}$. ■

Ejercicios de Autocontrol

1. Muestre que $\rho(x, y) = \arctan |x' - y|$ es una métrica en $\mathbb{R}$.

¿Es esta métrica equivalente a la métrica usual? ¿Es el espacio $(\mathbb{R}, \rho)$ completo?

2. Demuestre que todo espacio métrico discreto es completo.
3. Sea $C' [a, b]$ el espacio de todas las funciones reales continuamente derivables sobre $[a, b]$.

Demuestre que

$$\rho_1(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)| + \sup_{x \in [a, b]} |f'(x) - g'(x)| \text{ y}$$

$$\rho_2(f, g) = \sup_{x \in [a, b]} \left(|f(x) - g(x)| + \sup_{x \in [a, b]} |f'(x) - g'(x)| \right)$$

son métricas topológicamente equivalentes sobre $C' [a, b]$ y que provisto de cualquiera de ellas $C' [a, b]$ es un espacio métrico completo.

4. Sean F y G funciones reales y continuas fijas definidas sobre el intervalo $[a, b]$ ($F, G \in C [a, b]$ y tales que $F(x) \leq G(x)$ para todo $x \in [a, b]$).

Demuestre que el subespacio de $C [a, b]$:

$$\{f \in C [a, b] : F(x) \leq f(x) \leq G(x)\}$$

es completo provisto de la métrica inducida por d_∞

3.3 ALGUNOS RESULTADOS IMPORTANTES EN LOS ESPACIOS MÉTRICOS COMPLETOS

En este epígrafe veremos teoremas clásicos obtenidos en los espacios métricos completos, los cuales utilizaremos en el 3.5 para obtener algunas aplicaciones en diversas ramas de la matemática.

El teorema de Baire (llamado con frecuencia teorema de categoría de Baire) constituye el instrumento de trabajo principal en los espacios métricos completos, sus aplicaciones se ven fundamentalmente en el estudio de los operadores en espacios de Banach (vea el capítulo 6) y en la demostración de teoremas de existencia. En la demostración del teorema de Baire juega un papel importante el teorema de Cantor, denominado también *principio de bolas encajadas*.

Antes de pasar a estudiar el teorema de Baire, veremos el llamado teorema del punto fijo de Banach, el cual resulta de gran importancia en el análisis de la existencia de soluciones de ecuaciones funcionales en espacios métricos completos.

Para ello, necesitamos algunos conceptos.

Una aplicación f de un espacio métrico (X, d) se dice que es contractante de razón k , si $k < 1$ y

$$d(f(x), f(y)) < kd(x, y) \quad \forall x, y \in X.$$

Claramente una aplicación tal es uniformemente continua.

Dada una aplicación $g : (X, d) \rightarrow (X, d)$, a cada punto $x \in X$ tal que $g(x) = x$ se le llama un punto fijo de g .

Teorema 3.3.1 (del punto fijo). Cada aplicación contractante f de razón k definida en un espacio métrico completo (X, d) tiene un punto fijo único x_0 caracterizado por la propiedad siguiente: para cada punto $x \in X$ la sucesión de valores iterados

$$[f^n(x)] = \underbrace{f \circ f \dots \circ f(x)}_{n\text{-veces}}$$

converge al punto x_0 con el error de aproximación:

$$d(f^n(x), x_0) \leq \frac{k^n}{1-k} d(x_0, f(x_0)) \quad (9)$$

Demostración: Dividiremos la demostración en tres partes:

i) Para todo $x \in X$; $(f^n(x))$ converge a un cierto punto x_0 con el error de aproximación (9).

En efecto, del hecho que

$$d(f(x), f(y)) \leq kd(x, y) \quad \forall x, y \in X$$

se deduce sucesivamente:

$$d(f(x), f^2(x)) \leq kd(x, f(x))$$

$$d(f^n(x), f^{n+1}(x)) \leq k^n d(x, f(x)),$$

y por tanto, de la desigualdad triangular se obtiene:

$$d(f^n(x), f^{n+p}(x)) \leq k^n(1 + k + k^2 + \dots + k^{p-1})$$

$$d(x, f(x)) \leq \frac{k^n}{1-k} d(x, f(x)) \quad (10)$$

Esto prueba que $(f^n(x))$ es una sucesión de Cauchy, y por tanto, convergente a un cierto punto $x_0 \in X$ para el cual, según (10), se cumple:

$$d(f^n(x), x_0) \leq \frac{k^n}{1-k} d(x, f(x))$$

ii) $f(x_0) = x_0$.

Si sustituimos x por x_0 en la mayoría del error tendremos:

$$d(f^{n+1}(x_0), f(x_0)) \leq kd(f^n(x_0), x_0) \leq \frac{k^{n+1}}{1-k} d(x_0, f(x_0)),$$

lo cual implica que la sucesión $(f^n(x_0))$ converge a x_0 y $f(x_0)$ o sea $x_0 = f(x_0)$.

iii) Unicidad del punto fijo.

Si y_0 fuera otro punto fijo de f , entonces:

$$d(x_0, y_0) = d(f(x_0), f(y_0)) \leq kd(x_0, y_0)$$

lo cual es una contradicción. ■

El teorema siguiente nos muestra, dentro de las caracterizaciones teóricas de un espacio métrico completo, una de las más usadas en las aplicaciones. Este teorema tiene gran importancia en los espacios métricos completos, similar a la que tiene el teorema sobre los intervalos encajados en el Análisis.

Proposición 3.3.2 (Cantor). Un espacio métrico (X, d) es completo si y solo si toda sucesión decreciente (F_n) de subconjuntos cerrados no vacíos con diámetro $(\delta(F_n)) \rightarrow 0$ tiene intersección no vacía. Más exactamente $\bigcap_{n \in \mathbb{N}} F_n$ se deduce a un punto.

Demostración: Probemos primeramente que la condición es necesaria. Supongamos que el espacio (X, d) es completo y sea la sucesión decreciente de cerrados

$$F_1 \supset F_2 \supset \dots \supset F_n \supset \dots$$

tal que $\delta(F_n) \rightarrow 0$.

Seleccionemos, para cada n , un elemento $x_n \in F_n$. La sucesión es de Cauchy en (X, d) . En efecto:

$$d(x_n, x_m) < \delta(F_{\min\{m, n\}}), \text{ y por lo tanto,}$$

$$m, n \rightarrow \infty \Rightarrow d(x_n, x_m) \rightarrow 0.$$

Por ser el espacio (X, d) completo, existe $\lim(x_n) = x \in X$.

Como $x_m \in F_n$ para todo $m \geq n$, entonces $x \in \overline{F_n} = F_n$.

Por la arbitrariedad de n se tiene que $x \in \bigcap_{n \in \mathbb{N}} F_n$, luego $\bigcap_{n \in \mathbb{N}} F_n \neq \emptyset$.

Probemos ahora que $\bigcap_{n \in \mathbb{N}} F_n$ se reduce a un solo punto.

Supongamos que $x, y \in \bigcap_{n \in \mathbb{N}} F_n$, entonces $x, y \in F_n$ para todo $n \in \mathbb{N}$ y por tanto $d(x, y) \leq \delta(F_n)$ para todo $n \in \mathbb{N}$.

Haciendo tender $n \rightarrow \infty$ en esta desigualdad, obtendremos $d(x, y) = 0$, luego $x = y$.

Demostraremos ahora que la condición enunciada en la proposición es también suficiente. Sea (x_n) una sucesión de Cauchy en (X, d) y denotemos por S_n la selección final

$$S_n = \{x_n, x_n + 1, \dots\}.$$

La sucesión decreciente de conjuntos cerrados $F_n = S_n$ satisface las condiciones del teorema.

En efecto, la condición $\delta(F_n) \longrightarrow 0$ se obtiene de la propiedad de Cauchy para la sucesión (x_n) . Luego existe $x \in X$ tal que

$$\{x\} = \bigcap_{n \in \mathbb{N}} F_n.$$

Pero $x \in \overline{S_n}$ para todo n , y por lo tanto, x es un punto adherente de la sucesión (x_n) , de donde aplicando la observación que aparece después de la proposición 3.1.4, se deduce que $(x_n) \xrightarrow{d} x$. ■

Observación: Se puede obtener una caracterización equivalente a la proposición anterior sustituyendo los cerrados F_n por bolas cerradas.

En el capítulo siguiente veremos que la condición necesaria de la proposición 3.3.2 se cumple también en los espacios métricos compactos sin la hipótesis adicional de que $\delta(F_n) \longrightarrow 0$.

La diferencia entre ambas clases de espacios, los compactos y los completos, se refleja en la proposición 3.3.2 mediante dicha hipótesis complementaria, la cual es necesaria para los espacios métricos completos.

El hecho de que la completitud sea una *propiedad uniforme* y no topológica, se manifiesta en esta hipótesis que depende esencialmente de la posibilidad que brinda la métrica d de introducir un concepto de uniformidad.

Ejemplo 10. En $\mathbb{R}$ con la métrica usual la sucesión de intervalos $F_n = [n, \infty[$ ($n = 1, 2, \dots$), satisface las condiciones de la proposición 3.3.2 con excepción de la relacionada con el diámetro de F_n ya que $\delta(F_n) = \infty$ para todo n .

Sabemos que $\mathbb{R}$ es completo, pero $\bigcap_{n \in \mathbb{N}} [n, \infty[= \emptyset$

La siguiente condición necesaria para que un espacio métrico sea completo, es un resultado muy importante y tiene una gran aplicación en análisis, sobre todo, en la demostración de teoremas de existencia.

Teorema 3.3.3. En un espacio métrico completo la intersección de cualquier familia numerable de conjuntos abiertos densos es densa.

Demostración: Sean (X, d) un espacio métrico completo y (D_n) una sucesión de conjuntos densos en (X, d) y abiertos.

Probemos que $U \cap \left(\bigcap_{n \in \mathbb{N}} D_n \right) \neq \emptyset$ para todo abierto no vacío U en (X, d) .

En efecto, por ser $U \cup D_1 \neq \emptyset$ existe una bola abierta B_1 tal que $\overline{B_1} \subset U \cup D_1$ y $\delta(\overline{B_1}) \leq 1$. Por inducción, se puede construir una sucesión (B_n) de bolas abiertas tales que:

$$\overline{B}_n \subset \overline{B}_{n-1} \cap D_n \text{ y } \delta(\overline{B}_1) \leq 1 \setminus n.$$

Entonces,

$$\bigcap_{n \in \mathbb{N}} \overline{B}_n \subset \left(U \cap \left(\bigcap_{n \in \mathbb{N}} D_n \right) \right),$$

pero las $\overline{B}_n$ satisfacen las condiciones del teorema de Cantor, y por tanto,

$$\bigcap_{n \in \mathbb{N}} \overline{B}_n \neq \emptyset$$

con lo cual queda probado el teorema. ■

Definición 3.3.4. Sea (x, d) un espacio métrico. Un conjunto $E \subset X$ se llama nunca denso si su clausura $\overline{E}$ tiene interior vacío ($\overline{E} = \emptyset$). Un conjunto de *primera categoría* en X es un conjunto que puede representarse mediante la unión de una familia numerable de conjuntos nunca densos. Cada subconjunto de X que no sea de primera categoría se llama de *segunda categoría*. Los espacios métricos de segunda categoría se llaman también *espacios* de Baire.

Corolario 1. Todo espacio métrico completo es de segunda categoría en sí mismo (o sea, es de Baire).

Demostración: Si $\{E_n\}$ es una familia de subconjuntos nunca densos del espacio métrico completo (X, d) , y V_n es el complemento de $\overline{E}_n$, entonces cada uno de los conjuntos V_n es abierto y denso en X ; el teorema de Baire nos muestra que

$$\bigcap_{n \in \mathbb{N}} V_n \neq \emptyset \text{ Por consiguiente } X \neq \bigcup_{n \in \mathbb{N}} E_n. \blacksquare$$

Ejemplo 11. Se dice que un espacio métrico (x, d) es topológicamente completo, si existe una métrica d' sobre X topológicamente equivalente a d tal que (X, d') es completo. Según el teorema de Baire todo espacio métrico topológicamente completo es de segunda categoría. La propiedad necesaria para que un espacio métrico sea topológicamente completo. Es fácil ver, siguiendo este criterio, que el espacio métrico $\mathbb{Q}$ de los números racionales con métrica usual inducida por el módulo no es topológicamente completo. En efecto, basta observar que $\mathbb{Q}$ es la unión numerable de sus conjuntos unitarios que son conjuntos cerrados, para que $\mathbb{Q}$ no sea de Baire.

Ejercicios de Autocontrol

1. Demuestre que todo espacio métrico que posea al menos un punto aislado, es un conjunto de segunda categoría.
2. Demuestre que en todo espacio métrico completo, el complemento de un conjunto de primera categoría es un conjunto de segunda categoría.

3. Diga de qué categoría en $\mathbb{R}^2$ es el conjunto E de los puntos, cuyas coordenadas son irracionales. ¿De qué categoría es su complemento?
4. Demuestre que el conjunto E de todos los polinomios de grado $\leq N$, donde N es un número fijo, es nunca denso en $(C[0, 1], d_\infty)$.
5. Cite un ejemplo de un espacio métrico completo (X, d) y $f : (X, d) \rightarrow (X, d)$ que muestre, que en el teorema del punto fijo, la propiedad de contracción de f no puede ser sustituida por la condición $d(f(x), f(y)) < d(x, y)$ para $x \neq y$.
6. Sean (X, d) completo y $f : (X, d) \rightarrow (X, d)$ una aplicación tal que f^n ($n > 0$) es una contracción. Demuestre que f tiene un único punto fijo en X .
7. Encuentre una isometría de un espacio métrico completo sin punto fijo.

3.4 COMPLETAMIENTO DE UN ESPACIO MÉTRICO

Evidentemente no todo espacio métrico es completo. En el ejemplo 11 vimos que $\mathbb{Q}$ no era completo para ninguna distancia topológicamente equivalente a la usual. Sin embargo, es por todos conocido el proceso de *completamiento* de $\mathbb{Q}$ el cual da lugar a los números reales. Veremos que $\mathbb{R}$, en cierto sentido, es el menor espacio métrico completo a $\mathbb{Q}$; este proceso mediante el cual obtenemos $\mathbb{R}$ a partir de $\mathbb{Q}$, no es más que un tipo particular de *extensión* la cual conserva no solo las propiedades topológicas, sino también, las propiedades uniformes de $\mathbb{Q}$. Hay que destacar esta propiedad porque existen otras *extensiones topológicas* que no se comportan de la misma forma.

Definición 3.4.1. Sean (X, d) y (Y, d') dos espacios métricos:

i) $f : (X, d) \rightarrow (Y, d')$ se llama una *inmersión topológica (uniforme)* de (X, d) en (Y, d') , si $f : (X, d) \rightarrow (f[X], d_{f[X]})$ es un isomorfismo topológico, o sea, si f es biyectiva y tanto f como f^{-1} son continuas (es un isomorfismo uniforme, o sea, si f es biyectiva y tanto f como f^{-1} son uniformemente continuas).

ii) El par $(f, (Y, d'))$ se denomina *extensión topológica (uniforme)* de (x, d) , si $f : (X, d) \rightarrow (Y, d')$ es una inmersión topológica (uniforme) y $f[X]$ es denso en (Y, d') .

En este epígrafe nos interesan solamente aquellas extensiones uniformes que son completas.

Definición 3.4.2. Dado un espacio métrico (X, d) se dice que la extensión

$(f, (Y, d'))$ es un *completamiento* de (X, d) si (Y, d') es un espacio métrico completo.

Para facilitar el lenguaje en determinadas ocasiones, convendremos en llamar completamiento de (X, d) solamente al espacio (Y, d') .

Exigiremos del completamiento de un espacio que conserve no solo las propiedades uniformes, sino también las propiedades métricas del espacio. Para esto veremos el teorema siguiente:

Teorema 3.4.3 (completamiento de un espacio métrico). Todo espacio métrico (X, d) tiene un completamiento $(i, (Y, d))$ donde (Y, d) es un espacio métrico completo e $i : X \rightarrow i[X]$ es una isometría. Además, el completamiento de un espacio métrico es único salvo isometrías.

Demostración: Demostraremos primeramente la unicidad del completamiento. Sea $(i_1, (Y_1, d_1))$ y $(i_2, (Y_2, d_2))$ dos completamientos de (X, d) (fig. 61).

Evidentemente $i = i_2 \circ i_1^{-1} : i_1[X] \rightarrow i_2[X]$ es una isometría, y por tanto, un isomorfismo uniforme. Según la proposición 3.2.6 i puede extenderse a un isomorfismo uniforme $\bar{i} : (Y_1, d_1) \rightarrow (Y_2, d_2)$ el cual por continuidad es claramente una isometría.

Pasemos ahora a probar la existencia del completamiento.

Sea $x_0 \in X$ fijo y definamos la aplicación:

$$i : X \rightarrow B(X, \mathbb{R}) \quad (11)$$

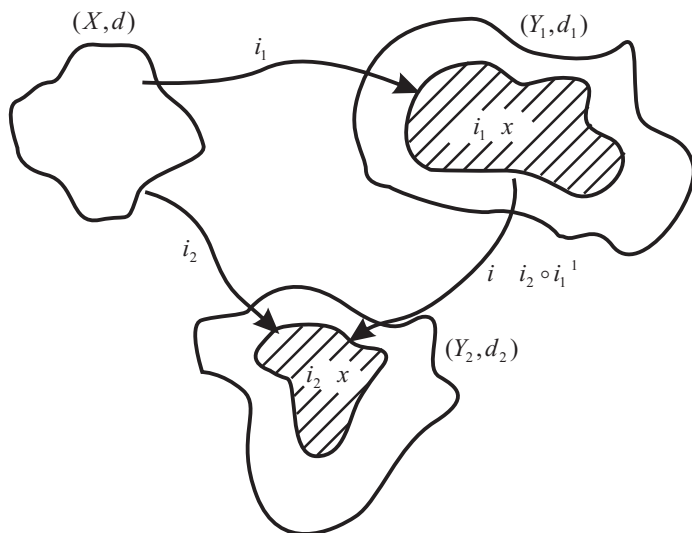


Figura 61

$$g_x(y) = d(y, x) - d(y, x_0) \quad (12)$$

Es claro que $g_x \in B(X, \mathbb{R})$ pues

$$\sup_{y \in X} |g_x(y)| = \sup_{y \in X} |d(y, x) - d(y, x_0)| = d(x, x_0) \quad (13)$$

Además, i es una isometría de X sobre $i[X]$ pues

$$d_\infty(g_x, g_x) = \sup_{y \in X} |d(y, x) - d(y, x')| = d(x, x') \quad (14)$$

Por tanto, el par $(i, \overline{i[X]})$ es un completamiento de (X, d) . ■

Ejercicios de Autocontrol

1. Defina la aplicación $\rho : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}_+$ mediante:

$$\rho(x, y) = |\arctan x - \arctan y|.$$

Demuestre que ρ es una métrica y que $(\mathbb{R}, \rho)$ no es completo.

Describa el completamiento de $(\mathbb{R}, \rho)$.

2. Describa el completamiento del conjunto de los polinomios de una variable real, con respecto a la métrica

$$d(f, g) = \max_{[-1, 1]} |f(x) - g(x)| + |f'(0) - g'(0)|.$$

3.5 APLICACIONES

Aplicación 1: Existencia y unicidad de solución para ecuaciones diferenciales de primer orden.

Expondremos ahora un teorema clásico relativo a la existencia de solución única para ciertos tipos de ecuaciones diferenciales de primer orden. El teorema que demostraremos no es más que una aplicación del teorema del punto fijo desarrollado anteriormente.

Teorema 3.5.1. Sea f una función continua con valores reales definida sobre un subconjunto abierto U de $\mathbb{R}^2$.

Supongamos que existe un número positivo M tal que:

$$|f(x, y_1) - f(x, y_2)| \leq M |y_1 - y_2|.$$

Dado $(x_0, y_0) \in U$, sean b, c y K tales que

$$S = [x_0 - c, x_0 + c] \times [y_0 - b, y_0 + b] \subset U$$

$$y \ K = \sup \{|f(x, y)| \mid (x, y) \in S\} + 1.$$

Entonces, para todo $0 < a < \min\{c, 1/M, b/K\}$ existe una función

$$y : [x_0 - a, x_0 + a] \longrightarrow \mathbb{R}$$

tal que:

$$y(x_0) = y_0 \text{ y} \tag{15}$$

$$y'(x) = f(x, y(x)) \ ; \ \forall x \in [x_0 - a, x_0 + a] \tag{16}$$

Además, y es la única función que satisface (15) y (16) sobre el intervalo $[x_0 - a, x_0 + a]$.

Demostración: Haremos algunas observaciones preliminares antes de proceder con la demostración.

Para probar este teorema, consideraremos la solución de la siguiente ecuación integral para y

$$y(x) = y_0 + \int_{x_0}^x f(t, y(t)) dt \tag{17}$$

El lector, puede fácilmente comprobar que una función y satisface (17) si y sólo si satisface (15) y (16).

Nuestra estrategia será hallar un espacio completo G de funciones y una aplicación contractante ψ , $\psi : G \rightarrow G$ que tenga la siguiente propiedad: si existe una solución y de (17), entonces $y \in G$ y además y es solución de (17) si y sólo si $\psi(y) = y$. Para complementar esto, notemos primeramente que cualquier solución posible y debe ser tal que $(x, y(x)) \in U$ para todo x en el dominio de y de manera que $f(x, y(x))$ esté definido.

Otras consideraciones relativas a la elección del espacio G quedarán claras al lector en el desarrollo de la demostración.

Sea $(x_0, y_0) \in U$. Por ser U abierto existen b y c tales que el rectángulo:

$$S = [x_0 - c, x_0 + c] \times [y_0 - b, y_0 + b] \subset U.$$

Sea a un número cualquiera tal que:

$$0 < a < \min \left\{ c, \frac{1}{M}, \frac{b}{K} \right\}.$$

Notemos que

$$[x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b] \subset U.$$

El espacio $(C([x_0 - a, x_0 + a], \mathbb{R}), d_\infty)$ es completo por ser cerrado en: $B([x_0 - a, x_0 + a], \mathbb{R})$, pero este no es el espacio adecuado para nuestros propósitos.

Definamos G como el subespacio de $C([x_0 - a, x_0 + a], \mathbb{R})$ formado por todos los elementos y tales que:

$$y(x) \in [y_0 - b, y_0 + b]; \forall x \in [x_0 - a, x_0 + a].$$

Evidentemente G es completo por ser cerrado en:

$C([x_0 - a, x_0 + a], \mathbb{R})$ y si $y \in G$ entonces $(x, y(x)) \in U$ para todo $x \in [x_0 - a, x_0 + a]$.

Definiremos ψ sobre G de la siguiente forma:

$$\psi(g)(x) = y_0 + \int_{x_0}^x f(t, g(t)) dt$$

El lector puede comprobar que $\psi[G] \subset G$. Por otra parte $\psi : G \rightarrow G$ es una contracción.

En efecto, si $g_1, g_2 \in G$

$$\begin{aligned} |\psi(g_1)(x) - \psi(g_2)(x)| &= \left| \int_{x_0}^x (f(t, g_1(t)) - f(t, g_2(t))) dt \right| \\ &\leq M \int_{x_0}^x |g_1(t) - g_2(t)| dt \\ &\leq M_a d_\infty(g_1, g_2). \end{aligned}$$

Como la desigualdad se cumple independientemente de $x \in [x_0 - a, x_0 + a]$, se obtiene:

$$d_\infty(\psi(g_1), \psi(g_2)) \leq M_a d_\infty(g_1, g_2)$$

lo cual implica, que ψ es contractante pues de la definición de a , $M_a = k < 1$. Según el teorema del punto fijo existe una función y única en G que cumple

$$\psi(y) = y$$

o sea y satisface (15) y (16).

Si y_1 es otra función sobre $[x_0 - a, x_0 + a]$ que satisface (15) y (16) entonces:

$$y_1(x) = y_0 + \int_{x_0}^x f(t, y_1(t)) dt \quad \forall x \in [x_0 - a, x_0 + a],$$

y por tanto:

$$|y_1(x) - y(x)| = \left| \int_{x_0}^x f(t, y_1(t)) - f(t, y_2(t)) dt \right| \leq$$

$$\leq M_a d_\infty(y_1, y_2) \quad \forall x \in [x_0 - a, x_0 + a]$$

de donde

$$d_\infty(y_1, y_2) < d_\infty(y_1, y_2)$$

lo cual es una contradicción. ■

Aplicación 2: Cálculo aproximado de las raíces de funciones reales (Métodos de Newton)

Consideremos una función real continuamente derivable definida sobre un intervalo $[a, b]$:

$$f : [a, b] \longrightarrow \mathbb{R},$$

y además,

$$\text{i) } f[a, b] \subset [a, b],$$

$$\text{ii) } \sup_{a \leq x \leq b} |f'(x)| = q < 1$$

Entonces la ecuación $f(x) = x$ tiene una única solución en $[a, b]$. En efecto, notando que $[a, b]$ como subespacio de $\mathbb{R}$ es completo por ser cerrado (3.2.3), para poder aplicar el teorema del punto fijo a este caso, es suficiente demostrar que f es una aplicación contractante lo cual se deduce de ii) y del teorema de valor intermedio:

$$\sim f(x) - f(y) = f'(\xi)(x - y), \text{ donde } b \geq x > (\xi) > y > a.$$

Luego partiendo de cualquier punto $x_0 \in [a, b]$, la sucesión de iteradas $f^{(n)}(x_0) = f \circ f \circ \dots \circ f(x_0)$, n veces converge a la única solución de la ecuación $f(x) = x$ sobre $[a, b]$.

En la práctica la dificultad aparece al tratar de encontrar un intervalo $[a, b]$ donde se satisfagan las condiciones i) y ii); pero pueden existir también otro tipo de situaciones en que el problema sea aún más complejo. Supongamos que queremos hallar una raíz de la ecuación $f(x) = 0$ en el intervalo $[a, b]$ o al menos demostrar que en este intervalo hay una raíz de dicha ecuación. Para ello, podemos auxiliarnos de la ecuación equivalente.

$$x - af(x) = x$$

donde a es una constante no nula que elegiremos convenientemente a *posteriori*.

Para que satisfaga las condiciones i) y ii) la constante debe ser elegida de manera tal que

- i') $x - af(x) \in [a, b]$ para todo $x \in [a, b]$,
 ii') $\left| f'(x) - \frac{1}{a} \right| < \frac{1}{|a|}$ para todo $x \in [a, b]$.

Esto resulta a veces muy difícil si se fija de antemano el intervalo $[a, b]$ y por ello el problema se plantea en la forma más general: hallar un intervalo $[a, b]$ para el cual se pueda encontrar un número real a que satisfaga i') y ii').

En la solución de este último problema, influyen determinantemente las propiedades de la función f , si logramos darle respuesta, con ello obtendremos que en el intervalo $[a, b]$ hallado existe una solución única de la ecuación $f(x) = 0$, que se obtiene partiendo de cualquier punto $x_0 \in [a, b]$ mediante iteraciones de la función $x - af(x)$.

Uno de los pasos que generalmente no se logra en estos análisis es el i').

Veamos un método aproximado de solución de ecuaciones del tipo $f(x) = 0$ que no utiliza la propiedad i') aunque se basa también en el teorema del punto fijo. Este método es el llamado *método de Newton o método de las tangentes*, el cual consiste en que para resolver la ecuación $f(x) = 0$ se buscan las aproximaciones sucesivas a la raíz de acuerdo con la fórmula recurrente

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (19)$$

(por aproximación nula x_0 se toma un punto arbitrario del segmento donde está definida f .)

Notemos que

$$y - f(x_n) = f'(x_n)(x - x_n)$$

es la ecuación de la recta tangente a la curva definida por f en el punto en que esta recta intersecta al eje OX .

Si por algún otro método conocemos *a priori* que la ecuación $f(x) = 0$ tiene una única raíz x^* en el segmento $[a, b]$, si la función $f(x)$ tiene en este segmento derivada no nula y segunda derivada acotada y con signo constante, entonces se puede probar que existe una vecindad $[x - \delta, x^* + \delta]$ de la raíz x^* , tal que si el punto x_n se toma en esta vecindad, la sucesión (19) converge hacia x^* . En efecto, basta probar que la función

$$g(x) = x \frac{f(x)}{f'(x)} \quad (20)$$

satisface las condiciones i) y ii) en algún intervalo $[x^*, x^* + \delta_1]$.

De forma análoga, podría obtenerse un resultado equivalente en algún intervalo de la forma $[x^* - \delta_2, x^*]$ y para demostrar la afirmación postulada anteriormente bastaría tomar $\delta = \min \{\delta_1, \delta_2\}$.

Supongamos, sin pérdida de generalidad, que $f' > 0$ y $f'' \geq 0$ en una vecindad de x^* . En cualquier otro caso es suficiente realizar uno de los cambios $f(x) \rightsquigarrow -f(x)$, $f(x) \rightsquigarrow f(-x)$ o $f(x) \rightsquigarrow -f(-x)$ para caer en el caso por nosotros analizado.

En nuestro caso, tanto f como f' serán funciones crecientes y $f \geq 0$ a la derecha de x^* . Luego, en cualquier intervalo $[x^*, x^* + \delta]$ ya que

$$0 < \frac{f(x)}{f'(x)} \leq \delta.$$

Esta última desigualdad se obtiene del hecho que f' es creciente y

$$\frac{f(x)}{f'(x)} = \frac{f'(\xi)}{f'(x)} (x - x^*), \quad x^* < \xi < x.$$

Por otra parte,

$$g'(x) = \frac{f'(x) f(x)}{(f'(x))^2} = \frac{f''(x) f'(\xi)}{(f'(x))^2} (x - x^*)$$

para algún $x^* < \xi < x$.

Como f' es continua y estrictamente positiva en $[x^*, x^* + \delta]$, entonces existe $m = \inf f'(x) : m > 0$.

$$x^* \leq x \leq x^* + \delta$$

Además, hemos supuesto que $\sup |f''(x)| \leq M < \infty$, luego

$$|g'(x)| \leq \frac{M}{m^2} |f'(\xi) - f'(x)| + \frac{M}{m} \delta. \quad (21)$$

Por continuidad de f' (de la cual se deduce su continuidad uniforme sobre $[a, b]$), dado $0 < \epsilon < \frac{m^2}{4M}$ se puede hallar $\eta > 0$ tal que si $x, \xi \in [x^*, x^* + \eta]$ entonces $|f'(\xi) - f'(x)| < \epsilon$.

$$\text{Tomemos } \delta \leq \min \left\{ \eta, \frac{M}{4m} \right\}.$$

De (21) se deduce que $|g'(x)| \leq 1/2$ para todo $x \in [x^*, x^* + \delta]$, donde concluimos nuestra afirmación aplicando el teorema del valor medio a la función g en $[x^*, x^* + \delta]$. Es evidente que un punto fijo para g es una raíz de la ecuación $f(x) = 0$, la cual se obtiene mediante el límite de la sucesión (19) con $x_0 \in [x^*, x^* + \delta]$

Como hemos podido ver los teoremas del punto fijo son muy importantes para la demostración de teoremas de existencia de solución de ecuaciones y obtener algoritmos numéricos para el cálculo de dichas soluciones en diversas ramas de la matemática.

Existen muchos otros teoremas del punto fijo cuyas aplicaciones fundamentalmente se hallan en la teoría de ecuaciones diferenciales.

Aplicación 3: Teorema de Goursat

Este teorema es de gran importancia en el desarrollo de la teoría de las funciones holomorfas de una variable compleja; la demostración se basa en el teorema de Cantor.

Teorema (de Goursat). Si f es holomorfa en un abierto U de $\mathbb{C}$, entonces la forma diferencial $f(z) dz$ es localmente exacta en U .

Demostración: Para probar el teorema basta demostrar que la integral de $f(z) dz$ a lo largo del borde de cada rectángulo con lados paralelos a los ejes coordenados es nula (fig. 62).

Sea R un tal rectángulo y sea $\Gamma(R)$ el camino en U formado recorriendo los lados de R en sentido contrario a las manecillas del reloj.

Sea $\alpha(R) = \int_{\Gamma(R)} f(z) dz$. Debemos probar que $\alpha(R) = 0$. Dividiendo R en cuatro rectángulos tales que $\Gamma(R) = \Gamma(R^{(1)}) + \Gamma(R^{(2)}) + \Gamma(R^{(3)}) + \Gamma(R^{(4)})$ deducimos que existe al menos uno de ellos, llamémosle $R^{(i)}$ tal que $|\alpha(R^{(i)})| \geq 1/4|\alpha(R)|$. Repitiendo este procedimiento se construye una sucesión (R_n) de rectángulos encajados en U con lados paralelos a los ejes tales que $|\alpha(R_n)| \geq 1/4^n |\alpha(R)|$ (fig. 62).

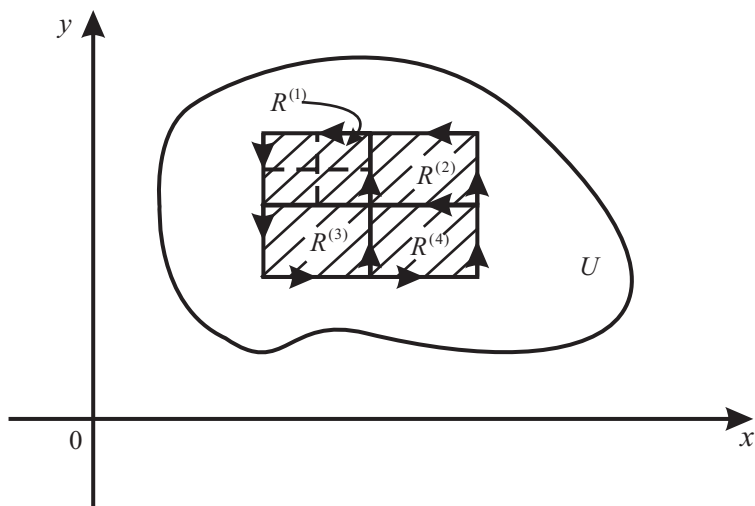


Figura 62

Según el teorema de Cantor existe un punto z_0 único común a todos los

rectángulos R_n .

De la derivabilidad de f en z_0 se deduce que dado $\epsilon > 0$, existe un entorno $V(z_0)$ de z_0 tal que:

$$|f(z) - f(z_0) - (z - z_0)f'(z_0)| < \epsilon |z - z_0| \forall z \in V(z_0).$$

Para n suficientemente grande $R_n \subset V(z_0)$.

Las formas diferenciales dz y $zdz = d(z^2/2)$, siendo exactas en U tienen integral nula sobre $\Gamma(R_n)$, luego:

$$\begin{aligned} |\alpha(R_n)| &= \left| \int_{\Gamma(R_n)} f(z) dz \right| \leq \epsilon \int_{\Gamma(R_n)} |z - z_0| dz \leq \\ &\leq \epsilon L_n \sup_{z \in \Gamma(R_n)} |z - z_0| \end{aligned}$$

donde L_n es la longitud de (R_n) .

Pero si d y P son las longitudes de la diagonal y el perímetro de $\Gamma(R_n)$ entonces $L_n = 2^{-n}P$ y $\sup_{z \in \Gamma(R_n)} |z - z_0|$ es menor que la diagonal de R_n cuya longitud es $2^{-n}d$.

Finalmente $|\alpha(R_n)| \leq \epsilon 4^{-n}Pd$. Pero por construcción de los rectángulos $R_n : |\alpha(R_n)| \geq 4^{-n}|\alpha(R)|$ de lo cual se deduce $|\alpha(R)| \leq \epsilon Pd$. Siendo $\epsilon > 0$ arbitrario el teorema queda probado. ■

Aplicación 4: Existencia de funciones en $C([0, 1], \mathbb{R})$ que no son derivables en ningún punto

Como ejemplo de aplicación del teorema de Baire al análisis, veremos el siguiente teorema de existencia.

Teorema 3.5.3. Existen funciones reales y continuas sobre $[0, 1]$ que no son derivables en ningún punto de $[0, 1]$.

Demostración: Para cada $n = 1, 2, \dots$ definamos $N_n \subset C([0, 1], \mathbb{R})$ mediante:

$$\begin{aligned} N_n &:= \{f \in C([0, 1], \mathbb{R}) \mid \exists x \in [0, 1 - 1/n], \\ &\forall h \in [0, 1/n], \left| \frac{f(x+h) - f(x)}{h} \right| \leq n\}. \end{aligned} \tag{22}$$

Las únicas funciones de $C([0, 1], \mathbb{R})$ que tienen derivada en algún punto, son aquellas que pertenecen a algún N_n . Probemos que

$$C([0, 1], \mathbb{R}) \setminus \bigcup_{n \in \mathbb{N}} N_n \neq \emptyset$$

y para ello, demostraremos que cada N_n es denso en ninguna parte, y como $C([0, 1], \mathbb{R})$ es completo, no podrá ser igual a $\bigcup_{n \in \mathbb{N}} N_n$.

1. Cada N_n es cerrado

Es fácil comprobar que la aplicación evaluación $W : C([0, 1], \mathbb{R}) \times I \rightarrow \mathbb{R}$ es continua (demuéstrelo).

Para cada punto fijo $h_0 \in]0, \frac{1}{n}]$ se tiene que las aplicaciones de

$$C([0, 1], \mathbb{R}) \times [0, 1 - \frac{1}{n}] \rightarrow \mathbb{R}$$

dadas por

$$(f, x) \rightsquigarrow \frac{1}{h} \cdot f(x + h_0) \text{ y}$$

$$(f, x) \rightsquigarrow \frac{1}{h} f(x)$$

son también continuas, luego también lo será

$$(f, x) \rightsquigarrow \left| \frac{f(x+h_0)-f(x)}{h_0} \right|$$

y por tanto

$$\left\{ (f, x) \left| \frac{f(x+h_0)-f(x)}{h_0} \right| \leq n \right\}$$

es cerrado en $C([0, 1], \mathbb{R}) \times [0, 1 - \frac{1}{n}]$.

La proyección de este conjunto paralela al eje $[0, 1 - \frac{1}{n}]$ es entonces cerrada en $C([0, 1], \mathbb{R})$ (demuéstrelo) y es precisamente:

$$N(h_0) = \left\{ f \mid \exists x \in [0, 1 - 1/n] : \left| \frac{f(x+h_0)-f(x)}{h_0} \right| \leq n \right\}$$

$$\text{Como } \mathbb{N}_n = \cap \{N(h_0) \mid h_0 \in]0, 1/n]\}.$$

2. Cada N_n tiene interior vacío

Si $f \in N_n$ probemos que en cada bola $B(f, \epsilon)$ (tomando las bolas con respecto a la distancia d_∞), existe $g \in C([0, 1], \mathbb{R}) \setminus N_n$, lo que significa que

$$\text{i) } d_\infty(f, g) < \epsilon$$

$$\text{ii) } \forall x \in [0, 1 - \frac{1}{n}] \exists h_x \in]0, 1/n] : \left| \frac{g(x+h_x)-g(x)}{h_x} \right| > n$$

Por el teorema de Weierstrass de (aproximación por polinomios), que el lector puede encontrar detallado en el capítulo 2, existe un polinomio p tal que $d_\infty(f, p) < \epsilon/3$; de esta manera necesitamos una función que esté a una distancia menor que $\epsilon/3$ de p y teniendo la propiedad ii).

Como la derivada p' es continua sobre I tenemos:

$$M = \text{máx}\{|p'(x)| \mid 0 \leq x \leq 1\} \tag{24}$$

y definamos $S(x)$ como una función continua no negativa sobre I cuyo gráfico consta de segmentos de recta, siendo el valor absoluto de la pendiente de cada segmento $M + n + 1$. La gráfica nunca se eleva más de $\epsilon/3$ unidades por encima del eje x .

Entonces: $d_\infty(p + s, p) < \epsilon/3$ y $p + s$ es la función buscada ya que

$$\left| \frac{p(x+h) + s(x+h) - s(x) - p(x)}{h} \right| \geq \left| \frac{s(x+h) - s(x)}{h} \right| -$$

$$- \left| \frac{p(x+h) - p(x)}{h} \right| \geq ||| - |||$$

y para cada $x \in [0, 1 - \frac{1}{n}]$ podemos evidentemente hallar un $h \in [0, \frac{1}{n}]$ tal que el lado derecho sea mayor o igual que $M + n + 1 - M = n + 1$. ■

Aplicación 5: Integral de Lebesgue

Hasta el momento hemos visto cómo utilizando el concepto *límite de sucesiones*, puede introducirse la integral de Cauchy para funciones continuas. La necesidad de ampliar el campo de funciones a las cuales fuera posible calcular su integral, dio lugar al método de integración de Riemann, que coincide para las funciones continuas con el método de integración de Cauchy, el cual puede ser definido mediante el límite de una red.

El objeto de estudio del Análisis funcional son los espacios de funciones provistos de una estructura topológica, así como los operadores y funcionales lineales definidos sobre ellos. Existen dos clases muy importantes de espacios funcionales que resultan de gran utilidad en las aplicaciones: *los espacios de funciones continuas y los espacios de funciones integrables*. Hasta ahora, hemos estudiado algunas propiedades importantes de $C_a(X, Y)$ y en particular demostraremos que este espacio provisto de la métrica de la convergencia uniforme es completo si lo es el espacio métrico Y .

Esto, sin embargo, no ocurre para el espacio $C([a, b], \mathbb{R})$ provisto de alguna de las métricas integrales d_1 o d_2 mediante un cierto concepto de integral que garantiza la integral de Riemann y que se denomina integral de Lebesgue. Precisamente el completamiento de $(C([a, b], \mathbb{R}), d_1)$ que se denota usualmente por $L_1[a, b]$ está constituido por aquellas funciones definidas sobre $[a, b]$ que son integrales en el sentido de Lebesgue.

CAPÍTULO 4

COMPACIDAD EN ESPACIOS MÉTRICOS

Los subconjuntos cerrados y acotados de $\mathbb{R}^n$ tienen dos propiedades fundamentales en las cuales se basan algunos de los teoremas más importantes del análisis:

Propiedad 1: Todo cubrimiento abierto $(U_\alpha)_{\alpha \in I}$ de C contiene un subcubrimiento finito $(U_{\alpha_1}, \dots, U_{\alpha_r})$ de C .

Propiedad 2: Toda sucesión (x_n) de puntos de C tiene un punto adherente en C (es decir, contiene una subsucesión $(x_{\varphi(n)})$ convergente a un punto de C).

Los pioneros del análisis funcional como Ascoli y Volterra reconocieron la gran importancia fundamental que tienen estas propiedades para la investigación de conjuntos de funciones. En efecto, el descubrimiento de que toda una serie de teoremas clásicos sobre subconjuntos de $\mathbb{R}^n$, pueden trasladarse a conjuntos de funciones que verifican estas propiedades, fue uno de los motivos centrales para el desarrollo del análisis funcional.

Teniendo en cuenta la gran utilidad de las propiedades 1 y 2 en el análisis y en análisis funcional, es natural que nos dediquemos a estudiar la clase de todos los espacios métricos que verifican estas mismas propiedades; estos espacios se llaman *espacios métricos compactos*. Históricamente el desarrollo de la noción de compacidad en espacios topológicos generales (que no veremos en este texto), se complicó por la equivalencia de criterios en los espacios $\mathbb{R}^n$ que no son equivalentes en un espacio topológico cualquiera.

En 4.1 se presentan las propiedades básicas de la clase de los espacios métricos compactos y se demuestra que al igual que en los espacios $\mathbb{R}^n$, en cualquier espacio métrico, son equivalentes las diferentes nociones de compacidad. El epígrafe 4.2 está dedicado a algunos otros conceptos relacionados con la noción de compacidad como son la acotación total y la propiedad de Lindelöf. En 4.3 se analizan las funciones continuas sobre espacios compactos y se generaliza toda una serie de teoremas conocidos del análisis. En 4.4 se demuestran los teoremas importantes de Tijonov y Arzelá y se discuten algunas de sus consecuencias, las cuales son fundamentales para las aplicaciones en el análisis funcional, las ecuaciones diferenciales y la teoría de las funciones

analíticas. Los espacios clásicos $\mathbb{R}^n$ y $\mathbb{Q}^n$ no son compactos, sin embargo, cada uno de sus puntos posee un sistema fundamental de vecindades compactas; tales espacios se llaman *localmente compactos*. En 4.5 se estudian los espacios compactos, en él se muestran cómo una gran variedad de espacios métricos pueden ser extendidos por diversos métodos a espacios métricos compactos, ofreciéndose algunas propiedades de tales *compactificaciones*, especialmente para las asociadas a los espacios localmente compactos.

4.1 ESPACIOS Y SUBCONJUNTOS COMPACTOS

Para espacios métricos en general, las dos propiedades mencionadas en la introducción son equivalentes. Más exactamente, se puede demostrar el teorema siguiente:

Teorema 4.1.1 (criterios de compacidad). En cualquier espacio métrico (X, d) las siguientes propiedades son equivalentes:

- (C₁) Todo cubrimiento abierto de X contiene un subcubrimiento finito.
- (C₂) Toda familia de cerrados de X de intersección vacía, contiene una subfamilia finita también de intersección vacía.
- (C₃) Toda familia de cerrados de X en la que cualquier subfamilia finita es de intersección no vacía, es a su vez de intersección no vacía.
- (C₄) Todo suceso en X posee una subsucesión convergente (es decir toda sucesión de elementos de X posee un punto adherente en X).

Observación: Notemos que los tres primeros enunciados que hablan de familias de subconjuntos, constituyen un grupo de propiedades correspondientes a la propiedad 1.

Demostración: $(C_1) \Leftrightarrow (C_2)$, ya que (C_1) y (C_2) son enunciados duales. En efecto si $(U_\alpha)_{\alpha \in I}$ es un cubrimiento abierto de X , entonces $(X \setminus U_\alpha)_{\alpha \in I}$ es una familia de cerrados de intersección vacía e inversamente, si $(C_\alpha)_{\alpha \in I}$ es una familia de cerrados en (X, d) de intersección vacía, entonces $(X \setminus C_\alpha)_{\alpha \in I}$ es una familia de abiertos que recubre X . $(C_2) \Leftrightarrow (C_3)$, ya que (C_3) no es más que otra versión de (C_2) .

Queda por probar la equivalencia entre el grupo de propiedades (C_1) , (C_2) , (C_3) , (C_4) .

$(C_3) \Leftrightarrow (C_4)$. Sea (x_n) una sucesión en (X, d) y consideremos para cada n la sección final:

$$S_n = \{x_m : m \geq n\}.$$

Es claro que $\{\overline{S_n}\}_{n \in \mathbb{N}}$ es una familia decreciente de cerrados, y por tanto, cualquier subfamilia finita tiene intersección no vacía. Luego por (C_3) se tendrá que existe

$$x \in \bigcap_{n \in \mathbb{N}} \overline{S_n}.$$

Pero entonces, el punto x es necesariamente un punto adherente de la sucesión (x_n) y como ya sabemos, en este caso existe una subsucesión (x_{φ_n}) de la sucesión (x_n) tal que $(x_{\varphi_n}) \longrightarrow x$.

Para demostrar la implicación $(C_4) \Rightarrow (C_1)$ necesitaremos un lema auxiliar pero primeramente daremos las definiciones siguientes:

Definición 4.1.2. Se dice que el cubrimiento $\mathbf{V} = \{V_j\}_{j \in J}$ es más fino que el cubrimiento $U = \{U_\alpha\}_{\alpha \in I}$ si para todo $j \in J$ existe $\alpha \in I$ tal que $V_j \subset U_\alpha$

Definición 4.1.3. Un subconjunto finito $F = \{x_1, x_2, \dots, x_n\}$ del espacio métrico (X, d) se llama una ϵ -red, si las bolas $\{B(x_i, \epsilon) : i = 1, \dots, n\}$ constituyen un cubrimiento de X . expresado de otra forma, $d(x, F) < \epsilon$, para todo $x \in X$ (fig. 63).

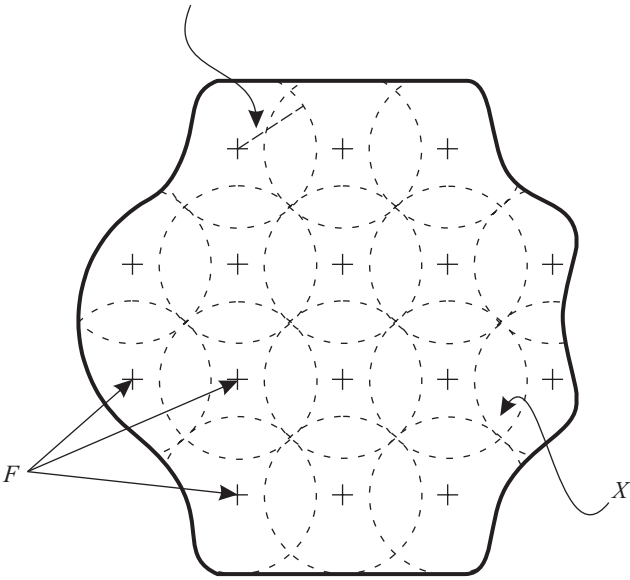


Figura 63

Lema 4.1.4. En un espacio métrico (X, d) con la propiedad (C_4) se cumple:

- i) Para todo cubrimiento abierto $U = \{U_\alpha\}_{\alpha \in I}$ de X , existe un número $\epsilon > 0$ (llamado número de Lebesgue) tal que el cubrimiento $\{B(x, \delta) : x \in X\}$ formado por todas las bolas de radio δ es más fino que U .
- ii) Para todo $\epsilon > 0$ existe una ϵ -red en X .

Demostración: i) Supongamos por el absurdo que existe un cubrimiento abierto U de X para el cual no existe algún número de Lebesgue. Esto significa que para todo $n \in \mathbb{N}^*$ existe un punto $x_n \in X$ tal que la bola $B(x_n, 2^{-n})$ no está contenida en algún $U \in \mathbf{U}$.

Por hipótesis (x_n) tiene un punto de adherencia x . Sea $x \in U \in \mathbf{U}$. Existe un $k \in \mathbb{N}^*$ tal que $B(x, 2^{-k}) \subset U$. Entonces, según la desigualdad triangular, se cumple para todo $y \in B(x, 2^{-k-1})$,

$$B(y, 2^{-k-1}) \subset B(x, 2^{-k}) \subset U,$$

Esto implica que $B(y, 2^{-n}) \subset U$ para todo $n \geq k + 1$. Luego, según la definición de (x_n) , ninguno de los puntos con $n \geq k + 1$ pueden pertenecer a $B(x, 2^{-k-1})$; pero esto contradice la suposición de que x es un punto adherente de (x_n) .

ii) Supongamos por el absurdo que ii) es falso. Entonces hay un $\epsilon > 0$ tal que X no puede cubrirse por un número finito de bolas de radio ϵ . Por inducción puede seleccionarse una sucesión de puntos $x_n \in X$ tal que

$$x_{n+1} \in X \setminus \bigcup_{i=1}^n B(x_i, \epsilon) \quad (n \in \mathbb{N}).$$

Esto implica que $d(x_n, x_m) \geq \epsilon$ para $m, n \in \mathbb{N}$ con $m \neq n$.

Concluimos que para todo $x \in X$, la vecindad $B(x, \epsilon/2)$ contiene a lo sumo un punto de la sucesión (x_n) . Luego, (x_n) no tiene ningún punto de adherencia en X , con lo cual llegamos a una contradicción. ■

Continuemos ahora la demostración de la implicación $(C_4) \Rightarrow (C_1)$ del teorema 4.1.1.

$(C_4) \Rightarrow (C_1)$ sea U un cubrimiento de X . Según i) del lema 4.1.4, existe un número de Lebesgue δ para U . Según ii) existe una sucesión finita $x_1, \dots, x_n$ tal que $X = \bigcup_{i=1}^n B(x_i, \delta)$.

Para $k = 1, \dots, n$ existen abiertos $U_{i_k} \in U$ que contienen a las bolas respectivas $B(x_k, \delta)$. Concluimos de esta forma, que $\{U_{i_1}, \dots, U_{i_n}\}$ es un subcubrimiento finito de U con lo cual se obtiene (C_1) . ■

Observación: Hemos probado que de las propiedades i)-ii) del lema 4.1.4,

se deduce cualquiera de las propiedades (C_1) – (C_4) del teorema 4.1.1. El lector puede comprobar que en realidad las propiedades i)–ii) son equivalentes a cada una de las (C_i) ($i = 1, \dots, 4$).

Definición 4.1.5. Un espacio métrico (X, d) se dice compacto si cumple alguna de las propiedades equivalentes (C_1) – (C_4) . Decimos que un *subconjunto* C de (X, d) es *compacto* si lo es el subespacio métrico (C, d_C) .

La definición anterior, es una de las definiciones más importantes de la topología general.

Observación: La noción de compacidad no depende intrínsecamente de la métrica ya que es una propiedad topológica; es decir, si el espacio (X, d) es compacto y d' es otra métrica en X topológicamente equivalente a d , entonces (X, d') también es compacto. En efecto, si toda sucesión (x_n) en (X, d') tiene una subsucesión convergente $x_{\varphi(n)} \xrightarrow{d} x$, entonces, $x_{\varphi(n)} \xrightarrow{d'} x$ y, por lo tanto, (X, d') satisface la propiedad (C_4) .

Es conveniente tener presente en el trabajo diario todas las caracterizaciones (C_1) – (C_4) de los espacios métricos compactos, puesto que en diferentes aplicaciones puede resultar más útil una que otra. Por la misma razón, es conveniente reescribir la definición de subconjunto compacto. De acuerdo con el concepto de *subespacio métrico*, C en (X, d) es compacto si y sólo si para toda familia de abiertos $(A_\alpha)_{\alpha \in I}$ de (X, d) tal que

$$C \subset \bigcup_{\alpha \in I} A_{\alpha_k} (\Leftrightarrow C = \bigcup_{\alpha \in I} (C \cap A_\alpha))$$

existe una subfamilia finita $(A_{\alpha_k})_{1 \leq k \leq n}$ para la cual también

$$C \subset \bigcup_{k=1}^n A_{\alpha_k} (\Leftrightarrow C = \bigcup_{k=1}^n (C \cap A_{\alpha_k}))$$

Esta forma es más *práctica por lo que se utiliza más*. ¿De qué manera se reescriben (C_2) – (C_4) ?

Observación: La definición de subconjunto compacto se establece de tal modo que resulta invariante por sobreespacios, es decir, si (X, d) es un subespacio métrico de (Y, d') se tiene que C es un subconjunto compacto de (X, d) si y sólo si $C \subset X$ es un subconjunto compacto de (Y, d') , lo que es claro pues $(C, d_C) = (C, d'_C)$.

Ejemplo 1. Un espacio métrico discreto es compacto si y sólo si es finito.

Ejemplo 2. Sea (X, d) un espacio métrico y sea (x_n) una sucesión en (X, d) que converge a un punto $x \in X$. Entonces el subconjunto $C = \{x_n : n \in$

$\mathbb{N}\} \cup \{x\}$ es compacto en (X, d) ya que si (U_α) es un cubrimiento abierto de C y si $x \in U_\beta$ para un cierto $\beta \in I$, entonces casi todos los x_n están contenidos en U_β y los puntos x_n que se quedan fuera pueden cubrirse por un número finito de U_α .

Ejemplo 3. La recta real $\mathbb{R}$ no es compacta, pues el cubrimiento abierto $\{]-n, n[: n \in \mathbb{N}^*\}$ no contiene un subcubrimiento finito de $\mathbb{R}$. Sin embargo, como ya sabemos, todo intervalo cerrado y acotado $[a, b]$ es compacto. Analicemos brevemente la demostración que se basa en el criterio (C_1) .

Sea $(U_x)_{x \in I}$ un cubrimiento abierto de $[a, b]$ en $\mathbb{R}$, y sea s el supremo de todos los números $r \in [a, b]$ tales que los subintervalos $[a, r]$ pueden cubrirse por un número finito de los $U_x (x \in I)$. Obviamente s está contenido en algún $U_x (x \in I)$. Puesto que U_x es abierto en $\mathbb{R}$ existe $\epsilon > 0$ tal que $[s - \epsilon, s + \epsilon] \subset U_x$. El subintervalo $[a, s - \epsilon]$ puede cubrirse por una subfamilia $(U_{x_1}, \dots, U_{x_n})$ de abiertos. Por lo tanto la subfamilia $(U_{x_1}, \dots, U_{x_n}, U_x)$ de $(U_x)_{x \in I}$ recubre $[a, s + \epsilon]$. Si $s < b$ esto significaría una contradicción, por lo que concluimos que $s = b$ y que $[a, b]$ puede cubrirse por un número finito de $U_x (x \in I)$.

Si queremos verificar el criterio (C_1) , podemos limitarnos a cubrimientos formados por los abiertos de una subbase. Esto facilita en muchos casos, la verificación de la compacidad. Analicemos brevemente la demostración.

Teorema 4.1.6 (teorema de Alexander). Sea B una base topológica del espacio métrico (X, d) . Entonces (X, d) es compacto si todo cubrimiento de X por elementos de B contiene un subcubrimiento finito.

Demostración: Supongamos por el absurdo que (X, d) no es compacto. Entonces existe una sucesión (x_n) en X que no tiene algún punto adherente. Por lo tanto, para todo $x \in X$ se puede hallar un abierto $U_x \in B$ tal que U_x contenga a lo sumo un elemento de la sucesión (x_n) (si $x = x_n$ para algún $n \in \mathbb{N}$).

Obviamente el cubrimiento,

$$U = \{U_x : x \in X, U_x \in B\}$$

no tiene algún subcubrimiento finito, ya que cualquier subfamilia finita de elementos de U deja fuera infinitos puntos de la sucesión (x_n) .

La elaboración detallada de la demostración, se deja de ejercicio al lector.

■

A continuación estudiaremos algunas propiedades elementales de los sub-

conjuntos compactos

Ejemplo 4. Si no consideramos en el compacto $[0, 1]$ el punto 1, se obtiene un subespacio que no es compacto. La sucesión $(1 - 1/n)$ no tiene algún punto de adherencia en $[0, 1[$

Este ejemplo nos muestra que la compacidad no se transmite a subespacios cualesquiera, sin embargo tenemos:

Proposición 4.1.7. En un espacio métrico compacto, todo subconjunto cerrado es también compacto.

Demostración: Sea C un subconjunto cerrado de un espacio métrico (X, d) , apliquemos el criterio (C_4) . Sea (x_n) una sucesión en C , por ser (X, d) compacto, (x_n) tiene un punto adherente x en X . Como en particular C es cerrado, el punto x pertenece a C . Luego C cumple el criterio (C_4) y por lo tanto es compacto. ■

Recíprocamente se tiene:

Proposición 4.1.8. Todo subconjunto compacto de un espacio métrico es cerrado y acotado.

Demostración: Sea C un subconjunto compacto del espacio métrico (X, d) y sea $x \in \overline{C}$. Entonces existe una sucesión (x_n) en C tal que $(x_n) \rightarrow x$. Luego, el único punto adherente de la sucesión (x_n) es el punto x y por ser C compacto se tiene que $x \in C$ de donde se deduce que $\overline{C} \subset C$.

Por otra parte, si fijamos un punto $x \in C$, la familia de abiertos formada por las bolas $\{B(x, n)\}_{n \in \mathbb{N}^*}$ es un cubrimiento abierto de C , y por ser C compacto, de dicho cubrimiento se podrá extraer un subcubrimiento finito $\{B(x, n_k)\}_{k=1}^m$. De aquí se obtiene que $C \subset B(x, \max_{1 \leq k \leq m} n_k)$, y por tanto, C es acotado. ■

Corolario 1. Sea (X, d) un espacio métrico compacto y $C \subset X$. Entonces C es compacto si y sólo si C es cerrado.

La demostración se deja al lector.

Veamos mediante un ejemplo que un subconjunto acotado de un espacio métrico no es necesariamente compacto.

Ejemplo 5. La bola unitaria en el espacio métrico $(C([0, 1], \mathbb{R}), d)$ no es un subconjunto compacto, ya que la sucesión $f_n(x) = x^n, n = 1, 2, \dots$, no tiene alguna subsucesión convergente en $C([0, 1], \mathbb{R})$. En efecto, para cualquier suce-

sión (n_k) de números naturales, la sucesión de funciones $f_{n_k}(x) = x^{n_k}$ converge puntualmente a 0 si $x \in [0, 1[$ y a 1 en el punto $x = 1$. Si f_{n_k} convergiera uniformemente a una cierta función f , también convergería puntualmente a f , luego necesariamente

$$f(x) = \begin{cases} 0, & \text{si } x \in [0, 1[\\ 1, & \text{si } x = 1 \end{cases}$$

y por tanto, f no es continua.

Ejemplo 6 (caracterización de los subconjuntos compactos en $\mathbb{R}$).

Todo subconjunto cerrado y acotado en $\mathbb{R}$ es compacto, puesto que es un subconjunto cerrado del intervalo compacto $[\inf C, \sup C]$. De los resultados obtenidos en la proposición 4.1.8 y en el corolario 1, concluimos la conocida caracterización del análisis para los subconjuntos compactos de $\mathbb{R}$:

C es compacto en $\mathbb{R} \Leftrightarrow C$ es cerrado y acotado.

De 4.1.7 y del corolario 1 se deduce que en un espacio métrico la intersección de una familia cualquiera de compactos es compacto. Como se observa, existe una analogía formal con la propiedad básica (C_3) de los cerrados. El análogo de la propiedad (C_2) se cumple también para los subconjunto compactos en un espacio métrico cualquiera.

Proposición 4.1.9. En todo espacio métrico la unión finita de subconjuntos compactos es compacto.

La demostración se deja al lector.

En resumen, en todo espacio métrico, la colección de los subconjuntos compactos es estable para la intersección (finita o infinita) y la unión finita. En los ejercicios de autocontrol 3, 4 y 5 al final del epígrafe, se muestran propiedades de las familias de compactos que no se cumplen en general para familias de cerrados.

Para terminar esta epígrafe solo nos queda introducir una noción técnica que no presenta ninguna dificultad conceptual: *los subconjuntos relativamente compactos*.

En $\mathbb{R}$ la clausura de un acotado es compacto. Toda sucesión acotada en $\mathbb{R}$ tiene un punto adherente en $\mathbb{R}$. Esta propiedad de los acotados en $\mathbb{R}$ puede generalizarse a espacios métricos cualesquiera.

Definición 4.1.10. Un subconjunto C de un espacio métrico (X, d) se llama *relativamente compacto*, si su clausura en (X, d) es compacto.

Observación: Esta noción depende esencialmente de la estructura topológica

correspondiente al sobreespacio y no solo de la definida por la métrica inducida. El intervalo $]0, 1[$ es relativamente compacto en $[0, 1]$ pero no en $[0, 1[$; por lo tanto, es necesario expresar siempre con claridad cuál es el espacio respecto al cual se toma la clausura.

Según 4.1.7 en un espacio métrico compacto, todo subconjunto es relativamente compacto. Además, en cualquier espacio métrico todo subconjunto compacto es cerrado y por lo tanto relativamente compacto.

En $\mathbb{R}$ un subconjunto es relativamente compacto si y solo si es acotado. Esto nos sugiere generalizar el teorema clásico de que toda sucesión acotada en $\mathbb{R}$ tiene un punto de adherencia a los conjuntos relativamente compactos.

Proposición 4.1.11. Sea (X, d) un espacio métrico y C un subconjunto relativamente compacto de X . Entonces toda sucesión de puntos de C tiene un punto adherente en X .

La demostración se deja al lector.

La analogía formal entre los subconjuntos acotados de un espacio métrico y los subconjuntos relativamente compactos, pueden expresarse diciendo que ambos forman una bornología. Este concepto ha sido introducido para tratar axiomáticamente los conjuntos acotados precindiendo de una métrica. En el presente texto no vamos estudiar este concepto.

No queremos desaprovechar la oportunidad para destacar una vez más la gran variedad de estructuras matemáticas: la estructura topológica, la estructura uniforme y la estructura bornológica, las cuales pueden ser introducidas gracias al concepto de *métrica*. Es importante que el lector comprenda, al menos intuitivamente, que existen diferencias entre los diferentes tipos de estructuras y sobre todo cuáles de las propiedades estudiadas corresponden a cada una de ellas.

Ejercicios de Autocontrol

1. Demuestre los enunciados 4.1.9 y 4.1.11 y el corolario 1 de las proposiciones 4.1.7 y 4.1.8.
2. Sin utilizar la normalidad de los espacios métricos, demuestre que dos conjuntos compactos disjuntos en un espacio métrico, pueden ser separados por vecindades disjuntas. Demuestre, además, que las vecindades se pueden tomar de manera tal que la distancia entre ellas sea estrictamente positiva.
3. Demuestre que cualquiera de las tres propiedades siguientes es equivalente a la compacidad del espacio métrico $(X, d) : (\text{NC}_1)$ Todo cubrimien-

to abierto numerable de X contiene un subcubrimiento finito. (NC₂) Toda sucesión decreciente de cerrados no vacíos en X tiene intersección no vacía.

(NC₃) Toda familia numerable de cerrados en la cual cualquier subfamilia finita es de intersección no vacía, es a su vez de intersección no vacía.

4. Sea (C_n) una sucesión decreciente de compactos no vacíos en el espacio métrico (X, d) y sea $C = \bigcap_{n=1}^{\infty} C_n$. Pruebe que si U es una vecindad de C , entonces existe $n \in \mathbb{N}$ tal que $C_n \subset C$.
5. Muestre mediante ejemplos que, en general, la familia de los subconjuntos cerrados de un espacio métrico cualquiera no satisface necesariamente las propiedades (NC₂) y (NC₃) del ejercicio 3 ni la propiedad expuesta en el ejercicio 4.
6. Demuestre que un espacioseudométrico también son equivalentes las propiedades (C₁–C₄).
7. Si definimos la compacidad en los espaciosseudométricos mediante las propiedades equivalentes (C₁–C₄), enuncie varios resultados obtenidos para los subconjuntos compactos y relativamente compactos en los espacios métricos que no puedan ser trasladados a los espaciosseudométricos.

4.2 ALGUNOS CONCEPTOS RELACIONADOS CON LA NOCIÓN DE COMPACIDAD

La caracterización (NC₂) de la compacidad mostrada en el ejercicio de autocontrol 3 del epígrafe 4.1, tiene una expresión similar a la dada en el teorema de Cantor, donde se expone una condición equivalente a la completitud. En (NC₂) solamente se exige una condición adicional de uniformidad, ya que la sucesión de los diámetros $\delta(F_n)$ debe tender a cero (donde $\{F_n\}$ es una sucesión decreciente de subconjuntos cerrados no vacíos). Este nos demuestra que en los espacios métricos existe una cierta relación entre los conceptos *compacidad* y *completitud*.

En efecto, de (NC₂) y el teorema de Cantor se deduce:

Todo espacio métrico compacto es completo. (1)

En el presente epígrafe explicaremos exhaustivamente la relación entre la compacidad y completitud en los espacios métricos, lo cual puede hacerse mediante la noción de acotación total.

Finalmente introduciremos el concepto de espacio de Lindelöf que se define, al igual que la compacidad, a partir de cubrimientos por conjuntos abiertos, lo cual resulta de gran importancia en las aplicaciones de la Topología General.

Definición 4.2.1. El espacio métrico (X, d) se llama *totalmente acotado* si para todo $\epsilon > 0$ hay una ϵ -red (4.1.1) en X . Diremos que el subconjunto A de X es *totalmente acotado* si lo es el subespacio (A, d_A) .

Observación: En algunos textos se llama precompactos a los espacios totalmente acotados.

No es difícil comprobar que el conjunto A es totalmente acotado en (X, d) si y solo si para cada $\epsilon > 0$ existe un subconjunto finito $F = \{x_1, \dots, x_n\} \subset A$ tal que

$$A \subset \bigcup_{i=1}^n B(x_i, \epsilon).$$

Del lema 4.1.4 (ii) se obtiene:

$$\text{Todo espacio métrico compacto es totalmente acotado} \quad (2)$$

Veamos algunas propiedades elementales de los subconjuntos totalmente acotados.

Proposición 4.2.2 (propiedades de los subconjuntos totalmente acotados). Sea A un subconjunto totalmente acotado del espacio métrico (X, d) ; entonces:

- i) A es acotado;
- ii) $\overline{A}$ es totalmente acotado
- iii) A es separable.

Demostración: i) Se deja al lector.

ii) Sea $\epsilon > 0$; por ser A totalmente acotado, existe una $\epsilon/2$ -red, tal que $F = \{x_1, \dots, x_n\}$ está contenida en A . Entonces F es una ϵ -red de $\overline{A}$. En efecto,

$$d\left(A, X \setminus \bigcup_{i=1}^n B(x_i, \epsilon) \geq \epsilon/2\right)$$

ya que

$$A \subset \bigcup_{i=1}^n B(x_i, \epsilon/2).$$

Como la distancia de cualquier punto del complemento de $\bigcup_{i=1}^n B(x_i, \epsilon)$ al conjunto A es estrictamente positiva, tendremos que ninguno de los puntos mencionados puede estar en $\overline{A}$ ya que para todo $x \in \overline{A}$, $d(x, A) = 0$. Luego,

$$\overline{A} \subset \bigcup_{i=1}^n B(x_i, \epsilon),$$

por lo tanto, F es una ϵ -red para $\overline{A}$. Por la arbitrariedad de ϵ se tiene que $\overline{A}$ es totalmente acotado.

iii) Por ser A totalmente acotado para cada $n \in \mathbb{N}^*$ existe una $\frac{1}{n}$ -red F_n de A en (X, d) .

El subconjunto

$$F = \bigcup_{n=1}^{\infty} F_n, \text{ donde } F_n = \{x_i^{(n)}, \dots, x_{k(n)}^{(n)}\}$$

es un subconjunto numerable denso de A .

En efecto, si $x \in A$, entonces para cada n se puede encontrar $x_i^{(n)} \in F$ ($1 \leq i \leq k(n)$) tal que $d(x_i^{(n)}, x) < 1/n$. Si denotamos $x_i^{(n)} = a_n$, entonces la sucesión (a_n) converge a x , por lo tanto, F es denso en A , de donde concluimos que A es separable. ■

Ejemplo 7. En el espacio euclideo n dimensional $\mathbb{R}^n$, la acotación total coincide con la acotación corriente, es decir, con la posibilidad de sumergir el conjunto dado dentro de un cubo suficientemente grande. En efecto, si para cada $\epsilon > 0$ dividimos este cubo en cubos con aristas de longitud ϵ , entonces el subconjunto finito constituido por los vértices de estos últimos formarán una $\frac{\sqrt{n}}{2}\epsilon$ -red en el cubo inicial y, por consiguiente, en cualquier conjunto, contenido en este cubo.

Ejemplo 8. La esfera unitaria S del espacio $l^2(\mathbb{N})$ de las sucesiones de cuadrado sumable, ofrece un ejemplo de un conjunto acotado, pero no totalmente acotado. En efecto, consideremos los siguientes puntos de S .

$$\begin{aligned} e_1 &= (1, 0, 0, \dots), \\ e_2 &= (0, 1, 0, \dots), \\ &\dots\dots\dots \\ e_n &= (0, 0, 0, \dots, 1, 0, \dots) \\ &\dots\dots\dots \end{aligned}$$

La distancia entre dos cualesquiera de estos puntos e_n y e_m ($n \neq m$) es igual a $\sqrt{2}$. De aquí se desprende que en S no puede existir una ϵ -red finita

para ningun $\epsilon < \frac{\sqrt{2}}{2}$

Ejemplo 9. Consideremos en $l^2(\mathbb{N})$ el conjunto H de puntos

$$x = (x_1, x_2, \dots, x_n, \dots)$$

que verifican las siguientes condiciones: $|x_1| \leq 1, |x_2| \leq 1/2, \dots, |x_n| \leq 1/2^{n-1}, \dots$ Este conjunto se llama *paralelepipedo fundamental* (o tambien, *cubo de Hilbert*) del espacio $l^2(\mathbb{N})$, y representa un ejemplo de un conjunto totalmente acotado de dimension infinita. Para demostrar que es totalmente acotado podemos proceder de la forma siguiente:

Sea $\epsilon > 0$ dado. Escojamos n de manera que $\frac{1}{2^{n-1}} < \frac{\epsilon}{2}$. A todo punto $x = (x_1, x_2, \dots, x_n, \dots)$ de H haremos corresponder el punto $x^* = (x_1, x_2, \dots, x_n, 0, 0, \dots)$ de este mismo conjunto.

En este caso

$$\rho(x, x^*) = \sqrt{\sum_{k=n+1}^{\infty} |x_k|^2} \leq \sqrt{\sum_{k=n}^{\infty} \frac{1}{4^k}} < \frac{1}{2^{n-1}} < \frac{\epsilon}{2}.$$

El conjunto H_n^* formado por los puntos de H del tipo x^* , es totalmente acotado por ser un subconjunto acotado de un espacio de dimension finita n . Tomemos en H_n^* una $\epsilon/2$ -red finita. (Esta claro que sera al mismo tiempo una ϵ -red para H .)

Hemos demostrado que todo espacio compacto es necesariamente acotado. No resulta difcil comprobar que esta condicion no es suficiente; por ejemplo, el conjunto de puntos racionales del intervalo $[0, 1]$ con la metrica inducida por el modulo, es un espacio totalmente acotado y sabemos que no es compacto. Sin embargo tiene lugar el teorema siguiente:

Teorema 4.2.3. Para que un espacio (X, d) sea compacto, es necesario y suficiente que sea:

i) totalmente acotado;

ii) completo.

Demostración: La necesidad de las condiciones i) y ii) ya la hemos asegurado en (1) y (2). Probemos ahora que si se cumplen las condiciones i) y ii) entonces (X, d) es compacto. Para ello es suficiente demostrar que toda sucesion (x_n) de puntos de X tiene al menos un punto adherente.

Sea F_1 una ϵ -red en (X, d) . Las bolas cerradas de radio $\epsilon/2$ centradas en los puntos de F_1 constituyen un cubrimiento finito de X . Luego, al menos una de estas bolas, llamemosla B_1 , contiene una subsucesion infinita $x_1^{(1)}$,

$\dots, x_n^{(1)}, \dots$ de la sucesión $\{x_n\}$. Escojamos ahora en B_1 una $1/2$ -red F_2 y construyamos alrededor de todo punto de esta red una bola cerrada de radio $1/2$. Al menos una de estas bolas, llamémosla B_2 , contiene una subsucesión infinita $x_1^{(2)}, \dots, x_n^{(2)}, \dots$, de la sucesión $(x_n^{(1)})$.

Así sucesivamente, por inducción, podemos encontrar para cada $k \in \mathbb{N}$ una bola cerrada B_k de radio $1/2^{k-1}$ que contiene una subsucesión infinita $(x_n^{(k)})$ de la sucesión $(x_n^{(k-1)})$. Es fácil ver que la familia (B_k) es una sucesión decreciente de conjuntos cerrados con diámetro que tiende a cero y por el teorema de Cantor, su intersección $\bigcap_{k=1}^{\infty} B_k$ consta de un solo punto x_0 . Este punto es un punto de adherencia de la sucesión inicial (x_n) , ya que toda vecindad suya contiene a alguna bola B_k y, por consiguiente, una subsucesión infinita $(x_n^{(k)})$ de la sucesión (x_n) . ■

Corolario 1. Un subconjunto A de un espacio métrico completo (X, d) es relativamente compacto si y solo si es totalmente acotado.

Demostración: En efecto, se tiene la siguiente cadena de equivalencias:

A es relativamente compacto $\Leftrightarrow \overline{A}$ es compacto $\Leftrightarrow \overline{A}$ es completo y totalmente acotado.

Pero $\overline{A}$ es siempre completo por ser cerrado en el espacio métrico completo (X, d) , de donde finalmente se obtiene el resultado deseado. ■

Hemos visto que la compacidad es una propiedad más exigente que la completitud en los espacios métricos. Por lo que es de esperar que algunos resultados obtenidos en los espacios métricos completos adquieran una formulación más amplia en los espacios compactos. A modo de ejemplo veamos el siguiente principio fuerte de las aplicaciones contraídas en los espacios métricos compactos.

Proposición 4.2.4 (principio fuerte de aplicaciones contraídas). Sea (X, d) un espacio métrico compacto y f una aplicación de (X, d) en sí mismo tal que

$$d(f(x), f(y)) < d(x, y) \text{ para } x \neq y.$$

Entonces la aplicación f tiene en X un punto fijo único.

Demostración: Para la demostración se recomienda al lector seguir las indicaciones expuestas en el ejercicio 4 de autocontrol de 4.3. ■

Introduzcamos un último concepto en este epígrafe que no es otra versión de la compacidad, sino un nuevo concepto de propia importancia. Lindelöf de-

mostró a finales del siglo pasado que todo cubrimiento abierto de $\mathbb{R}$ contiene un subcubrimiento numerable. Este resultado, sorprendente en su tiempo, ha promovido la introducción de una noción topológica relacionada simultáneamente con la compacidad y los axiomas de numerabilidad.

Definición 4.2.5. Un espacio métrico (X, d) se dice que cumple el axioma (L) o que es un espacio de *Lindelöff*, cuando todo cubrimiento abierto de X contiene un subcubrimiento numerable.

De esta definición se deduce inmediatamente que todo espacio métrico compacto es también de *Lindelöff* aunque no se cumple el inverso según nos muestra el caso de $\mathbb{R}$.

Recordemos que en un espacio métrico todo punto tiene un sistema fundamental numerable de vecindades; se dice entonces que todo espacio métrico satisface el *primer axioma de numerabilidad*-AN1; mas, no todo espacio métrico tiene una base topológica numerable, es decir, no todo espacio métrico satisface el *segundo axioma de numerabilidad*-AN2. Sin embargo, se tiene el teorema siguiente:

Teorema 4.2.6 (Teorema de Lindelöff). Sea (X, d) un espacio métrico; entonces son equivalentes:

- i) (X, d) satisface el segundo axioma de numerabilidad;
- ii) toda base topológica de (X, d) contiene una base numerable;
- iii) (X, d) tiene un subconjunto denso numerable (es separable);
- iv) (X, d) es de Lindelöff

Demostración: La equivalencia i) $\Leftrightarrow$ iii) fue ya probada en el capítulo 3.

Probemos que i) $\Leftrightarrow$ iv).

Para ello veamos primeramente que si $\mathbf{B}$ es una base topológica de (X, d) entonces (X, d) es de Lindelöff cuando todo cubrimiento de X por elementos de $\mathbf{B}$ contiene un subcubrimiento numerable.

En efecto, sea $\mathcal{U}$ otro cubrimiento abierto cualquiera de X . Para todo $x \in X$ existe un abierto $U_x \in \mathcal{U}$ y un elemento de la base topológica $B_x \in \mathbf{B}$ tal que $x \in B_x \subset U_x$. Puesto que los $B_x (x \in X)$ recubren X , existe una familia numerable (x_n) tal que $X = \bigcup_{n \in \mathbb{N}} B_{x_n}$, luego (U_{x_n}) es un subcubrimiento de X .

De aquí, se deduce evidentemente que si (X, d) tiene una base numerable, entonces es de Lindelöff.

Probemos ahora que iv) $\Rightarrow$ i).

Para el espacio métrico de Lindelöf (X, d) , consideremos los cubrimientos abiertos, $\{B(x, 1/k) : x \in X\}$. Para cada $k \in \mathbb{N}^*$ existe un subcubrimiento numerable $\{B(x_n^{(k)}, 1/k) : k \in \mathbb{N}^*, (n \in \mathbb{N})\}$ que es una base topológica para (X, d) . Sea $U \in \theta_d$ y $x \in U$. Existe $\epsilon > 0$ tal que $B(x, \epsilon) \subset U$. Escojamos $k \in \mathbb{N}$ tal que $k > 2/\epsilon$; como $\{B(x_n^{(k)}, 1/k) : n \in \mathbb{N}\}$ es un cubrimiento de X , existe $n \in \mathbb{N}$, tal que $x \in B(x_n^{(k)}, 1/k)$.

Por otra parte si $y \in B(x_n^{(k)}, 1/k)$ entonces

$$d(x, y) \leq d(x, x_n^{(k)}) + d(x_n^{(k)}, y) < 2/k < \epsilon$$

luego

$$B(x_n^{(k)}, 1/k) \subset B(x, \epsilon) \subset U.$$

Nos queda por probar la equivalencia $i) \Leftrightarrow ii)$ de la cual la implicación $ii) \Rightarrow i)$ es evidente.

Probemos entonces que $i) \Rightarrow ii)$. Sean $\mathbf{B} = \{B_n : n \in \mathbb{N}\}$ una base topológica numerable de (X, d) y $\mathbf{B}'$ otra base topológica cualquiera. Todo $B_n (n \in \mathbb{N})$ es cubierto por los abiertos $U \in \mathbf{B}'$ tal es que $U \subset B_n$.

B_n , como subespacio, cumple (AN 2), luego por la implicación $i) \Rightarrow iv)$ se tiene que es de Lindelöf. Por lo tanto, existe un subcubrimiento numerable $\mathbf{B}'_n \subset \mathbf{B}'$ de B_n . Concluimos que $\bigcup_{n \in \mathbb{N}} \mathbf{B}'_n$ es una base numerable de (X, d) contenida en $\mathbf{B}'$. ■

Ejercicios de Autocontrol

1. Demuestre el principio fuerte de las aplicaciones contraídas 4.2.4.
2. Pruebe que la unión de un número finito de conjuntos totalmente acotados es un conjunto totalmente acotado.
3. Demuestre que la unión de una familia numerable de subespacios de Lindelöf en un espacio métrico es un subespacio de Lindelöf.
4. En un espacio métrico discreto, diga cuales son los subconjuntos totalmente acotados, compactos y de Lindelöf.

4.3 COMPACTOS Y APLICACIONES CONTINUAS

A partir del criterio (C_1) resulta inmediato que la compacidad es una propiedad topológica, o sea, es una propiedad invariante mediante homeomorfismos; en

realidad se cumple mas: *la compacidad es invariante ante aplicaciones continuas.*

Teorema 4.3.1. La imagen de un espacio métrico (subconjunto) compacto por una aplicación continua es un compacto.

Demostración: Sea (X, d) un espacio métrico compacto y f una aplicación continua de (X, d) en otro espacio métrico (Y, d') . Verificaremos que $f[X] \subset Y$ satisface el criterio (C_1) . Sea $(U_\alpha)_{\alpha \in I}$ un cubrimiento de $f[X]$, abierto en (Y, d') . Entonces $(f^{-1}[U_\alpha]_{\alpha \in I})$ es un cubrimiento abierto del compacto (X, d) . Según (C_1) existe un subcubrimiento finito de X es decir

$$X \subset \bigcup_{k=1}^r f^{-1}[U_{\alpha_k}] = f^{-1} \left[\bigcup_{k=1}^r U_{\alpha_k} \right]$$

de lo que podemos concluir que

$$f[X] \subset \bigcup_{k=1}^r U_{\alpha_k},$$

luego $f[X]$ cumple el criterio (C_1) . ■

Observación: Como muestra toda aplicación constante sobre un espacio no compacto, la imagen recíproca de un compacto por una aplicación continua en general no es compacto.

Ademas de ser un medio frecuentemente utilizado para probar la compacidad de un espacio dado, el teorema 4.3.1 es fundamental por los numerosos corolarios que tienen gran importancia tanto en la teoría como en las aplicaciones.

Corolario 1. Sea (X, d) compacto, (Y, d') metrico y $f : X \longrightarrow Y$ una aplicación continua; entonces se cumple:

- i) f es una aplicación cerrada.
- ii) si f es sobreyectiva, entonces los subconjuntos cerrados en (Y, d') son exactamente las imagenes de conjuntos cerrados en (X, d) , es decir, $f\{\tau_d\} = \tau_{d'}$.
- iii) Si f es inyectiva, entonces f es una inmersión, o sea, $f[X]$ es homeomorfo a X .

Demostración:

- i) Sea C cerrado en (X, d) . Por 4.1.7, C es compacto; por tanto, $f[C]$ también lo es, luego es cerrado por 4.1.8.
- ii) Si A es cerrado en X , entonces por i) se tiene que $f[A]$ es cerrado en Y . Recíprocamente si B es cerrado en Y , $f^{-1}[B]$ es cerrado en X y $B = f[f^{-1}[B]]$ por ser f sobreyectiva, luego B es la imagen por f del cerrado $f^{-1}[B]$ en X .
- iii) Por ser f inyectiva, la aplicación

$$f : (X, d) \longrightarrow (f[X], d'_{f[X]})$$

es biyectiva, pero además por i) es cerrado, por lo tanto, es un homeomorfismo.

■

Ejemplo 10. La función $f : t \rightsquigarrow f(t) = e^{i2\pi t}$ del intervalo $[0, 1[$ en $\mathbb{C}$ es una biyección continua de $[0, 1[$ sobre el círculo unidad T en $\mathbb{C}$ (el toro unidimensional). Pero f no es un homeomorfismo. En efecto, T no puede ser homeomorfo a $[0, 1[$ pues es compacto mientras que $[0, 1[$ no lo es.

Los dos ejemplos siguientes muestran aplicaciones del corolario 1 del teorema 4.3.1 en el análisis.

Ejemplo 11. Sea $I \subset \mathbb{R}$ un intervalo compacto (cerrado y acotado). Del análisis sabemos que una función continua $f : I \longrightarrow \mathbb{R}$ es inyectiva si y sólo si es estrictamente monótona (creciente o decreciente). $I' := f[I] \subset \mathbb{R}$ será también un intervalo compacto. Aplicando el corolario anterior, tenemos el resultado siguiente:

Si $f : I \longrightarrow \mathbb{R}$ es una función continua estrictamente monótona sobre un intervalo acotado y cerrado, entonces su recorrido $I' = f[I]$ es un intervalo cerrado y acotado, y su función inversa $f^{-1} : I' \longrightarrow I$ es también continua. Este resultado se utiliza en el análisis por ejemplo para demostrar (por pedazos) la continuidad de la función arcosen a partir de la continuidad del sen.

Ejemplo 12. La existencia de las curvas de Peano muestra que el intervalo $[0, 1]$ puede aplicarse continuamente sobre el rectángulo $[0, 1]^2$. Del corolario 1 iii) se deduce que no pueden existir curvas de Peano inyectivas ya que $[0, 1]$ no es homeomorfo a $[0, 1]^2$.

Un subconjunto $C \subset \mathbb{R}^2$ se llama curva de Jordan si es imagen de $[0, 1]$ por una inyección continua $f : [0, 1] \longrightarrow \mathbb{R}^2$. Según el corolario anterior, una tal inyección f es un homeomorfismo de $[0, 1]$ sobre la curva de Jordan. Concluimos que no existe una curva de Jordan que llene todo el rectángulo $[0, 1]^2$ puesto que éste no es homeomorfo a $[0, 1]$.

Si aplicamos el corolario 1 a la biyección idéntica sobre un mismo conjunto provisto de diferentes métricas, obtenemos el corolario siguiente:

Corolario 2. Sea (X, d) un espacio métrico compacto; entonces se tiene:

i) La familia θ_d de los abiertos es minimal con respecto a la inclusión en el conjunto

$$\{\theta_{d'} : d' \text{ es una métrica sobre } X\}$$

Es decir, si d' es otra métrica sobre X tal que $\theta_{d'} \subset \theta_d$ (o lo que es lo mismo, la identidad $\mathbb{I}_x : (X, d) \longrightarrow (X, d')$ es continua), entonces $\theta_{d'} = \theta_d$ (o sea, d y d' definen los mismos conjuntos abiertos en X , y por tanto, la misma convergencia).

ii) θ_d es maximal en el conjunto ordenado por la inclusión

$$\{\theta_{d'} : (X, d') \text{ es compacto}\}, \text{ o sea si } (X, d') \text{ es compacto}$$

y $\theta_d \subset \theta_{d'}$, entonces $\theta_d = \theta_{d'}$.

Demostración:

i) Sea d' una métrica sobre X tal que $\theta_{d'} \subset \theta_d$, entonces $\mathbb{I}_x : (X, d) \longrightarrow (X, d')$ es continua sobre el compacto (X, d) , luego, es un homeomorfismo.

ii) Se demuestra de forma análoga. ■

Este es un resultado realmente inesperado y muy importante en las aplicaciones para demostrar la equivalencia topológica entre dos métricas.

Ejemplo 13. Sean (X, d) un espacio métrico compacto y $(f_i)_{i=1}^n$ una familia finita de funciones reales continuas sobre X ($f_i : X \longrightarrow \mathbb{R}$) que separa los puntos de X , es decir, para todo par de puntos $x, y \in X$; $x \neq y$ se tiene que $f_i(x) \neq f_i(y)$ para algún $i \in \{1, \dots, n\}$.

Sea además,

$$g : \mathbb{R}^n \longrightarrow \mathbb{R}_+$$

una función con las propiedades siguientes:

i) g es continua;

$$\text{ii) } g(-x_1, \dots, -x_n) = g(x_1, \dots, x_n);$$

$$\text{iii) } g(x_1, \dots, x_n) = 0 \Leftrightarrow x_1 = x_2 = \dots = x_n = 0;$$

$$\text{iv) } g(x_1 + y_1, \dots, x_n + y_n) \leq g(x_1, \dots, x_n) + g(y_1, \dots, y_n).$$

Entonces podemos definir para $x, y \in X$ que

$$\rho(x, y) = g(f_1(x) - f_1(y), \dots, f_n(x) - f_n(y))$$

donde ρ define una métrica en X topológicamente equivalente a la métrica d con respecto a la cual las funciones f_i son continuas. En efecto, la desigualdad triangular para ρ se deduce inmediatamente de iv) mientras que el resto de los axiomas de métrica se verifican gracias a ii), iii) y la propiedad de separabilidad de la familia $(f_i)_{i=1}^n$ (Verifique el resto).

Como $\rho : X \times X \longrightarrow \mathbb{R}_+$ puede ser expresada como una composición de aplicaciones continuas, tendremos que

$$\mathbb{I}_x : (X, d) \longrightarrow (X, \rho)$$

es continua, y aplicando el corolario 2 del teorema 4.3.1i) se obtiene que d y ρ son topológicamente equivalentes.

Es fácil comprobar en particular, que las métricas

$$\rho_p(x, y) = \sqrt[p]{\sum_{n=1}^n |f_i(x) - f_i(y)|^p} \text{ y}$$

$$\rho_\infty(x, y) = \max_{1 \leq i \leq n} |f_i(x) - f_i(y)|$$

son topológicamente equivalentes a d , es decir, para toda sucesión (x_k) en X se tiene:

$$d(x_k, x) \longrightarrow 0 \Leftrightarrow |f_i(x_k) - f_i(x)| \longrightarrow 0$$

para todo $i = 1, \dots, n$.

Si consideramos como caso especial funciones continuas con valores reales definidas sobre un espacio métrico compacto, obtendremos el siguiente resultado fundamental, que es el prototipo de los numerosos teoremas de existencia, los cuales son de gran utilidad en las aplicaciones, y que se demuestran utilizando un argumento de compacidad.

Teorema 4.3.2 (teorema del máximo y del mínimo). Toda aplicación real continua definida sobre un espacio métrico compacto alcanza sus valores máximo y mínimo.

Demostración: En efecto, la imagen de una tal aplicación es un subconjunto compacto de $\mathbb{R}$, luego, es cerrado y acotado, y por lo tanto, contiene sus extremos. Por ello, si $f : X \longrightarrow \mathbb{R}$ existen puntos $x_0, y_0 \in X$, tales que

$$\inf_{x \in X} f(x) = f(x_0) \in f[X] \subset \mathbb{R} \text{ y}$$

$$\sup_{x \in X} f(x) = f(y_0) \in f[X] \subset \mathbb{R}. \blacksquare$$

Ejemplo 14. La condición de que (X, d) sea compacto no es solamente para que se cumpla la tesis del corolario 2 del teorema 4.3.1, sino en cierta medida es también necesaria. En efecto si X no es compacto pueden existir funciones continuas sobre X que no alcancen su valor máximo o mínimo en X . Por ejemplo, basta considerar la aplicación identidad de $]0, 1[$ en $\mathbb{R}$, la cual no alcanza ni su valor máximo 1 ni el mínimo 0 en el intervalo $]0, 1[$.

Más generalmente si (X, d) no es compacto, entonces se puede hallar una función real continua sobre X que no alcanza su máximo en X . Hagamos la siguiente construcción: por ser (X, d) no compacto existe una sucesión (x_n) en X que no tiene alguna subsucesión convergente. Esto quiere decir que (x_n) no tiene puntos adherentes en X , y por lo tanto, para cada $n \in \mathbb{N}$ se puede hallar una bola cerrada $B_n = B(x_n, \epsilon_n)$ tal que

$$B_n \cap B_m = \emptyset \text{ si } n \neq m.$$

Por la regularidad de los espacios métricos (vea el capítulo 2) para cada n existe una función continua

$$f_n : X \longrightarrow [0, 1 - \frac{1}{n}]$$

tal que

$$f_n(x) = \begin{cases} 1 - 1/n, & \text{si } x \in B(x_n, \epsilon_n/2) \\ 0, & \text{si } x \notin B(x_n, \epsilon_n). \end{cases}$$

La función $f : X \longrightarrow [0, 1]$ definida por

$$f(x) := \begin{cases} f_n(x) & \text{si } x \in B(x_n, \epsilon_n) \\ 0, & \text{si } x \notin \bigcup_{n=1}^{\infty} B(x_n, \epsilon_n) \end{cases}$$

es continua (demuéstrello) y

$$\sup_{x \in X} f(x) = 1,$$

pero el valor 1 no es alcanzado nunca por f en X .

Con esto, hemos probado la equivalencia siguiente:

(X, d) es compacto $\Leftrightarrow$ toda función real continua sobre X alcanza su valor máximo en X .

En la práctica resulta a veces importante tener un teorema similar al 4.3.2 para funciones reales semicontinuas.

Teorema 4.3.3. Sea (X, d) un espacio métrico compacto; entonces toda función numérica $f : X \rightarrow \overline{\mathbb{R}}$ semicontinua superiormente (inferiormente) alcanza su máximo (mínimo) sobre X .

Demostración: Sea f semicontinua superiormente y sea $s = \sup f[X] \in \overline{\mathbb{R}}$.

Definamos $S_n := [s - 1/n, \infty[$, ($n \in \mathbb{N}^*$). Para todo $n \in \mathbb{N}^*$, $f^{-1}[S_n]$ es cerrado en (X, d) por ser f semicontinua superiormente, luego $(f^{-1}[S_n])$ es una sucesión decreciente de compactos no vacíos en (X, d) . De la propiedad (NC₂) del ejercicio 3 de autocontrol del epígrafe 4.1, se deduce que la intersección $C := \bigcap_{n \in \mathbb{N}} f^{-1}[S_n]$ no es vacía. Evidentemente en todo punto $x \in C$ se alcanza el máximo de f sobre X . ■

La demostración para las funciones semicontinuas inferiormente es análoga y por ello, se deja al lector.

Para las funciones reales definidas sobre espacios métricos compactos tiene lugar el siguiente teorema, que generaliza el teorema estudiado en el curso elemental de Análisis acerca de las funciones reales continuas sobre un intervalo.

Teorema 4.3.4. Toda función real continua definida sobre un espacio métrico compacto es uniformemente continua.

Demostración: Supongamos que $f : (X, d) \rightarrow \mathbb{R}$, donde (X, d) es compacto, es una función continua pero no uniformemente continua. Entonces, para cierto ϵ positivo y cualquier n natural, existirán x_n y x'_n de X tales que $d(x_n, x'_n) \leq 1/n$, mientras que $|f(x_n) - f(x'_n)| \geq \epsilon$. De la sucesión (x_n) se puede extraer, debido a la compacidad de X , una subsucesión (x_{n_k}) convergente a un punto $x \in X$. En este caso (x'_{n_k}) también converge a x pero para cada k debe cumplirse al menos una de las desigualdades $|f(x) - f(x'_{n_k})| \geq \epsilon/2$, $|f(x) - f(x_{n_k})| \geq \epsilon/2$ y esto contradice la continuidad de la función f en el punto x . ■

Para finalizar este epígrafe veremos algunos resultados importantes en calidad de aplicación de los resultados anteriores.

Ejemplo 15 (separación de subconjuntos compactos en un espacio métrico). i) Sabemos que dos subconjuntos cerrados disjuntos arbitrarios en un espacio métrico pueden ser separados por vecindades disjuntas (propiedad de normalidad). Comencemos por analizar la separación entre un punto x y

un subconjunto C del espacio métrico (X, d) , donde $x \notin C$. Recordemos que la distancia de un punto x al subconjunto C está definida por

$$d(x, C) := \inf\{d(x, z) : z \in C\}.$$

Si C es un compacto, existe un punto y en C más cercano a x , lo cual se muestra en la (fig. 64a).

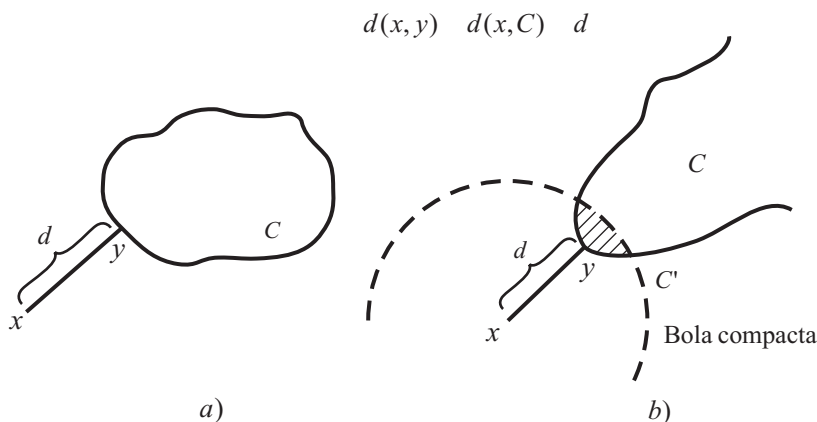


Figura 64

En efecto, la función real $z \mapsto d(x, z)$ es continua sobre C , luego alcanza su mínimo, d , en un punto $y \in C$. El mismo razonamiento puede aplicarse a cerrados cualesquiera C en (X, d) si todas las bolas cerradas de (X, d) son compactas (fig. 64b). Como sabemos esto ocurre, por ejemplo, en los espacios $\mathbb{R}^n$, y veremos en el capítulo 6 que lo mismo sucede en cualquier espacio euclideo de dimensión finita. Sea x un punto, C un cerrado en tal espacio y $d = d(x, C)$. $C' := C \cap B(x, d+1)$ es compacto según hipótesis, luego, existe un $y \in C'$ tal que $d(x, y) = d(x, C')$. El punto y es también un punto más cercano de x en todo C , pues para todo $z \in C \setminus C'$ se cumple $d(x, z) \geq d+1 > d(x, y)$.

Si para cada $n \in \mathbb{N}$, denotamos por ρ_n el subespacio de $C([0, 1], \mathbb{R}) = C[0, 1]$ sustituido por los polinomios algebraicos con coeficientes reales y de grado menor o igual a n , entonces ρ_n es homeomorfo al espacio euclideo $\mathbb{R}^{n+1}$ y por ello, las bolas cerradas en ρ_n con la métrica inducida por d_∞ , son compactas. Luego, para cada función continua $f \in C[0, 1]$ existe un polinomio P_n en ρ_n que aproxima mejor a f con respecto a d_∞ , es decir,

$$d_\infty(f, \rho_n) = \inf_{p \in \rho_n} d_\infty(f, p) = E_\infty(f, \rho_n).$$

Por el teorema de Weierstrass sabemos que toda función continua $f \in C[0, 1]$ puede ser aproximada uniformemente por polinomios. Luego, para la sucesión de las mejores aproximaciones $(E_\infty(f, \rho_n))$ se tiene

$$\lim_{n \rightarrow \infty} E_\infty(f, \rho_n) = 0.$$

- ii) Sabemos que existen cerrados disjuntos K y C en $\mathbb{R}^2$ tales que $d(K, C) = 0$. Esto implica que en $K \times C$ no existe un par de puntos de distancia minimal tales que $d(x, y) = d(K, C)$.

Probaremos ahora que dos compactos disjuntos cualesquiera K y C en (X, d) , siempre tienen una distancia estrictamente positiva, y contienen un par de puntos de distancia minimal.

La función real $x \mapsto d(x, C)$ es continua sobre el compacto K , luego alcanza su mínimo $d = \min\{d(x, C) : x \in K\} = d(K, C)$ en un punto x_0 . Puesto que $x_0 \notin C$ y C es cerrado tenemos que $d = d(x_0, C) > 0$. Según i), existe $y_0 \in C$ tal que

$$d(x_0, y_0) = d(K, C) > 0.$$

Si todas las bolas cerradas de (X, d) son compactas, el mismo resultado puede probarse para el caso más general en que K es compacto y C cerrado. Esta situación ocurre en las aplicaciones, a veces, en la forma siguiente: $K \subset X$ es un compacto y U es una vecindad abierta de K , entonces necesariamente $d = d(K, X \setminus U) > 0$ (siempre que las bolas cerradas de (X, d) sean compactas) (fig. 65).

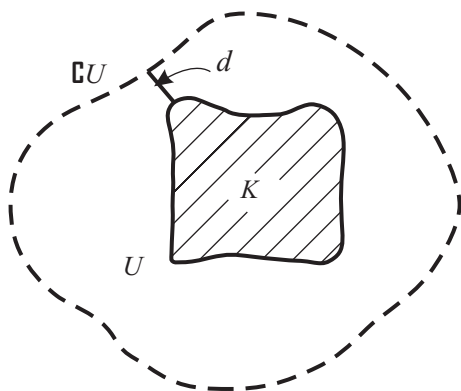


Figura 65

Ejemplo 16 (funciones sobre espacios metricos). Si X es un conjunto cuyos elementos son funciones definidas sobre un cierto dominio D , entonces

toda aplicación $F : X \longrightarrow \mathbb{C}$, donde $\mathbb{C}$ es el campo de los números complejos se llama un funcional sobre X . Se dice que el funcional F es real si $F[X] \subset \mathbb{R}$.

Veamos algunos ejemplos de funcionales en espacios de funciones reales $x(t)$ definidas sobre el segmento $[0, 1]$:

$$F_0(x) = \int_0^{1/2} x(t) dt - \int_{1/2}^1 x(t) dt;$$

$$F_1(x) = \sup x(t);$$

$$F_2(x) = \inf x(t);$$

$$F_3(x) = x(t_0), \text{ donde } t_0 \in [0, 1]$$

$$F_4(x) = \varphi[x(t_0), x(t_1), \dots, x(t_n)],$$

donde $t_i \in [0, 1]$ y la función $\varphi[s_0, \dots, s_n]$ está definida para todo s_i real;

$$F_5(x) = \int_0^1 \varphi[t, x(t)] dt, \text{ donde,}$$

$\varphi(t, s)$ está definida y es continua para todo $0 \leq t \leq 1$ y todo s real;

$$F_6(x) = x'(t_0), \text{ donde } t_0 \in [0, 1];$$

$$F_7(x) = \int_0^1 \sqrt{1 + x'^2(t)} dt;$$

$$F_8(x) = \int_0^1 |x'(t)| dt.$$

Si el dominio X del funcional en cuestión, es además un espacio métrico, entonces tiene sentido hablar de la continuidad de dicho funcional.

Tomemos $(X, d) = ((C[0, 1], \mathbb{R}), d_\infty)$; los funcionales F_0, F_1, F_2, F_3, F_4 y F_5 están definidos sobre este espacio métrico que por comodidad denotaremos por $C[0, 1]$ y además, no es difícil comprobar que

$$F_i : C[0, 1] \longrightarrow \mathbb{R}$$

es continuo para todo $i = 0, 1, \dots, 5$. (Demuéstrelo.) Sin embargo, F_6 está definido sobre el subespacio de $C[0, 1]$ constituido por las funciones diferenciables en el punto t_0 : F_7 para aquellas funciones de $C[0, 1]$ para las cuales la expresión $\sqrt{1 + x'^2(t)}$ es integrable y F_8 para las funciones tales que $|x'(t)|$ es integrable.

El funcional F_6 es discontinuo en todo punto de $C[0, 1]$ donde está definido. En efecto, sea $x(t)$ una función tal que $x'(t_0) = 1$ y $|x(t)| < \epsilon$ y sea $y = x_0 + x$. Entonces $y'(t_0) = x'_0(t_0) + 1$, mientras que $d_\infty(y, x_0) < \epsilon$. Sin embargo, no resulta difícil comprobar que F_6 es continuo sobre el espacio $C^{(1)}[0, 1]$ (de las funciones que tiene derivada continua) si lo proveemos de la métrica.

$$\rho_\infty(x, y) = \sup_{t \in [0, 1]} \{|x(t) - y(t)| + |x'(t) - y'(t)|\}.$$

El funcional F_7 es también discontinuo sobre el espacio $C[0, 1]$. En efecto, sean $x_0(t) = 0$ y $x_n(t) = \frac{1}{n} \sin 2\pi n t$, entonces $d_\infty(x_n, x_0) = \frac{1}{n} \rightarrow 0$; sin embargo, $F_7(x_n) = 1$. Por consiguiente, $F_7(x_n)$ no tiende a $F_7(x_0)$ cuando $x_n \rightarrow x_0$. El mismo ejemplo sirve para demostrar que el funcional F_8 es también discontinuo en el espacio $C[0, 1]$. Puede probarse que ambos funcionales F_7 y F_8 son continuos en el espacio $C^{(1)}[0, 1]$. (Pruébalo.)

Hemos visto que el funcional F_0 es continuo sobre $C[0, 1]$. Se verifica fácilmente que el supremo de f sobre la bola unidad de $C[0, 1]$:

$$B = \{x \in C[0, 1] : \sup_{t \in [0, 1]} |x(t)| \leq 1\}$$

es 1. Sencillamente hay que aproximar la función constante por trozos

$$t \rightsquigarrow \text{sig}(1/2 - t) \quad (0 \leq t \leq 1)$$

sucesivamente por funciones continuas x que estén en B (fig. 66).

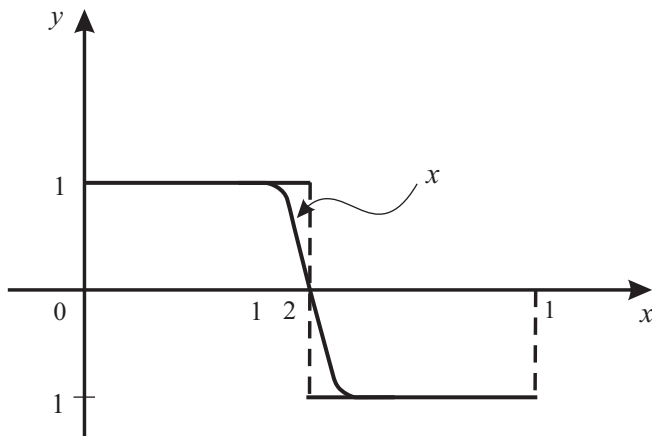


Figura 66

Sin embargo, evidentemente para toda función $x \in B$ se tiene que $|F_0(x)| < 1$, luego, F_0 no alcanza su supremo sobre B . Concluimos una vez más que la

bola unidad $C[0, 1]$ no es compacta. Como podemos observar esta parece ser una característica de algunos espacios métricos de dimensión infinita. Efectivamente, veremos en el capítulo 6 un resultado que relaciona la estructura vectorial con las propiedades topológicas en aquellos espacios en que la métrica es *compatible* con la estructura algebraica.

Notemos que aún cuando restringjamos algún funcional F_i ($i = 0, 1, \dots, 5$) a un subespacio compacto K de $C[0, 1]$ solamente se puede asegurar la existencia de dos funciones $x_0, y_0 \in k$ tales que

$$\sup_{x \in k} |F_i(x)| = |F_i(x_0)| \quad \text{y} \quad (3)$$

$$\inf_{x \in k} |F_i(x)| = |F_i(y_0)|, \quad (4)$$

pero no se puede decir nada acerca de la naturaleza de x_0 y y_0 . La información sobre las condiciones necesarias y suficientes que deben satisfacer las funciones x_0 y y_0 en (3) y (4) puede ser obtenida como resultado de un análisis más profundo que envuelve una generalización del concepto de diferenciación para funciones numéricas y que veremos en el capítulo 7.

Ejercicios de Autocontrol

1. Culmine los detalles en el desarrollo de los ejemplos 13 y 14.
2. Demuestre que un espacio métrico (X, d) es compacto si y solo si cada función real continua sobre X es acotada.
3. Sea $\{V_i\}_{i=1}^n$ un cubrimiento abierto finito del espacio métrico compacto (X, d) . Suponga que $\varphi_i(x) = d(x, X \setminus V_i)$ y sea

$$\lambda_i(x) = \frac{\varphi_i(x)}{\varphi_1(x) + \dots + \varphi_n(x)}$$

Demuestre que los λ_i son funciones continuas y además,

- a) $0 < \lambda_i(x) < 1$
- b) $\lambda_i(x) = 0$ fuera de V_i
- c) $\lambda_i(x) + \dots + \lambda_n(x) = 1$ para todo $x \in X$.

4. Demuestre el principio fuerte de las aplicaciones contraídas en espacios compactos (ver 4.2.4).

Indicación: Considere la aplicación $\varphi : X \longrightarrow \mathbb{R}$, definida mediante $\varphi(x) = d(f(x), x)$ y utilice el teorema 4.3.8.

4.4 COMPACIDAD EN ESPACIOS PRODUCTO Y ESPACIOS FUNCIONALES

En este epígrafe estudiaremos las propiedades de los subconjuntos compactos en espacios de funciones con respecto a la convergencia puntual (espacios producto) y a la convergencia uniforme.

Como ya hemos visto en el capítulo 2, la convergencia puntual en el conjunto Y^X de las funciones de X en el espacio métrico (Y, d) , corresponde a la convergencia segun la métrica producto cuando X es numerable. Por ello, comenzaremos por estudiar la compacidad en espacios producto arbitrarios $\prod_{n \in \mathbb{N}} X_n$, donde (X_n, d_n) es un espacio métrico.

En 1935 Tijonov publicó su célebre teorema de que la compacidad es invariante ante productos, el cual expondremos a continuación, en una forma limitada, solamente para productos numerables.

Teorema 4.4.1 (Teorema de Tijonov). Sea (X_n, d_n) $n \in \mathbb{N}$ una familia de espacios métricos no vacíos. Entonces $\prod_{n \in \mathbb{N}} X_n$ es compacto (con respecto a la métrica producto) si y sólo si lo es cada espacio (X_n, d_n) $(n \in \mathbb{N})$.

Demostración: $\Rightarrow$: Por el axioma de elección $\prod X_n \neq \emptyset$. Como las proyecciones $\pi_n : \prod X_k \longrightarrow X_n$ son sobreyectivas y continuas, la compacidad de $\prod X_k$ implica la de cada espacio (X_n, d_n) $(n \in \mathbb{N})$.
 $\Leftarrow$: Sea (x_n) una sucesión en $\prod X_k$ donde $(x_n^{(1)}, x_n^{(2)}, \dots, x_n^{(k)})$ con $x_n^{(k)} \in X_k$. Para cada k fijo, $x_n^{(k)} = \pi_k x_n$ es una sucesión en X_k , y según el criterio (C4), tiene una subsucesión convergente en X_k .

Realicemos ahora la demostración por inducción en la forma siguiente:

- Paso 1:** La sucesión $(x_n^{(1)})$ tiene una subsucesión $x_{n_1}^{(1)}$ convergente a un punto $x^{(1)}$ en X_1 .
- Paso 2:** La sucesión $x_{n_1}^{(2)}$ tiene una subsucesión $x_{n_2}^{(2)}$ convergente al punto $x^{(2)}$ en X_2 .

.....

Paso k : La sucesión $(x_{n_{k-1}}^{(k)})$ tiene una subsucesión $x_{n_k}^{(k)}$ convergente al punto $x^{(k)}$ en X_k .

Por la construcción hecha no es difícil ver que para cada k fijo la sucesión $(x_{p_p}^{(k)})$ es una subsucesión de la sucesión $(x_{k_k}^{(k)})$ a partir del k -ésimo ($p \geq k$) y, por lo tanto, $(x_{p_p}^{(k)})$ converge a $(x^{(k)})$ cuando $p \longrightarrow \infty$

Luego, la sucesión (x_{p_p}) en $\prod X_k$, donde $x_{p_p} = (x_{p_p}^{(1)}, x_{p_p}^{(2)}, \dots, x_{p_p}^{(k)}, \dots)$, es una subsucesión de la sucesión (x_n) , además, es convergente ya que converge por coordenadas y sabemos que esta propiedad caracteriza la convergencia en los espacios producto. ■

El teorema de Tijonov en su expresion mas general que no veremos aqui (vea D. Hinrichsen y otros; *Topologia General*), es decir, cuando el conjunto de indices puede ser no numerable, es uno de los medios mas poderosos para verificar la compacidad de ciertos espacios (especialmente espacios funcionales). No obstante, la versión del teorema aqui expuesta (cuando el conjunto de indices es numerable), resulta una herramienta muy util de trabajo, lo cual ilustraremos en los ejemplos siguientes:

Ejemplo 17 (caracterización de los compactos en $\mathbb{R}^n$). Según 4.1.8, todo subconjunto compacto en un espacio métrico es necesariamente cerrado y acotado. Veremos ahora que en el espacio métrico $\mathbb{R}^n$ se cumple también el inverso.

Sea $C \subset \mathbb{R}^n$ cerrado y acotado. Entonces C está contenido en un producto de intervalos cerrados y acotados $I_1 \times \dots \times I_n$ (fig. 67).

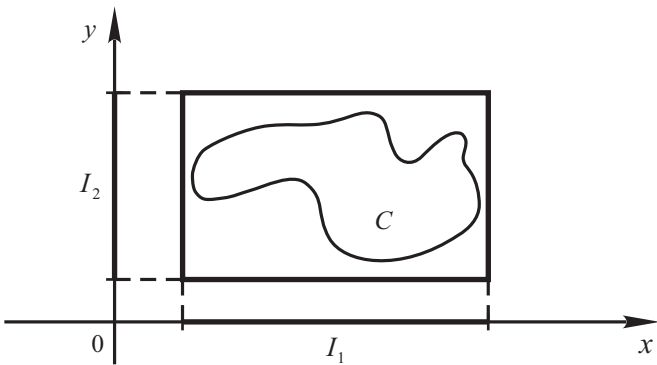


Figura 67

Según el teorema de Tijonov, el producto $I_1 \times \dots \times I_n \subset \mathbb{R}^n$ es compacto. Podemos concluir por 4.1.7 que el subconjunto cerrado C también es compacto.

Con esto hemos demostrado de nuevo un teorema fundamental del análisis matemático.

Un subconjunto $C \subset \mathbb{R}^n$ es compacto si y sólo si es acotado y cerrado.

Ejemplo 18. Todo cubo I^T (potencia finita o numerable del intervalo compacto $I = [0, 1]$) es compacto.

Se llama cubo de Hilbert, al conjunto de las sucesiones de números reales $(x_1, x_2, \dots, x_n, \dots)$ tales que $|x_n| \leq 1/n$. Precisamente el cubo de Hilbert, coincide con el espacio producto $\prod_{n \in \mathbb{N}^*} [-1/n, 1/n]$

y por lo tanto, provisto de la métrica producto (la cual define la convergencia por coordenadas) es un espacio compacto.

Ejemplo 19 (discontinuo de Cantor). Sean $A_n = \{0, 2\}$ para cada $n \in \mathbb{N}$. Entonces $\prod_{n \in \mathbb{N}} A_n = \{0, 2\}^{\mathbb{N}}$ es el conjunto de todas las sucesiones (x_n) tales que $x_n \in \{0, 2\}$. Se ve fácilmente que la aplicación $g : \prod_{n \in \mathbb{N}} A_n \longrightarrow [0, 1]$ definida por

$$g((x_n)) = \sum_{n=1}^{\infty} x_n/3 \text{ es inyectiva.}$$

La imagen de g en $[0, 1]$ se llama *conjunto de Cantor* (o también, *discontinuo de Cantor*), el cual geoméricamente puede ser descrito como sigue:

Dividamos $[0, 1]$ en tres partes iguales y retiremos el intervalo intermedio $]1/3, 2/3[$; esto elimina todos los números reales en $[0, 1]$ que requieren $x_1 = 1$ en su desarrollo en base tres (representación triádica).

Como segundo paso retiremos los intervalos intermedios de cada uno de los intervalos restantes, $[0, 1/3]$, $[2/3, 1]$, eliminando así todos los números reales en $[0, 1]$ que requieren $x_2 = 1$ en su desarrollo triádico.

Procediendo análogamente retiramos, en el n -ésimo paso, la unión M_n de los intervalos intermedios a los 2^{n-1} intervalos presentes.

El conjunto $C = [0, 1] \setminus \bigcup_{n \in \mathbb{N}} M_n$ es el llamado conjunto de Cantor y está formado por todos los números reales en $[0, 1]$ que no requieren del uso de $\langle\langle 1 \rangle\rangle$ en su desarrollo triádico.

Como el desarrollo (no utilizando 1) de cada elemento de $[0, 1]$ es único, g es una biyección de $\prod_{n \in \mathbb{N}} A_n$ en C . Si ahora definimos sobre cada A_n la métrica discreta, A_n se convierte en un espacio métrico compacto, y por el teorema Tjonov, el espacio producto $\prod_{n \in \mathbb{N}} A_n$ provisto de la métrica producto será también compacto. Si consideramos sobre C la métrica inducida por la métrica

usual de $\mathbb{R}$, no es difícil verificar que g es continua (demuéstrelo) y por 4.3.1 C será compacto. De la biyectividad de

$$g : \prod_{n \in \mathbb{N}} A_n \longrightarrow C$$

y del corolario 1 iii) del teorema 4.3.1, se deduce además que g es un homeomorfismo.

Ejemplo 20. Sea T un subconjunto finito o numerable cualquiera y consideremos el espacio $\mathbb{R}^T$ de todas las funciones reales definidas sobre T , provisto de la métrica de la convergencia puntual (métrica producto). Un conjunto $C \subset \mathbb{R}^T$ se llama *débilmente acotado* si para todo $t \in T$ la proyección

$$\pi_t[C] = \{x(t) : x \in C\} \subset \mathbb{R}$$

es acotado en $\mathbb{R}$. Por lo tanto, si C es débilmente acotado, existen intervalos compactos $I_t = [a_t, b_t] \subset \mathbb{R}$ ($t \in T$) tales que

$$C \in \prod [a_t, b_t].$$

En este caso C es un subconjunto compacto $\prod_{t \in T} I_t$. Concluimos que todo subconjunto cerrado y débilmente acotado en $\mathbb{R}^T$ es compacto.

El inverso se cumple trivialmente. Todo subconjunto compacto de $\mathbb{R}^T$ es cerrado y débilmente acotado puesto que las proyecciones son continuas. Así vemos que la caracterización de los compactos expuesta en el ejemplo 17 se generaliza a los espacios producto $\mathbb{R}^T$ de dimensión infinita si se sustituye *acotado* por *débilmente acotado*.

Observación: No obstante hay que destacar una diferencia esencial entre los espacios $\mathbb{R}^n$ y $\mathbb{R}^T$ (T infinito).

En $\mathbb{R}^n$ toda bola abierta nos brinda un ejemplo de un abierto acotado no vacío, mientras que en $\mathbb{R}^T$ no existe un abierto no vacío, que sea débilmente acotado (vea las propiedades de los abiertos para la métrica producto), por tanto, todos los compactos en $\mathbb{R}^T$ tienen el interior vacío.

Según el ejemplo 17 todo subconjunto acotado de $\mathbb{R}^n$ es relativamente compacto. Esto no se cumple en espacios métricos cualesquiera. Incluso probaremos en el capítulo 6, que un espacio normado que contiene una bola relativamente compacta es necesariamente de dimensión finita. En los ejemplos 5 y 16 ya hemos visto ejemplos de espacios normados donde la bola unidad (cerrada) no es compacta.

Ejemplo 21. Sea T un conjunto a lo más numerable, y consideremos el espacio producto $\mathbb{R}^T$ identificado con el espacio de las aplicaciones de T en $\mathbb{R}$.

No es difícil demostrar que la convergencia puntual en $\mathbb{R}^T$ coincide con la convergencia uniforme si y solo si T es finito, es decir, para una sucesión de funciones

$$\begin{aligned} f_n : T &\longrightarrow \mathbb{R} \ (n \in \mathbb{N}), \text{ se tiene:} \\ (f_n) &\longrightarrow f \text{ (puntualmente en } T) \Leftrightarrow (f_n) \longrightarrow f \\ &\text{(uniformemente en } T). \end{aligned}$$

Expresado de otra forma:

La métrica d_∞ de la convergencia uniforme es topológicamente equivalente a la métrica producto sobre $\mathbb{R}^T$ si y solo si, T , es un conjunto finito, o sea, si $\mathbb{R}^T$ coincide con un espacio euclideo $\mathbb{R}^n$.

Hemos visto en el ejemplo 20 una caracterización de los subconjuntos compactos en $\mathbb{R}^T$ con respecto a la métrica producto. Sin embargo, como nos muestra el ejemplo 5 tales caracterizaciones son más difíciles de obtener para los subconjuntos compactos en espacios de funciones, con respecto a la métrica de la convergencia uniforme.

En este capítulo no nos dedicaremos al estudio de los subconjuntos compactos en espacios de funciones generales provistos de la métrica de la convergencia uniforme, ya que esto no presenta interés práctico, sino nos limitaremos a examinar los subconjuntos compactos en espacios de funciones continuas.

Como es conocido, los espacios de funciones continuas son uno de los tipos de espacios métricos de mayor importancia en las aplicaciones del análisis y el análisis funcional. El teorema fundamental que demostraremos es el llamado teorema de Arzelá-Ascoli, pero para poder enunciarlo necesitamos introducir los conceptos siguientes:

Definición 4.4.2. Sean (X, d) y (Y, d') espacios métricos. Un subconjunto $H \subset Y^X$ se llama *equicontinuo* en $x \in X$ si para todo $\epsilon > 0$ existe $\delta > 0$ tal que para todo $f \in H$ se cumple

$$d(y, x) < \delta \Rightarrow d'(f(y), f(x)) < \epsilon. \quad (5)$$

H se llama *equicontinuo* si es equicontinuo en todo $x \in X$.

Hemos identificado el espacio producto $Y^X = \prod_{x \in X} Y$ con el conjunto de todas las aplicaciones $f : X \longrightarrow Y$, y por ello, en lo adelante lo denotaremos también mediante $F(X, Y)$.

Para el conjunto de todas las funciones continuas de (X, d) en (Y, d') utilizaremos la notación $C(X, Y)$.

La definición 4.4.2 muestra que todas las funciones de un conjunto equicontinuo $H \subset F(X, Y)$ son continuas, sin embargo, el inverso no se cumple, o sea,

los subconjuntos $H \subset C(X, Y)$ en general no son equicontinuos.

En efecto, si $H \subset C(X, Y)$ y $x_0 \in X$, entonces para cada función $f \in H$ existe individualmente un δ_f tal que se cumple (5), sin embargo, puede ocurrir que $\inf \delta_f = 0$, y por tanto, no se pueda encontrar un $\delta > 0$ común para todas las $f \in H$.

Ejemplo 22. Todo conjunto finito $G \subset C(X, Y)$ es equicontinuo.

Ejemplo 23. Sean (X, d) y (Y, d') espacios métricos. Una función $f : X \longrightarrow Y$ se dice lipchitziana de razón $k > 0$ si $d'(f(x), f(y)) \leq kd(x, y)$ para todo $x, y \in X$. Todo conjunto de funciones lipchitzianas de una misma razón $k > 0$ es equicontinuo.

Ejemplo 24. En la figura 68 se muestran gráficos de conjuntos $H \subset C(\mathbb{R}, \mathbb{R})$ que ilustran bien la diferencia entre conjuntos equicontinuos y conjuntos de aplicaciones continuas que no son equicontinuos.

- i) $H = \{f_b \in C(\mathbb{R}, \mathbb{R}) : f_b(x) = ax + b, b \in \mathbb{R}\}$ para a prefijado es equicontinuo (fig. 68a)
- ii) $H = \{f_a : f_a(x) = ax + b, a \in \mathbb{R}\}$ para b prefijado, no es equicontinuo (fig. 68b).
- iii) $H = \{f_a : f_a(x) = ax + b, a \in \mathbb{R}, |a| \leq M\}$ para cualquier número fijo $M > 0$, es equicontinuo (fig. 68c).

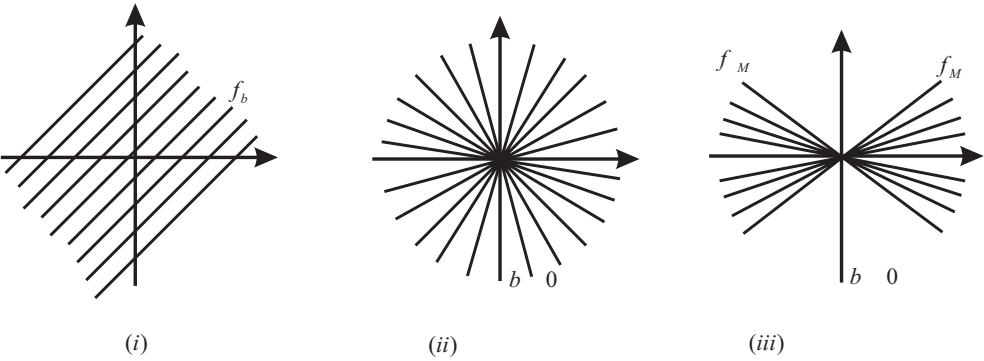


Figura 68

Observemos que la equicontinuidad es una propiedad definida independientemente de una métrica sobre H , y solo depende de la métrica de X y Y . Como las métricas de uso práctico sobre H están generalmente relacionadas de alguna

manera con las de X y Y , existe de hecho una relación entre la equicontinuidad y dichas métricas. Para ver como esto se pone de manifiesto con respecto a la métrica de la convergencia uniforme d_∞ , introduciremos además la definición siguiente:

Definición 4.4.3. Sea X un conjunto y (Y, d') un espacio métrico. Un subconjunto $H \subset F(X, Y)$ se llama *equiacotado* en el punto $x \in X$

$$H(x) = \{f(x) : f \in H\} \subset Y$$

es relativamente compacto en (Y, d') .

H se llama *equiacotado* si es equiacotado en todo punto $x \in X$.

Ejemplo 25. Sabemos que un subconjunto de $\mathbb{R}^n$ es relativamente compacto si y solo si es acotado, luego, si $Y = \mathbb{R}^n$ un conjunto $H \subset F(X, \mathbb{R}^n)$ es equiacotado si y solo si para cada $x \in X$ existe una constante $M_x > 0$, tal que para todo $x \in X$

$$|f(x)| \leq M_x,$$

donde $|\cdot|$ denota el módulo en $\mathbb{R}^n$.

Ejemplo 26. Es conocido que todo subconjunto de un espacio métrico compacto es siempre relativamente compacto. Luego, si (Y, d') es compacto, cualquier conjunto $H \subset F(X, Y)$ es equiacotado.

Antes de pasar a demostrar el resultado fundamental (teorema de Arzelá-Ascoli) veamos los lemas siguientes:

Lema 4.4.4. Sean, (X, d) un espacio métrico compacto y (Y, d') un espacio métrico cualquiera. Entonces son equivalentes:

- i) $H \subset C(X, Y)$ es equicontinuo.
- ii) Para todo $\epsilon > 0$ existe $\delta > 0$ tal que si $d(x, y) < \delta$ entonces $d'(f(x), f(y)) < \epsilon$, cualesquiera que sean $f \in H; x, y \in X$.

Observación: Notemos que en la definición 4.4.2 de conjunto equicontinuo el número δ no depende de $f \in H$ pero sí depende del punto x en cuestión. En el lema 4.4.4 se muestra que si X es compacto, se puede elegir δ independiente de $f \in H$ y $x \in X$. Podíamos haber dado originalmente esta definición y limitarnos al trabajo en espacios métricos compactos (X, d) , pero hemos preferido no hacerlo así ya que en primer lugar en la práctica es más fácil de comprobar la definición 4.4.2, aún en el caso en que (X, d) sea compacto y además, en

el epígrafe siguiente veremos una generalización importante del teorema de Arzelá-Ascoli donde utilizaremos las definiciones generales 4.4.2 y 4.4.3 aun cuando (X, d) no es compacto sino solo localmente compacto.

Demostración: La implicación $\text{ii}) \Rightarrow \text{i})$ es evidente, y por tanto, nos limitaremos a demostrar $\text{i}) \Rightarrow \text{ii})$.

Por ser $H \subset C(X, Y)$ equicontinuo, para cualesquiera $\epsilon > 0$ y $x \in X$, existe $\delta(x) > 0$, tal que para todo $f \in H$ se cumple:

$$d(x, y) < \delta(x) \Rightarrow d'(f(x), f(y)) < \epsilon/2. \quad (6)$$

Consideremos el cubrimiento de (X, d) por las bolas abiertas $\{B(x, \delta(x))\}$, $x \in X$. Por el lema 4.1.4 existe un número δ de Lebesgue con respecto a este cubrimiento, es decir un $\delta > 0$ tal que para todo $x \in X$ existe $z \in X$ que cumple

$$B(x, \delta) \subset B(z, \delta(z)). \quad (7)$$

Sean ahora $x, y \in X$ arbitrarios tales que $d(x, y) < \delta$. Sabemos por (7) que existe $z \in X$ tal que $x, y \in B(z, \delta(z))$, es decir,

$$\begin{aligned} d(x, z) &< \delta(z), \\ d(x, y) &< \delta(z). \end{aligned}$$

Pero entonces, por (6) tendremos:

$$\begin{aligned} d'(f(x), f(z)) &< \epsilon/2, \\ d'(f(y), f(z)) &< \epsilon/2. \end{aligned}$$

para todo $f \in H$.

Luego, si $d(x, y) < \delta$ se tendrá que $d'(f(x), f(y)) \leq d'(f(x), f(z)) + d'(f(z), f(y)) < \epsilon/2 + \epsilon/2 = \epsilon$, para todo $f \in H$, con lo cual queda probado $\text{ii})$. ■

Lema 4.4.5. Sean (X, d) compacto, (Y, d) métrico y $H \subset C(X, Y)$ equicontinuo. Entonces son equivalentes:

- i) H es equiacotado;
- ii) $\bigcup_{x \in X} H(x)$ es relativamente compacto en (Y, d')

Demostración: La implicación $\text{ii}) \Rightarrow \text{i})$ es evidente ya que cualquier subconjunto de un conjunto relativamente compacto es también relativamente compacto.

Probemos que $i) \Rightarrow ii)$.

Para ello debemos demostrar que toda sucesión (y_n) de elementos de $\bigcup_{x \in X} H(x)$, tiene una subsucesión que converge a un punto $y_0 \in Y$. Por definición, existen puntos $x_n \in X$ y funciones $f_n \in H$ tales que

$$y_n = f_n(x_n).$$

Por ser (X, d) compacto, la sucesión (x_n) contiene una subsucesión $(x_{\varphi(n)})$ que converge a un punto $x_0 \in X$. Por ser $H(x_0)$ relativamente compacto, la sucesión $(f_{\varphi(n)}(x_0))$ de $H(x_0)$ tiene una subsucesión $(f_{\psi(n)}(x_0))$ que converge a un punto $y_0 \in Y$. Evidentemente la sucesión $(x_{\psi(n)})$ converge al punto $x_0 \in X$, por ser una subsucesión de la sucesión convergente $(x_{\varphi(n)})$.

Demostremos entonces que la subsucesión $(f_{\psi(n)}(x_{\psi(n)}))$ de la sucesión (y_n) converge al punto $y_0 \in Y$. En efecto, por ser H equicontinuo, dado $\epsilon > 0$ existe $\delta > 0$, tal que

$$d(x, y) < \delta \Rightarrow d(f_{\psi(n)}(x), f_{\psi(n)}(y)) < \epsilon/2. \quad (8)$$

Como $(x_{\psi(n)})$ converge a x_0 existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$,

$$d(x_{\psi(n)}, x_0) < \delta,$$

y por tanto, por (8) si $n \geq n_0$, entonces,

$$d(f_{\psi(n)}(x_{\psi(n)}), f_{\psi(n)}(x_0)) < \epsilon/2. \quad (9)$$

Pero por la desigualdad triangular

$$d(f_{\psi(n)}(x_{\psi(n)}), y_0) \leq d(f_{\psi(n)}(x_{\psi(n)}), f_{\psi(n)}(x_0)) + d(f_{\psi(n)}(x_0), y_0). \quad (10)$$

Del hecho que $f_{\psi(n)}(x_0)$ converge a y_0 se deduce que existe $n_1 \in \mathbb{N}$ tal que si $n \geq n_1$, entonces

$$d(f_{\psi(n)}(x_0), y_0) < \epsilon/2. \quad (11)$$

Tomando $n \geq \max(n_0, n_1)$ de (9), (10) y (11) se concluye que

$$d(f_{\psi(n)}(x_{\psi(n)}), y_0) < \epsilon$$

y por la arbitrariedad de $\epsilon > 0$, obtenemos que la sucesión $(f_{\psi(n)}(x_{\psi(n)}))$ converge a y_0 . ■

Pasemos finalmente a caracterizar los subconjuntos relativamente compactos de $C(X, Y)$ con respecto a la métrica d_∞ de la convergencia uniforme, cuando el espacio métrico (X, d) es compacto.

Teorema 4.4.6 (Teorema de Arzelá-Ascoli para (X, d) compacto). Sean (X, d) compacto y (Y, d') métrico, y consideremos el espacio $C(X, Y)$ provisto de la métrica d_∞ de la convergencia uniforme sobre X :

$$d_\infty(f, g) = \sup_{x \in X} d'(f(x), g(x)).$$

Entonces un conjunto $H \subset C(X, Y)$ es relativamente compacto si y solo si es equicontinuo y equiacotado.

Demostración: Supongamos que H es relativamente compacto y demostremos que es equiacotado y equicontinuo.

Consideremos el espacio producto $C(X, Y) \times X$ provisto de la métrica producto $d_\infty \times d$. Del método de demostración del lema 4.5 se deduce que la aplicación evaluación

$$\begin{matrix} E: C(X, Y) \\ (f, x) \rightsquigarrow f(x) \end{matrix} \times X \longrightarrow Y$$

es continua. Si H es relativamente compacto en $C(X, Y)$, entonces por el teorema de Tijonov $H \times \{x\}$ es relativamente compacto en $C(X, Y) \times X$ y por ser E continua

$$E[H \times \{x\}] = H(x) \quad (\text{compruébe esta igualdad}),$$

es relativamente compacto en Y . Luego, H es equiacotado.

Para demostrar la equicontinuidad de H basta ver que esta propiedad es equivalente a la continuidad uniforme de E restringida a cada conjunto del tipo $H \times \{x\}$ (con $x \in X$ (demuéstrelo)).

Ya hemos dicho que $H \times \{x\}$ es relativamente compacto en $C(X, Y) \times X$ y como E es continua, por el teorema 4.3.4 será también uniformemente continua sobre $H \times \{x\}$. Luego, queda probada la equicontinuidad de H .

Demostremos ahora que la equiacotación y la equicontinuidad son condiciones suficientes para que H sea relativamente compacto en $C(X, Y)$. Para ello, denotemos por Y_0 a la clausura de $\bigcup_{x \in X} H(x)$ en Y ; entonces si H es relativamente compacto en $C(X, Y_0)$, también lo será en $C(X, Y)$. Por este motivo y por el lema 4.4.5 es suficiente demostrar que H es relativamente compacto en $C(X, Y_0)$ donde X y Y_0 son ambos espacios métricos compactos. Para esto sumerjamos $C(X, Y_0)$ en el espacio $F(X, Y_0)$ provisto de la métrica d_∞ , y demostremos la compacidad relativa de H en $F(X, Y_0)$. Puesto que $C(X, Y_0)$ es cerrado en $F(X, Y_0)$, ya que el límite de una sucesión uniformemente convergente de aplicaciones continuas es también una aplicación continua, la compacidad relativa del conjunto H en $F(X, Y_0)$ implica su compacidad relativa en $C(X, Y_0)$.

Tomemos $\epsilon > 0$ arbitrariamente y escojamos δ de manera que de

$$d(x, y) < \delta,$$

se tenga

$$d'(f(x), f(y)) < \epsilon,$$

para todo $f \in H$ y todo $x, y \in X$.

Sea ahora $\{x_1, x_2, \dots, x_n\}$ una $\delta/2$ -red en X y consideremos

$$E_i = B(x_i, \delta) \setminus \bigcup_{j < i} B(x_j, \delta).$$

No es difícil comprobar que los conjuntos $\{E_i\}_{i=1}^n$ constituyen una partición de X tal que de $x, y \in E_i$ se deduce $d(x, y) < \delta$. Consideremos en el compacto Y_0 una ϵ -red $\{y_1, y_2, \dots, y_n\}$ y denotemos mediante L la colección de las funciones $g \in F(X, Y)$ que toman los valores y_j sobre los conjuntos E_i . El número de estas funciones es, evidentemente, finito. Probemos que ellas forman una 2ϵ -red para H en $F(X, Y_0)$. En efecto, sea $f \in H$, para todo punto x_i de $\{x_1, \dots, x_n\}$ existe un punto $y_j \in \{y_1, \dots, y_m\}$ tal que

$$d'(f(x_i), y_j) < \epsilon.$$

Sea la función $g \in L$ seleccionada de manera que $g(x_i) = y_j$; entonces:

$$\begin{aligned} d'(f(x), g(x)) &\leq d'(f(x), f(x_i)) + d'(f(x_i), g(x_i)) + \\ &+ d'(g(x_i), g(x)) \end{aligned}$$

y si para cada $x \in X$ tomamos el valor de i tal que x pertenezca a E_i , tendremos

$$d'(g(x_i), g(x)) = 0.$$

y por tanto,

$$d'(f(x), g(x)) < 2\epsilon.$$

De aquí se desprende que $d_\infty(f, g) < 2\epsilon$ y, por consiguiente, la compacidad relativa de H en $F(X, Y_0)$, con lo cual queda probada también en $C(X, Y_0)$, y por lo tanto, en $C(X, Y)$. ■

Del teorema 4.4.6 y utilizando el ejemplo 25 y los lemas 4.4.4 y 4.4.5 se deduce el corolario siguiente:

Corolario 1. Una familia H de funciones de $C([a, b], \mathbb{R})$ es relativamente compacta con respecto a la métrica de la convergencia uniforme, si y solo si satisface las dos condiciones siguientes:

i) Existe un número real K tal que

$$|\varphi(x)| < K \text{ para todo } x \in [a, b] \text{ y } \varphi \in H.$$

ii) Para cualquier $\epsilon > 0$ existe $\delta > 0$ tal que

$|\varphi(x) - \varphi(y)| < \epsilon$ para todas las funciones $\varphi \in H$ y para todo par de puntos x, y en $[a, b]$ tales que $|x - y| < \delta$.

La demostración se deja como un ejercicio al lector.

Veamos las dos aplicaciones siguientes que hemos extraído del libro de A.N. Kolmogorov y S.V.Fomin.

Aplicación 1 (teorema de Peano) El teorema de Arzelá tiene múltiples aplicaciones. Como primera aplicación veamos el siguiente teorema de existencia para ecuaciones diferenciales ordinarias cuyo miembro es continuo. Recordemos que en el capítulo 3 vimos un teorema similar para cuando el miembro derecho satisfacía una cierta condición de Lipschitz. A diferencia de este caso, en el capítulo 3 pudimos demostrar no solo la existencia sino además la unicidad de la solución, utilizando el teorema del punto fijo.

El resultado siguiente es un teorema de existencia local.

Teorema 4.4.7 (teorema de Peano). Sea $\frac{dy}{dx} = f(x, y)$ una ecuación diferencial dada. Si la función real f es continua en un dominio compacto G de $\mathbb{R}^2$, entonces por cada punto interior (x_0, y_0) de G pasa al menos una curva integral de dicha ecuación. Es decir, para cada punto (x_0, y_0) perteneciente al interior de G , se puede hallar un número $\epsilon > 0$ y una función continuamente derivable $\varphi : [x_0 - \epsilon, x_0 + \epsilon] \longrightarrow [y_0 - \epsilon, y_0 + \epsilon]$ tal que

$$\varphi(x_0) = y_0 \text{ y}$$

$$\frac{d\varphi}{dx} = f(x, \varphi(x)) \text{ para todo } x \in [x_0 - \epsilon, x_0 + \epsilon].$$

Demostración: Puesto que la función f es continua en el compacto G , es acotada:

$$|f(x, y)| < M = \text{constante.}$$

Tracemos por el punto (x_0, y_0) las rectas con pendientes $M_y - M$. Tracemos además las rectas verticales $x = a$ y $x = b$ de manera que los dos triángulos con vértice común en (x_0, y_0) que ellas determinan pertenezcan íntegramente al interior de la región de G .

Construyamos ahora para la región dada las llamadas quebradas de Euler del modo siguiente:

Tracemos por el punto (x_0, y_0) la recta con pendiente $f(x_0, y_0)$. Tomemos en esta un punto (x_1, y_1) y tracemos, a través de él una recta de pendiente $f(x_1, y_1)$. En esta recta tomemos un punto (x_2, y_2) y tracemos, a través de él otra recta de pendiente $f(x_2, y_2)$, y así sucesivamente (fig. 69).

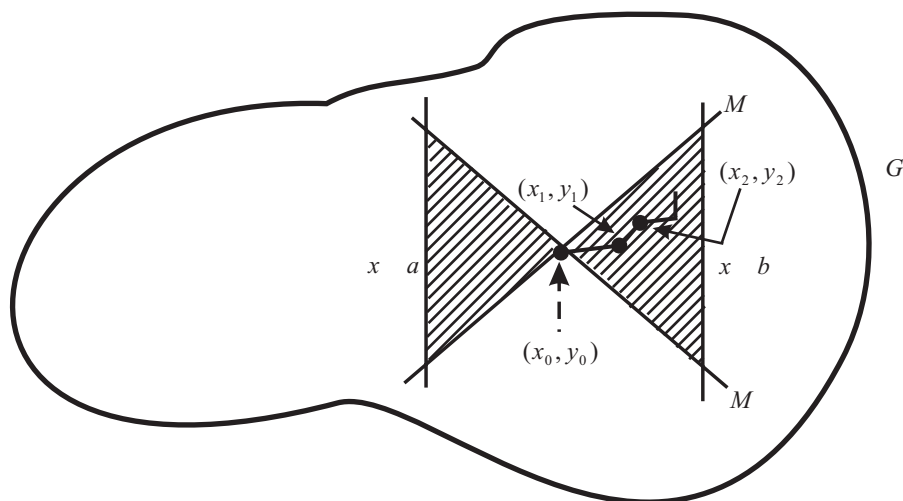


Figura 69

Consideremos ahora la sucesión de quebradas de Euler $L_1, L_2, \dots, L_n, \dots$, que pasan por el punto (x_0, y_0) y tales que la longitud del mayor de los eslabones de la quebrada L_k tiende a cero cuando $k \rightarrow \infty$.

Sea φ_k la función, cuya gráfica es la línea L_k , no es difícil verificar que las funciones continuas $\varphi_1, \varphi_2, \dots, \varphi_n, \dots$ poseen las propiedades siguientes:

- 1) están definidas sobre el segmento $[a, b]$;
- 2) son equiacotadas;
- 3) son equicontinuas.

En virtud del teorema de Arzelá-Ascoli, de la sucesión φ_k se puede extraer una subsucesión uniformemente convergente que denotaremos por $\varphi^{(1)}, \varphi^{(2)}, \dots, \varphi^{(n)}, \dots$

Consideremos $\varphi(x) = \lim_{k \rightarrow \infty} \varphi^{(k)}(x)$. Es evidente que $\varphi(x_0) = y_0$. Luego nos queda por probar que sobre el segmento $[a, b]$ la función φ verifica la ecuación diferencial dada. Para ello, es suficiente demostrar que para cualquier $\epsilon > 0$

$$\left| \frac{\varphi(x'') - \varphi(x')}{x'' - x'} - f(x', \varphi(x')) \right| < 2\epsilon \quad (12)$$

siempre que $|x'' - x'|$ sea una cantidad suficientemente pequeña. Supongamos que x'' y x' están fijos y tenemos k suficientemente grande de manera que $d_\infty(\varphi^{(k)}, \varphi) < \frac{\epsilon}{2} |x'' - x'|$. Entonces,

$$\left| \frac{\varphi^{(k)}(x'') - \varphi^{(k)}(x')}{x'' - x'} - \frac{\varphi(x'') - \varphi(x')}{x'' - x'} \right| < \epsilon,$$

y por lo tanto, aplicando la desigualdad triangular

$$\left| \frac{\varphi(x'') - \varphi(x')}{x'' - x'} - f(x', \varphi(x')) \right| \leq \epsilon +$$

$$+ \left| \frac{\varphi^{(k)}(x'') - \varphi^{(k)}(x')}{x'' - x'} - f(x', \varphi(x')) \right|$$

Luego, para probar (12) es suficiente demostrar que existe k_0 tal que si $k > k_0$ se tiene

$$\left| \frac{\varphi^{(k)}(x'') - \varphi^{(k)}(x')}{x'' - x'} - f(x', \varphi(x')) \right| < \epsilon, \quad (13)$$

siempre que la diferencia $|x'' - x'|$ sea suficientemente pequeña.

Puesto que f es continua en la región G , para cualquier $\epsilon > 0$ se puede encontrar $\eta > 0$ tal que

$$f(x', y') - \epsilon < f(x, y) < f(x', y') + \epsilon,$$

donde

$$y' = \varphi(x'),$$

cuando

$$|x - x'| < 2\eta, |y - y'| < 4M\eta.$$

El conjunto de los puntos $(x, y) \in G$ que verifican estas dos desigualdades constituye un rectángulo Q . Sea ahora k_0 tan grande que para todo $k \geq k_0$ se tenga

$$|\varphi(x) - \varphi^{(k)}(x)| < 4\eta$$

y además, todos los eslabones de la quebrada L_k tengan longitud menor que η . Entonces si $|x - x'| < 2\eta$, todas las quebradas de Euler $\varphi^{(k)}$ correspondientes a $k \geq k_0$ se encuentran completamente incluidas en el interior de $\overline{Q}$.

Sean ahora $(a_0, b_0), (a_1, b_1), \dots, (a_{n+1}, b_{n+1})$ los vértices de la quebrada $\varphi^{(k)}$, donde

$$a_0 \leq x' < a_1 < a_2 < \dots < a_n < x'' \leq a_{n+1}.$$

(Supongamos, para concretar, que $x'' > x'$ ya que el caso $x'' < x'$ se considera análogamente).

Entonces,

$$\varphi^{(k)}(a_1) - \varphi^{(k)}(x') = f(a_0, b_0)(a_1 - x')$$

$$\varphi^{(k)}(a_{i+1}) - \varphi^{(k)}(a_i) = f(a_i, b_i)(a_{i+1} - a_i)$$

$$= 1, 2, \dots, n-1,$$

$$\varphi^{(k)}(x'') - \varphi^{(k)}(a_n) = f(a_n, b_n)(x'' - a_n).$$

De aquí, para $|x'' - x'| < \eta$, obtenemos

$$(f(x', y') - \epsilon)(a_1 - x') < \varphi^{(k)}(a_1) - \varphi^{(k)}(x') <$$

$$< (f(x', y') + \epsilon)(a_1 - x');$$

$$(f(x', y') - \epsilon)(a_{i-1} - a_i) < \varphi^{(k)}(a_{i+1}) - \varphi^{(k)}(a_i) <$$

$$< (f(x', y') + \epsilon)(a_{i+1} - a_i), i = 1, 2, \dots, n-1$$

$$(f(x', y') - \epsilon)(a_{i-1} - a_i) < (x'' - a_n) <$$

$$< \varphi^{(k)}(x'') - \varphi^{(k)}(a_n) < (f(x', y') + \epsilon)(x'' - a_n).$$

Sumando estas desigualdades encontramos que

$$(f(x', y') - \epsilon)(x'' - x') < \varphi^{(k)}(x'') - \varphi^{(k)}(x') <$$

$$< (f(x', y') + \epsilon)(x'' - x')$$

que es lo que queríamos demostrar.

Notemos que diferentes subsucesiones de la sucesión de quebradas de Euler, pueden converger a diferentes soluciones de la ecuación $\frac{dy}{dx}(fx, y)$. Por eso, la solución que hemos obtenido no es, en general, la única solución de la ecuación dada y que pasa por el punto (x_0, y_0) .

Aplicación 2 (existencia de curvas geodésicas)

Sea dada una aplicación continua

$$P = f(t)$$

del segmento $a \leq t \leq b$ en un espacio métrico (X, d) . Cuando t “recorre” el segmento desde a hasta b , el punto correspondiente P “recorre” un cierto subconjunto en el espacio X que se llama *curva continua en X parametrizada por f* . Se considera que el orden en el que se recorren los puntos de la curva es una propiedad intrínseca de la propia curva, o sea, la definición de una curva no depende solamente de la imagen de f , sino también de sus propiedades como función del parámetro t . Por ejemplo, un mismo conjunto, como el representado en la figura 70, recorrido en las dos direcciones señaladas, será considerado como diferentes curvas.

En calidad de un segundo ejemplo, consideremos la función real definida sobre el segmento $[0, 1]$ la cual está representada en la figura 71. Dicha figura representa la gráfica de una curva en el segmento $[0, 1]$ del eje Y y distinta de este segmento recorrido una vez desde el punto 0 hasta el 1, ya que el segmento $[A, B]$ se recorre tres veces (dos hacia arriba y una hacia abajo).

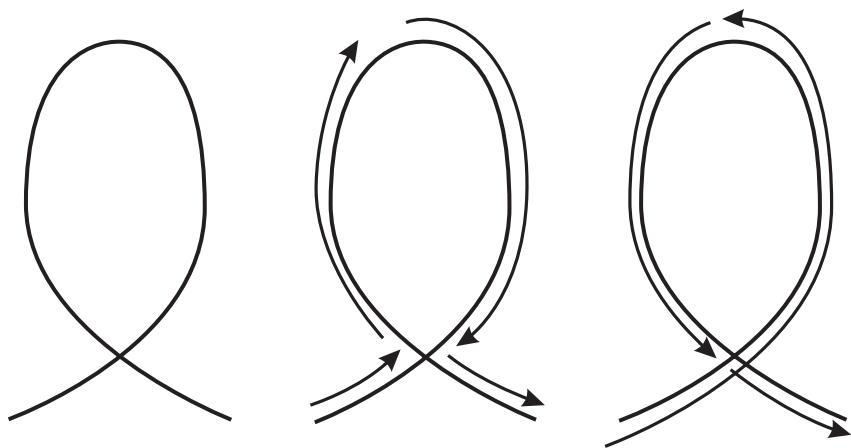


Figura 70

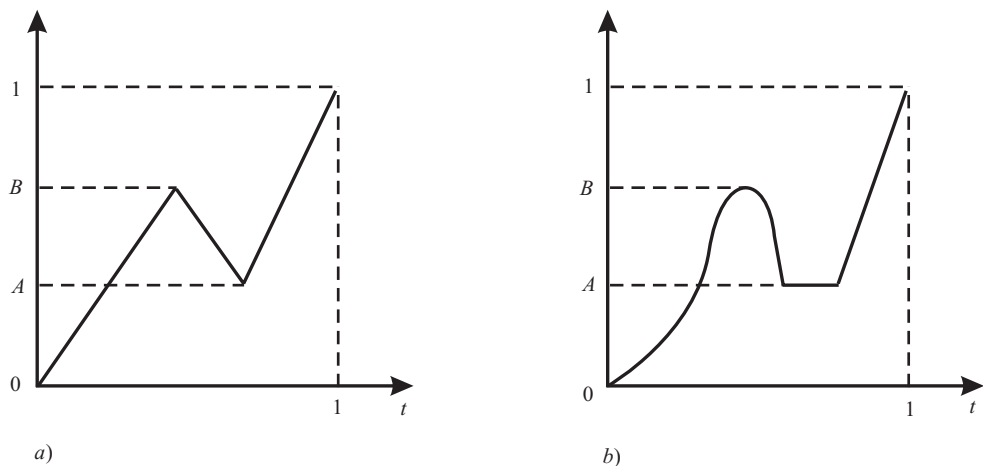


Figura 71

Sin embargo, si los puntos del espacio se recorren en el mismo orden, consideraremos que la selección del parámetro t no es esencial. Por ejemplo, las funciones representadas en la figura 71 determinan una misma curva, situada sobre el eje Y , aún cuando los valores del parámetro t , correspondientes a algún punto de la curva en $[0, 1]$, resulten diferentes. Por ejemplo, en el caso (a) de la figura 71, al punto A le corresponden sobre el eje T dos puntos aislados, mientras que en el segundo caso corresponden sobre el eje T un punto aislado y un segmento situado a su derecha (cuando t recorre este segmento el punto de la curva se mantiene fijo).

Hablando en términos un poco más físicos, no distinguiremos entre dos curvas recorridas por partículas que siguen la misma trayectoria aunque la velocidad del desplazamiento sea diferente. En efecto, como veremos más adelante a nosotros nos interesa tan solo la longitud total de la trayectoria recorrida.

Para formalizar los conceptos introducidos pasemos a las definiciones siguientes:

Dos funciones continuas

$$p' = f'(t') \text{ y } p'' = f''(t'')$$

definidas respectivamente sobre los segmentos

$$a' \leq t' \leq b' \text{ y } a'' \leq t'' \leq b''$$

y con valores en un mismo espacio métrico (X, d) se llaman *equivalentes*, cuando existen un par de funciones continuas no decrecientes

$$t' = \varphi'(t) \text{ y } t'' = \varphi''(t)$$

definidas sobre un segmento

$$a \leq t \leq b$$

y que poseen las propiedades

- 1) $\varphi'(a) = a', \varphi'(b) = b'$
- 2) $\varphi''(a) = a'', \varphi''(b) = b''$
- 3) $f'[\varphi'(t)] = f''[\varphi''(t)]$, para todo $t \in [a, b]$.

Es fácil ver que esta relación es de equivalencia, y por ello, todas las funciones continuas del tipo considerado se dividen en clases de funciones. Cada una de estas clases define una curva continua en el espacio (X, d) .

Es fácil ver que para toda función $p = f'(t')$, definida sobre un segmento $[a', b']$, existe una función equivalente a ella definida sobre el segmento $[a'', b''] = [0, 1]$. En efecto, es suficiente tomar

$$t' = \varphi'(t) = (b' - a')t + a',$$

$$t'' = \varphi''(t) = t.$$

Por consiguiente, podemos suponer que toda curva viene dada en forma paramétrica mediante una función definida sobre el segmento $[0, 1]$.

Por eso resulta oportuno introducir el espacio $C(I, \mathbb{R})$ de las aplicaciones continuas del segmento $I = [0, 1]$ en el espacio (X, d) con la métrica

$$d_{\infty}(f, g) = \sup_{t \in [0, 1]} d(f(t), g(t)).$$

Diremos que la sucesión de curvas $L_1, L_2, \dots, L_n, \dots$ converge a la curva L , cuando las curvas L_n pueden ser representadas paraméricamente en la forma

$$P_n = f_n(t), \quad 0 \leq t \leq 1,$$

y la curva L , en la forma

$$P = f(t), \quad 0 \leq t \leq 1$$

de manera que $d_{\infty}(f, f_n) \rightarrow 0$ cuando $n \rightarrow \infty$.

Aplicando el teorema de Arzelá-Ascoli, es fácil demostrar el teorema siguiente:

Teorema 4.4.8. Si la sucesión de curvas $L_1, L_2, \dots, L_n, \dots$ situadas en un compacto K , se puede representar paraméricamente mediante funciones equicontinuas sobre el segmento $[0, 1]$, se puede extraer de ella una subsucesión convergente.

Definamos ahora la longitud de una curva, que en forma paramétrica viene dada por la función

$$P = f(t), a \leq t \leq b,$$

como el extremo superior de las sumas del tipo

$$\sum_{i=1}^n d(f(t_{i-1}), f(t_i))$$

donde los puntos t_i están solamente sujetos a las condiciones

$$a = t_0 \leq t_1 \leq \dots \leq t_n = b.$$

Es fácil ver que la longitud de una curva no depende de su representación paramétrica. Si nos limitamos a las representaciones paramétricas por medio de funciones, definidas sobre el segmento $[0, 1]$ es fácil demostrar que la longitud de una curva es una funcional semicontinua inferiormente de f (en el espacio $C(I, \mathbb{R})$). En el lenguaje geométrico este resultado puede ser enunciado en la forma del siguiente teorema sobre semicontinuidad:

Teorema 4.4.9. Si la sucesión de curvas L_n converge a la curva L , la longitud de esta curva no es mayor que el límite inferior de las longitudes de las curvas L_n .

Consideremos ahora especialmente las curvas de longitud finita. Supongamos que la curva está definida por la función paramétrica

$$P = f(t), a \leq t \leq b.$$

La función f , considerada solamente en el segmento $[a, T]$, donde $a \leq T \leq b$, define el *segmento* inicial de la curva desde el punto

$$P_a = f(a)$$

hasta el punto

$$P_T = f(T).$$

Sea

$$s = \varphi(t)$$

su longitud. Es fácil probar que

$$P = g(s) [f\varphi^{-1}(s)]$$

es una nueva representación paramétrica de la misma curva. Aquí s recorre el segmento $0 \leq s \leq S$, donde S es la longitud de toda la curva considerada. Esta representación verifica la exigencia

$$d(g(s_1), g(s_2)) \leq |s_2 - s_1|$$

(la longitud del arco es mayor o igual que la longitud entre dos puntos de la cuerda).

Pasando al segmento $[0, 1]$, obtenemos la representación paramétrica

$$P = F(\tau) = g(s), \quad \tau = \frac{s}{S}$$

que verifica la condición de Lipschitz

$$d(F(\tau_1), F(\tau_2)) \leq S |\tau_1 - \tau_2|.$$

Veamos, por consiguiente, que para todas las curvas de longitud $S \leq M$, donde M es un constante, es posible una representación paramétrica, mediante funciones equicontinuas, definidas sobre el segmento $[0, 1]$. Por lo tanto, a estas es aplicable el teorema 4.4.8.

Mostremos el alcance de los resultados generales obtenidos aplicándolos a la demostración de la siguiente proposición importante, la cual se refiere a la existencia de curvas geodésicas, es decir, curvas de longitud mínima que unen dos puntos de un cierto conjunto.

Teorema 4.4.10. Si dos punto A y B de un compacto K pueden unirse por medio de una curva continua de longitud finita, entonces entre esas curvas existe una de longitud mínima.

Demostración: En efecto, sea I el extremo inferior de las longitudes de las curvas que unen los puntos A y B del compacto K . Supongamos que (L_n) es un sucesión de curvas en el compacto que unen los puntos A y B , tales que la sucesión (I_n) de sus longitudes converge a I . De acuerdo con el teorema 4.4.8, de la sucesión (L_n) se puede extraer una convergente. De acuerdo con el teorema 4.4.9, la curva límite de esta subsucesión no puede tener longitud mayor que I . ■

Observemos que incluso en el caso en que K es una superficie cerrada suave (diferenciable suficiente número de veces) del espacio euclidean de tres dimensiones, este teorema no se desprende directamente de los resultados que se establecen en el curso de Geometría Diferencial, donde se considera, generalmente, solo el caso de puntos A y B suficientemente próximos.

Todo lo expuesto adquiriría mayor claridad si hubiésemos provisto de una estructura de espacio métrico, al conjunto de todas las curvas del espacio métrico dado (X, d) . Esto se puede hacer definiendo las distancias entre las curvas L_1 y L_2 mediante la fórmula

$$d_\infty(L_1, L_2) = \inf d_\infty(f_1, f_2),$$

donde el extremo inferior se toma respecto a todos los pares de representaciones paramétricas de la curva L_1 por medio de la función

$$P = f_1(t), 0 \leq t \leq 1$$

y de la curva L_2 por medio de la función

$$P = f_2(t), 0 \leq t \leq 1.$$

La demostración de que esta distancia verifica los axiomas de métrica es sencilla, a excepción de un detalle ya que ofrece ciertas dificultades demostrar que de

$$d_\infty(L_1, L_2) = 0$$

se deduce la identidad de las curvas L_1 y L_2 . Este resultado es consecuencia directa del hecho de que el extremo inferior en la fórmula, mediante la cual hemos definido la distancia $d_\infty(L_1, L_2)$, se alcanza si se escogen adecuadamente las representaciones paramétricas f_1 y f_2 pero la demostración de esta última proposición tampoco es sencilla.

Ejercicios de Autocontrol

1. Demuestre que el cubo de Hilbert es homeomorfo a $I^{\mathbb{N}}$, provisto de la métrica producto, donde $I = [0, 1]$.
2. Precise los detalles en el desarrollo del ejemplo 19.
3. Demuestre las afirmaciones de los ejemplos 22 y 23
4. Pruebe que

i) $H = \{f_\alpha \in C([1, \infty[, \mathbb{R}) : f_\alpha(x) = x^\alpha; \alpha \leq 1\}$

es equicontinuo

ii) $H = \{f_\alpha \in C([1, 0[, \mathbb{R}) : f_\alpha(x) = x^\alpha; \alpha > 1\}$

no es equicontinuo.

Utilice la orientación gráfica del ejemplo 24.

5. Demuestre que la equicontinuidad del conjunto $H \subset C(X, Y)$ es equivalente a la continuidad uniforme de la aplicación evaluación:

$$E \quad (f, x) \mapsto f(x) \longrightarrow Y$$

4.5 ESPACIOS LOCALMENTE COMPACTOS. COMPACTIFICACIONES

Los espacios clásicos $\mathbb{R}, \mathbb{C}, \mathbb{R}^n, \mathbb{C}^n$ no son compactos, pero cada uno de sus puntos posee un sistema fundamental de vecindades compactas. Es decir, localmente ellos son compactos. Espacios con esta propiedad juegan un papel importante en la matemática, por ejemplo, en la teoría de integración de Radon.

Definición 4.5.1. Decimos que un espacio métrico (X, d) es *localmente compacto* si cada uno de sus puntos posee un sistema fundamental formado por vecindades compactas.

Observación: Algunos autores llaman localmente compacto a todo espacio métrico cuyos puntos tengan cada uno al menos una vecindad compacta. Esto no coincide con la noción intuitiva de lo que debe ser una propiedad local. En efecto, una propiedad local debe transmitirse a todo subconjunto abierto, para lo cual, hace falta dar la definición en término de sistemas fundamentales de vecindades y no de una sola vecindad. Sin embargo, veremos más tarde que en los espacios métricos ambas definiciones son equivalentes, es decir, el hecho de existir una vecindad compacta implica la existencia de un sistema fundamental de vecindades compactas.

Ejemplo 27. Todo espacio métrico discreto (X, d) es localmente compacto, puesto que el conjunto unitario $\{\{x\}\}$ formado por el compacto $\{x\}$ es un sistema fundamental de vecindades de x para todo $x \in X$.

Ejemplo 28. Todos los espacios $\mathbb{R}, \mathbb{C}, \mathbb{R}^n, \mathbb{C}^n$ son localmente compactos.

Ejemplo 29. El subespacio $\mathbb{Q} \subset \mathbb{R}$ con la métrica inducida por el módulo, no es localmente compacto ya que ningún intervalo de números racionales es compacto en $\mathbb{Q}$.

Ejemplo 30. Todo espacio métrico compacto (X, d) es localmente compacto, puesto que las vecindades cerradas de $x \in X$ forman un sistema fundamental de vecindades (compactas) de x .

Proposición 4.5.2 (caracterización de los espacios métricos localmente compactos). En un espacio métrico son equivalentes las proposiciones siguientes:

- i) Todo punto posee un sistema fundamental de vecindades compactas.

- ii) Todo subconjunto compacto posee un sistema fundamental de vecindades compactas.
- iii) Existe una base de la topología formada por abiertos relativamente compactos.
- iv) Todo punto posee una vecindad compacta.

Demostración: ii) $\Rightarrow$ i) es evidente ya que cada punto es, en particular, un subconjunto compacto.

Probemos inversamente i) $\Rightarrow$ ii). En efecto, C es compacto y $U \supset C$ abierto, entonces todo $x \in C$ posee una vecindad compacta $V_x \subset U$. Las vecindades $\{V_x\}_{x \in C}$ recubren C , por lo tanto, existe un recubrimiento finito, es decir

$C \subset \bigcup_{i=1}^n V_{x_i}$. Evidentemente $\bigcup_{i=1}^n V_{x_i}$ es una vecindad compacta de C

i) $\Rightarrow$ iii) Si las vecindades compactas V de x forman un sistema fundamental de vecindades, entonces lo hacen también las vecindades abiertas relativamente compactas $\overset{\circ}{V}$. Obviamente, la unión de todos estos sistemas fundamentales de vecindades relativamente compactas de x , cuando x recorre todo el espacio, constituye una base topológica formada por abiertos relativamente compactos.

iii) $\Rightarrow$ iv) es trivial por lo que no lo demostraremos.

Probemos que iv) $\Rightarrow$ i)

Sea V una vecindad compacta de x y sea $k \in \mathbb{N}$ tal que $\overline{B}(x, 1/k) \subset V$ (donde $\overline{B}(x, 1/k)$ denota la bola cerrada); entonces $\{\overline{B}(x, 1/k)\}_{n \geq k}$ constituye un sistema fundamental de vecindades compactas de x ya que $\overline{B}(x, 1/k) \subset V$ y todo subconjunto cerrado de un conjunto compacto es también compacto.



Corolario 1. Todo espacio métrico localmente compacto y AN2, posee una base topológica numerable de abiertos relativamente compactos.

La demostración se desprende del criterio iii) y del teorema de Lindelöf 4.2.6 ii).

Los espacios euclidianos $\mathbb{R}^n$ y $\mathbb{C}^n$ además de ser localmente compactos, tienen la importante propiedad de poder expresarse como la unión de una familia numerable de conjuntos compactos.

Tales espacios se llaman $\sigma(\text{sigma})$ -compactos y resultan de gran utilidad en el análisis (vea teorema de Arzelá-Ascoli para espacios σ -compactos y su aplicación al estudio de las familias normales de funciones holomorfas).

Definición 4.5.3. Un espacio métrico localmente compacto se llama σ -compacto si se puede expresar como la unión de una familia numerable de sub-

conjuntos compactos.

Proposición 4.5.4 (caracterización de los espacios σ -compactos). Para un espacio métrico localmente compacto los enunciados siguientes son equivalentes:

- i) (X, d) es AN2;
- ii) (X, d) es σ -compacto;
- iii) X es la unión de una familia numerable de abiertos relativamente compactos (U_n) tales que $\overline{U}_n \subset U_{n+1}$.

Demostración: i) $\Rightarrow$ ii) Si (X, d) es AN2, entonces por el corolario 1 de la proposición 4.5.2, posee una base numerable $\{V_n\}$ de abiertos relativamente compactos, luego, X puede expresarse como la unión de la familia numerable de subconjuntos compactos $\{\overline{V}_n\}$.

ii) $\Rightarrow$ iii) Sea $X = \bigcup_{n \in \mathbb{N}} K_n$. Definimos $C_n = \bigcup_{j=0}^n K_j$ construyamos (U_i) por inducción. Supongamos construidas $U_1, U_2, \dots, U_n$. Para cada $x \in \overline{U}_n \cup C_{n+1}$ existen vecindades abiertas relativamente compactas. Un número finito de ellas cubre el compacto $\overline{U}_n \cup C_{n+1}$. Consideremos U_{n+1} igual a la unión de estas vecindades, tal sucesión es la deseada.

iii) $\Rightarrow$ ii) es trivial.

ii) $\Rightarrow$ i) Sea $X = \bigcup_{n=1}^{\infty} K_n$ y sea $\{U_\alpha\}_{\alpha \in I}$ un cubrimiento abierto de X . Por ser K_n compacto, para cada n existe un subconjunto finito I_n de I tal que la familia $\{U_\alpha\}_{\alpha \in I_n}$ es un cubrimiento de K_n .

Obviamente la familia $\{U_\alpha\}_{\alpha \in I}$, donde $J = \bigcup_{n=1}^{\infty} I_n$ es un subcubrimiento numerable de X . Luego, (X, d) es de Lindelöf, y por lo tanto, AN2. ■

Por definición, todo subconjunto abierto de un espacio localmente compacto (X, d) es a su vez localmente compacto (como subespacio). Por otro lado, es fácil ver que también todo subconjunto cerrado (X, d) es localmente compacto. Si nos preguntamos cuál es la forma más general de un subconjunto localmente compacto de (X, d) , obtenemos la siguiente caracterización interesante.

Proposición 4.5.5 (caracterización de los subconjuntos localmente compactos).

- i) En un espacio métrico (X, d) todo subconjunto localmente compacto es intersección de un abierto y de un cerrado.
- ii) Si (X, d) es localmente compacto toda intersección de un cerrado y de un abierto en (X, d) es localmente compacto.
- iii) En un espacio localmente compacto, un subconjunto es localmente compacto si y solo si puede representarse como intersección de un abierto y de un cerrado en (X, d) .

Demostración:

- i) Sea $A \subset X$ localmente compacto. Para todo $x \in A$ existe una vecindad V_x en (X, d) tal que $A \cap V_x$ es compacto, y por lo tanto cerrado en (X, d) .

Sea $U := \bigcup_{x \in A} V_x$; U es abierto y contiene a A . Todo subconjunto $\overset{\circ}{V}_x \cap A = \overset{\circ}{V}_x \cap (V_x \cap A)$ es cerrado en el abierto $\overset{\circ}{V}_x$; por lo tanto, $A = U \cap A = \bigcup_{x \in A} (\overset{\circ}{V}_x \cap A)$ es cerrado en U (demuéstrello). Concluimos que existe un cerrado $C \subset X$ tal que $A = U \cap C$.

- ii) Sea A un abierto y C un cerrado en (X, d) y sea $x \in A \cap C$ y U una vecindad abierta de x . Hay que buscar una vecindad compacta de X en el subespacio $A \cap C$ que esté contenida en U . Por hipótesis, existe una vecindad compacta W de x contenida en la vecindad abierta $A \cap U$. $V := W \cap C = W \cap (A \cap C) \subset U$ es una vecindad compacta de x en $A \cap C$.

$V : W \cap C = W \cap (A \cap C) \subset U$ es una vecindad compacta de x en $A \cap C$.

- iii) Es una consecuencia inmediata de i) y ii). ■

Ejemplo 31. Toda variedad algebraica en $\mathbb{R}^n$, es decir, el conjunto de puntos $\{(x_1, \dots, x_n) : \rho\{(x_1, \dots, x_n) = 0\}\}$, donde ρ es un polinomio, es localmente compacta porque es cerrada

Corolario 1. Sea (X, d) un espacio de métrico; entonces la intersección de dos subconjuntos localmente compactos en (X, d) es localmente compacto.

La demostración es consecuencia inmediata de la caracterización iii) de la proposición 4.5.5.

A diferencia del corolario 1 de la proposición 4.5.5, la unión de dos subconjuntos localmente compactos en un espacio métrico, no necesariamente es localmente compacto.

Una diferencia importante con la compacidad es que la compacidad local no se transmite a las imágenes continuas.

Ejemplo 32. Sea $f : [0, \infty[\rightarrow \mathbb{R}^2$ definida por

$$f(x) = \begin{cases} (x, \sin 1/x), & \text{si } 0 < x \leq 1/n \\ (2/\pi - x,) & \text{si } 1/\pi \leq x \end{cases}$$

cuyo gráfico se observa en la (fig. 72).

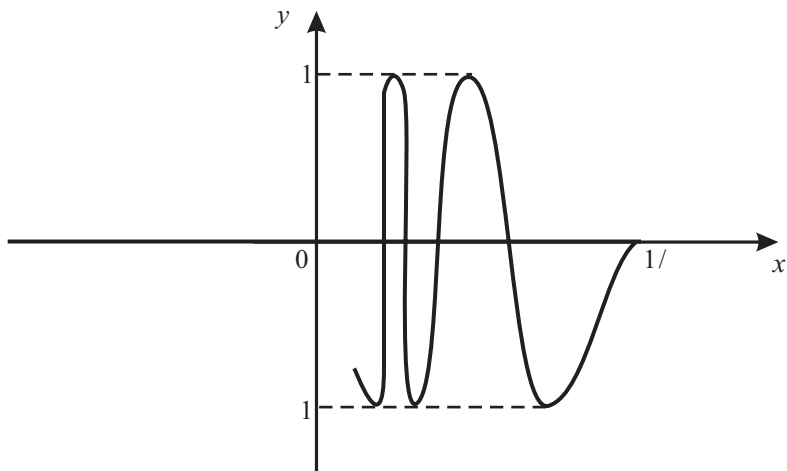


Figura 72

Es fácil verificar que f es continua, pero $f[0, \infty[$ no es localmente compacto aunque $[0, \infty[$ sí lo es. Obsérvese que $(0, 0) \in f[0, \infty[$ no posee ninguna vecindad compacta pues ella tendría que ser cerrada en $\mathbb{R}^2$ lo que no es posible porque toda vecindad de $(0, 0)$ en $f[0, \infty[$ contiene una sucesión de puntos que converge a un punto $(0, \epsilon) \notin f[0, \infty[$ ($\epsilon > 0$). Es de esperar que esto no pueda ocurrir si f es además abierta.

Proposición 4.5.6. La compacidad local es invariante por aplicaciones que sean a la vez continuas y abiertas.

Demostración: Sea (X, d) localmente compacto, (Y, d') un espacio métrico y $f : X \rightarrow Y$ una aplicación continua y abierta. Debemos mostrar que una vecindad arbitraria W de un punto $y = f(x)$ en $f[X]$ contiene a una vecindad (relativa) compacta de y . Pero esto es evidente, porque $f^{-1}[W]$ es una vecindad

de x y contiene, por lo tanto, una vecindad compacta V de x cuya imagen $f[V] \subset W$ es una vecindad compacta de y en $f[X]$ (por ser f abierta). ■

Concluimos que la compacidad local es una propiedad topológica. Incluso podemos deducir más. La compacidad local es invariante ante homeomorfismos locales, porque todo homeomorfismo local es a la vez abierto y continuo.

Otra diferencia con la compacidad es que la compacidad local no se transmite a productos cualesquiera. En efecto, supongamos que V sea una vecindad compacta de un punto $x = (x_n)$ en un espacio producto $\prod X_n$; entonces las proyecciones $\pi_n[V]$ son todas compactas. Pero sabemos que casi todas estas proyecciones son iguales a los espacios coordenados X_n . Concluimos que si existe en un espacio producto, al menos una vecindad compacta, entonces casi todos los espacios coordenados, excepto un número finito, son compactos.

Proposición 4.5.7 (caracterización de los espacios producto localmente compactos). Sea $((X_n, d_n))_{n \in \mathbb{N}}$ una familia numerable de espacios métricos. $\prod_{n \in \mathbb{N}} X_n$ es localmente compacto si y solo si todo espacio coordinado es localmente compacto, y a partir de un cierto n_0 todos los espacios X_n ($n \geq n_0$) son compactos.

La demostración se deja al lector.

Pasemos ahora a la caracterización de las familias relativamente compactas en $C(X, Y)$ con respecto a la métrica de la convergencia uniforme sobre cada compacto, cuando (X, d) es localmente compacto en lugar de compacto, y (Y, d') un espacio métrico cualquiera. Primeramente precisemos la métrica que vamos a considerar sobre $C(X, Y)$. Para ello, consideraremos que el espacio (X, d) es σ -compacto. En este caso, según la proposición 4.5.4 (iii), X puede expresarse como la unión de una familia numerable de abiertos relativamente compactos (U_n) tales que $\overline{U}_n \subset U_{n+1}$. La convergencia uniforme sobre cada subconjunto compacto $\overline{U}_n$ viene determinada por la pseudométrica

$$d_{\infty, n}(f, g) = \sup_{x \in \overline{U}_n} d'(f(x), g(x)); \quad (14)$$

$$f, g \in C(X, Y).$$

Definamos

$$\rho_{\infty, n} = \frac{d_{\infty, n}}{1 + d_{\infty, n}}, \quad (15)$$

$$\rho_{\infty} = \sum_{n=1}^{\infty} \frac{1}{2^n} \rho_{\infty, n}. \quad (16)$$

Evidentemente $d_{\infty, n}$ y $\rho_{\infty, n}$ son pseudométricas topológicamente equivalentes para cada $n \in \mathbb{N}$, y no es difícil comprobar que ρ_{∞} es una métrica sobre $C(X, Y)$ cuya convergencia asociada corresponde a la convergencia uniforme sobre cada uno de los compactos $\overline{U}_n$.

Más aún, veremos que la definición de esta métrica no depende de la elección de la familia (U_n) , siempre que satisfaga la propiedad (iii) de la proposición 4.5.4.

Proposición 4.5.8 (métrica de la convergencia uniforme sobre los subconjuntos compactos). Sean (X, d) σ -compactos y (X, d') métrico.

Sea (U_n) una sucesión de subconjuntos abiertos en (X, d) que satisface la propiedad (iii) de 4.5.4 y sea ρ_{∞} la métrica asociada a (U_n) según (16).

Entonces se cumple:

- i) El espacio $(C(X, Y), \rho_{\infty})$ es completo
- ii) La sucesión (f_n) de funciones de $C(X, Y)$ converge a f según ρ_{∞} si y solo si (f_n) converge a f uniformemente sobre cada compacto K de X . Es decir, para cada K en X

$$\sup_{x \in K} d'(f_n(x), f(x)) \longrightarrow 0. \quad (17)$$

Demostración: La demostración de i) se deja de ejercicio al lector.

ii) Sea (f_n) una sucesión de funciones en $C(X, Y)$ que converge a f según ρ_{∞} y sea K un subconjunto compacto de X . La sucesión (U_n) es un cubrimiento abierto de K , y por tanto, contiene un subcubrimiento finito $\{U_{i_1}, \dots, U_{i_k}\}$, donde $i_1 < i_2 < \dots < i_k$. Por ser la sucesión (U_n) creciente se tiene que

$$K \subset U_{i_k}.$$

Luego,

$$\sup_{x \in K} d'(f_n(x), f(x)) \leq d_{\infty, i_k}(f_n, f). \quad (18)$$

Pero si $\rho_{\infty}(f_n, f) \xrightarrow{n \rightarrow \infty} 0$, entonces $\rho_{\infty, i}(f_n, f) \xrightarrow{n \rightarrow \infty} 0$ para cada i , y como $\rho_{\infty, i}$ y $d_{\infty, i}$ son pseudométricas topológicamente equivalentes, se tendrá que

$d_{\infty, i_k}(f_n, f) \xrightarrow{n \rightarrow \infty} 0$, donde por (18), se deduce que $\sup_{x \in K} d'(f_{(x)_n}, f(x)) \xrightarrow{n \rightarrow \infty} 0$, o sea (f_n) converge uniformemente a f sobre K .

Si ahora suponemos que (f_n) converge a f uniformemente sobre cada compacto, en particular tendremos que se verifica la convergencia uniforme sobre cada conjunto $\overline{U}_k$, es decir, $d_{\infty, k}(f_n, f) \xrightarrow{n \rightarrow \infty} 0$, y por tanto, $\rho_{\infty, k}(f_n, f) \xrightarrow{n \rightarrow \infty} 0$. Pero sabemos que entonces (vea capítulo 2) (f_n) converge a f con respecto a ρ_{∞} . ■

Observación: De la proposición anterior se deduce que todas las métricas ρ_{∞} , construídas a partir de sucesiones de abiertos (U_n) que satisfagan la propiedad (iii) de 4.5.4, son topológicamente equivalentes sobre $C(X, Y)$ cuando el espacio (X, d) es σ -compacto. El límite definido por cada de estas métricas ρ_{∞} en $C(X, Y)$ es precisamente la convergencia uniforme sobre los subconjuntos compactos de X .

Teorema 4.5.9 (Teorema de Arzela-Ascoli para espacios σ -compactos). Sean (X, d) σ -compacto, (Y, d) métrico y proveamos al espacio $C(X, Y)$ de una métrica ρ_{∞} definida a partir de una sucesión (U_n) de abiertos en X que satisfaga la condición (iii) de 4.5.4. Entonces son equivalentes:

- i) $H \subset C(X, Y)$ es relativamente compacto en $(C(X, Y), \rho_{\infty})$.
- ii) H es equicontinuo y equiacotado.

Demostración: i) $\Rightarrow$ ii). Definamos para cada n la aplicación restricción R_n de f a $\overline{U}_n$. Evidentemente:

$$R_n : (C(X, Y), \rho_{\infty}) \longrightarrow (C(\overline{U}_n, Y), d_{\infty, n})$$

donde,

$$R_n f = f|_{\overline{U}_n} \text{ es continua (¿por qué?),}$$

Sea x elemento arbitrario de X y demostremos que H es equicontinuo y equiacotado en x . El punto x pertenece a U_n para algún n . Consideremos la correspondiente aplicación R_n ; por ser R_n continua, $R_n[H]$ es relativamente compacto en $C(\overline{U}_n, Y)$.

No es difícil comprobar que $R_n[H](x) = H(x)$ (evaluaciones en el punto x de las funciones de H).

Por ser $\overline{U}_n$ compacto, aplicando el teorema de Arzelá-Ascoli antes demostrado, tendremos entonces que H es equicontinuo y equiacotado en el punto x .

ii) $\Rightarrow$ i). Sea (f_n) una sucesión en H y demostremos que de ella se puede extraer una subsucesión convergente en $C(X, Y)$ con respecto a ρ_∞ . Para ello, basta demostrar la existencia de una subsucesión, que converge uniformemente a una cierta función continua f sobre cada compacto $\overline{U}_n$.

Es fácil ver que si H es equicontinuo y equiacotado en $C(X, Y)$, entonces también lo será $R_k[H]$ en $C(\overline{U}_k, Y)$ para cada $k \in \mathbb{N}$.

Por el teorema de Arzelá-Ascoli para espacios compactos, se tiene que para cada $k \in \mathbb{N}$ la sucesión restringida $(f_n/\overline{U}_k)$ tiene una subsucesión convergente en $C(\overline{U}_k, Y), d_{\infty, k}$.

Apliquemos de nuevo *método de diagonalización* que ya fue utilizado en la demostración del teorema de Tijonov. Para mayor comodidad denotemos la restricción

$f_n|_{U_k}$ de f_n a U_k mediante $f_n^{(k)}$

Paso 1: Para $k = 1$, sea $(f_{n_1}^{(1)})$ una subsucesión de $(f_n^{(k)})$ que converge a $f^{(1)}$ en $C(\overline{U}_1, Y)$.

Paso 2: Para $k = 2$, sea $(f_{n_2}^{(2)})$ una subsucesión de $(f_{n_1}^{(2)})$ que converge a $f^{(2)}$ en $C(\overline{U}_2, Y)$.

.....

Paso k : Sea $(f_{n_k}^{(k)})$ una subsucesión de $(f_{n_{k-1}}^{(k)})$ que converge a $f^{(n)}$ en $C(\overline{U}_k, Y)$.

De esta manera por inducción hemos construido una *sucesión diagonal* (f_{k_k}) que es una subsucesión de la sucesión original (f_n) .

Por la propia construcción de la subsucesión (f_{k_k}) se puede ver que para cada $p \in \mathbb{N}$, la sucesión $(f_{k_k}^p)$ $k \geq p$ (restricciones de f_{k_k} a $\overline{U}_p$) es una subsucesión de la sucesión $(f_{n_p}^p)$, y por lo tanto, $(f_{k_k}^{(p)})$ converge a $(f^{(p)})$ cuando $k \longrightarrow \infty$ en $(C(\overline{U}_p, Y), d_{\infty, p})$.

Por otra parte, para cada k , la sucesión $(f_{n_k}^{(k-1)})$ es una subsucesión de $(f_{n_{k-1}}^{(k-1)})$, entonces $f^k/\overline{U}_{k-1} = f^{(k-1)}$ lo cual nos permite definir unívocamente una función $f \in C(X, Y)$ mediante.

$f(x) = f^k(x), \text{ si } x \in \overline{U}_k.$

Como para cada $p \in \mathbb{N}$ sabemos que la sucesión (f_{k_k}) restringida a $\overline{U}_p$ converge uniformemente a f^p , entonces (f_{k_k}) converge con respecto a ρ_∞ a

la función f , con lo cual queda probado que H es relativamente compacto en $(C(X, Y), \rho_\infty)$. ■

Veamos ahora dos importantes aplicaciones del teorema de Arzelá-Ascoli en espacios σ -compactos.

Aplicación 1 (teorema de D'Alembert)

El teorema siguiente es también conocido como *teorema Fundamental del Algebra*.

Teorema 4.5.10 (teorema de D'Alembert). Todo polinomio $p(z) = a_0 + a_1 z + \dots + a_n z^n$, de coeficientes complejos a_i y de grado $m > 0$, tiene una raíz compleja.

Demostración: $\mathbb{C}$ es localmente compacto. Es fácil ver que la función $f : z \mapsto p(z)$ tiende al infinito si $|z| \rightarrow \infty$. Por tanto, alcanza su mínimo μ en un punto z_0 sobre $\mathbb{C}$. Supongamos por el absurdo que $p(z_0) \neq 0$. Consideremos el desarrollo de Taylor de la función polinomial p en z_0 :

$$p(z) = p(z_0) + c_k(z - z_0)^k + \dots + c_m(z - z_0)^m \quad \text{donde } c_k \neq 0.$$

Existe una circunferencia $C = \{z \in \mathbb{C} : |z - z_0| = \rho\}$

de centro z_0 tal que $|c_{k+1}(z - z_0)^{k+1} + \dots + c_m(z - z_0)^m| < |c_k(z - z_0)^k| = |c_k|\rho^k$

Si z recorre toda la circunferencia C , entonces $c_k(z - z_0)^k$ recorre toda la circunferencia de centro $p(z_0)$ y de radio $|c_k|\rho^k$. Existe entonces un punto $z_1 \in \mathbb{C}$ tal que $p(z_0) + c_k(z - z_0)^k$ pertenece al segmento $[0, p(z_0)]$ en el plano complejo. Esto significa que:

$$|p(z_0) + c_k(z_1 - z_0)^k| = \mu - |c_k|\rho^k.$$

Luego tenemos la desigualdad

$$|p(z_1)| \leq |p(z_0) + c_k(z_1 - z_0)^k| +$$

$$+ |c_{k+1}(z_1 - z_0)^{k+1} + \dots + c_m(z_1 - z_0)^m| <$$

$$< (\mu - |c_k|\rho^k) + |c_k|\rho^k = \mu,$$

que contradice a la definición de $\mu = \min\{|p(z)| : z \in \mathbb{C}\}$ (fig. 73). ■

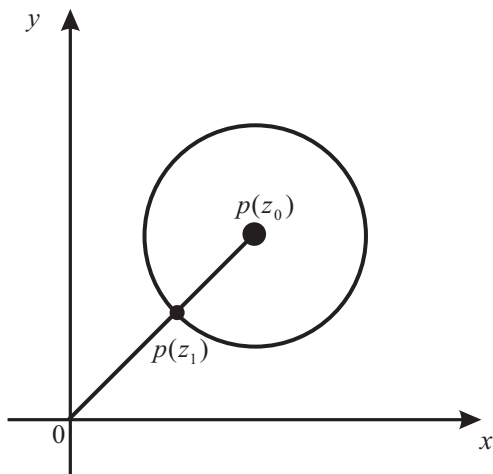


Figura 73

Del análisis complejo se sabe que toda función polinomial $p : z \rightsquigarrow p(z)$ de $\mathbb{C}$ en $\mathbb{C}$ es abierta. Si presuponemos este resultado, podemos evitar todos los cálculos en la demostración precedente. En efecto, no es difícil verificar que la función $z \rightsquigarrow |p(z)|$ admite su mínimo en un punto del plano complejo. Por otro lado, se puede probar que una función $z \rightsquigarrow |p(z)|$ no puede admitir un mínimo diferente de 0 si la función $z \rightsquigarrow p(z)$ es abierta.

Aplicación 2 (caracterización de las familias normales de funciones analíticas) Sea Ω un subconjunto abierto de $\mathbb{R}^2 \approx \mathbb{C}$ y consideremos el subespacio $H(\Omega, \mathbb{C})$ de $C(\Omega, \mathbb{C})$ constituido por todas las funciones holomorfas definidas sobre Ω .

Todo abierto Ω de $\mathbb{C}$ es localmente compacto y por eso, tiene sentido proveer a $C(\Omega, \mathbb{C})$ de la métrica ρ_∞ de la convergencia uniforme sobre cada compacto de Ω . En cualquier libro sobre teoría de funciones de variable compleja (vea por ejemplo Ahlfors), se demuestra que el límite uniforme sobre cada compacto de Ω de una sucesión de funciones holomorfas Ω , es también una función holomorfa y por lo tanto, $H(\Omega, \mathbb{C})$ es un subespacio cerrado de $(C(\Omega, \mathbb{C}), \rho_\infty)$. Como $(C(\Omega, \mathbb{C}), \rho_\infty)$ es completo, entonces $H(\Omega, \mathbb{C})$ con la métrica inducida por ρ_∞ es también un espacio métrico completo.

Pasemos ahora a caracterizar los conjuntos relativamente compactos en $(H(\Omega, \mathbb{C}), \rho_\infty)$.

Definición 4.5.11. Se llama *familia normal* sobre $\Omega \subset \mathbb{C}$ a todo subconjunto relativamente compacto de $(H(\Omega, \mathbb{C}), \rho_\infty)$.

Teorema 4.5.12 (caracterización de las familias normales). El sub-

conjunto F de $H(\Omega, \mathbb{C})$ es una familia normal si y solo si está uniformemente acotado sobre cada compacto de Ω . Es decir, si para cada compacto $K \subset \Omega$ existe una constante $M_K > 0$ tal que para toda $f \in F$ y todo $z \in K$ se cumple $|f(z)| \leq M_K$.

Demostración: Por ser $H(\Omega, \mathbb{C})$ cerrado en $C(\Omega, \mathbb{C})$, tendremos que F es relativamente compacto en $H(\Omega, \mathbb{C})$ si y solo si lo es en $C(\Omega, \mathbb{C})$. Luego, por el teorema 4.5.9, F es una familia normal si y solo si es equicontinua y equiacotada.

Si F está uniformemente acotada sobre los compactos de Ω entonces, evidentemente, está equiacotada. Demostraremos que, en este caso, F es también equicontinua. Sea z_0 un punto cualquiera de Ω y sea $K = K(z_0, r)$ un cuadrado de centro z_0 y de lado $4r > 0$ tal que $K \subset \Omega$. Entonces para dos puntos $z', z'' \in K$ y para toda función $f \in F$ se cumple, según la fórmula de Cauchy que.

$$f(z'') - f(z') = \frac{1}{2\pi i} \int_{\partial K} \frac{f(\xi)}{\xi - z''} d\xi - \frac{1}{2\pi i} \int_{\partial K} \frac{f(\xi)}{\xi - z'} d\xi.$$

Si $M_K = \sup \{|f(z)| : z \in \partial K \wedge f \in F\}$

y z', z'' varían en el disco $D = \{z \in \mathbb{C} : |z - z_0| \leq r\}$,

entonces

$|\xi - z'| \geq r$ y $|\xi - z''| \geq r$ para todo $\xi \in \partial K$, luego,

$$|f(z'') - f(z')| \leq \frac{16M_K}{\pi r} |z'' - z'|$$

de donde se concluye la equicontinuidad de F (fig. 74).

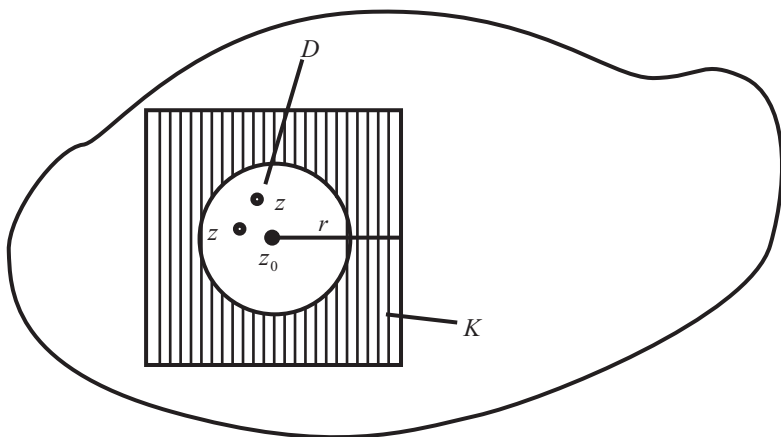


Figura 74

La condición impuesta sobre la acotación uniforme de F sobre compactos de Ω es evidentemente necesaria ya que si F es normal, la evaluación $E : (x, f) \rightsquigarrow f(x)$ de $\Omega \times F$ en $\mathbb{C}$ es continua, luego la imagen $E[K \times F] = \cup \{F(x) : x \in K\}$ es relativamente compacto en $\mathbb{C}$ para todo compacto K en Ω .

Notemos que un subconjunto F de $C(\Omega, \mathbb{C})$ está uniformemente acotado sobre los compactos de Ω si y solo si es acotado con respecto a la métrica ρ_∞ . ■

El teorema 4.5.12 nos dice que para todo subconjunto F de $(H(\Omega, \mathbb{C}), \rho_\infty)$ se tiene:

F es compacto $\Leftrightarrow F$ es cerrado y acotado,

lo cual nos recuerda la caracterización de los subconjuntos compactos en los espacios producto (vea los ejemplos 17 y 20).

El trabajo sobre espacios compactos es muy cómodo dado el gran número de propiedades a las cuales puede recurrirse. Por lo tanto, si queremos investigar un espacio métrico dado (X, d) no compacto, puede resultar útil tener una inmersión de (X, d) en un espacio compacto (Y, d') de forma tal que podamos trabajar en (Y, d') para estudiar X . Por ejemplo, podríamos utilizar en este caso la propiedad de que toda sucesión en X tendría una subsucesión convergente a un punto de Y .

Más generalmente, si tenemos un espacio (X, d) difícil de analizar, es natural buscar una inmersión f en un espacio (Y, d') que tenga propiedades cómodas o que ya conocemos mejor.

Este es el llamado *problema de inmersión* que se presenta en diferentes ramas de la Topología, en Topología Algebraica, así como en Topología General. Mencionemos solamente dos ejemplos:

Primero: En Topología diferencial se estudia el problema de cuales variedades diferenciables pueden considerarse como superficies sumergidas en $\mathbb{R}^n$.

Segundo: El problema de la construcción de los números reales consiste esencialmente en hallar una inmersión de $\mathbb{Q}$ en un cuerpo ordenado más amplio que cumpla el axioma del supremo o equivalentemente, el teorema de Cauchy (toda sucesión de Cauchy es convergente).

Ahora bien, sea $f : (X, d) \longrightarrow (Y, d')$ una inmersión. La estructura de los abiertos en el subespacio abierto $Y \setminus \overline{f[X]}$ no depende absolutamente de una

métrica d . Luego, si variamos la métrica d' sobre $Y \setminus \overline{f[X]}$, f sigue siendo una inmersión. Por lo tanto, parece natural limitarnos al subespacio $Y_0 = \overline{f[X]}$.

Definición 4.5.13. Sean (X, d) , (Y, d') espacios métricos y $f : X \longrightarrow Y$ una inmersión. El par $(j, (Y, d'))$ se llama *extensión* de (X, d) si $f[X]$ es denso en (Y, d') . La extensión $(f, (Y, d'))$ se dice *compactificación* de (X, d) si (Y, d') es compacto.

No siempre el problema de extensión va a sustituir al problema de inmersión. Por ejemplo, si queremos utilizar estructuras algebraicas sobre Y , estas estructuras en general no están definidas sobre el subconjunto $\overline{f[X]}$.

Ejemplo 33. Sea t la inyección canónica de $\mathbb{Q}$ en $\mathbb{R}$; entonces el par $(i_{\mathbb{Q}}, \mathbb{R})$ es una extensión de $\mathbb{Q}$ que no es una compactificación.

Ejemplo 34. El par $(\arctan, [-\pi/2, \pi/2])$ es una compactificación de $\mathbb{R}$.

A diferencia del completamiento que existe para cualquier espacio métrico, puede demostrarse que existen espacios métricos cuya compactificación no puede realizarse mediante un espacio métrico (Y, d') . Sin embargo, nosotros no ilustraremos este hecho para no hacer demasiado compleja nuestra exposición.

Nos dedicaremos a estudiar aquellos espacios métricos que tienen una compactificación y para ello, continuaremos con el ejemplo siguiente:

Ejemplo 35 (proyección estereográfica). Denotemos por E^1 la esfera de radio 1 con centro en $(0, 1)$ de $\mathbb{R}^2$ y sea s_1 la proyección estereográfica de dicha esfera (círculo unidad) sobre $\mathbb{R}$. No es difícil verificar que el par (s_1^{-1}, E^1) es una compactificación de $\mathbb{R}$.

Siguiendo esta idea de *unir los puntos extremos* de $\mathbb{R}$ para obtener un *círculo* (compactificación de $\mathbb{R}$), puede obtenerse también una compactificación interesante de $\mathbb{R}^2$.

En efecto consideremos la proyección estereográfica s_2 de la esfera $E^2 \subset \mathbb{R}^3$, con polo en el punto w sobre el plano $\mathbb{R}^2$. Sabemos que s_2 aplica la esfera punteada $E^2 \setminus \{w\}$ de manera homeomorfa sobre $\mathbb{R}^2$. Por tanto, el par (s_2^{-1}, E^2) es una compactificación de $\mathbb{R}^2 = \mathbb{C}$ (fig. 75) que se obtiene considerando al plano $\mathbb{R}^2$ como una hoja de papel de infinitas dimensiones, que se recoge formando una “bolsa” en la cual se ha identificado el *infinito* del plano complejo con el *punto al infinito* w .

Se dice que E^2 es la *esfera* de *Riemann* obtenida por adjunción del punto infinito w al plano $\mathbb{R}^2 = \mathbb{C}$.

Las vecindades abiertas U del punto $w \in E^2$, corresponden por el homeomorfismo s_2 a los complementos de los compactos K en $\mathbb{R}^2$. Se dice que una

sucesión (x_n) en $\mathbb{R}^2$ converge al infinito si en la compactificación E^2 de $\mathbb{R}^2$, la sucesión correspondiente $(s_2^{-1}(x_n))$ converge al punto infinito w . Esto significa que para todo compacto K en $\mathbb{R}^2$ casi todos los términos x_n ($n \in \mathbb{N}$) se quedan fuera de K .

La idea de compactificar $\mathbb{R}^2$ por adjunción de un punto infinito, cuyas vecindades sean los complementos de compactos, puede trasladarse a cualquier espacio σ -compacto, y en el contexto de una teoría más amplia como es la teoría de los espacios topológicos, esta idea puede ampliarse aún a espacios mucho más generales obteniéndose así la llamada compactificación de Alexandrovff (vea *Topología General*, D. Hinrichsen y otros 8.6).

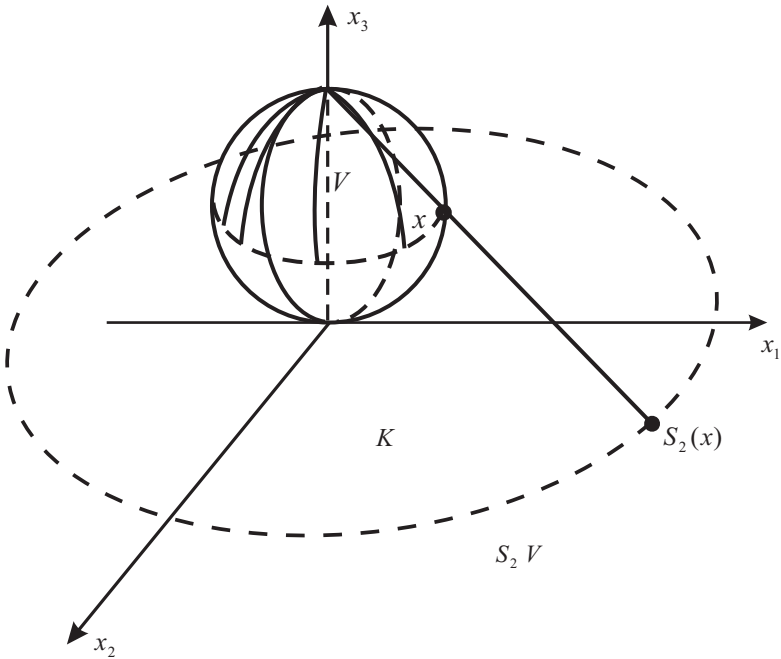


Figura 75

Nosotros no veremos aquí este tipo de compactificación ya que la construcción de una métrica en dicha compactificación presenta ciertas dificultades. Estas dificultades están relacionadas con la forma en que debe variarse la métrica en el espacio siendo topológicamente equivalente a la métrica dada, y que a su vez *mida* la cercanía al punto infinito w que se adjunta tal y como ocurre para la métrica sobre la esfera de Riemann con respecto al punto infinito.

Sin embargo, veremos otra forma muy instructiva en que pueden ser compactificados una gran clase de espacios métricos que incluye a los espacios

σ -compactos.

Teorema 4.5.14 (teorema de Urysohn). Todo espacio métrico AN2 tiene una inmersión en el espacio producto $I^{\mathbb{N}}$, donde $I = [0, 1]$.

Demostración: Por ser (X, d) normal, para cada $x \in X$ y cada $V \in \mathbf{V}(x)$ existe una función continua $\varphi_{x,V} : X \rightarrow [0, 1]$ tal que $\varphi_{x,V}(x) = 0$ y $\varphi_{x,V}|_{X \setminus V} = 1$, de donde

$$x \in \varphi_{x,V}^{-1}[[0, 1[\subset V.$$

Por lo tanto:

$$B = \{\varphi^{-1}[[0, 1[: \varphi \in C(X, [0, 1])\}$$

constituye una base de (X, d) .

Como (X, d) es AN2, según el teorema de Lindelöf, B contiene una base topológica numerable

$$B_0 = \{\varphi_n^{-1}[[0, 1[/ n \in \mathbb{N}\}.$$

Consideremos la aplicación

$$f : X \rightarrow I^{\mathbb{N}}$$

definida por

$$f : x \rightarrow (\varphi_n(x))_{n=1}^{\infty}$$

Como vimos en el capítulo 2, f es continua por ser la función producto de las aplicaciones continuas φ_n . Veamos que f es inyectiva. Si $x, y \in X$, $x \neq y$, entonces existe un abierto $\varphi_k^{-1}[[0, 1[$ que contiene a x y no contiene a y , luego

$$[0, 1[\ni \varphi_k(x) \neq \varphi_k(y) = 1,$$

y por tanto,

$$f(x) = (\varphi_n(x)) \neq (\varphi_n(y)) = f(y).$$

Falta por probar que toda imagen por f de un abierto de la forma $\varphi_k^{-1}[[0, 1[$, es abierto en $f[X]$. Del esquema de la figura 76 se deduce que

$$\begin{aligned} f[\varphi_k^{-1}[[0, 1[)] &= [[0, 1[\\ &= [f(\pi_k \circ f)^{-1}[[0, 1[) = \\ &= f[f^{-1}[\pi_k^{-1}[[0, 1[)] = \\ &= \pi_k^{-1}[[0, 1[\cap f[X] \end{aligned}$$

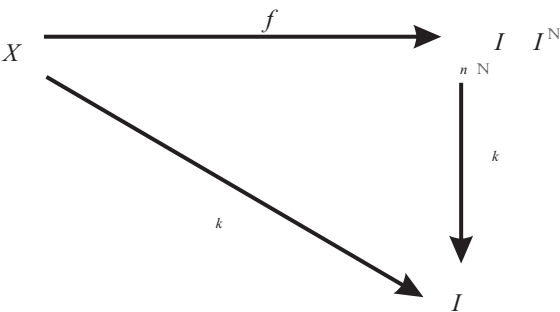


Figura 76

Sabemos que $\pi_k^{-1}[[0, 1]]$ es abierto en $I^{\mathbb{N}}$, de donde se tiene que $\pi_k^{-1}[[0, 1]] \cap \overline{f[X]}$ es abierto en $f[X]$ con lo cual se concluye el resultado deseado.■

Corolario 1 (compactificación de espacios AN2). Todo espacio métrico AN2 (en particular todo espacio σ –compacto) admite una compactificación.

Demostración: Sabemos que $I^{\mathbb{N}}$ es un espacio métrico por ser un producto numerable de espacios métricos. Además, por el teorema de Tijonov I es compacto, luego el par $(f, \overline{f[X]})$ del teorema anterior constituye una compactificación de (X, d) .■

La compactificación hallada en el corolario anterior se llama compactificación de Stone-Cech. Es posible demostrar, aunque nosotros no lo haremos, que todas las posibles compactificaciones de un espacio métrico σ –compacto se hallan “entre” la de Alexandroff y la de Stone-Cech. Más exactamente, en la familia de las compactificaciones de un espacio métrico dado, puede establecerse un orden con respecto al cual la compactificación de Alexandroff es la *menor* posible mientras que la de Stone-Cech es la *mayor*. Recalcamos que hay compactificaciones diferentes de estas dos y que precisamente las necesidades incidentales en diversos problemas particulares, son las que conducen a la búsqueda de compactificaciones que se adapten idóneamente. (Para más información al respecto vea Topología general de D. Hinrichsen y otros 8.6)

Ejercicios de Autocontrol

1. Pruebe que todo espacio métrico localmente compacto es de Baire.
2. De un ejemplo de dos subconjuntos localmente compactos en $\mathbb{R}^2$ cuya unión no sea localmente compacta.

3. Demuestre la proposición 4.5.7.
4. Demuestre que el espacio $(C(X, Y), \rho_\infty)$ definido en 4.5.8 es completo.
5. Demuestre el enunciado del ejemplo 34.

CAPÍTULO 5

CONEXION EN ESPACIOS MÉTRICOS

En el análisis clásico uno de los resultados fundamentales en una dimensión, lo constituye el teorema de Bolzano que dice:

Si $F : [a, b] \longrightarrow \mathbb{R}$ es continua y $f(a)f(b) < 0$, entonces existe al menos un punto $c \in [a, b]$ tal que $f(c) = 0$.

Para la demostración de este teorema sobre la existencia de raíces de funciones reales continuas, se necesita una propiedad topológica muy importante de los intervalos. En este capítulo estudiaremos esta propiedad llamada *conexión*.

Algunos espacios métricos pueden *dividirse* de una *manera natural* en dos o mas partes. Por ejemplo, un espacio formado por dos rectas que no se cortan puede ser dividido en las dos rectas.

Otro ejemplo puede ser el complemento de una circunferencia en el plano que ésta constituido por dos partes: la parte interior a la circunferencia y la parte exterior.

El conjunto $\mathbb{Q}$ de los números racionales puede dividirse en dos partes de muchas maneras. Cada número irracional x da lugar a una división de $\mathbb{Q}$ en los racionales mayores y menores que x . El conjunto de los números irracionales puede dividirse por cada racional de una manera semejante. Precisamente, una forma simple de construir un espacio métrico a partir de una familia dada de espacios métricos es “pegando” dichos espacios de manera que el análisis de la estructura topológica del espacio obtenido, se reduzca de manera trivial al análisis topológico de cada uno de los espacios componentes.

Es natural preguntar cuando un espacio métrico es coherente en el sentido de que no admite una tal reducción, es decir, ¿cuando no puede ser expresado como “suma de componentes separadas”? En efecto, al menos intuitivamente podemos ver que hay ciertos conjuntos que no admiten una división en partes de manera natural; esto ocurre, por ejemplo, para una recta, un segmento, un plano, un disco circular, etcétera; tales espacios métricos llamados *conexos* serán estudiados en el primer epígrafe 5.1. Analizaremos en este epígrafe en particular, el fundamento topológico del llamado Teorema del valor medio o también Teorema de Bolzano, del cual hablamos al inicio, y estudiaremos además algunas aplicaciones de su generalización a espacios conexos cualesquiera.

De manera natural se plantea el problema de si un espacio métrico no conexo permite siempre una descomposición en partes conexas, descubriéndose algunas técnicas sencillas para identificar o distinguir diferentes espacios métricos, o sea, decidir si son homeomorfos (topológicamente equivalentes) o no. Con este objetivo en el 5.2 se introduce el concepto *componente conexa*, estudiándose además, aquellos espacios métricos cuya estructura topológica puede reconstruirse a partir de la de sus componentes conexas. Estos son precisamente los espacios localmente conexos.

El concepto *conexo* es la noción más elemental para analizar la coherencia de un espacio métrico desde el punto de vista topológico. Para muchos problemas del análisis se necesita una propiedad más fuerte. Esta noción que será estudiada en el 5.3, es importante también porque es el punto de partida para la llamada Teoría de homotopía, la cual estudia invariantes más sutiles de los espacios métricos asociándoles los llamados grupos de homotopía. Nosotros no nos dedicaremos al estudio de la teoría de homotopía. Al lector interesado en este contenido le recomendamos consultar el libro de E.H.Spanier: Algebraic Topology, Mc.Graw-Hill, 1966.

5.1 ESPACIOS MÉTRICOS CONEXOS. COMPONENTES CONEXAS

Comencemos por precisar el concepto coherencia de un espacio métrico al que nos referíamos de manera intuitiva en nuestra introducción.

Definición 5.1.1. Un espacio métrico (X, d) se llama *conexo* si no existe una partición de X formada por dos abiertos no vacíos. $A \subset (X, d)$ se dice conexo si el subespacio (A, d_A) es conexo.

Ejemplo 1. Ningún espacio métrico discreto de más de un punto es conexo. Un ejemplo muy importante, se da en la siguiente caracterización de los subconjuntos conexos de la recta real.

Teorema 5.1.2 (caracterización de los subconjuntos conexos de $\mathbb{R}$). Un subconjunto C de $\mathbb{R}$ es conexo si y sólo si es un intervalo.

Observación: Cuando se habla de $\mathbb{R}$ sin hacer mención a la métrica subyacente nos referimos a la métrica usual definida por el módulo. El ejemplo 1 nos muestra que si se provee a $\mathbb{R}$ de la métrica discreta, los únicos subconjuntos conexos son los reducidos a un punto.

Demostración: Recordemos que $C \subset \mathbb{R}$ es un intervalo si y sólo si cumple

para todo $x, y, z \in \mathbb{R}$:

$$x, y \in C; \quad x < z < y \Rightarrow z \in C \quad (1)$$

Supongamos que C no es un intervalo; entonces existen $x, y \in C$ y $z \in \mathbb{R} \setminus C$ tales que $x < z < y$. Los subconjuntos abiertos no vacíos $A = \{c \in C : c < z\}$, $B = \{c \in C : c > z\}$ de C forman una partición de C , luego C no es conexo.

Sea inversamente C un intervalo, veamos que es conexo. Para ello supongamos por el absurdo que existen dos subconjuntos abiertos no vacíos A y B del espacio C tales que $A \cup B = C$ y $A \cap B = \emptyset$. Intercambiando A y B ; si es necesario, existen $a \in A$, $b \in B$ tales que $a < b$. Sea $z = \sup \{A \cap [a, b]\}$, $z \in [a, b] \subset C$ es punto adherente de A en C . Puesto que A es cerrado en C (por ser el complemento del abierto B), tenemos que $z \in A$, y por lo tanto, $z < b$. Por otra parte, A es también abierto en C ; por ello, existe $\epsilon > 0$ tal que $[z, z + \epsilon] \subset A$ y $z + \epsilon < b$. Pero esto contradice a la definición $z = \sup \{A \cap [a, b]\}$; por lo que concluimos que C es conexo. ■

En la práctica es conveniente tener algunas caracterizaciones equivalentes de los espacios y subconjuntos conexos. Para ello, precisemos primeramente el concepto intuitivo de separación.

En la figura 77a se muestra una representación de conjuntos separados y en la 77b una representación de conjuntos no separados.

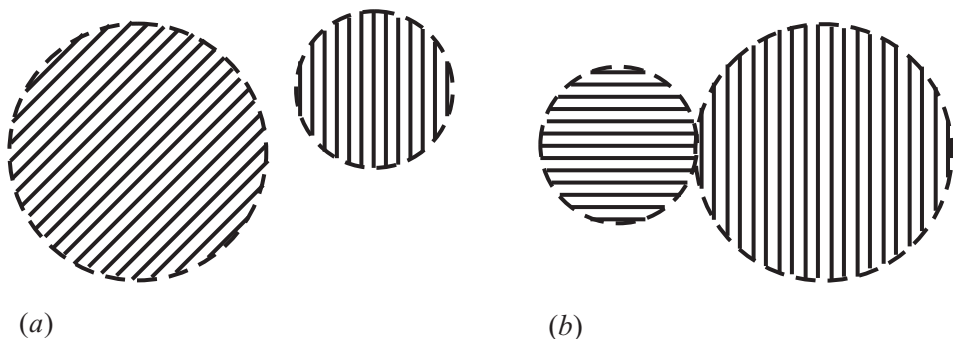


Figura 77

Definición 5.1.3. Sea (X, d) un espacio métrico y $C \subset X$. Se dice que dos subconjuntos $A, B \subset X$ son *separados* en (X, d) si $\overline{A} \cap B = \emptyset = A \cap \overline{B}$.

Se dice que A y B se *tocan* si no son separados.

Un par de subconjuntos separados (A, B) se llama *una separación* de C en X , si $C \subset A \cup B$; $C \cap A \neq \emptyset$ y $C \cap B \neq \emptyset$.

Si (X, d) tiene una partición (A, B) de abiertos no vacíos, el par (A, B) es una separación de X e inversamente.

Proposición 5.1.4 (caracterizaciones de los espacios métricos conexos). Un espacio métrico (X, d) es conexo si y sólo si cumple una de las condiciones equivalentes siguientes:

- i) No existe una separación de X en (X, d) .
- ii) No existe una partición $(X_\alpha)_{\alpha \in I}$ de X formada por más de un abierto no vacío.
- iii) No existe una partición de (X, d) formada por un número finito ≥ 2 de cerrados no vacíos.
- iv) Los únicos subconjuntos que son a la vez abiertos y cerrados en (X, d) son X y $\emptyset$.
- v) No existe una sobreyección continua de (X, d) sobre el espacio discreto con dos elementos.
- vi) Toda aplicación continua de (X, d) en un espacio discreto es constante.

La demostración de esta proposición resulta un ejercicio interesante, por ello se deja al lector.

Los tres últimos criterios son probablemente los más útiles en la práctica. Como se verá en las aplicaciones siguientes, ellos permiten hacer más económicas las demostraciones de diversas proposiciones sobre espacios conexos.

Ejemplo 2. Consideremos el espacio vectorial $M_n(\mathbb{R})$ de todas las matrices reales de tamaño $n \times n$.

Por la identificación que existe entre $M_n(\mathbb{R})$ y el espacio euclideo $\mathbb{R}^{n^2}$, deducimos que $M_n(\mathbb{R})$ es conexo; sin embargo, el subconjunto $O_n(\mathbb{R})$ de todas las matrices ortogonales no es conexo, ya que la aplicación continua

$$\det : M_n(\mathbb{R}) \rightarrow \mathbb{R},$$

que asocia a cada matriz su determinante, aplica $O_n(\mathbb{R})$ sobre el espacio discreto $\{-1, 1\}$.

Ejemplo 3 (convergencia uniforme y diferenciación). Consideremos la sucesión de funciones reales

$$f_n(x) = \frac{\sin nx}{\sqrt{n}}, \quad n \in \mathbb{N}.$$

Es claro que la sucesión (f_n) converge uniformemente sobre $\mathbb{R}$ a la función $f = \langle 0 \rangle$ ($f(x) \equiv 0$); sin embargo, la sucesión de las derivadas

$$f'_n(x) = \sqrt{n} \cos nx$$

no converge uniformemente hacia $f = \langle 0 \rangle$ ($f'(x) \equiv 0$) ya que por ejemplo

$$f'_n(0) = \sqrt{n} \xrightarrow{n \rightarrow \infty} +\infty$$

mientras que $f'(0) = 0$.

Este ejemplo nos muestra que la convergencia uniforme de (f_n) no implica nada sobre la convergencia de (f'_n) . Obviamente, se necesitan hipótesis más rigurosas para asegurar que $(f'_n) \xrightarrow{d_\infty} f'$ si $(f_n) \xrightarrow{d_\infty} f$.

El siguiente resultado puede considerarse como una condición necesaria para poder intercambiar la operación de límite uniforme con la diferenciación, es decir:

$$\frac{d}{dx} \lim = \lim \frac{d}{dx}.$$

Supongamos que (f_n) es una sucesión de funciones reales diferenciables en $[a, b]$ (no necesariamente continuamente diferenciables), tales que la sucesión de números reales $(f_n(x_0))$ converge para algún $x_0 \in [a, b]$ y además (f'_n) converge uniformemente sobre $[a, b]$. Entonces se tiene:

- i) La sucesión (f_n) converge uniformemente sobre $[a, b]$ hacia una cierta función f derivable.
- ii) La sucesión (f'_n) converge uniformemente a f' sobre $[a, b]$.

Veamos cómo se demuestra este resultado aplicando la caracterización iv) expuesta en la proposición 5.1.4.

Demostración: Denotemos por

$$g = \lim_{n \rightarrow \infty} f'_n$$

y definamos

$$A = \{x \in [a, b] : (f_n(x)) \text{ converge}\}.$$

Demostremos que el conjunto A es abierto y cerrado en $[a, b]$. Como $A \neq \emptyset$ puesto que $x_0 \in A$ y por ser $[a, b]$ conexo (5.1.2) de la caracterización iv) en la proposición 5.1.4, se deduce entonces que $A = [a, b]$. En efecto, sea $x \in A$ y supongamos por comodidad que $x \neq a$ y $x \neq b$. Para obtener que A es abierto en $[a, b]$ debemos probar la existencia de un número $\delta > 0$ tal que $]x - \delta, x + \delta[\subset A$. Con este objetivo, utilizaremos otro conocido teorema del valor medio para funciones derivables aplicado a la función $f_m(x) - f_n(x)$. Para el par de puntos $x, y \in \mathbb{R}$, existe $\xi \in [x, y]$ (suponiendo $x < y$) tal que

$$\begin{aligned} (f_m(y) - f_n(y)) - (f_m(x) - f_n(x)) &= \\ &= (f'_m(\xi) - f'_n(\xi))(y - x). \end{aligned} \tag{2}$$

Por la convergencia uniforme de (f'_n) a g , existe n_0 tal que si $n \geq n_0$

$$|f'_m(\epsilon) - f'_n(\epsilon)| < \epsilon/2 \text{ para todo } \xi \in \mathbb{R}. \quad (3)$$

Además, por ser $x \in A$, la sucesión $(f_n(x))$ es convergente en $\mathbb{R}$ y también se tendrá la existencia de un $n_1 \in \mathbb{R}$, tal que si $n \geq n_1$

$$|f_m(x) - f_n(x)| < \epsilon/2. \quad (4)$$

Luego, para $\delta \leq 1$, de (2), (3) y (4) se deduce que la sucesión de números reales $(f_n(y))$ es de Cauchy para cualquier $y \in]x - \delta, x + \delta[$.

Por ser $\mathbb{R}$ completo, esta sucesión es convergente para cada $y \in]x - \delta, x + \delta[$, de donde se concluye que A es abierto. De (2) se obtiene fácilmente que A es cerrado en $[a, b]$ (demuestrelo).

Por lo tanto, concluimos que $A = [a, b]$, es decir, existe el límite puntual $\lim_{n \rightarrow \infty} f_n(x)$ para cada $x \in \mathbb{R}$, al cual denotaremos por $f(x)$:

$$\lim_{n \rightarrow \infty} f_n(x) = f(x).$$

Pero ahora en (2) considerando $y = x + h$ y haciendo tender n a ∞ , después de dividir por h , se tiene que existe $f'(x) = g(x)$ para todo $x \in [a, b]$.

Finalmente, pasando al límite cuando $n \rightarrow \infty$ en (2) y dejando x fijo se obtiene que (f_m) converge a f uniformemente. ■

Observación: Para la demostración de la existencia de f , podríamos haber utilizado la completitud de $(C([a, b], \mathbb{R}), d_\infty)$ pero hemos preferido utilizar argumentos de conexidad en este caso.

Ejemplo 4 (residuo logarítmico de funciones analíticas). Sea $f(z)$ una función analítica en una región $\Omega \subset \mathbb{C}$ (Ω abierto y conexo) sin “huecos” y sea Γ una curva en Ω diferenciable, cerrada y sin autointersecciones. Para cada $w \notin f[\Gamma] = \{f(z) : z \in \Gamma\}$ se define el residuo logarítmico de la función $f(z) - w$ con respecto a la curva Γ , como la integral

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z)}{f(z) - w} dz = R_{\Gamma}(w).$$

No es difícil demostrar que la función $R_{\Gamma}(w)$ es continua en cada punto $w \notin f[\Gamma]$. Por otra parte, en el curso elemental de Variable Compleja, se demuestra que $R_{\Gamma}(w)$ es precisamente igual al número de veces que la función f alcanza el valor w en el interior de Γ . Luego, R_{Γ} puede considerarse como una función continua definida sobre $\mathbb{C} \setminus f[\Gamma]$ y con valores en $\mathbb{Z}$.

Por lo tanto, aplicando la caracterización iv) de la proposición 5.1.4, tendremos que R_{Γ} es constante sobre cualquier subconjunto conexo de $\mathbb{C} \setminus f[\Gamma]$, es

decir, todos los valores w pertenecientes a un subconjunto conexo de $\mathbb{C} \setminus f[\Gamma]$ son alcanzados por f el mismo número de veces en el interior de Γ .

Caractericemos ahora los subconjuntos conexos de un espacio métrico (X, d) por propiedades relativas a todo el espacio. Según la definición 5.1.1, un subconjunto $C \subset X$ es conexo si y sólo si no existen dos abiertos (cerrados) $A, B \subset X$ tales que $A \cap C \neq \emptyset$

$$B \cap C \neq \emptyset, C \subset A \cup B \text{ y } A \cap B \cap C = \emptyset. \quad (5)$$

Proposición 5.1.5. Un subconjunto C de (X, d) es conexo si y sólo si no existe una separación de C en (X, d) .

Demostración: Si (A, B) es una separación de C en (X, d) , entonces $(A \cap C, B \cap C)$ forma una separación de C en (C, d_C) . Luego, según 5.1.4i), C no puede ser conexo si posee una separación en (X, d) .

Supongamos inversamente que C no es conexo, entonces existe una separación (A, B) de C en (C, d_C) . Puesto que A y B son cerrados relativos en C , se tiene que $\overline{A} \cap B = \overline{A} \cap (C \cap B) = \overline{A}_C \cap B = A \cap B = \emptyset$, de igual forma se tiene que $\overline{B} \cap A = \emptyset$; luego (A, B) es una separación de C en (X, d) . ■

Corolario 1. Sea (X, d) un espacio métrico $C \subset X$ y (A, B) una separación de un superconjunto $D \supset C$ en (X, d) . Si C es conexo entonces $C \subset A$ o $C \subset B$.

Demostración: Si $C \cap A \neq \emptyset$ y $C \cap B \neq \emptyset$, entonces (A, B) formaría una separación de C . ■

Después de estas caracterizaciones de los espacios y subespacios conexos probaremos una serie de teoremas por medio del criterio v) en 5.1.4.

Proposición 5.1.6. Sea A un subconjunto de un espacio métrico (X, d) ; entonces todo conjunto $B \subset X$, tal que $A \subset B \subset \overline{A}$, es conexo. En particular, la clausura $\overline{A}$ es conexo.

Demostración: Sea f una aplicación de B en el espacio discreto $\{1, 2\}$. Sabemos que $f[A]$ es unitario, luego $f[B] \subset f[\overline{A}] \subset f[\overline{A}]$ es unitario, por lo que concluimos que f no es sobreyectiva. ■

Proposición 5.1.7. Sea $(C_\alpha)_{\alpha \in I}$ una familia de subconjuntos conexos de (X, d) que tocan todos un mismo $C_\beta (\beta \in I)$. Entonces $C = \bigcup_{\alpha \in I} C_\alpha$ es conexo.

Demostración: Sea (A, B) una separación de C . Según el corolario 1 del

teorema 5.1.5 tenemos $C_\alpha \subset A$ o $C_\alpha \subset B$ para todo $\alpha \in I$. Supongamos en particular, que el conjunto C_β que se toca con todos los demás conjuntos C_α ($\alpha \in I$) está incluido en A . Puesto que $B \cap C \neq \emptyset$, existe $\gamma \in I$ tal que $C_\gamma \subset B$. Pero A y B no se tocan, luego C_β y C_γ no pueden tocarse. ¡Contradicción! No existe una separación de C . ■

Corolario 1. Sea $(C_\alpha)_{\alpha \in I}$ una familia de subconjuntos conexos de (X, d) tal que $\bigcap_{\alpha \in I} C_\alpha \neq \emptyset$. Entonces $\bigcup_{\alpha \in I} C_\alpha$ es conexo.

Corolario 2. Sea (C_n) una sucesión de conjuntos conexos en (X, d) tales que C_n toca a C_{n+1} para todo $n \in \mathbb{N}$. Entonces $\bigcup_n C_n$ es conexo.

Demostración: Por inducción se demuestra que los conjuntos $C'_n = \bigcup_{i \leq n} C_i$ son conexos. Después sólo hay que aplicar la proposición 5.1.7 a la familia (C'_n) que tiene la misma unión que los (C_n) . ■

La intersección de dos conjuntos conexos en general no es conexo (fig. 78). Incluso, la intersección de una sucesión decreciente de conjuntos conexos no es necesariamente conexo: Un contraejemplo está dado por el espacio $X = [-1, 1]^2 \setminus \{(0, 0)\}$ y los subconjuntos conexos

$$A_n = \{(x, y) \in X : |y| \leq 1/n\} \quad (n \in \mathbb{N}^*)$$

cuya intersección es el segmento punteado $\{(x, 0) : 0 < |x| \leq 1\}$ (fig. 79).

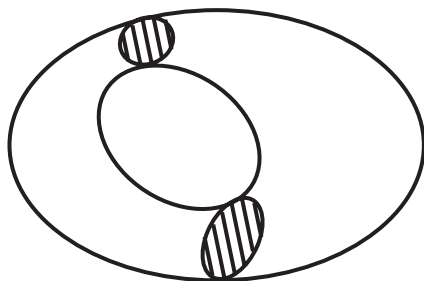


Figura 78

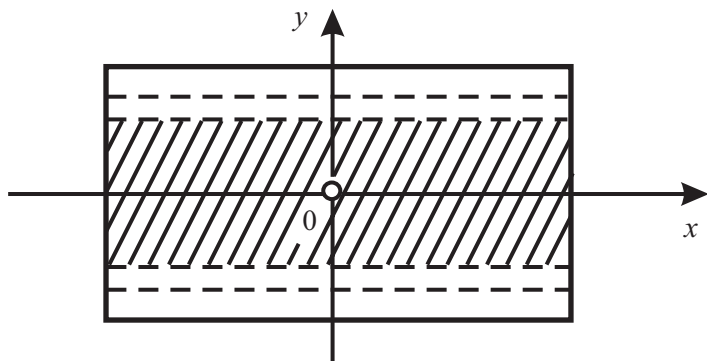


Figura 79

Sin embargo, la conexión se mantiene bajo aplicaciones continuas.

Teorema 5.1.8. Sea f una aplicación continua de (X, d) en (Y, d') . Si (X, d) es conexo, entonces $f[X]$ es conexo en (Y, d') .

Demostración: Sea g una aplicación continua del subespacio $f[X]$ en el espacio discreto $\{1, 2\}$. Entonces $g \circ f$ es una aplicación continua de X en $\{1, 2\}$, luego g no puede ser sobreyectiva. ■

A pesar de su sencillez, el teorema anterior tiene una importancia fundamental como muestran los corolarios siguientes:

Corolario 1. La conexión es una propiedad topológica.

La demostración es evidente.

Corolario 2 (Teorema del valor medio generalizado). Sea (X, d) un espacio métrico conexo y f una función real continua sobre (X, d) . Si $x, y \in X$ y $\alpha \in \mathbb{R}$ tal que $f(x) \leq \alpha \leq f(y)$, entonces existe $z \in X$ con $f(z) = \alpha$.

Demostración: Por 5.1.8 $f[X]$ es un subconjunto conexo de $\mathbb{R}$ y por 5.1.2 $f[X]$ es un intervalo, de donde se deduce la tesis del corolario. ■

Ejemplo 5. Sean (X, d) , (Y, d') espacios métricos y sea f una aplicación continua de un subconjunto conexo A de (X, d) en (Y, d') . Entonces el grafo

$$G_f = \{(x, f(x)) : x \in A\}$$

es conexo en $X \times Y$, ya que la aplicación $F : x \rightsquigarrow (x, f(x))$ de A sobre $G_f \subset X \times Y$ es continua, y por lo tanto, podemos aplicar 5.1.8.

Ejemplo 6. El grafo $G \subset \mathbb{R}^2$ de la función $x \rightsquigarrow \sin \frac{1}{x}$ sobre $]0, \infty[$ es conexo

según el ejemplo anterior. Es fácil comprobar que $\overline{G} = G \cup \{(0, y) \in \mathbb{R}^2 : -1 \leq y \leq 1\}$ en $\mathbb{R}^2$. De la proposición 5.1.6 concluimos que cualquier subconjunto S de $\mathbb{R}^2$ tal que $G \subset S \subset G \cup \{(0, y) \in \mathbb{R}^2 : -1 \leq y \leq 1\}$ es también conexo (fig. 80). El grafo G_* de la misma función $x \rightsquigarrow \sin \frac{1}{x}$ sobre $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ ya no es conexo, pues $H_+ := \{(x, y) \in \mathbb{R}^2 : x > 0\}$ y $H_- := \{(x, y) \in \mathbb{R}^2 : x < 0\}$ forman una separación de G_* en $\mathbb{R}^2$. Esto nos muestra que la conexidad de (X, d) es, en general, necesaria para obtener la conexidad de $f[X]$ en el teorema 5.1.8. Sin embargo, si se añade un solo punto de $\{(0, y) \in \mathbb{R}^2, -1 \leq y \leq 1\} = \{0\} \times [-1, 1]$ a G_* , según 5.1.7 obtenemos un subconjunto conexo.

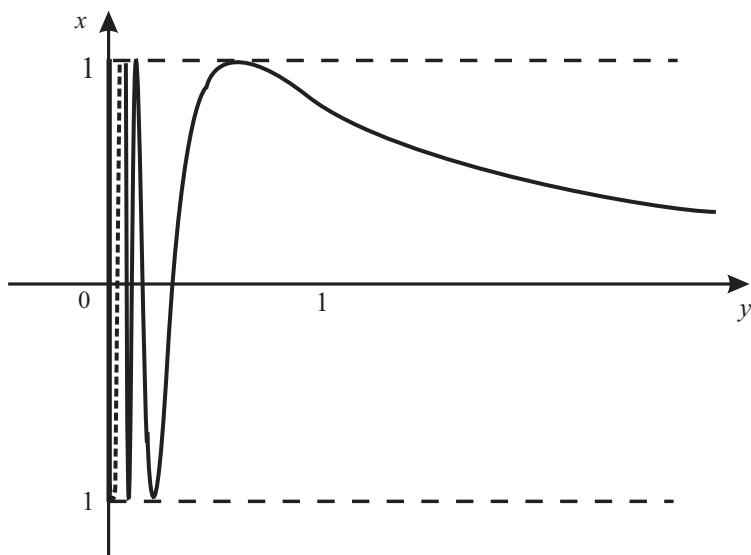


Figura 80

Los ejemplos siguientes nos muestran las múltiples aplicaciones del teorema generalizado del valor medio.

Ejemplo 7. Supongamos que buscamos soluciones de una ecuación del tipo

$$f(x) = \alpha \quad (x \in X), \quad (6)$$

donde f es una aplicación real continua sobre un espacio conexo (X, d) . Para probar que existe una solución es suficiente, según el corolario 2 del teorema 5.1.8, hallar dos puntos $x_1, x_2 \in X$ tales que $f(x_1) \leq \alpha \leq f(x_2)$. Incluso, en este caso puede decirse más: todo subconjunto conexo A de (X, d) que contiene x_1 y x_2 debe contener una solución de (6). Por ejemplo, si (X, d) es un espacio

euclideo, entonces se halla una solución de (6) sobre el segmento $[x_1, x_2]$ que conecta x_1 con x_2 .

Ejemplo 8 (distinción entre espacios métricos). Una técnica usual para examinar una homeomorfía entre dos espacios métricos (X, d) , (Y, d') se basa sobre la observación siguiente:

Si $h : X \rightarrow Y$ es un homeomorfismo entonces para un subconjunto cualquiera A de X los subespacios $X \setminus A$ y $Y \setminus h[A]$ deben ser homeomorfos y luego tener los mismos invariantes topológicos.

A continuación daremos dos aplicaciones sencillas de esta técnica:

Aplicación 1. En $\mathbb{R}$ ninguno de los tres intervalos $[0, 1]$, $[0, 1[$, $]0, 1[$ es homeomorfo al otro, pues del primero podemos sustraer a lo más dos puntos (0 y 1), del segundo uno (0) y del tercero ningún punto sin violar la conexión de estos subespacios.

Aplicación 2. La recta real no es homeomorfa a ningún $\mathbb{R}^n$ tal que $n > 1$; pues $\mathbb{R}^n$ ($n \geq 2$) sigue siendo conexo si le quitamos un punto, lo cual no ocurre para $\mathbb{R}$.

El teorema general donde se demuestra que $\mathbb{R}^n$ no es homeomorfa $\mathbb{R}^m$ para $m \neq n$ es mucho más profundo. Para su demostración se utilizan invariantes topológicos más sofisticados relacionados con las técnicas de la teoría de homología.

Ejemplo 9 (aplicación a un problema geométrico). Probaremos que sobre cada curva cerrada y rectificable en $\mathbb{R}^3$ existen cuatro puntos que están en un mismo plano y que la dividen en partes de igual longitud. Sea $x : s \rightsquigarrow x(s)$ una representación paramétrica de la curva que toma la longitud de la curva como parámetro: $0 \leq s \leq l$ (donde l es la longitud total de la curva) y $x(0) = x(l)$. Para $s \in [0, l/4]$ sea $V(s)$ el volumen orientado del tetraedro con los vértices en $x(s), x(s+l/4), x(s+l/2), x(s+3/4l)$. $V : [0, l/4] \rightarrow \mathbb{R}$ es continua. El volumen orientado del tetraedro $[x(0), x(l/4), x(l/2), x(3/4l)]$ es, con signo opuesto, igual al volumen del tetraedro $[x(l/4), x(l/2), x(3/4l), x(l)]$ puesto que $x(l) = x(0)$. Concluimos que $-V(0) = V(l/4)$; luego existe $s_0 \in [0, l/4]$ tal que $V(s_0) = 0$. Los puntos $x(s_0), x(s_0+l/4), x(s_0+l/2), x(s_0+3/4l)$ están entonces en el mismo plano a la vez que dividen la curva en partes de igual longitud.

Veamos otros ejemplos que nos muestran la potencialidad del teorema del valor medio para el estudio de las funciones reales continuas.

Ejemplo 10 (teorema del punto fijo). El siguiente teorema del punto fijo

no se basa en la noción de completitud como el expuesto en el capítulo 3, sino en el concepto de conexión:

Toda aplicación continua de un segmento real $[a, b]$ en sí mismo tiene al menos un punto fijo

En efecto, la función auxiliar

$$g : [a, b] \longrightarrow \mathbb{R},$$

definida por

$$g(x) = f(x) - x$$

cumple

$$g(a) \geq 0 \text{ y } g(b) \leq 0$$

luego, según el corolario 2 del teorema 5.1.8. existe $x_0 \in [a, b]$ con $g(x_0) = 0$, es decir $f(x_0) = x_0$.

Ejemplo 11 (aplicaciones de una circunferencia en una recta). Una circunferencia tiene la propiedad notable siguiente:

Cada aplicación real continua definida sobre una circunferencia de $\mathbb{R}$, aplica algún par de puntos diametralmente opuestos en el mismo punto imagen.

En efecto, sin pérdida de generalidad, podemos suponer que la circunferencia C está centrada en el punto $(0, 0)$ de $\mathbb{R}^2$, en cuyo caso el punto diametralmente opuesto a $z \in C$ es $-z \in C$.

Sea $f : C \longrightarrow \mathbb{R}$, la aplicación dada y construyamos la función auxiliar $g : C \longrightarrow \mathbb{R}$ definida por

$$g(z) = f(z) - f(-z).$$

Obviamente g es continua por serlo f y además, $g(z) = -g(-z)$ para cada $z \in C$, luego, $g(z)$ y $g(-z)$ tienen signos opuestos para todo $z \in C$, de donde se deduce que existe $z_0 \in C$ tal que

$$g(z_0) = 0$$

y por lo tanto,

$$f(z_0) = f(-z_0).$$

No es difícil verificar un resultado similar si se sustituye la circunferencia C por una elipse.

El resultado expuesto en el ejemplo 11 se utiliza para resolver algunos problemas geométricos interesantes como el siguiente:

Ejemplo 12 (problemas de la tarta). Los dos problemas de la tarta a que nos referiremos pueden plantearse aproximadamente como sigue:

- a) Supongamos que tenemos dos trozos de tarta de forma irregular sobre la misma bandeja; demuéstrese que es posible dividir a ambos en dos partes exactamente iguales con un solo corte de cuchillo.

b) Supongamos que sobre la bandeja se encuentra un solo trozo de tarta; demuéstrese que es posible cortar este trozo en cuatro partes de igual área mediante dos cortes perpendiculares.

Los resultados matemáticos precisos correspondiente a estos enunciados son los siguientes:

- a')** Si A y B son dos regiones (abiertos y conexos) acotados de un plano, entonces existe una recta del plano que divide simultáneamente a cada región en dos de la misma área. (Este resultado se cumple aún cuando las regiones A y B se intersectan).
- b')** Si A es una región acotada del plano, existen dos rectas perpendiculares que dividen a A en cuatro partes de igual área.

Veamos una idea de la demostración de la proposición a').

Puesto que A y B son acotados podemos elegir un disco C con centro en el origen que incluya a $A \cup B$ (fig. 81).

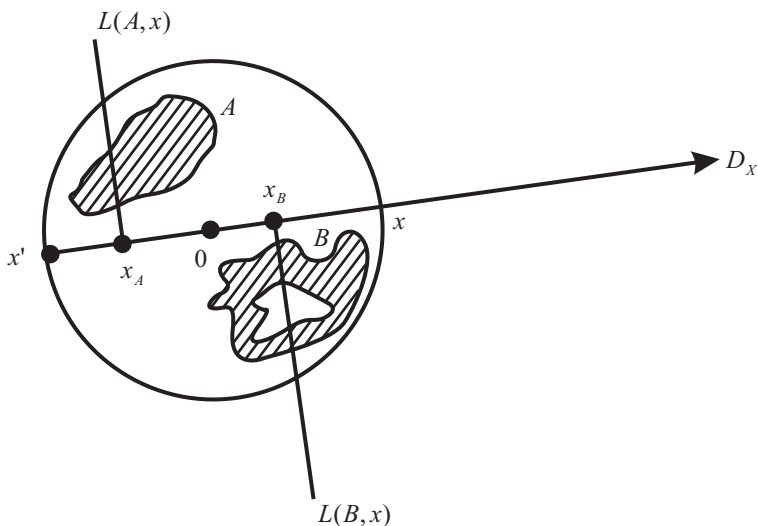


Figura 81

Para cualquier $x \in C$, sea $x' = -x$ su punto diametralmente opuesto en C y D_x el diámetro de extremos x y x' . No es difícil comprobar (verifíquelo) que en la familia de todas las rectas perpendiculares a D_x existe un único par de rectas $L(A, x)$ y $L(B, x)$ que dividen a A y B respectivamente en dos partes

de igual área. Designemos por x_A y x_B los puntos donde D_x corta a $L(A, x)$ y $L(B, x)$ respectivamente. Los puntos x_A y x_B están definidos de manera única para cada x y además, no resulta difícil verificar que las aplicaciones g_A y g_B de C en $\mathbb{R}$ definidas por $g_A(x) = x_A$ y $g_B(x) = x_B$ son continuas, si identificamos con $x_A(x_B)$ la distancia dirigida del propio punto $x_A(x_B)$ al centro del círculo, tomándola positiva cuando el punto $x_A(x_B)$ esté al mismo lado de x y negativa cuando esté en el lado opuesto.

Luego, la función $h : C \rightarrow \mathbb{R}$,

$$h(x) = g_A(x) - g_B(x)$$

es también una función continua. No es difícil ver que una propiedad esencial de h es que sus valores en dos puntos diametralmente opuestos de C son iguales en valor absoluto, pero de signo opuesto:

$$h(x) = -h(x') = -h(-x) \text{ para todo } x \in C.$$

Esto se demuestra observando que $D_x = D_{x'}$, así como $L(A, x) = L(A, x')$ y $L(B, x) = L(B, x')$, puesto que $x'_A = x_A$ y $x'_B = x_B$. Sin embargo, para las coordenadas en $D_{x'}$ la dirección positiva es opuesta a la de D_x , de aquí que $g_A(x') = -g_A(x)$ y $g_B(x') = -g_B(x)$, y por tanto,

$$h(x') = g_A(x') - g_B(x') = -g_A(x) + g_B(x) = -h(x).$$

Aplicando el resultado del ejemplo 21 existe un punto $x \in C$ tal que $h(x) = h(x')$. Para este punto x se verifica que

$$h(x) = h(x'),$$

$$h(x) = -h(x')$$

y por lo tanto, $h(x) = 0$, lo cual implica que $x_A = x_B$ y la recta $L(A, x) = L(B, x)$ divide a A y a B en partes de igual área respectivamente.

La demostración de b') se deja de ejercicio al lector.

Ejemplo 13 (aplicaciones de un plano en sí mismo). Nuestro propósito con este ejemplo es indicar un teorema de existencia de solución de pares de ecuaciones. No es nuestro objetivo su demostración, la cual requiere la introducción de conceptos más elaborados aunque están estrechamente relacionados con la noción de conexión. El concepto fundamental que se usa es el referente al número de vueltas de una curva cerrada plana alrededor de algún punto del plano que no pertenezca a la curva. Daremos la siguiente definición intuitiva de número de vueltas. Supongamos que $\varphi : [a, b] \rightarrow \mathbb{R}^2$ es una curva cerrada plana y $z \in \mathbb{R}^2$ un punto que no pertenece a la curva. Para cada $t \in [a, b]$, designemos por L_t la semirrecta de origen z , y que pasa por $\varphi(t)$. Cuando t varía de a hasta b , el punto $\varphi(t)$ recorre la curva y la semirrecta sobre el punto fijo z . Puesto que la curva es cerrada, L_t vuelve a su posición inicial, es decir $L_b = L_a$. Por tanto, durante su movimiento, la semirrecta da un número en-

tero de vueltas sobre z . Este número se llama el *número de vueltas* de la curva cerrada alrededor del punto z . Por conveniencia, se asigna el signo positivo a las rotaciones en sentido contrario a las agujas del reloj, y el signo negativo a las rotaciones en el mismo sentido que las agujas del reloj.

Como puede notar el lector, esta definición no es muy precisa y solamente podemos considerarla por su carácter intuitivo. Para una definición más rigurosa habría que comenzar por demostrar, por ejemplo, que el número de vueltas de una curva alrededor de un punto fijo z no depende de la parametrización φ tomada para representar dicha curva. No obstante, haciendo uso de este concepto de número de vueltas aún con sus limitaciones (no es nuestro objetivo profundizar en este aspecto), podemos formular el resultado principal de este ejemplo en la forma siguiente:

Sea $f : D \longrightarrow \mathbb{R}^2$ una aplicación continua de un disco en el plano y sea el C el círculo frontera de D . Si el punto $z \in \mathbb{R}^2$ no pertenece a $f[C]$ y el número de vueltas de la curva $f[C]$ es distinto de cero, entonces $z \in f[D]$, es decir, existe un punto $x \in D$ tal que $f(x) = z$.

No realizaremos una demostración de este resultado, sino que lo utilizaremos para deducir las aplicaciones que veremos en los ejemplos siguientes:

Ejemplo 14 (sobre el teorema del punto fijo). En el ejemplo 10 hemos probado que una aplicación continua de un segmento de recta en sí mismo tiene al menos un punto fijo, y para ello, hemos utilizado en cierta medida la propiedad de conexión de los intervalos.

Utilizando la propiedad de conexión de un disco $D \subset \mathbb{R}^2$, reflejada mediante el resultado enunciado en el ejemplo 13, se puede demostrar una proposición análoga para el disco. Antes de pasar a enunciar nuestro resultado fundamental, veremos intuitivamente la formulación de algunos resultados auxiliares. Consideremos el efecto de fijar un disco delgado de goma flexible por su contorno al tablero de una mesa. Si queremos ver lo que está bajo el disco, no conseguiremos nada estirando o retorciendo la goma, en tanto su contorno permanezca pegado. Esta idea podemos formularla formalmente de la manera siguiente:

- a) Sea $D \subset \mathbb{R}^2$ un disco y $f : D \longrightarrow \mathbb{R}^2$ una aplicación continua que deja fijo cada punto del círculo frontera C de D ; entonces la imagen $f[D]$ contiene a todo D .

En efecto, si f deja fijo cada punto de C , entonces la curva $f[C]$ es el propio círculo C recorrido una vez. Luego, el número de vueltas de la curva $f[C]$ alrededor de cualquier punto interior a D es igual a 1, y aplicando el resultado principal del ejemplo 13, cada punto de D pertenece a $f[D]$. Como una

consecuencia inmediata de este resultado tenemos el siguiente:

- b) No existe ninguna aplicación continua f de un disco D en su contorno C que deje fijos todos los puntos de C .

Por supuesto, puede aplicarse un rectángulo y su interior sobre uno de sus lados de modo que quede fijo. Justamente esto es lo que se hace al enrollar una persiana; mediante esta aplicación f , se verifica que $f(x) = x$ para todo x que pertenece al palo en que se enrolla. El lado que representa dicho palo se denomina la *retracción* de la región rectangular (persiana). El resultado b) nos afirma que no se puede enrollar de manera continua (es decir, sin cortar ni pegar) el interior de un círculo sobre su contorno; o sea, la circunferencia no es la retracción del círculo.

Con ayuda de la proposición b) podemos enunciar el siguiente *resultado fundamental* (Teorema del punto fijo en el disco):

Toda aplicación continua $f : D \rightarrow D$ deja fijo al menos un punto de D , es decir, $f(x) = x$ para algún punto $x \in D$.

En efecto, si suponemos que se cumple lo contrario, entonces para cada $x \in D$, los puntos $f(x)$ y x son diferentes. De aquí concluimos que para cada $x \in D$ se puede contruir una semirrecta L_x con origen en $f(x)$ y que pasa por x . Sea $g(x)$ el punto de C en el cual L_x corta a C . En el caso de que $f(x)$ pertenezca C , $g(x)$ es el otro punto en que L_x corta a C ; y si $x \in C$, $g(x) = x$. La figura 82 muestra varias de las posibilidades.

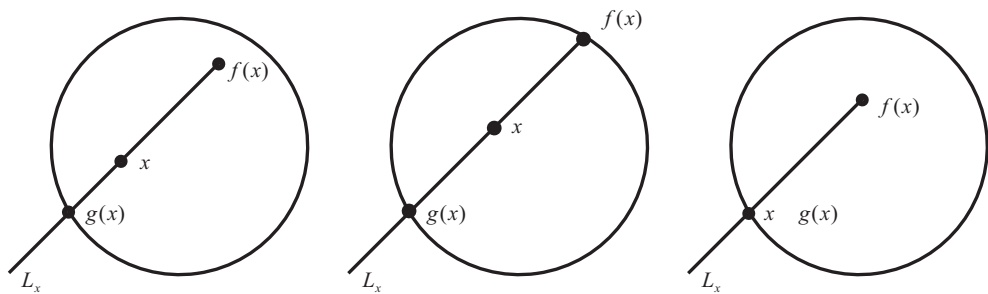


Figura 82

Se deja al lector demostrar que la función $g : D \rightarrow C$ así definida es continua.

Como d deja fijos todos los puntos de C su existencia está en contradicción con el resultado b) del ejemplo 14. Precisamente g pudimos definirla suponién-

do que los puntos x y $f(x)$ no coinciden para ningún $x \in C$. Luego, nuestra suposición fue errónea.

Ejemplo 15 (existencia de puntos antipodales en la tierra con igual presión y temperatura). El lector interesado en comprobar su capacidad de creatividad puede demostrar el siguiente resultado que se deduce del teorema del punto fijo expuesto en el ejemplo 14.

Cada aplicación continua $f : S \longrightarrow \mathbb{R}^2$ de una esfera $S \subset \mathbb{R}^2$ en el plano, aplica algún par de puntos antipodales de S en el mismo punto; es decir, para al menos un par x, x' de antípodas $f(x) = f(x')$.

Como aplicación, consideremos que la superficie de la tierra S es esférica, y que en cualquier instante la presión del aire $p(x)$ y la temperatura $t(x)$ son funciones continuas de $x \in S$.

Para cada $x \in S$ definamos mediante $f(x)$ el punto de $\mathbb{R}^2$ cuyas coordenadas son $(p(x), t(x))$. Puesto que p y t son funciones continuas, se deduce que $f : S \longrightarrow \mathbb{R}^2$ es continua. Aplicando el resultado anterior a esta función f , obtenemos el resultado siguiente:

En cada instante de tiempo existe un par de puntos antipodales en la superficie de la tierra donde son iguales la temperatura y la presión.

Como vemos, las propiedades físicas de la presión y temperatura no influyen en la conclusión; p y t pueden ser dos funciones reales continuas cualesquiera definidas sobre S .

Obsérvese también que si consideramos una sola función, por ejemplo la temperatura, el resultado expuesto en el ejemplo 11 nos dice que en cada circunferencia de perímetro máximo sobre la superficie terrestre existe un par de antípodas donde las temperaturas son iguales.

Para terminar este epígrafe hablemos algo de la hereditabilidad de la noción de conexidad. Está claro que la propiedad de ser conexo no se hereda a los subespacios. Aclaremos ahora qué ocurre con los espacios producto.

Proposición 5.1.19 (producto de conexos). Sea $\{(X_n, d_n)\}_{n \in \mathbb{N}}$ una familia a lo más numerable de espacios métricos no vacíos. Entonces, si proveemos a $\prod_{n \in \mathbb{N}} X_n$ de la métrica producto se tiene:

$\prod_{n \in \mathbb{N}} X_n$ es conexo si y solo si (X_n, d_n) es conexo para todo $n \in \mathbb{N}$.

Observación: El caso en que haya solo un número finito de X_n no vacíos es un caso particular de la proposición anterior.

Demostración: La implicación $\Rightarrow$ resulta del teorema 5.1.8, puesto que los

X_n son las imágenes continuas de $\prod_{n \in \mathbb{N}} X_n$ por las proyecciones π_n ($n \in \mathbb{N}$).

$\Leftarrow$: Sean inversamente todos los X_n ($n \in \mathbb{N}$) espacios conexos. Demostremos primeramente que si el número de espacios es finito su producto es conexo. Para ello solo tenemos que probar que si (X, d) y (Y, d') son conexos, $X \times Y$ es conexo. Pero esto se deduce evidentemente de 5.1.7, puesto que

$$X \times Y = \left[\bigcup_{x \in X} \{x\} \times Y \right] \cup [X \times \{y_0\}]$$

para un elemento cualquiera $y_0 \in Y$. Todos los subespacios $\{x\} \times Y$ y $X \times \{y_0\}$ son conexos por ser homeomorfos a Y y X respectivamente que son conexos. Además, por hipótesis, todos estos subespacios tienen una intersección no vacía con $X \times \{y_0\}$.

Consideremos ahora el caso general.

Sea $x = \{x_n\} \in \prod_{n \in \mathbb{N}} X_n$ un elemento fijo y C la unión de todos los subconjuntos conexos de $\prod_{n \in \mathbb{N}} X_n$ que contienen a x . C es conexo según el corolario 1 de la proposición 5.1.7 y contiene a todos los subconjuntos conexos de la forma

$$\prod_{j=1}^r X_{n_j} \times \prod_{m \neq n_1, \dots, n_r} \{x_m\}$$

que son isomorfos a $\prod_{j=1}^r X_{n_j}$

luego, C es denso en $\prod_{n \in \mathbb{N}} X_n$. De aquí concluimos según 5.1.6 que $\overline{C} = \prod_{n \in \mathbb{N}} X_n$ es conexo. ■

Ejercicios de Autocontrol

1. Pruebe el enunciado siguiente. Si $C \subset (X, d)$ es conexo y si C contiene puntos interiores y puntos exteriores de un conjunto $A \subset X$, entonces también contiene puntos de la frontera de A .
2. Pruebe que en un espacio métrico conexo (X, d) , cada subconjunto propio no vacío tiene una frontera no vacía.
3. Demuestre el enunciado b') del ejemplo 12.
4. Demuestre que la función g definida en el ejemplo 14, es continua.
5. Demuestre el resultado enunciado en el ejemplo 15.
6. Si $C \subset (X, d)$ es conexo ¿Es necesariamente conexo el interior $\overset{\circ}{C}$?

7. Sea $x : t \rightsquigarrow x(t)$ ($0 \leq t \leq 1$) una curva cerrada. Demuestre que existe un rectángulo circunscribiendo la curva x (Indicación: considere $a(\varphi) - b(\varphi)$ en la fig. 83).
8. Sea A un subconjunto conexo en un espacio métrico conexo (X, d) . Sea B abierto y cerrado en $X \setminus A$. Pruebe que $A \cup B$ es conexo.

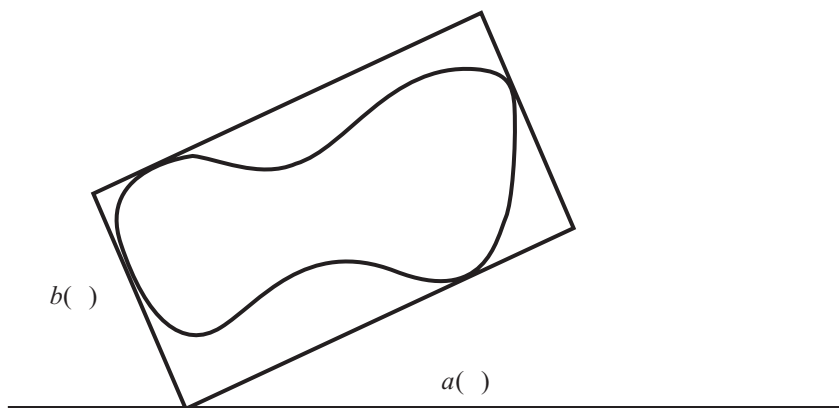


Figura 83

9. Sea C un subconjunto del espacio métrico (X, d) . Pruebe que C no es conexo si y solo si existen conjuntos abiertos $A, B \subset X$ tales que

- i) $A \cap B = \emptyset$
- ii) $C \subset A \cup B$
- iii) $A \cap C \neq \emptyset$ y $B \cap C \neq \emptyset$.

10. Sea $A \subset \mathbb{R}^n$ numerable. Pruebe que $\mathbb{R}^n \setminus A$ es conexo si $n \geq 2$

5.2 COMPONENTES CONEXAS. ESPACIOS LOCALMENTE CONEXOS

Si un espacio métrico no es conexo, se plantea el problema de si puede descomponerse en sus partes conexas de manera tal que el estudio de su estructura topológica se reduzca al estudio de los abiertos en cada uno de estos subespacios. Para investigar esta cuestión, parece natural considerar los subespacios conexos maximales $(C_\alpha)_{\alpha \in I}$ de (X, d) y nos interesa analizar en qué caso U es abierto en (X, d) si y solo si lo es en cada C_α . Si C y D son subconjuntos conexos maximales en (X, d) y $C \cap D \neq \emptyset$, entonces $C = D$, porque $C \cup D$

es conexo y contendría estrictamente a C y a D si fuera $C \neq D$, con lo cual aparecería una contradicción por la maximalidad supuesta.

Esto nos muestra que diferentes conexos maximales en (X, d) son disjuntos.

Según el corolario 1 de la proposición 5.1.7 para todo $x \in X$, la unión C_x de todos los subconjuntos conexos $C \subset X$ que contienen a x es conexa, luego C_x es el mayor conexo que contiene a x . Concluimos que para todo $x \in X$ existe un subespacio conexo maximal que le contiene: los conexos maximales en (X, d) forman una partición de X .

Definición 5.2.1. Sea (X, d) un espacio métrico. El mayor subconjunto conexo en (X, d) que contiene a x se llama *componente conexas* de x en (X, d) y se denota por C_x .

Puesto que las componentes forman una partición de X , pueden interpretarse como clases de una relación de equivalencia, a saber de la relación:

$$xRy \iff \text{existe un subconjunto conexo } C \subset (X, d) \text{ que "conecta" (contiene)} \quad (7)$$

Evidentemente se tiene:

$$xRy \iff C_x = C_y \iff y \in C_x \iff x \in C_y. \quad (8)$$

Tenemos ahora dos nuevos criterios sencillos de conexión:

(X, d) es conexo $\iff (X, d)$ tiene una única componente conexas $\iff$ Para todos los puntos $x, y \in X$ existe un conexo C que conecta x con y .	(9)
--	-----

Proposición 5.2.2. Sea (X, d) un espacio métrico; entonces:

- i) Toda componente conexas en (X, d) es cerrada.
- ii) Cada dos componentes conexas diferentes son separadas.
- iii) Si $U \subset X$ es abierto y cerrado, y contiene a x , entonces $C_x \subset U$.

Demostración:

- i) Si C_x es un conexo maximal, entonces $\overline{C_x}$ es conexo y contiene a C_x , luego $\overline{C_x} = C_x$.
- ii) Dos cerrados disjuntos son siempre separados.

iii) $U \cap C_x \neq \emptyset$ es abierto y cerrado en C_x , luego igual a C_x según 5.1.4 (iv). ■

Si $f : (X, d) \longrightarrow (Y, d')$ es una aplicación continua, entonces aplica toda componente conexa C de (X, d) en un subconjunto conexo de (Y, d') porque $f[C]$ es conexo. En particular, todo homeomorfismo $h : (X, d) \longrightarrow (Y, d')$ induce una biyección entre las componentes conexas de (X, d) y de (Y, d') . Luego, el cardinal de las componentes conexas de (X, d) es un invariante topológico de (X, d) . Si aplicamos esta observación a ciertos subespacios $X \setminus A$ de (X, d) según el método explicado en el ejemplo 8, obtenemos criterios eficientes para distinguir diferentes espacios topológicos. En el caso más sencillo se considera el complemento de un conjunto unitario en (X, d) .

Definición 5.2.3. Sea (X, d) un espacio conexo, entonces $x \in X$ se llama *punto de intersección* de orden $k \in \mathbb{N}$ si $X \setminus \{x\}$ tiene k componentes conexas.

Puesto que todo homeomorfismo $h : (X, d) \rightarrow (Y, d')$ aplica un punto de intersección de orden k en (X, d) a un punto de intersección de orden k en (Y, d') , el cardinal de los puntos de intersección de orden k ($k \in \mathbb{N}$) es un invariante topológico de (X, d) .

¿Cómo puede reformularse el razonamiento del ejemplo 8, mediante el nuevo concepto?

Ejemplo 16. Los espacios “uno-dimensionales” de la figura 84 tienen diferentes números de puntos de intersección de orden 1, 2 y 3.

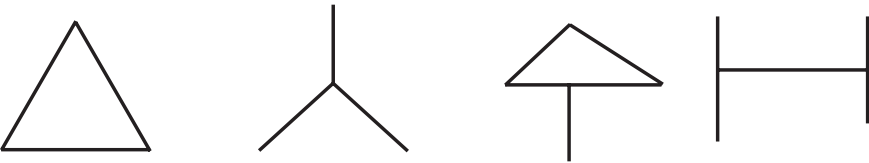


Figura 84

Luego, no son homeomorfos. Este método de distinguir entre distintos espacios topológicos ya lo hemos utilizado intuitivamente en el capítulo 2.

Ejemplo 17. El método de considerar el número de las componentes conexas de $X \setminus A$ para conjuntos unitarios $A = \{x\}$ no permite distinguir entre los espacios de la (fig. 85)

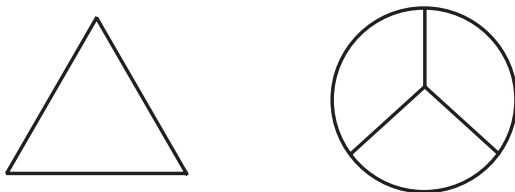


Figura 85

En este ejemplo es suficiente considerar subconjuntos A de dos puntos. Quitando dos puntos del primer espacio obtenemos siempre dos componentes conexas, mientras no ocurre lo mismo para el segundo.

Pero aún no hemos resuelto el problema que planteamos inicialmente. Hasta ahora hemos obtenido una partición de todo espacio métrico formada por cerrados conexos y mutuamente separados. Sin embargo, esto no permite en general reconstruir los abiertos de (X, d) a partir de los abiertos de sus componentes conexas como nos muestra el ejemplo siguiente:

Ejemplo 18. Consideremos $X = \mathbb{Q}$ provisto de la topología usual. Si $A \subset X$ tiene al menos dos puntos $r, s (r \neq s)$, entonces A no es conexo, ya que existe un número irracional z tal que $r < z < s$ y luego los dos conjuntos $\{x \in \mathbb{Q} : x < z\}$ y $\{y \in \mathbb{Q} : y > z\}$ forman una separación de A . Concluimos que las componentes conexas son los conjuntos unitarios. Por supuesto, los abiertos de $\mathbb{Q}$ con respecto a la métrica usual no pueden reconstruirse partiendo de los abiertos de sus subconjuntos unitarios.

Los abiertos que se obtienen mediante la unión de abiertos en las componentes conexas de $\mathbb{Q}$, son los que corresponden a la métrica discreta sobre $\mathbb{Q}$.

Los espacios métricos (X, d) con la propiedad de que para cada $x \in X$, la componente conexa C_x se reduce al conjunto puntual x , se llaman *totalmente desconexos*. Evidentemente todo espacio métrico discreto es totalmente desconexo. A continuación daremos otro ejemplo de un espacio totalmente desconexo no discreto que ha jugado un papel importante en la historia del análisis.

Ejemplo 19. En el capítulo 4 introducimos el llamado conjunto de Cantor C constituido por todos los números reales en $[0, 1]$, que no requieren del uso de “1” en su desarrollo triádico.

Recordemos que C es homeomorfo al espacio producto $\{0, 2\}^{\mathbb{N}}$, donde en cada componente $\{0, 2\}$ se tiene la métrica discreta. Puesto que todo producto de espacios totalmente desconexos es totalmente desconexo (verifíquelo), el

discontinuo de Cantor también lo es.

Después de esta digresión volvamos a analizar nuestro problema inicial. Los espacios totalmente desconexos no discretos, como son $\mathbb{Q}$ y el discontinuo de Cantor C , nos sirven como ejemplos extremos de espacios cuyos abiertos no pueden reconstruirse a partir de los abiertos de sus componente conexas.

Definición 5.2.4. Se dice que el espacio métrico (X, d) es suma topológica de los espacios (subespacios) métricos $((X_\alpha, d_\alpha))_{\alpha \in I}$ si $X = \bigcup_{\alpha \in I} X_\alpha$ y además, U es abierto en $(X, d) \Leftrightarrow$ Para todo $\alpha \in I$, $U \cap X_\alpha$ es abierto en (X_α, d_α) .

Ejemplo 20. Evidentemente, si (X, d) es la suma topológica de $((X_\alpha, d_\alpha))_{\alpha \in I}$, entonces cada X_α es abierto en (X, d) ($X_\alpha \in \mathcal{O}_d$). Recíprocamente, si $(U_\alpha)_{\alpha \in I}$ es una partición arbitraria de (X, d) formada por subconjuntos abiertos, y si denotamos por d_α la restricción de la métrica d a U_α , entonces (X, d) es la suma topológica de los subconjuntos (U_α, d_α) , $\alpha \in I$.

Del ejemplo 20 deducimos que un espacio métrico es la suma topológica de sus componentes conexas si y solo si estas son todas abiertas. Introduciremos ahora una condición suficiente para eso.

Definición 5.2.5. Un espacio métrico (X, d) se llama *localmente conexo* si para $x \in X$ existe un sistema fundamental de vecindades conexas.

Proposición 5.2.6. En un espacio métrico localmente conexo toda componente conexa es abierta.

Demostración: Sea C una componente conexa. Para todo $y \in C$ existe una vecindad conexa V , por lo que concluimos que $V \subset C$, o sea, que C es vecindad de cada uno de sus puntos. ■

Corolario 1. Todo espacio localmente conexo es la suma de sus componentes conexas.

La demostración es consecuencia inmediata del ejemplo 20.

Si (X, d) es localmente conexo, evidentemente todo subconjunto abierto $U \subset X$ forma un subespacio localmente conexo. De aquí se concluye la siguiente ampliación de la proposición 5.2.6.

Corolario 2. Para un espacio métrico (X, d) las condiciones siguientes son equivalentes:

(LC1) (X, D) es localmente conexo.

(LC2) Las componentes conexas de todo subespacio abierto (A, d_A) son abiertas.

(LC3) Los subconjuntos abiertos conexos de (X, d) forman una base topológica de (X, d) .

Demostración: (LC1) $\Rightarrow$ (LC2) según el razonamiento anterior.

(LC2) $\Rightarrow$ (LC3): Todo abierto $A \subset \theta_d$ es la unión de sus componentes conexas. ■

La demostración de la implicación (LC3) $\Rightarrow$ (LC2) es obvia.

Ejemplo 21. Los espacios euclidianos $\mathbb{R}^n (n \geq 1)$ son localmente conexos, puesto que las bolas $B(x, \rho) \subset \mathbb{R}^n$ son conexas como veremos en el epígrafe siguiente: $\mathbb{Q}$ y $\mathbb{R} \setminus \mathbb{Q}$ no son localmente conexos.

Ejemplo 22. Todos los espacios discretos son localmente conexos. Todo espacio discreto con más de un punto da un ejemplo para un espacio que es localmente conexo pero no conexo. El inverso es también posible como mostrarán los ejemplos siguientes:

Ejemplo 23. Sea $X \subset \mathbb{R}^2$ el grafo de la función

$$g : x \rightsquigarrow \begin{cases} \sin 1/x & , \text{ si } x \neq 0 \\ 0 & , \text{ si } x = 0 \end{cases}$$

X es conexo según el ejemplo 6, sin embargo, no es localmente conexo: $(0, 0) \in X$ no posee un sistema fundamental de vecindades conexas (fig. 86).

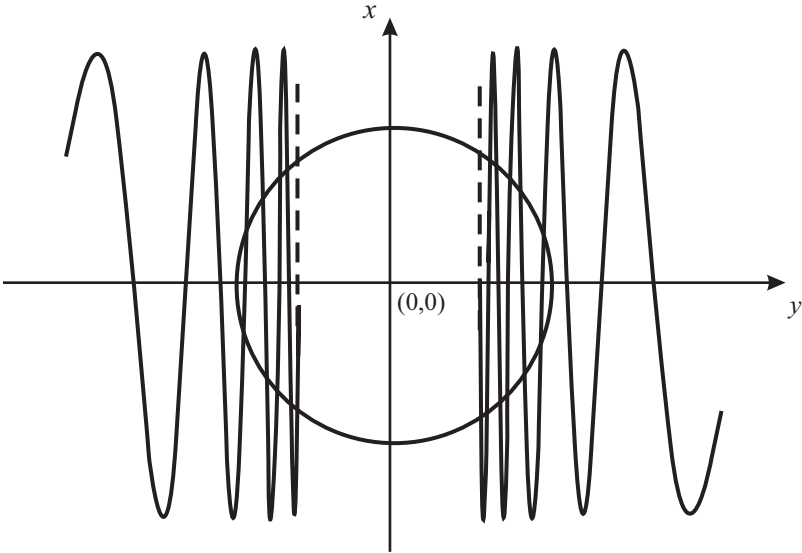


Figura 86

La imagen de un espacio localmente conexo por una aplicación continua no es localmente conexa en general. Si esto fuese cierto, todo espacio métrico (X, d) sería localmente conexo, pues es imagen continua del espacio X provisto de la métrica discreta (luego, localmente conexo).

Pasemos a analizar desde un punto de vista puramente métrico la relación entre conexión y conexión local de un espacio métrico. Para ello introduciremos la siguiente definición

Definición 5.2.7. Sea (X, d) un espacio métrico, decimos que una sucesión finita $(A_0, ..., A_k)$ de subconjuntos $A_i \subset X$ es una ϵ -cadena de longitud $k(\epsilon > 0, k \in \mathbb{N})$ si $A_j \cap A_{j-1} \neq \emptyset$ para $j = 1, ..., k$ y si el diámetro $\delta(A_j) < \epsilon$ para todo $j = 0, ..., k$. Se dice que tal cadena conecta dos puntos x y y de X , si $x \in A_0$ y $y \in A_k$ (fig. 87).

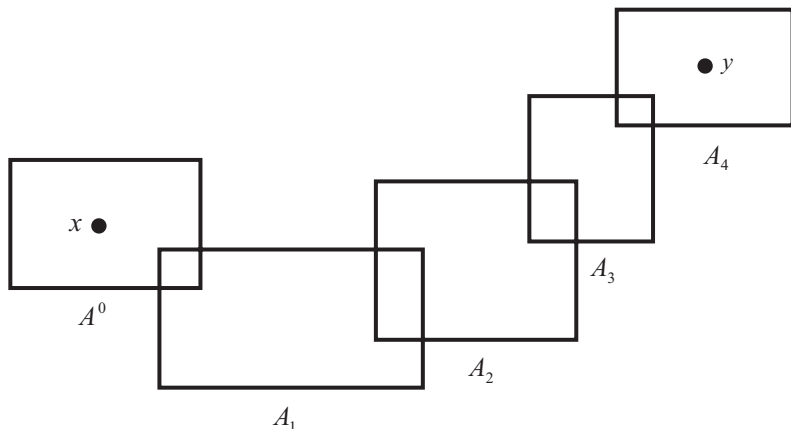


Figura 87

Proposición 5.2.8. Un espacio métrico localmente conexo (X, d) es conexo si y sólo si dos puntos cualesquiera $x, y \in X$ pueden conectarse por una ϵ -cadena de conexos (de bolas, si las bolas en (X, d) son conexas).

Demostración: Si $(A_0, \dots, A_k)$ es una ϵ -cadena de conexos, entonces $\bigcup_{j=0}^k A_j$ es conexo según el corolario 2 de la proposición 5.1.7. Luego, la condición es suficiente.

Sea inversamente (X, d) conexo. Consideremos la relación de equivalencia sobre $X : xRy \Leftrightarrow x$ y y pueden conectarse por una ϵ -cadena de conexos (de bolas).

Las clases de equivalencia correspondientes son abiertas, puesto que para todo $x \in X$ existe una vecindad conexa V (una bola de centro x) que tiene un diámetro menor que ϵ . Todos los puntos de V pueden conectarse con x por una ϵ -cadena de conexos (de bolas) de longitud 1. Por lo tanto, sólo puede existir una única clase de equivalencia. ■

Ejercicios de Autocontrol

1. Pruebe que el espacio producto $\prod_{n \in \mathbb{N}} X_n$ de los espacios métricos: $(X_n, d_n)_{n \in \mathbb{N}}$, es localmente conexo si y solo si todos los (X_n, d_n) son localmente conexos y casi todos excepto un número finito son además conexos.
2. Sea $A \subset (X, d)$ un subconjunto conexo que es a la vez abierto y cerrado. Pruebe que A es una componente conexa.

3. Pruebe que si un espacio métrico tiene sólo un número finito de componentes conexas, entonces es la suma topológica de estas componentes.
4. Determinar cuáles de los espacios siguientes son homeomorfos (fig. 88)

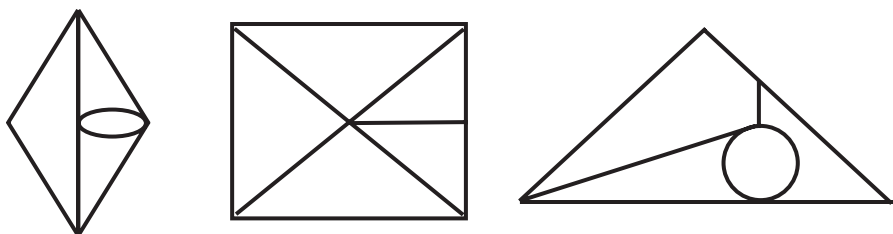


Figura 88

5. Sea (X, d) localmente conexo, $A \subset X$ y sea C una componente conexa de A . Pruebe:

- i) $\overset{\circ}{C} = C \cap \overset{\circ}{A}$

- ii) $Fr C \subset Fr A$

- iii) $A = \overline{A} \Rightarrow Fr C = C \cap Fr A$

5.3 ESPACIOS MÉTRICOS CONEXOS POR ARCOS

Tratemos, por un momento, de definir intuitivamente lo que es un conjunto conexo $C \subset \mathbb{R}^2$ sin utilizar conceptos topológicos introducidos por la métrica. Parece natural decir que C es conexo si dos puntos cualesquiera pueden conectarse por una curva que no sale del conjunto. En efecto, esta fué la primera definición precisa de un conjunto conexo (C. Weierstrass). En la figura 89a) se muestra un subconjunto no conexo y en la 89b) un subconjunto conexo según Weierstrass.

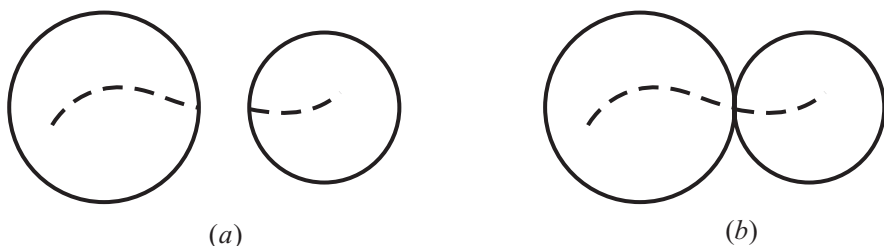


Figura 89

De hecho, este concepto de conexión es más exigente que el concepto introducido en los epígrafes anteriores. En efecto, exige que dos puntos cualesquiera del espacio puedan conectarse por un conjunto conexo especial, a saber una curva (vea (9)). Para precisar el nuevo concepto empezaremos por distinguir entre unas nociones cercanas.

Definición 5.3.1. Sea (X, d) un espacio métrico; entonces una aplicación continua f del intervalo $[0, 1]$ en (X, d) se llama *arco* en (X, d) . Si $f(0)$ es el origen del arco f y $f(1)$ su extremo, se dice que f conecta los puntos $f(0)$ y $f(1)$ en (X, d) . $C = f[[0, 1]] \subset X$ se llama *curva* determinada por f ; f se dice una representación paramétrica de esta curva.

Para una curva C existen diferentes representaciones paramétricas (arcos), por ejemplo, los dos arcos distintos $t \rightsquigarrow t$ y $t \rightsquigarrow t^2$ ($0 \leq t \leq 1$) tiene la misma curva.

Introduciremos dos operaciones para arcos.

Si $f : [0, 1] \rightarrow (X, d)$ es un arco, el arco inverso, denotado por f^- , se define por $f^- : t \rightsquigarrow f(1 - t)$

donde f^- define la misma curva que f , pero $f(1)$ es su origen y $f(0)$ su extremo (fig. 90).

Si $f : t \rightsquigarrow a \in X$ es un arco constante, entonces $f^- = f$. Tales arcos se llaman 0-arcos. La curva de un 0-arco se reduce a un punto en X , (notación para 0-arcos: $\langle a \rangle$)

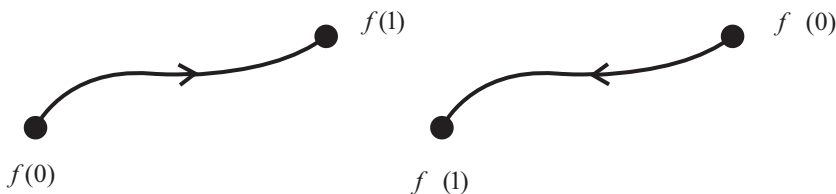


Figura 90

Si f, g son dos arcos en (X, d) tales que $f(1) = g(0)$, entonces se define la composición $f \top g$ por:

$$f \top g : t \rightsquigarrow \begin{cases} f(2t) & \text{si } 0 \leq t \leq 1/2 \\ g(2t - 1) & \text{si } 1/2 \leq t \leq 1 \end{cases}$$

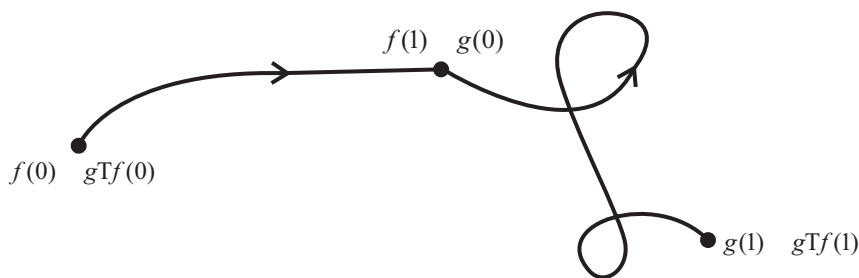


Figura 91

La operación $\top$ no es conmutativa: en general, si el arco $g \top f$ está definido, $f \top g$ puede no estar definido.

Por medio de las operaciones introducidas es fácil comprobar que la relación sobre X :

$$a \sim b \Leftrightarrow a \text{ puede conectarse con } b \text{ por un arco en } (X, d) \quad (10)$$

es una relación de equivalencia

En efecto, $a \in X$ se conecta consigo mismo por un 0-arco (reflexividad); si a se conecta con b por el arco f , entonces el arco f^- conecta b con a (simetría) y si f conecta a con $b \in X$ y si g conecta b con c en (X, d) , entonces $f \top g$ conecta a con c (transitividad).

Definición 5.3.2. Las clases de equivalencia de un espacio topológico (X, d) según la relación (10) se llaman *componentes conexas por arcos* de (X, d) .

(X, d) se llama *conexo por arcos*, si tiene una única componente conexa por arcos, o sea, si todo $x \in X$ puede conectarse con todo $y \in X$ por un arco.

Un subconjunto $C \subset X$ se dice *conexo por arcos*, si lo es el subespacio (C, d_C) .

(X, d) se llama *localmente conexo por arcos* si todo punto $x \in X$ posee un sistema fundamental de vecindades, formado por conjuntos conexos por arcos.

Puesto que toda curva es un subconjunto conexo según 5.1.8, en particular, toda componente conexa por arcos de (X, d) está incluida en una componente conexa de (X, d) .

Proposición 5.3.3. Todo espacio conexo por arcos es conexo.

Los ejemplos siguientes muestran que la conexión de muchos espacios métricos se demuestra fácilmente probando que dos puntos cualesquiera pueden conectarse por un arco.

Ejemplo 24. Consideremos el espacio euclideo $\mathbb{R}^n$ y sean $x, y \in \mathbb{R}^n$. El

conjunto

$$[x, y] := [tx + (1 - t)y : 0 \leq t \leq 1] \quad (11)$$

se llama segmento (cerrado) entre x y y .

La aplicación

$$f : \begin{matrix} [0, 1] \longrightarrow \mathbb{R}^n \\ t \longmapsto tx + (1 - t)y \end{matrix} \quad (12)$$

es un arco cuya curva asociada es el segmento $[x, y]$, el cual está completamente contenido en $\mathbb{R}^n$. Luego, $\mathbb{R}^n$ es conexo por arcos y por lo tanto conexo.

Ejemplo 25. Veamos cómo se puede generalizar la situación del ejemplo 24. En el capítulo 6 estudiaremos métricas definidas sobre espacios vectoriales X , que hacen compatible la estructura topológica con la algebraica, en el sentido de que la operación de suma y producto por un escalar son continuas con respecto a dicha métrica. En tales espacios, para cada par de puntos x, y puede definirse el segmento $[x, y]$ como en (11), el cual se encuentra completamente contenido en X y además, la aplicación (12) es continua por lo cual define un arco en X que conecta los puntos x y y . Concluimos que tales espacios métricos son conexos por arcos.

Si bien un espacio conexo por arcos es conexo (5.3.3), los espacios conexos no son necesariamente conexos por arcos como nos muestra el ejemplo siguiente:

Ejemplo 26. Sea $J := \{0\} \times [-1, 1] \subset \mathbb{R}^2$ y X la clausura $\overline{G} = G \cup J$ del grafo G de la función $x \rightsquigarrow \text{sen} 1/x$ sobre $\mathbb{R}^*$.

Según el ejemplo 5, sabemos que $\overline{G}$ es conexo; sin embargo, $\overline{G}$ no es conexo por arcos. Supongamos por el absurdo que $f = (f_1, f_2) : I = [0, 1] \rightarrow \overline{G} \subset \mathbb{R}^2$ sea un arco que conecta $(0, 0)$ con $(1/\pi, 0)$. Sea $I_0 = \{t \in I \mid f_1[[t, 1]] \subset]0, 1/\pi[\}$ y $t_0 := \inf I_0$. El intervalo I_0 es un intervalo abierto en I de la forma $[t_0, 1]$. Necesariamente $f_1(t_0) = 0$, luego $f_1[[t_0, 1]] =]0, 1/\pi[$. Sea (t_n) una sucesión en $]t_0, 1]$ tal que $t_n \rightarrow t_0$ y $f_1(t_n) = 2/n\pi$ (tal sucesión existe según el teorema del valor medio). Entonces se tiene que $f_2(t_0) = \lim(\text{sen}(1/f_1(t_n)))$; pero la sucesión oscila entre los valores $+1, -1$, luego no tiene límite. ¡Contradicción!

El lector puede probar fácilmente que $X = \overline{G}$ tiene exactamente tres componentes conexas por arcos, a saber

$$G_+ := \{(x, \sin 1/x) \mid x > 0\}, \quad J,$$

$$G_- := \{(x, \sin 1/x) \mid x < 0\},$$

Una serie de propiedades de los espacios conexos por arcos se deducen de manera análoga a las de los espacios conexos.

Si f es un arco que conecta los puntos x, y en (X, d) y si $\varphi : (X, d) \rightarrow (Y, d')$ es una aplicación continua, podemos concluir que el arco $\varphi \circ f$ conecta los puntos $f(x), f(y)$ en $f[X]$.

Proposición 5.3.4. La imagen continua de un espacio conexo por arcos es conexo por arcos. En particular, la conexión por arcos es una propiedad topológica.

La proposición siguiente es trivial.

Proposición 5.3.5. Sea $(C_\alpha)_{\alpha \in I}$ una familia de subespacios conexos por arcos de (X, d) tal que existe un C_β ($\beta \in I$) que tiene una intersección no vacía con todos los C_α ($\alpha \in I$). Entonces $\bigcup_{\alpha \in I} C_\alpha$ es conexo por arcos.

Proposición 5.3.6. Todo producto de espacios conexos por arcos es conexo por arcos.

La demostración se deja de ejercicio al lector. La proposición análoga a la 5.1.6 no se cumple.

Como muestra el ejemplo anterior existen subconjuntos conexos por arcos ($G_+ \subset \mathbb{R}^2$) cuya clausura ($\overline{G}_+ = G_+ \cup J$) no es conexo por arcos. El mismo ejemplo muestra también que las componentes conexas por arcos no son necesariamente cerradas: G_+ no es cerrada en $G_+ \cup J$.

Ejercicios de Autocontrol

1. Considere un subconjunto de puntos $x_0, \dots, x_k$ en el espacio euclideo $\mathbb{R}^n$. La unión de segmentos $\bigcup_{i=1}^k [x_{i-1}, x_i]$ se llama polígono conectando x_0 y x_k y se denota por $(x_0, \dots, x_k)$. Demuestre que todo polígono define una curva en $\mathbb{R}^n$ y que además, cada par de puntos de un abierto conexo A en $\mathbb{R}^n$ pueden conectarse por un polígono formado por segmentos paralelos a los ejes coordenados y completamente contenido en A .
2. Demuestre que para un espacio métrico (X, d) son equivalentes:
 - a) Todas las componentes conexas por arcos en (X, d) son abiertas.
 - b) Todo punto $x \in X$ posee una vecindad conexa por arcos en (X, d) .
3. Pruebe que un espacio métrico (X, d) es conexo por arcos si y sólo si es conexo y todo punto $x \in X$ posee una vecindad conexa por arcos.

CAPÍTULO 6

ESPACIOS NORMADOS Y EUCLIDEANOS

En este capítulo nos dedicaremos fundamentalmente al estudio de las propiedades métricas y topológicas de los espacios normados y euclidianos, sin entrar, salvo en casos en que el contexto lo requiera, en problemas más relacionados con el Análisis Funcional, es decir, con el estudio de los operadores y funcionales en espacios de funciones.

De acuerdo con el objetivo fundamental del texto que es el estudio de los conceptos de *convergencia* y *continuidad*, así como de otras nociones topológicas, a través de la estructura métrica, hemos decidido dedicar el 6.1 a introducir algunos resultados auxiliares sobre espacios vectoriales y el 6.2 al estudio de la compatibilidad entre la estructura algebraica y la estructura métrica en ciertos espacios vectoriales métricos que dan lugar a los espacios normados.

En 6.3 se analiza cómo se manifiestan ciertas propiedades topológicas y uniformes tales como la compacidad, la completitud y la conexión en los espacios normados, así como la relación de estas propiedades con la estructura algebraica de dichos espacios. En este epígrafe se introduce la importante clase de los espacios normados completos, *espacios de Banach*, y se estudian algunas de sus propiedades.

El 6.4 se dedica al estudio de la convergencia de series en espacios normados.

En el 6.5 se estudian los espacios euclidianos que dan lugar a una clase importante de espacios normados. El concepto fundamental que se introduce en este epígrafe es el de ortogonalidad, mediante el cual se establece una estrecha relación entre la estructura métrica, la estructura algebraica y la estructura geométrica de dichos espacios. Este epígrafe se complementa con el 6.6 donde se introduce una clase importante de espacios euclidianos, los llamados *espacios de Hilbert*.

El 6.7 está dedicado a estudiar los operadores y funcionales lineales en espacios normados. Siendo consecuentes con nuestro planteamiento inicial, no abordaremos con profundidad temas tan importantes como:

el Análisis Funcional,
la Teoría Espectral de operadores,
la Teoría de Optimización,

a pesar de la relación estrecha que tienen estos temas con el objeto de estudio de este último epígrafe. Sin embargo, en este epígrafe veremos algunos resultados de gran utilidad en las ramas antes señaladas y que tienen afinidad con el contenido estudiado en el resto del libro. Tales resultados se refieren al límite puntual de sucesiones de operadores lineales continuos (consecuencia del teorema de Banach Steinhauss) y a la separación de conjuntos convexos en espacios normados (teorema de Hahn–Banach).

El lector interesado en complementar su información acerca de algunos de los teoremas más importantes del Análisis Funcional, puede consultar los siguientes resultados fundamentales, que no serán tratados en este texto:

Teorema de Banach de la aplicación abierta.

Teorema del grafo cerrado.

Teorema de Krein–Milman.

Teorema de Banach–Aloulou;

todos los cuales se encuentran desarrollados, por ejemplo, en el libro *Análisis Funcional* de Walter Rudin.

6.1 RESULTADOS AUXILIARES SOBRE ESPACIOS VECTORIALES

Comenzaremos este epígrafe introduciendo algunas nociones algebraicas. Consideremos un conjunto arbitrario sobre el cual hay definidas dos operaciones: suma (S -ley interna) y producto por escalares (P -ley externa). Estas dos operaciones se interpretan como aplicaciones

$$S : E \times E \longrightarrow E \text{ y} \quad (1)$$

$$P : K \times E \longrightarrow E, \quad (2)$$

donde K será eventualmente uno de los conjuntos numéricos $\mathbb{R}$ o $\mathbb{C}$; en otras palabras, a cada par de elementos $x, y \in E$ y cada número $\lambda \in K$ se les hace corresponder los elementos unívocamente determinados de E , $S(x, y)$ y $P(\lambda, x)$.

Para mayor comodidad denotaremos al elemento $S(x, y)$ por $x + y$:

$$S(x, y) = x + y \quad (3)$$

y le llamaremos *suma de los elementos x y y de E* . De la misma forma denotaremos al elemento $P(\lambda, x)$ por λx :

$$P(\lambda, x) = \lambda x \quad (4)$$

y le llamaremos *producto del elemento* x por el número λ .

A la suma (3) y el producto (4) le exigiremos determinadas relaciones naturales de compatibilidad. Es decir, si $x, y, z \in E$; $\lambda, \mu \in K$, entonces debe cumplir:

(E V 1) $x + y = y + x$ (conmutatividad de la adición);

(E V 2) $(x + y) + z = x + (y + z)$ (asociatividad de la adición);

(E V 3) $\lambda(x + y) = \lambda x + \lambda y$ (distributividad del producto $(\lambda + \mu)x = \lambda x + \mu x$ con respecto a la suma);

(E V 4) $\lambda(\mu x) = (\lambda\mu)x$ (asociatividad de la multiplicación);

(E V 5) existe un elemento $\theta \in E$, llamado elemento nulo, tal que $0x = \theta$ para cualquier $x \in E$, donde 0 es el cero de K (el producto de cualquier elemento de E por el número cero es el elemento nulo);

(E V 6) $1x = x$ para cualquier $x \in E$ (el producto de cualquier elemento de E por el número 1 de K es igual al propio elemento).

De esta forma hemos incluido en las propiedades exigidas a la suma y el producto por escalares en E , un grupo fundamental de leyes algebraicas que son válidas para las correspondientes operaciones en los espacios euclidianos de dimensión finita $\mathbb{R}^n$.

Definición 6.1.1. Un *espacio vectorial* sobre un cuerpo K , es un conjunto E provisto de dos operaciones una interna llamada suma (3) y otra externa llamada producto por escalares (4), las cuales satisfacen las propiedades (E V 1)-(E V 6).

El espacio vectorial se llama real si $K = \mathbb{R}$ y complejo si $K = \mathbb{C}$. Los elementos del espacio vectorial E se denominan *vectores*.

Un subconjunto F del espacio vectorial E se llama un *subespacio vectorial* de E , si es un espacio vectorial con respecto a las operaciones de suma y producto por escalares de E restringidas a sus elementos.

No resulta difícil verificar que un subconjunto F del espacio vectorial E es un subespacio vectorial de E , si para todos $\lambda, \mu \in K$ y $x, y \in F$ se cumple

$$\lambda x + \mu y \in F \quad (5)$$

Un mismo conjunto E puede ser provisto de diferentes estructuras vectoriales, o sea, sobre un mismo conjunto E se pueden definir distintas operaciones de suma y producto que lo conviertan en un espacio vectorial. Por este motivo, cuando digamos que el conjunto E es un espacio vectorial, se sobreentiende que es con respecto a ciertas operaciones de suma y producto preestablecidas.

Veamos, de manera algo informal y sin entrar en detalles, algunas definiciones y resultados relacionados con la estructura algebraica de los espacios vectoriales.

Dados los vectores $x_1, \dots, x_n$ del espacio vectorial E y los números: $\lambda_1, \dots, \lambda_n$ de K , al vector

$$\lambda_1 x_1 + \dots + \lambda_n x_n \quad (6)$$

se le llama *combinación lineal* de los vectores $\{x_i : i = 1, \dots, n\}$ con *coeficientes*

$\{\lambda_i : i = 1, \dots, n\}$.

La familia de vectores $F = \{x_\alpha\}_{\alpha \in I}$ de E se llama *linealmente dependiente* si es posible obtener una combinación lineal nula de alguna subfamilia finita de F con coeficientes no necesariamente iguales a cero; en caso contrario, se dice que la familia F es *linealmente independiente*. Luego, la familia F de vectores de E es linealmente independiente si:

$$\begin{aligned} & \{[x_{\alpha_1}, \dots, x_{\alpha_n} \in F] \wedge [\lambda_1, \dots, \lambda_n \in K] \wedge \\ & \wedge [[\lambda_1 x_{\alpha_1} + \lambda_2 x_{\alpha_2} + \dots + \lambda_n x_{\alpha_n} = 0]]\} \Rightarrow \\ & [\lambda_1 = \lambda_2 = \dots = \lambda_n = 0] \end{aligned} \quad (7)$$

Se dice que un subespacio vectorial F del espacio vectorial E es suma de los subespacios F_1 y F_2 (notación: $F = F_1 + F_2$), si todo elemento $x \in F$ se puede expresar de la forma $x = x_1 + x_2$, donde $x_i \in F_i$, $i = 1, 2$. Si la representación de x mediante la suma de elementos de F_1 y F_2 es única, se dice que *la suma* de F_1 y F_2 es *directa* y se denota por $F = F_1 \oplus F_2$ (suma directa de los subespacios F_1 y F_2).

No resulta difícil comprobar que la suma F_1 y F_2 es directa si y solo si $F_1 \cap F_2 = \{\theta\}$.

En general, dados dos subconjuntos A y B de un espacio vectorial E , se utiliza la notación $A + B$ y λA ($\lambda \in K$) para denotar los conjuntos suma y producto por un escalar respectivamente donde:

$$A + B = \{x + y : x \in A \wedge y \in B\} \quad (\text{suma de } A \text{ y } B)$$

$$\lambda A = \{\lambda x : x \in A\} \quad (\text{producto de } A \text{ por el escalar } \lambda)$$

Se puede demostrar (utilizando el axioma de elección) que toda familia F de vectores de E contiene una *subfamilia maximal* G (no necesariamente única) de vectores linealmente independientes, y que además, todas estas subfamilias maximales tienen la misma potencia (vea el capítulo 1), o dicho en sentido figurado: *tienen la misma cantidad de elementos*. Si en lugar de F se considera el propio espacio vectorial E , entonces a la potencia de cualquier subfamilia

maximal de vectores linealmente independientes en E se le llama *dimensión algebraica* de E o simplemente *dimensión* de E . No es difícil probar que la intersección de cualquier familia $\{F_\alpha\}_{\alpha \in I}$ de subespacios vectoriales de un espacio vectorial E , es también un subespacio vectorial de E ; luego, para toda familia F de vectores de E , existe un *menor subespacio vectorial* en E que contiene a F ; a este menor subespacio vectorial se le llama *generado por F* y se denota por $[F]$. El lector puede comprobar que $[F]$ coincide con el conjunto de todas las combinaciones lineales de elementos de F . Toda subfamilia maximal G de vectores linealmente independientes de F , constituye una *base algebraica* (o simplemente *base*) del subespacio vectorial $[F]$, es decir todo $x \in F$ puede ser expresado de forma única como combinación lineal de elementos de G .

De los resultados expuestos hasta ahora se deduce que todas las bases de un subespacio vectorial F tienen la misma potencia, la cual por definición, es la *dimensión* de F . Más adelante definiremos un concepto *dimensión topológica* para espacios normados, veremos, por ejemplo, que el espacio $C([a, b], \mathbb{R})$ como espacio normado tiene dimensión topológica igual a la potencia de $\mathbb{N}$ (base topológica numerable), mientras que su dimensión como espacio vectorial tiene la potencia del continuo. En efecto, puede demostrarse que la dimensión algebraica de $C([a, b], \mathbb{R})$ no coincide con la potencia de $\mathbb{N}$ pues esto equivaldría a decir que existe una sucesión (f_n) de funciones reales y continuas sobre $[a, b]$ tal que cualquier otra función de $C([a, b], \mathbb{R})$ se puede escribir como combinación lineal de una familia finita de funciones $(f_{n_i})_{i=1}^k$.

Por otra parte $C([a, b], \mathbb{R})$ está contenido en $\mathbb{R}^{[a, b]}$ cuya dimensión algebraica puede comprobarse que coincide con la del continuo; luego, si aplicamos el axioma del continuo obtendremos que las dimensiones algebraicas de $C([a, b], \mathbb{R})$ y $\mathbb{R}^{[a, b]}$ coinciden.

Establezcamos algunas consecuencias inmediatas de los axiomas (EV1)-(EV6) de espacio vectorial.

1) El elemento nulo de E es el cero para la adición en E .

En efecto, para cualquier $x \in E$ por (EV5), (EV6) y (EV3) $x + \theta = 1x + 0x = (1 + 0)x = 1x = x$; luego, θ es el cero para la adición.

2) El elemento nulo de E es único.

Si θ_1 fuera cualquier otro cero para la adición de E , entonces $\theta + \theta_1 = \theta$ pero de acuerdo con lo que hemos probado para θ , entonces $\theta_1 + \theta = \theta_1$. De aquí, por la conmutatividad de la adición $\theta_1 = \theta$.

3) Para todo x de E se define el elemento $-x$ mediante $-x = (-1)x$ y se llama *elemento opuesto* a x .

Entonces se tiene

$$x + (-x) = \theta. \quad (8)$$

Notemos que

$$-(-x) = x \quad (9)$$

Luego, para dos elementos x, y de E se puede definir su *diferencia mediante*

$$x - y = x + (-y). \quad (10)$$

De la definición de diferencia se deduce que

$$(x - y) + y = x \quad (11)$$

Esta notación nos sugiere que la sustracción es la operación inversa a la adición.

- 4) Para cualquier $x, y \in E$, la ecuación $y + z = x$ con la incógnita z tiene una única solución $z = x - y$. En otras palabras, se pueden pasar los sumandos de un miembro a otro de una igualdad, si los pasamos con el signo opuesto. Para cerciorarnos de esto, notemos primeramente como ya hemos visto, que $z = x - y$ es solución de la ecuación dada. Por otra parte, si un elemento z de E satisface dicha ecuación, entonces adicionando $-y$ a ambos miembros, tendremos

$$\begin{aligned} -y + y + z &= x - y, \text{ es decir,} \\ \theta + z &= x - y, \text{ o lo que es lo mismo} \\ z &= x - y. \end{aligned}$$

Se prueba de 3) que $y - y = 0$.

Inversamente se tiene:

- 5) Si $x - y = 0$, entonces $x = y$. En efecto, la igualdad $x - y = 0$ significa que x es una solución de la ecuación $x + (-y) = 0$. Luego, por 4)

$$x = \theta - (-y) = \theta + y = y$$

Examinemos ahora algunas consecuencias de los axiomas de un espacio vectorial relacionados con la operación de multiplicación. Notemos que las leyes de distributividad del producto (EV3) son ciertas también con respecto a la sustracción. En efecto,

$$\begin{aligned} \lambda(x - y) &= \lambda[x + (-y)] = \lambda x + \lambda(-1)y = \\ \lambda x + (-1)\lambda y &= \lambda x - \lambda y \\ (\lambda - \mu)x &= [\lambda + (-\mu)]x = \lambda x + (-\mu)x = \lambda x - \mu x. \end{aligned}$$

Se cumplen además, las siguientes propiedades cuya demostración se deja de ejercicio al lector.

- 6) $\lambda\theta = \theta$ para cualquier λ ;
- 7) Si $\lambda x = \theta$, entonces $\lambda = 0$ o $x = \theta$;
- 8) Si $\lambda x = \lambda y$, entonces $\lambda = 0$ o $x = y$;
- 9) Si $\lambda x = \mu x$, entonces $\lambda = \mu$ o $x = \theta$;

De esta forma, podemos ver que las operaciones algebraicas definidas en los espacios vectoriales, poseen muchas de las propiedades ordinarias de la adición y multiplicación de números.

Una gran cantidad de los espacios métricos considerados a través de este libro son espacios vectoriales. Entre ellos podemos citar:

- los espacios euclidianos $\mathbb{R}^n$ y $\mathbb{C}^n$,
- los espacios de funciones continuas,
- los espacios de sucesiones $l^p(\mathbb{N})$ $1 \leq p \leq \infty$,
- los espacios producto $\mathbb{R}^l$.

Se deja al lector comprobar que con respecto a las operaciones de suma y producto para números funciones y sucesiones introducidas de manera natural en el texto, los conjuntos dados poseen una estructura de espacio vectorial.

Ejercicios de Autocontrol

1. Demuestre que $(C[a, b], \mathbb{R})$ provisto de la suma y el producto usual por escalares es un espacio vectorial.
2. Demuestre que no existe ningún espacio vectorial sobre $\mathbb{R}$ o $\mathbb{C}$ con un número finito de elementos estrictamente mayor que uno.
3. Diga cuál es el subespacio vectorial de $C([a, b], \mathbb{R})$ generado por las funciones

$$\frac{1}{x-2}, \frac{1}{x-3}, \frac{1}{x+1}.$$

4. Demuestre que la intersección de una cantidad arbitraria de subespacios vectoriales de un espacio vectorial dado, es también un subespacio vectorial.
5. Sean M y N subespacios de un espacio vectorial. Pruebe que $[M \cup N] = M + N$.
6. Sea M un subespacio vectorial del espacio vectorial E . Definamos la siguiente relación en E :

$$xRy \Leftrightarrow x - y \in M$$

- a) Pruebe que R es una relación de equivalencia.
- b) Pruebe que esta relación de equivalencia es compatible con la estructura vectorial de E ; es decir, si $x_1, x_2, y_1, y_2 \in E$, $\lambda, \mu \in K$ ($K = \mathbb{R}$ o $\mathbb{C}$), entonces

$$\left. \begin{array}{l} x_1 R y_1 \\ x_2 R y_2 \end{array} \right\} \Rightarrow (\lambda x_1 + \mu x_2) R (\lambda y_1 + \mu y_2)$$

- c) Pruebe que en el conjunto cociente $E/R = E/M$, se puede introducir una estructura vectorial definiendo la suma y el producto por escalares como sigue:

$$[x] + [y] = [x + y]$$

$$\lambda [x] = [\lambda x]$$

- d) Si E tiene dimensión n y M tiene dimensión m , demuestre que la dimensión de E/M es $n - m$.

7. Sean E y F espacios vectoriales.

- a) Demuestre que el conjunto producto $E \times F$ provisto de la suma y el producto por escalares siguiente:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

$$\lambda(x_1, y_1) = (\lambda x_1, \lambda y_1)$$

es un espacio vectorial.

- b) Pruebe que si las dimensiones de E y F son n y m respectivamente, entonces $\dim(E \times F) = n + m$.

6.2 MÉTRICAS COMPATIBLES CON LA ESTRUCTURA ALGEBRAICA DE UN ESPACIO VECTORIAL

En el capítulo 2 hemos demostrado que para las métricas usuales d_1 , d_2 y d_∞ en los espacios euclidianos $\mathbb{R}^n$, las operaciones de suma y producto por escalares son continuas en estos espacios. Un resultado análogo lo probamos también con respecto a las métricas integrales y la métrica uniforme en el espacio de las funciones continuas. Sin embargo, un espacio vectorial puede ser provisto de una métrica para la cual las aplicaciones de suma y producto por escalares no sean continuas; tal es el caso de $\mathbb{R}^2$ con la métrica del bosque como se muestra en los ejercicios del capítulo 2.

Al estudiar estructuras métricas sobre espacios vectoriales, resulta natural excluir estos casos singulares de nuestra consideración.

Definición 6.1.2. Sea E un espacio vectorial sobre el cuerpo K , provisto de una métrica d . Se dice que la métrica d es *topológicamente compatible con la estructura vectorial* de E , si las aplicaciones de suma y de producto por escalares son continuas con respecto a d .

Teniendo en cuenta que la continuidad en un espacio métrico se puede caracterizar mediante la convergencia de sucesiones, y utilizando las propiedades del límite de sucesiones en los espacios producto, la definición 6.2.1 se expresa también de la forma equivalente siguiente:

Para todo par de sucesiones $(x_n), (y_n)$ en E y toda sucesión numérica (λ_n) de K tales que

$$x_n \xrightarrow{d} x$$

$$y_n \xrightarrow{d} y$$

$$\lambda_n \rightarrow \lambda \text{ (convergencia usual en } \mathbb{R} \text{ o } \mathbb{C}, \text{ según el caso)}$$

se tiene que

$$\begin{aligned} x_n + y_n &\xrightarrow{d} x + y \\ \lambda_n x_n &\xrightarrow{d} \lambda x. \end{aligned} \quad (12)$$

Veremos ahora que la estructura topológica (abiertos y cerrados) de un espacio vectorial E provisto de una métrica compatible, queda totalmente determinada si se conoce el comportamiento de dicha métrica en un entorno del elemento nulo θ de E . En efecto, de (12) se deduce que para toda métrica d topológicamente compatible con la estructura vectorial de E se cumple:

$$x_n \xrightarrow{d} x \Leftrightarrow x_n - x \xrightarrow{d} \theta \quad (13)$$

Luego, para el estudio de la convergencia con respecto a las métricas topológicamente compatibles con la estructura vectorial de E , no es necesario conocer la distancia entre dos elementos cualesquiera, sino basta con saber la distancia de cada elemento x de E al elemento nulo θ , es decir, es suficiente conocer el comportamiento de la función

$$N : \begin{matrix} E \longrightarrow \\ x \rightsquigarrow N(x)=d(x,\theta) \end{matrix} \mathbb{R}_+ \quad (14)$$

ya que

$$\begin{aligned} x_n \xrightarrow{d} x &\Leftrightarrow N(x_n - x) \rightarrow 0. \\ &\text{(convergencia usual en } \mathbb{R}) \end{aligned} \quad (15)$$

Más aún, se tiene el resultado siguiente:

Proposición 6.2.2 (invarianza de los abiertos y cerrados por traslaciones y homotecias). Sea E un espacio vectorial provisto de una métrica d compatible con su estructura algebraica.

Entonces las siguientes condiciones son equivalentes:

- i) A es abierto (cerrado) en E ;
- ii) $\forall a \in E, a + A$ es abierto (cerrado) en E ,
- iii) $\forall \lambda \in K, \lambda \neq 0, \lambda A$ es abierto (cerrado) en E .

Demostración: De las propiedades 4) y 9) de la suma y el producto por escalares expuestas en el 6.1, se deduce que para cada $a \in A$ y $\lambda \in K$, las aplicaciones

$$S_a : E \xrightarrow{x \rightsquigarrow a+x} E \quad (16)$$

$$P_\lambda : E \xrightarrow{x \rightsquigarrow \lambda x} E \quad (17)$$

son biyectivas, y sus inversos son respectivamente las aplicaciones S_{-a} y $P_{1/\lambda}$. Por la continuidad de la suma y el producto por escalares en E se tiene que S_a, S_{-a}, P_λ y $P_{1/\lambda}$ son continuas, y por lo tanto, definen homeomorfismos de E (aplicaciones biyectivas y bicontinuas de E en sí mismo). Como los conjuntos abiertos (cerrados) se mantienen abiertos (cerrados) al ser transformados por homeomorfismos y por ser

$$S_a(A) = a + A, \quad (18)$$

$$P_\lambda(A) = \lambda A, \quad (19)$$

obtenemos el resultado deseado. ■

De la proposición 6.2.2 se deduce inmediatamente el corolario siguiente:

Corolario 1. Sea (E, d) un espacio métrico como en la proposición 6.2.2 y sea $F_\theta = (V_\alpha)_{\alpha \in I}$ un sistema fundamental de vecindades del elemento nulo θ de E . Entonces:

- i) Para cada $a \in E, F_a = (a + V_\alpha)_{\alpha \in I}$ es un sistema fundamental de vecindades de a .
- ii) Si todos los $V_\alpha \in F_\theta$ son abiertos, entonces cada U abierto en (E, d) ($U \in \theta_d$) se puede expresar como la unión de conjuntos del tipo

$a + V_\alpha$ con $a \in E$ y $V_\alpha \in F_\theta$. La demostración se deja de ejercicio al lector.

Introduzcamos un nuevo concepto de compatibilidad entre la métrica y la estructura algebraica de un espacio vectorial, que incluye las métricas topológicamente compatibles y que permite describir la estructura topológica de (E, d) a partir de una sola vecindad del elemento nulo en lugar de un sistema fundamental de vecindades como nos muestra el corolario 1 de la proposición 6.2.2.

Para ello notemos, que si bien la función N definida en (14) nos permite obtener información sobre la estructura topológica de todo el espacio métrico (E, d) a partir de su estructura topológica *local* en un entorno del elemento nulo, esta no nos permite describir las propiedades de la métrica d en todo E a partir de sus propiedades tomadas con respecto al θ . Dicho de otra forma: si se conoce la distancia del cualquier elemento $x \in E$ al elemento nulo $\theta \in E$ esto en general no es suficiente para conocer la distancia de x a otro punto $y \in E$.

Para poder obtener esta propiedad tan útil e importante, es necesario imponer condiciones adicionales a la métrica d .

Definición 6.2.3. Sea E un espacio vectorial sobre el cuerpo K , provisto de una métrica d . Se dice que la métrica d es *métricamente compatible* con la estructura vectorial de E , si para todos $x, y, z \in E$ y $\lambda \in K$ se cumple:

$$d(x + z, y + z) = d(x, y), \quad (20)$$

$$d(\lambda x, \lambda y) = |\lambda| d(x, y), \quad (21)$$

donde $|\cdot|$ denota el módulo de un número (real o complejo).

Obviamente toda distancia d métricamente compatible con la estructura vectorial de E , es también topológicamente compatible.

En efecto, si $x_n \xrightarrow{d} x$, $y_n \xrightarrow{d} y$, $\lambda_n \xrightarrow{d} \lambda$ en K , entonces de (20) se tiene

$$d(x_n + y_n, x + y) = d(x_n - x, y - y_n) \leq$$

$$\leq d(x_n - x, \theta) + d(\theta, y - y_n) \leq$$

$$\leq d(x_n, x) + d(y_n, y)$$

y además, por (20) y (21)

$$d(\lambda_n x_n, \lambda x) = d(\lambda_n x_n - \lambda_n x, \lambda x - \lambda_n x) \leq$$

$$\leq d(\lambda_n x_n - \lambda_n x, \theta) + d(\theta, \lambda x - \lambda_n x) =$$

$$= |\lambda_n| d(x_n, x) + |\lambda - \lambda_n| d(\theta, x),$$

de donde obtenemos que $x_n + y_n \xrightarrow[d]{\quad} x + y$ y $\lambda_n x_n \xrightarrow[d]{\quad} \lambda x$ o sea, las operaciones vectoriales de suma y producto por escalares son continuas con respecto a d .

El lector puede comprobar, buscando un ejemplo, que en un espacio vectorial existen métricas topológicamente compatibles con su estructura algebraica y que no son métricamente compatible.

Consideremos ahora la función $N(x) = d(x, 0)$, definida en (14), correspondiente a una distancia d definida sobre el espacio vectorial E y métricamente compatible con su estructura algebraica.

Evidentemente,

$$N(x - y) = d(x, y). \quad (22)$$

De (20), (21), (22) y las propiedades de una métrica, se obtienen inmediatamente las siguientes propiedades para la correspondiente función N :

$$(N1) \quad N(x) = 0 \Leftrightarrow x = \theta;$$

$$(N2) \quad N(\lambda x) = |\lambda| N(x);$$

$$(N3) \quad N(x + y) \leq N(x) + N(y),$$

donde x y y son elementos de E arbitrarios y $\lambda \in K$.

Por la analogía de las propiedades (N1)-(N3) con las propiedades usuales del módulo para los números reales o complejos, se conviene utilizar la notación

$$N(x) = \|x\|, \quad (23)$$

que utilizaremos a partir de ahora para denotar la función N correspondiente a una distancia métricamente compatible con la estructura vectorial de E .

Con esta nueva notación las propiedades (N1)-(N3) se expresan en la forma siguiente:

$$(N1) \quad \|x\| = 0 \Leftrightarrow x = 0;$$

$$(N2) \quad \|\lambda x\| = |\lambda| \|x\|;$$

$$(N3) \quad \|x + y\| \leq \|x\| + \|y\|.$$

Notemos que de (N1) y (N3) se deduce que $\|x\| \geq 0$ para todo $x \in E$.

Definición 6.2.4. Sea E un espacio vectorial sobre el cuerpo K . Toda función definida sobre E con valores en $\mathbb{R}_+$ y que satisfaga las propiedades (N1)-(N3) se denomina una *norma* sobre E . Se dice que el espacio métrico (E, d) es *normado*, si la métrica d proviene de una norma, es decir, si existe una norma $\|\cdot\|$ sobre E tal que

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

Se llama *subespacio normado* del espacio normado (E, d) a cualquier subespacio vectorial de E provisto de la métrica inducida por d .

El espacio normado se llama real si $K = \mathbb{R}$ y complejo si $K = \mathbb{C}$.

Observación: Generalmente se conviene en llamar espacio normado al par $(E, \|\cdot\|)$ y a partir de la norma $\|\cdot\|$, se define una distancia d en E mediante (24) para la cual se demuestra posteriormente que satisface las propiedades (20) y (21). Nosotros utilizaremos esta notación a pesar de que a lo largo del desarrollo del texto hemos seguido el camino inverso y más natural para nuestro objetivo, es decir, hallamos las propiedades adicionales que se deben imponer a una métrica para que ella pueda ser descrita mediante una norma, y a partir de ellas, obtuvimos las propiedades características de la norma. Luego, cuando hagamos referencia a nociones métricas o topológicas en un espacio normado $(E, \|\cdot\|)$, nos estaremos refiriendo a dichas nociones con respecto a la métrica d asociada a $\|\cdot\|$ mediante (24).

Proposición 6.2.5. Un espacio vectorial métrico (E, d) es normado si y solo si la distancia d es métricamente compatible con la estructura algebraica de E .

La demostración se deja de ejercicio al lector.

Ejemplo 1. Las siguientes aplicaciones reales son normas

$$\text{en } \mathbb{R}^n \text{ o } \mathbb{C}^n : \|x\| = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \quad (25)$$

$$\|x\| = \sum_{i=1}^n |x_i| \quad (26)$$

$$\|x\| = \max_{1 \leq i \leq n} |x_i| \quad (27)$$

$$\text{en } l^1(\mathbb{N}) : \|x\| = \sum_{i=1}^n |x_i| \quad (28)$$

$$\text{en } l^2(\mathbb{N}) : \|x\| = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \quad (29)$$

$$\text{en } l^\infty(\mathbb{N}) : \|x\| = \sup_{i \in \mathbb{N}} |x_i| \quad (30)$$

$$\text{en } C([a, b], \mathbb{R}) : \|x\| = \max_{t \in [a, b]} |x(t)| \quad (31)$$

$$\|x\| = \int_a^b |x(t)| dt \quad (32)$$

$$\|x\| = \left(\int_a^b |x(t)|^2 \right)^{1/2} \quad (33)$$

Ejemplo 2. El conjunto $\mathbf{P}$ de los polinomios algebraicos $p(x)$ definidos sobre un intervalo $[a, b]$, constituye un subespacio normado del espacio normado:

$E = ((C[a, b], \mathbb{R}), \|\cdot\|_\infty)$, que es además denso en E según nos muestra el teorema de Weierstrass expuesto en el capítulo 2. No resulta difícil comprobar que las aplicaciones (31)-(33) definen normas sobre $C([a, b], \mathbb{R})$. Hemos demostrado que si hacemos $d(x, y) = \|x - y\|$ para cualquiera de las aplicaciones $\|\cdot\|$ del ejemplo 1, se obtiene una distancia que es métricamente compatible con la estructura vectorial de los respectivos espacios. Cuando no hay lugar a confusión, a las normas (26), (28) y (32) se les denota indistintamente por $\|\cdot\|_1$ y a (27), (30) y (31) por $\|\cdot\|_2$.

Ejemplo 3 (continuidad de la norma). Supongamos que tenemos un espacio vectorial métrico (E, d) cuya métrica proviene de una norma $\|\cdot\|$ mediante (24); entonces la aplicación $x \rightsquigarrow \|x\|$ es continua como aplicación de (E, d) en $\mathbb{R}_+$.

En efecto, si aplicamos la desigualdad triangular a los vectores $x, y - x$, tendremos

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

de donde

$$\|y\| - \|x\| \leq \|x - y\|.$$

Intercambiando x y y obtenemos

$$|||y\| - \|x\|| \leq \|x - y\|. \quad (34)$$

No resulta difícil comprobar que la desigualdad (34) es equivalente a (N3). Si ahora tomamos una sucesión convergente en $E : x_n \xrightarrow{d} x$; entonces poniendo x_n en lugar de y en (34) se tiene

$$|||x_n\| - \|x\|| \leq \|x_n - x\| = d(x_n, x).$$

Luego, $\|x_n\| \longrightarrow \|x\|$, es decir, la aplicación $x \rightsquigarrow \|x\|$ es continua.

Veamos ahora que el corolario 1 de la proposición 6.2.2 tiene una expresión más simple en los espacios normados.

Proposición 6.2.6. Sea V una vecindad acotada del elemento nulo en un espacio vectorial normado $(E, \|\cdot\|)$ y sea (λ_n) una sucesión de elementos de K que converge a cero. Entonces se tiene:

- i) La familia $(\lambda_n V)_{n \in \mathbb{N}}$ es un sistema fundamental de vecindades del elemento nulo de E .
- ii) Todo abierto U en $(E, \|\cdot\|)$ se puede expresar como la unión de conjuntos del tipo $a + \lambda_n V$, donde $a \in E$ y $n \in \mathbb{N}$.

Demostración:

- i) Por la proposición 6.2.2 se tiene que para cada $n \in \mathbb{N}$, $\lambda_n V$ es una vecindad de $\theta \in E$. Por ser V acotada existe un número $M > 0$ tal que

$$V \subset B(\theta, M) = \{x \in E : \|x\| < M\}.$$

Luego, aplicando las propiedades de la norma, se tiene que

$$\begin{aligned} \lambda_n V &\subset \lambda_n B(\theta, M) = \{x \in E : \|x\| < |\lambda_n| M\} \\ &= B(\theta, |\lambda_n| M). \end{aligned}$$

Como $\lambda_n \rightarrow 0$ en K , la sucesión de bolas $B(\theta, |\lambda_n| M)$ constituye un sistema fundamental de vecindades de θ en $(E, \|\cdot\|)$. Por lo tanto, de la inclusión $\lambda_n V \subset B(\theta, |\lambda_n| M)$, concluimos que $(\lambda_n V)_{n \in \mathbb{N}}$ es también un sistema fundamental de vecindades de θ .

- ii) Se deduce del corolario 1 de la proposición 6.2.2 y de i). ■

Observación: De la proposición 6.2.6 se deduce, en particular, que todo abierto en un espacio normado $(E, \|\cdot\|)$ puede ser reconstruido a partir de la bola unidad $\{x \in E : \|x\| < 1\}$. Como veremos en el próximo epígrafe, las propiedades de la bola unidad no solo son determinantes en la estructura topológica de los espacios normados, sino además, tienen una estrecha relación con su estructura algebraica (vea teorema 6.3.6).

Estudiemos a continuación algunas propiedades sobre la cerrabilidad de los subespacios normados, las cuales son de gran utilidad en las aplicaciones, fundamentalmente en la teoría de optimización.

Proposición 6.2.7 (clausura de subespacios). Sean $(E, \|\cdot\|)$ un espacio normado y F un subespacio vectorial de E ; entonces la clausura $\overline{F}$ de F es también un subespacio vectorial de E .

Demostración: Sean $x, y \in \overline{F}$, y $\lambda, M \in K$. Existen sucesiones (x_n) y (y_n) en F tales que $(x_n) \xrightarrow{d} x$ y $(y_n) \xrightarrow{d} y$, donde d es la métrica asociada a $\|\cdot\|$. Por la continuidad de la suma y el producto por escalares se tiene que

$$\lambda x_n + \mu y_n \xrightarrow{d} \lambda x + \mu y$$

Por la definición de clausura, como $\lambda x_n + \mu y_n \in F$, entonces $\lambda x + \mu y \in \overline{F}$, con lo cual queda probado que $\overline{F}$ es también un subespacio vectorial de E (ver (5)).■

Proposición 6.2.8. Todo subespacio vectorial de dimensión finita de un espacio normado, es cerrado.

Demostración: Sea F un subespacio vectorial de dimensión N en $(E, \|\cdot\|)$ y tomemos una sucesión (x_n) en F tal que $x_n \rightarrow x$ ($x \in E$). En el próximo epígrafe se demuestra que todo subespacio vectorial de dimensión finita de un espacio normado, es completo (como subespacio métrico) (6.3.11). Apliquemos este resultado a nuestro problema.

Por ser la sucesión (x_n) convergente en E , (x_n) será de Cauchy en F . Como F es completo (como subespacio), la sucesión (x_n) tiene un límite x' en F . De la unicidad del límite se obtiene que $x = x'$ por lo tanto, $x \in F$ de donde F es cerrado.■

Ejemplo 4 (subespacios vectoriales no cerrados). De la proposición 6.2.8 se deduce que todo subespacio vectorial de un espacio normado de dimensión finita es cerrado; esto no ocurre generalmente en los espacios normados de dimensión infinita. En efecto, la colección de todas las funciones continuamente diferenciables del espacio normado $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ constituye obviamente un subespacio vectorial que no es cerrado, ya que el límite de una sucesión uniformemente convergente de funciones continuamente derivables no es necesariamente una función continuamente derivable.

Proposición 6.2.9 (suma de subespacios vectoriales cerrados). Sean F y M subespacios vectoriales cerrados del espacio normado $(E, \|\cdot\|)$. Si F es de dimensión finita, entonces el subespacio vectorial suma $F + M$ es también cerrado en E .

Demostración:

Sea F generado por los vectores linealmente independientes $e_1, \dots, e_k$ ($k = \dim F$). Evidentemente basta considerar el caso en que la suma de F y M es directa ya que en caso contrario, existirán r vectores e_i ; digamos $\{e_1, \dots, e_r\}$ los cuales pertenecen a M y entonces

$$F + M = F_0 \oplus M,$$

donde F_0 es el subespacio vectorial generado por $\{e_{r+1}, \dots, e_k\}$. Supongamos que F y M están en suma directa; para obtener el resultado deseado, basta utilizar el principio de inducción con respecto a la dimensión de F . En efecto,

si pudiéramos demostrar que para todo subespacio F de dimensión 1 ($F = [e] = Ke$ con $e \in E$) que está en suma directa con M , $Ke \oplus M$ es cerrado, entonces por la hipótesis de inducción, tendríamos que suponer que $F^k \oplus M$ es cerrado para todo subespacio de dimensión k , y demostrar que esto también es válido al sustituir $F^{(k)}$ por un subespacio $F^{(k+1)}$ de dimensión $k+1$. Pero esto es obvio, ya que $F^{(k+1)}$ se puede expresar en la forma

$$F^{(k+1)} = F^{(k)} \oplus Ke_{k+1}, e_{k+1} \in E,$$

y por tanto

$$F^{(k+1)} \oplus M = Ke_{k+1} \oplus F^{(k)} \oplus M$$

$$= Ke_{k+1} \oplus M_k,$$

donde $M_k = F^{(k)} \oplus M$ sería cerrado por la hipótesis de inducción, y aplicando el resultado inicial para $k=1$, se tendría que $F^{(k+1)} \oplus M$ es cerrado.

Pasemos entonces a demostrar que $Ke \oplus M$ es cerrado cuando M es cerrado y $e \notin M$. Para ello, consideremos una sucesión convergente

$$\lambda_n e + x_n \in Ke \oplus M,$$

a un punto $y \in E$ y probemos que $y \in Ke \oplus M$. Comprobemos primeramente que la sucesión numérica (λ_n) está acotada en K .

En efecto, si esto no ocurriera, existiría una subsucesión $(\lambda_{\varphi(n)})$ de elementos no nulos de la sucesión (λ_n) tal que $|\lambda_{\varphi(n)}| \rightarrow \infty$ en $\mathbb{R}$. Como la sucesión $\lambda_n e + x_n$ es convergente en E , su norma está uniformemente acotada por una constante $A > 0$:

$$\|\lambda_n e + x_n\| \leq A. \quad (35)$$

Sustituyendo n por $\varphi(n)$ y dividiendo por $|\lambda_{\varphi(n)}|$ cada miembro de (35) llegamos a que

$$\left\| e + \frac{1}{\lambda_{\varphi(n)}} x_{\varphi(n)} \right\| \leq \frac{A}{|\lambda_{\varphi(n)}|}. \quad (36)$$

De (36) haciendo tender n a ∞ , se deduce que la sucesión $\left(\frac{1}{\lambda_{\varphi(n)}} x_n \right)$ de elementos de M converge al elemento $-e$, y por tanto, $e \in M$ (por ser cerrado). Pero esto es una contradicción con la suposición de que $e \notin M$, luego, la sucesión (λ_n) está acotada en K . Por ser K igual a $\mathbb{R}$ o $\mathbb{C}$, y por el teorema de Bolzano Weierstrass (capítulo 4), cualquier sucesión acotada en K contiene una subsucesión convergente. Sea $(\lambda_{\varphi(n)})$ una subsucesión convergente a un elemento λ de K . Como toda subsucesión de una sucesión convergente converge al mismo límite que la sucesión, se tiene que la sucesión

$$(\lambda_{\varphi(n)} e + x_{\varphi(n)}) \rightarrow y.$$

Pero $\lambda_{\varphi(n)}$ converge en E al elemento λe , y por lo tanto, de la continuidad de la suma se deduce que

$$x_{\varphi(n)} \longrightarrow y - \lambda e.$$

Denotemos $x = y - \lambda e$. Obviamente $x \in M$ por ser M cerrado, luego,

$$y = \lambda e + x \in Ke \oplus M,$$

de donde concluimos que $Ke \oplus M$ es cerrado. ■

Observación: Puede demostrarse, aunque no veremos aquí ningún ejemplo para no complicar la exposición, que en general la suma de dos subespacio vectoriales cerrados cualesquiera no es un subespacio vectorial cerrado.

Definición 6.2.10. Sea F una familia de vectores en el espacio normado $(E, \|\cdot\|)$. Se llama *envoltura lineal cerrada de la familia F* , al conjunto de todos los límites en E de combinaciones lineales de elementos de F . La familia F se llama *total* si su envoltura lineal cerrada coincide con todo E .

Observación: Si denotamos por $[F]$ al subespacio vectorial generado por F (ver 6.1), entonces la envoltura lineal cerrada de F coincide con la clausura de $[F]$; luego, por la proposición 6.2.7 $[F]$ es el menor subespacio vectorial cerrado de E que contiene a F .

Si la familia F es total en E , entonces para cada $x \in F$ y todo número real $\epsilon > 0$, se pueden encontrar vectores $x_1, \dots, x_n \in F$ y números $\lambda_1, \dots, \lambda_n$ tales que

$$\left\| \sum_{k=1}^n \lambda_k x_k - x \right\| < \epsilon.$$

Ejemplo 5. Del teorema de Weierstrass sobre la aproximación de toda función real continua uniformemente sobre $[a, b]$ por polinomios algebraicos, se deduce que la familia de funciones

$$\{1, x, x^2, \dots, x^n, \dots\}$$

es total en $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$. Luego este espacio tiene subfamilias totales numerables.

Pasemos a estudiar las condiciones que deben cumplir dos normas definidas sobre un espacio vectorial E para que definan la misma convergencia, es decir, la misma estructura topológica (abiertos y cerrados) en E .

Proposición 6.2.11. Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ definidas sobre el espacio vectorial E definen la misma convergencia en E si y solo si existen constantes c_1 y $c_2 > 0$ tales que

$$c_2 \|x\|_1 \leq \|x\|_2 \leq c_1 \|x\|_1, \quad (37)$$

para todo $x \in E$.

Demostración: . Evidentemente la condición (37) es suficiente para que am-

bas normas definan la misma convergencia de E . Demostremos que esta condición es también necesaria. En efecto, si las normas $\|\cdot\|_1$ y $\|\cdot\|_2$ definen la misma convergencia en E , entonces la identidad $\mathbb{I}_E : (E, \|\cdot\|_1) \longrightarrow (E, \|\cdot\|_2)$ es un homeomorfismo (biyectiva y bicontinua). Analicemos cómo se manifiesta la continuidad de $\mathbb{I}_E$ en el elemento nulo θ de E y para ello, denotemos por d_1 y d_2 las respectivas métricas asociadas a $\|\cdot\|_1$ y $\|\cdot\|_2$. Por ser $\mathbb{I}_E$ continua en θ para cada número real $\epsilon > 0$, existe un número real $\delta > 0$ tal que:

$$x \in B_{d_1}(\theta, \delta) \Rightarrow \mathbb{I}_E(x) \in B_{d_2}(\mathbb{I}_E(\theta), \epsilon). \quad (38)$$

Teniendo en cuenta que:

$$\begin{aligned} \mathbb{I}_E(x) &= x \text{ para todo } x \in E, \\ B_{d_1}(\theta, \delta) &= \{x \in E : \|x\|_1 < \delta\} \text{ y} \\ B_{d_2}(\theta, \epsilon) &= \{x \in E : \|x\|_2 < \epsilon\} \end{aligned}$$

la expresión (38) puede escribirse en la forma equivalente

$$\|x\|_1 < \delta \Rightarrow \|x\|_2 < \epsilon.$$

Luego, como para cada $x \in E$ el elemento $y = \frac{\delta}{2\|x\|_1}x$ cumple que $\|x\|_1 = \delta/2 < \delta$, entonces por (38) se tendrá, que $\|y\|_2 < \epsilon$.

De aquí se tiene que

$$\left\| \frac{\delta}{2\|x\|_1}x \right\|_2 < \epsilon$$

o de forma equivalente

$$\|x\|_2 < \frac{2\epsilon}{\delta} \|x\|_1,$$

con lo cual queda probada una de las desigualdades en (37) tomando $c_1 = \frac{2\epsilon}{\delta}$.

La desigualdad en sentido opuesto se obtiene intercambiando $\|x\|_1$ y $\|x\|_2$. ■

Definición 6.2.12. Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ sobre un espacio vectorial E son *equivalentes* si satisfacen las desigualdades (37) para algún par de constantes $c_1, c_2 > 0$.

Observación: La proposición 6.2.11 nos dice que las métricas d_1 y d_2 asociadas a las normas $\|\cdot\|_1$ y $\|\cdot\|_2$ son topológicamente equivalentes si y solo si son métricamente equivalentes, es decir, d_1 y d_2 definen la misma convergencia en $E \Leftrightarrow$ existen constantes c_1 y c_2 tales que $c_2 d_1 \leq d_2 \leq c_1 d_1$.

Recordemos que en el capítulo 2 demostramos que la implicación $\Leftarrow$ se cumple siempre, aunque no necesariamente se satisface la implicación en sentido opuesto.

Pasemos finalmente a demostrar una propiedad importante de los espacios normados de dimensión finita.

Teorema 6.2.13 (equivalencia de las normas en los espacios normados de dimensión finita). Todas las normas definidas sobre un espacio

vectorial de dimensión finita son equivalentes.

Demostración: Sea E un espacio vectorial de dimensión n sobre el cuerpo K y sea $\{e_1, \dots, e_n\}$ una base de E . Evidentemente la aplicación

$$f : K^n \longrightarrow E,$$

(donde K^n es el producto cartesiano de n copias del cuerpo K), definida por

$$f(\lambda_1, \dots, \lambda_n) = \lambda_1 e_1 + \dots + \lambda_n e_n,$$

es una biyección de K^n sobre E .

Sabemos que mediante

$$\max_{1 \leq i \leq n} |\lambda_i| = \|\lambda\|_\infty$$

donde $\lambda = (\lambda_1, \dots, \lambda_n)$ se define una norma en K^n . Luego, la biyección f nos permite definir una norma en E de la forma

$$\|x\| = \|f^{-1}(x)\|_\infty, \quad (39)$$

notando que si $x = \lambda_1 e_1 + \dots + \lambda_n e_n$, entonces $f^{-1}(x) = (\lambda_1, \dots, \lambda_n)$. Demostremos que cualquier otra norma $\|\cdot\|_1$ definida sobre E es equivalente a la norma $\|\cdot\|$ definida en (39), y para ello, probemos primeramente que

$$f : (K^n, \|\cdot\|_\infty) \longrightarrow (E, \|\cdot\|_1)$$

es continua en $\underbrace{(0, 0, \dots, 0)}_{n \text{ veces}} \in K^n$ y que su inversa f^{-1} es continua en $\theta \in E$.

Tengamos en cuenta que

$$f(0, 0, \dots, 0) = \theta \text{ y}$$

$$f^{-1}(\theta) = (0, 0, \dots, 0).$$

Entonces, si $\lambda(k) = (\lambda_1(k), \dots, \lambda_n(k)) \longrightarrow (0, 0, \dots, 0)$ en K^n , se obtiene

$$\|f(\lambda_1(k), \dots, \lambda_n(k))\|_1 \leq \left(\sum_{i=1}^n \|e_i\|_1 \right) \|\lambda(k)\|_\infty,$$

pero como $\sum_{i=1}^n \|e_i\|_1$ es una constante y $\|\lambda(k)\|_\infty \xrightarrow{\infty} 0$

concluimos que $f(\lambda_1^{(k)}, \dots, \lambda_n^{(k)}) \longrightarrow \theta = f(0, 0, \dots, 0)$ en $(E, \|\cdot\|_1)$.

Probemos ahora la continuidad de f^{-1} en $\theta \in E$. Siguiendo la idea de la demostración de la proposición 6.2.9, obtenemos que si $x^{(k)} = \lambda_1^{(k)} e_1 + \dots + \lambda_n^{(k)} e_n$ converge a θ en E , entonces $\lambda_i^{(k)} e_i$ converge a θ en E para todo

$i = 1, \dots, n$. De aquí se deduce que $\left| \lambda_i^{(k)} \right| \xrightarrow{R \rightarrow \infty} 0$ para todo $i = 1, \dots, n$ y por lo tanto, la convergencia a θ de la sucesión $(x^{(k)})$ en E implica la convergencia a cero de $\lambda^{(k)} = (\lambda_1^{(k)}, \dots, \lambda_n^{(k)}) = f^{-1}(x^{(k)})$ en K^n , con lo cual queda probada la continuidad de f^{-1} en $\theta \in E$. Notemos ahora que la propiedad de continuidad de f en $(0, 0, \dots, 0) \in K^n$ y de f^{-1} en $\theta \in E$ es equivalente a decir que para toda sucesión $(x^{(k)})$ en E se tiene:

$$\left\| x^{(k)} \right\| \longrightarrow 0 \Leftrightarrow \left\| x^{(k)} \right\|_1 \longrightarrow 0. \quad (40)$$

Por las propiedades de la norma (40), se obtiene que

$$(x^{(k)}) \longrightarrow x \text{ en } (E, \|\cdot\|) \Leftrightarrow (x^{(k)}) \longrightarrow x \text{ en } (E, \|\cdot\|_1),$$

luego, las normas $\|\cdot\|$ y $\|\cdot\|_1$ definen la misma convergencia en E y por la proposición 6.2.11 son equivalentes. ■

Si consideramos dos normas diferentes $\|\cdot\|_1$ y $\|\cdot\|_2$ definidas sobre el espacio de dimensión finita E , por la demostración del teorema 6.2.13 cada una de ellas será equivalente a la norma $\|\cdot\|$ definida en (39), y por lo tanto, equivalentes entre sí.

Observación: Como podemos concluir del teorema 6.2.13 todas las normas definidas sobre un espacio vectorial de dimensión finita, proveen a este de un mismo concepto de convergencia, y por lo tanto, de una misma estructura topológica.

El siguiente ejemplo nos muestra que este no es el caso cuando se trata de espacios vectoriales de dimensión infinita.

Ejemplo 6. En el espacio vectorial $C([0, 1], \mathbb{R})$ las normas $\|\cdot\|_\infty$ y $\|\cdot\|_1$ definidas por (31) y (32) no son equivalentes ya que no definen la misma convergencia. En efecto, la sucesión (x^n) converge a cero con respecto a $\|\cdot\|_1$ pero no converge con respecto a $\|\cdot\|_\infty$

Ejercicios de Autocontrol

1. Pruebe que las aplicaciones (25)-(33) definidas en el ejemplo 1 son normas en los correspondientes espacios vectoriales allí señalados.
2. Una función real $|x|$ definida en un espacio vectorial E sobre K se llama seminorma si cumple:

$$\text{S N 1 } |x| \geq 0 \text{ para todo } x \in E;$$

$$\text{S N 2 } |\alpha x| = |\alpha| \cdot |x|, x \in E, \alpha \in K;$$

$$\mathbf{S\ N\ 3} \quad |x + y| \leq |x| + |y|, x, y \in E.$$

Sea $M = \{x : |x| = 0\}$. Pruebe que M es un subespacio vectorial de E y que el espacio cociente E/M con

$$\|x\| = \inf_{m \in M} |x + m|$$

es un espacio normado.

3. Sea E el espacio de todas las funciones reales x definidas sobre $[0, 1]$ que se anulan excepto en un conjunto numerable (t_n) de puntos para los cuales

$$\|x\| = \sum_{n=1}^{\infty} |x(t_n)| < \infty.$$

Pruebe que E es un espacio normado no separable.

4. Pruebe el corolario 1 de la proposición 6.2.2.

5. Pruebe la proposición 6.2.5.

6. Demuestre que si U y A son subconjuntos de un espacio normado E , y U es abierto, entonces $U + A$ también es abierto.

6.3 CONEXION, COMPACIDAD Y COMPLETITUD EN ESPACIOS NORMADOS

Veamos como se manifiestan en los espacios normados las propiedades topológicas de conexión, compacidad y completitud.

Proposición 6.1.3 (conexión en espacios normados). Todo espacio vectorial normado es conexo por arcos y localmente conexo por arcos.

Observación: De este resultado se deduce que los espacios normados son además conexos y localmente conexos.

Demostración: Todo par de puntos x, y de $(E, \|\cdot\|)$ se pueden unir por el arco continuo

$$[0, 1] \xrightarrow{t \rightsquigarrow y + (1-t)x} E$$

llamado *segmento con origen en x y extremo en y* , luego, $(E, \|\cdot\|)$ es conexo por arcos. Si x, y pertenecen a una bola $B(a, \epsilon)$, entonces el segmento con

origen en x y extremo en y está completamente contenido en dicha bola, de donde se deduce que $(E, \|\cdot\|)$ es además localmente conexo por arcos.■

Definición 6.3.2. Un subconjunto C del espacio vectorial E se llama *convexo* si para todo par de puntos $x, y \in C$, el segmento con origen en x y extremo en y está completamente contenido en C . El conjunto C se llama *estrellado* con respecto al punto $a \in C$, si para cada $x \in C$ el segmento con origen en a y extremo en x está completamente contenido en C (fig. 92)

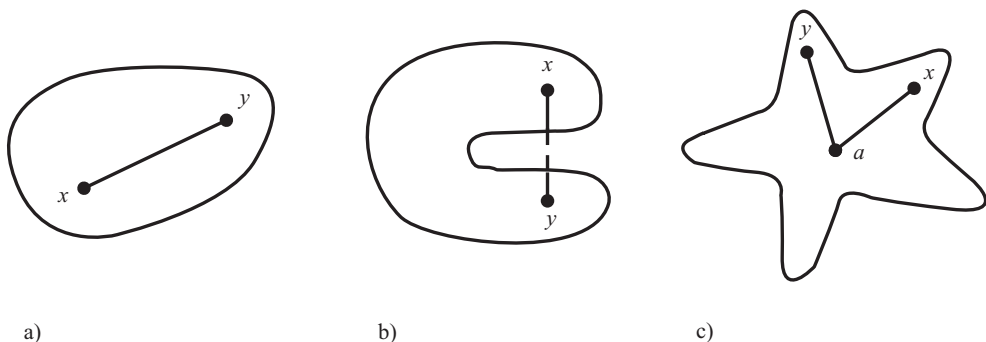


Figura 92

Ejemplo 7. Todo conjunto convexo es estrellado con respecto a cada uno de sus puntos.

Ejemplo 8. Cualquier bola en un espacio normado es un conjunto convexo.

El siguiente resultado es elemental, por lo que se deja de ejercicio al lector.

Proposición 6.3.3. Todo subconjunto convexo y todo subconjunto estrellado en un espacio normado es conexo por arcos.

Obviamente ningún espacio normado $(E, \|\cdot\|)$ puede ser compacto ya que no es acotado. En efecto, si fuera acotado, se tendría que existe un número $\rho > 0$ tal que

$$\|x\| \leq \rho, \text{ para todo } x \in E,$$

pero esto no puede ocurrir ya que en un espacio normado se pueden encontrar vectores con norma todo lo grande que se quiera, por ejemplo, si $x \neq 0$, el elemento $M \frac{x}{\|x\|}$ tiene norma igual a M .

Pasemos a demostrar un resultado no trivial con respecto a la compacidad en los espacios normados y de gran importancia en las aplicaciones. Para ello, es necesario primeramente el siguiente lema auxiliar.

Lema 6.3.4 (lema de Riesz). Sea $(E, \|\cdot\|)$ un espacio vectorial normado y

sea F un subespacio vectorial cerrado propio de E . Entonces para cada número α , $0 < \alpha < 1$ existe un vector x_α tal que $\|x_\alpha\| = 1$ y $\|x - x_\alpha\| \geq \alpha$ si $x \in F$.

Demostración: Seleccionemos $x_1 \in E/F$, y sea

$$\rho = \inf_{x \in F} \|x - x_1\|.$$

Como F es cerrado, tenemos que $\rho > 0$ y además, existe $x_0 \in F$ tal que $\|x_0 - x_1\| \leq \alpha^{-1}\rho$ (porque $\alpha^{-1}\rho > \rho$). Pongamos $x_\alpha = h(x_1 - x_0)$, donde $h = \|x_1 - x_0\|^{-1}$. Es claro que $\|x_\alpha\| = 1$. Si $x \in F$, entonces $h^{-1}x + x_0 \in F$, y por lo tanto,

$$\|x - x_\alpha\| = \|x - hx_1hx_0\| = h\|(h^{-1}x + x_0) - x_1\| \geq h\rho.$$

Pero $h\rho = \|x_1 - x_0\|^{-1}\rho \geq \alpha$ por la forma en que x_0 fué seleccionado. De aquí que $\|x - x_\alpha\| \geq \alpha$ si $x \in F$, lo cual completa la demostración. ■

Observación: El lema de Riesz puede ser enunciado de forma equivalente: Si F es un subespacio cerrado propio de E , entonces existen puntos sobre la superficie de la esfera unidad en E cuya distancia a F está tan cerca de 1 como se quiera.

Teorema 6.3.5. En un espacio vectorial normado $(E, \|\cdot\|)$ son equivalentes:

- i) $(E, \|\cdot\|)$ es de dimensión finita.
- ii) La bola unidad $B_1 = \{x : \|x\| \leq 1\}$ es compacta.
- iii) $(E, \|\cdot\|)$ es localmente compacto.

Demostración:

- i) $\Rightarrow$ iii) En la demostración de 6.2.13 vimos que todo espacio normado de dimensión finita n sobre el cuerpo K es homeomorfo al espacio producto $K^n (= \mathbb{R}^n \text{ o } \mathbb{C}^n)$.

Consideremos el homeomorfismo $f : K^m \longrightarrow X$ definido en 6.2.13. La bola unidad B_1 de E es un subconjunto cerrado en E' , y por lo tanto, $f^{-1}[B_1]$ es cerrado y acotado en K^n . Por la caracterización de los compactos en los espacios euclidianos de dimensión finita, se tendrá que $f^{-1}[B_1]$ es compacto en K^n . Luego, la imagen por la aplicación continua f del compacto $f^{-1}[B_1]$ será compacto en E . De aquí se deduce ii) teniendo en cuenta que $f[f^{-1}[B_1]] = B_1$.

- ii) $\Rightarrow$ iii) Si $a \in E$, entonces por 6.2.6 ii) y el resultado del ejercicio de auto-control 3 del 6.3, se tiene que $(a + \lambda_n B_1)_{n \in \mathbb{N}}$ es un sistema fundamental de vecindades compactas de a .

iii) $\Rightarrow$ ii) Sea V una vecindad compacta del elemento nulo $\theta \in E$. Evidentemente V contiene a una cierta bola $B(\theta, \rho) = \{x : \|x\| \leq \rho\}$ que también será compacta, luego por el resultado del ejercicio de autocontrol 3 del 6.3, tendremos que $\frac{1}{\rho}B(\theta, \rho) = B_1$ es también compacta.

ii) $\Rightarrow$ i) Supongamos por el absurdo que F no es de dimensión finita.

Seleccionemos $x_1 \in S_1 = \{x \in E : \|x\| = 1\}$ y sea F_1 el subespacio generado por x_1 . Entonces F_1 es un subespacio propio de E y además es cerrado según 6.2.8. Luego, por el lema de Riesz (6.3.4) existe $x_2 \in S_1$, tal que $\|x_2 - x_1\| \geq 1/2$. Sea F_2 el subespacio cerrado propio de E generado por x_1 y x_2 ; entonces, existe $x_3 \in S_1$ tal que $\|x_3 - x\| \geq 1/2$ si $x \in F_2$ (6.3.4). Siguiendo el procedimiento de inducción, obtenemos una sucesión infinita de elementos de S_1 tal que $\|x_n - x_m\| \geq 1/2$ si $m \neq n$. Evidentemente esta sucesión no puede contener ninguna subsucesión convergente lo cual contradice la compacidad de S_1 . Luego, E debe ser de dimensión finita. ■

Pasemos ahora a estudiar el problema de la mejor *aproximación* en los espacios normados. En muchos problemas de la matemática se considera el problema de la mejor (en cierto sentido) representación aproximada de una función dada por medio de funciones de otra clase determinada. En ocasiones es útil conocer el menos si esa mejor aproximación existe. Este problema de la existencia de la mejor aproximación fue ya considerado al estudiar los problemas de distancia mínima de un punto a un conjunto en un espacio métrico.

En el capítulo 4, se demostró que la distancia mínima de un punto x a un subconjunto compacto A de un espacio métrico (E, d) se alcanza en algún punto $x_0 \in A$;

$$d(x, x_0) = d(x, A).$$

El problema de existencia de la mejor aproximación puede ser tratado también en los espacios normados donde adquiere otras características, ya que aquí se nos plantea generalmente hallar la distancia mínima de un punto a un subespacio vectorial no trivial, que como ya sabemos no puede ser compacto. Introduzcamos la definición siguiente:

Definición 6.3.6. Sea A un subconjunto del espacio normado $(E, \|\cdot\|)$ y sea $x \in E$. Se llama *desviación mínima* del elemento x con respecto al conjunto A al número

$$\rho = \inf \|x - y\|, y \in A. \quad (41)$$

Si existe algún elemento $y_0 \in A$ tal que

$$\rho = \|x - y_0\|, \quad (42)$$

entonces a este y_0 se le llama un *vector de la mejor aproximación* en el conjunto A al elemento x . Notemos que el elemento de la mejor aproximación no tiene que ser necesariamente único. Se tiene el siguiente resultado de gran importancia en la teoría de aproximación.

Teorema 6.3.7 (existencia del elemento de la mejor aproximación).

Sea $(E, \|\cdot\|)$ un espacio vectorial normado y sea F un subespacio vectorial de dimensión finita de E . Entonces para todo $x \in E$, existe en F un vector de la mejor aproximación al elemento x .

Demostración: Obviamente es suficiente considerar el caso cuando $x \notin F$. Sea y^* un elemento fijo de F y denotemos su desviación de x por $\delta = \|y^* - x\|$. Es claro que un elemento de F de la mejor aproximación a x debe encontrarse entre aquellos $x \in F$ para los cuales $\|x - y\| \leq \delta$. Pero para estos elementos y se tiene que

$$\|y\| \leq \|y - x\| + \|x\| \leq \|x\| + \delta$$

Tomemos $M = \|x\| + \delta$ y denotemos por $B = \{y \in F : \|y\| \leq M\}$. En particular todo $y \in F$ para el cual $\|y - x\| \leq \delta$ pertenece a B y por ello es suficiente buscar un elemento de la mejor aproximación en el conjunto B . Este conjunto B , por ser una bola cerrada en F , es compacto según 6.3.5. Consideremos ahora la desviación $\|y - x\|$ como una función real de y definida sobre el subconjunto compacto B de F . Por la continuidad de la suma y de la norma, esta función es continua, y por lo tanto, alcanza su valor mínimo sobre B en algún punto y_0 . Evidentemente y_0 será un elemento en F de la mejor aproximación a x , con lo cual queda probado el teorema. ■

Supongamos que el subespacio vectorial F de $(E, \|\cdot\|)$ está generado por n vectores linealmente independientes $x_1, x_2, \dots, x_n$. Entonces el problema de la mejor aproximación se reduce a hallar los coeficientes $\lambda_1, \lambda_2, \dots, \lambda_n$ en K para los cuales

$$\left\| \sum_{k=1}^n \lambda_k x_k \right\| = x, \quad (43)$$

tenga el menor valor. En otras palabras, se pide hallar una combinación lineal de los vectores x_i ($1 \leq i \leq n$) cuya desviación a x sea mínima.

El teorema probado garantiza la existencia de una tal combinación lineal. Su unicidad puede ser probada solamente bajo algunas condiciones adicionales, pero no entraremos aquí a discutir este problema.

Considerando espacios funcionales específicos con distintas normas, se puede establecer la existencia de la mejor aproximación a una función dada por diferentes métodos al igual que la estimación de la exactitud de la aproxi-

mación.

Ejemplo 9 (sobre la mejor aproximación polinomial a una función continua). Los problemas clásicos de la teoría de aproximación tratan sobre la mejor aproximación uniforme de funciones continuas mediante polinomios. Este problema fue estudiado a mediados del siglo XIX por el famoso matemático ruso P.L. Chebishev (1821-1894).

Si denotamos por $\mathbf{P}_n$ al subespacio vectorial de $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ constituido por los polinomios algebraicos de grado $\leq n$, no es difícil demostrar que $\mathbf{P}_n$ está generado por las funciones linealmente independientes $1, t, t^2, \dots, t^n$, es decir, $\dim \mathbf{P}_n = n + 1$.

Luego, aplicando el teorema 6.3.7 se deduce que para cada función real continua f definida sobre (a, b) existe un polinomio $P(t) \in \mathbf{P}_n$ cuya desviación uniforme a f es mínima:

$$\max_{a \leq t \leq b} |f(t) - P(t)| = \inf_{Q \in \mathbf{P}_n} \|f - Q\|_\infty.$$

El propio Chebishev en sus trabajos, no consideró el problema de la existencia de la mejor aproximación polinomial a una función continua y estudió las propiedades de este polinomio asumiendo que existía. Más tarde, a principios del siglo XX, fue demostrada la existencia del polinomio de la mejor aproximación uniforme por el matemático francés E. Bord (1871-1956). En esta época el matemático belga De la Vallée-Poussin, una vez demostrada la unicidad del polinomio de la mejor aproximación, obtuvo toda una serie de propiedades características de este polinomio. Para tener una idea de la importancia del estudio de los problemas sobre la mejor aproximación citemos el siguiente teorema de D. Jackson obtenido en 1915 el cual no demostraremos:

Sea $0 < \alpha < 1$ y $f : \mathbb{R} \rightarrow \mathbb{R}$ una función continua periódica de período 2π . Entonces f pertenece a la clase de Lipschitz de orden α (ie. $|f(x) - f(y)| \leq M|x - y|^\alpha$) si y solo si la mejor aproximación uniforme de f por polinomios trigonométricos de grado $\leq n$, es del orden de C/n^α para todo $n \in \mathbb{N}$, donde C es una constante positiva que no depende de n .

Dicho de otra forma, si se denota

$$E_n(f) = \inf \{ \|f - T_n\|_\infty : T_n \text{ es un polinomio trigonométrico de grado } \leq n \},$$

donde,

$$T_n(t) = \alpha_0 + \sum_{k=1}^n \alpha_k \cos kt + b_n \sin kt,$$

entonces, f pertenece a la clase de Lipschitz de orden

$$\alpha \ (0 < \alpha < 1) \Leftrightarrow E_n(f) \leq C/n^\alpha.$$

Más adelante en el 6.6 veremos como se resuelve el problema de la mejor aproximación en un espacio de Hilbert, y como caso particular, la mejor

aproximación cuadrática (con respecto a la norma (33)) de funciones.

Pasemos ahora al estudio de la completitud en los espacios normados. Sabemos que el espacio euclideo finito dimensional K^n ($K = \mathbb{R}$ o $\mathbb{C}$) es completo con cualquiera de las normas (25)-(27). Por otra parte, no es difícil probar que si $(E, \|\cdot\|_1)$ es un espacio vectorial normado de dimensión finita n , entonces, al sustituir $\|\cdot\|_1$ por la norma equivalente $\|\cdot\|$ definida en (39), la aplicación auxiliar $f : (K^n, \|\cdot\|_\infty) \longrightarrow (E, \|\cdot\|)$ del teorema 6.2.13 se convierte en una isometría ya que

$$\|f(\lambda)\| = \|\lambda\|_\infty.$$

Pero como $(K^n, \|\cdot\|_\infty)$ es completo y la completitud es invariante ante isometrías, se deduce que $(E, \|\cdot\|)$ es completo. Finalmente por ser las normas $\|\cdot\|$ y $\|\cdot\|_1$ equivalentes sobre E , concluimos que $(E, \|\cdot\|_1)$ es completo.

El resultado obtenido puede enunciarse de la manera siguiente:

Proposición 6.3.8. Todo espacio vectorial normado de dimensión finita es completo. La demostración se obtiene del razonamiento anterior.

Es conocido que en los espacios normados de dimensión infinita no se cumple un resultado análogo al 6.3.8. En efecto, en el capítulo 3 vimos que el espacio normado $((C[a, b], \|\cdot\|_2)$ donde $\|\cdot\|_2$ viene definida por (33), no es completo. La importancia del concepto de completitud en los espacios métricos justifica la introducción de la definición siguiente:

Definición 6.3.9. Se llama *espacio de Banach* a todo espacio normado completo.

Observación: Evidentemente todo subespacio vectorial cerrado de un espacio de Banach es también un espacio de Banach.

Ejemplo 10. De la proposición 6.3.8 se deduce que todo espacio normado de dimensión finita es de Banach.

Ejemplo 11. Si $(E, \|\cdot\|)$ es un espacio normado sobre el cuerpo K , entonces, el espacio $B(E, \mathbb{R})$ de las funciones acotadas de E en $\mathbb{R}$ es un espacio vectorial para la suma usual de funciones y el producto usual por un escalar de K . Además, no es difícil comprobar que para $g \in B(E, \mathbb{R})$: $\|g\|_\infty = \sup_{x \in E} |g(x)|$

define una norma en $B(E, \mathbb{R})$ que provee a este conjunto de una estructura de espacio de Banach (comparar con los resultados obtenidos en el capítulo 3).

Proposición 6.3.10. Cada espacio normado tiene un completamiento que es

un espacio de Banach.

Observación: Recordemos que en el capítulo 3 fue demostrado que todo espacio métrico admite un completamiento. Como a todo espacio normado $(E, \|\cdot\|)$ se le puede asociar una distancia métricamente compatible con su estructura vectorial, mediante

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

este espacio (E, d) admite un completamiento $(\tilde{E}, d)$, es decir, existe una isometría i de (E, d) en un subespacio denso de $(\tilde{E}, d)$ y además el espacio métrico completo

$(\tilde{E}, d)$ es único salvo isometrías.

La proposición 6.3.10 nos dice que en la familia de todos los completamientos de un espacio normado existe uno que es un espacio de Banach. Es decir, el correspondiente E es un espacio vectorial y la distancia ρ asociada es métricamente compatible con la estructura algebraica de E , por lo cual ρ proviene de la norma

$$\|x\|_c = \rho(x, 0), x \in E.$$

Demostración: de la proposición 6.3.10. Siguiendo la demostración del teorema sobre completamiento expuesto en el capítulo 3 y considerando $(E, \|\cdot\|)$ en lugar de (X, d) y $x_0 = \theta$, tendremos que E se identifica con el subespacio vectorial $i[E]$ de $B(E, \mathbb{R})$ (funciones acotadas de E en $\mathbb{R}$ formado por las funciones g_x definidas por:

$$g_x(y) = \|y - x\| - \|y\|, x \in E.$$

Es claro que $g_x \in B(E, \mathbb{R})$, ya que

$$\|g_x\|_\infty = \sup_{y \in E} |g_x(y)| = \|x\|,$$

luego, la restricción de la norma de $\overline{i[E]} = \tilde{E}$ a $i[E]$ coincide con la norma de E ; pero $\|\cdot\| : i[E] \rightarrow \mathbb{R}$ es una aplicación uniformemente continua y por el teorema de prolongación de funciones uniformemente continuas (ver capítulo 3), esta aplicación tendrá una prolongación continua a $\tilde{E}$ que por unicidad tiene que coincidir con la norma $\|\cdot\|_c$ de dicho completamiento. ■

Ejemplo 12. En el capítulo 3 hemos visto que el espacio $C(E, \mathbb{R})$ no es de Banach para ninguna de las normas $\|\cdot\|_1$ y $\|\cdot\|_2$ definidas por (32) y (33). Los espacios normados

$$(C([a, b], \mathbb{R}), \|\cdot\|_1) \text{ y}$$

$$(C([a, b], \mathbb{R}), \|\cdot\|_2)$$

admiten sendos completamientos que se denotan mediante $L_1[a, b]$ y $L_2[a, b]$

respectivamente y que son espacios de Banach.

Un resultado interesante, y al que ya hicimos referencia en el capítulo 3, es que la norma en $L_1[a, b]$ y $L_2[a, b]$ puede expresarse mediante un concepto de integral más general llamada *integral de Lebesgue*, que coincide para las funciones continuas con la integral de Cauchy, a partir de la cual se definen $\|\cdot\|_1$ y $\|\cdot\|_2$.

Proposición 6.3.11 (criterio de Cauchy para la convergencia de sucesiones en un espacio de Banach). Una sucesión (x_n) de puntos de un espacio de Banach $(E, \|\cdot\|)$ converge si y solo si

$$\|x_n - x_m\| \xrightarrow{n, m \rightarrow \infty} 0. \quad (44)$$

La demostración se obtiene inmediatamente de la definición de sucesión de Cauchy y de (24).

Ejercicios de Autocontrol

1. Demuestre las siguientes propiedades de los conjuntos convexos:
 - a) La intersección de cualquier familia de conjuntos convexos en un espacio vectorial es convexo.
 - b) $C \subset E$ es convexo si y solo si se cumple:

$$[x_1, \dots, x_n \in C] \wedge [\lambda_1, \dots, \lambda_n \geq 0] \wedge \sum_{i=1}^n \lambda_i = 1 \Rightarrow \sum_{i=1}^n \lambda_i x_i \in C.$$

- c) Si A, B son convexos en E y $\alpha, \beta \in \mathbb{R}$, entonces,

$$\alpha A + \beta B = \{\alpha x + \beta y : x \in A \wedge y \in B\} \text{ es conexo.}$$
- d) $C \subset E$ se llama cono si cumple:

$$\lambda > 0 \Rightarrow \lambda C \subset C.$$

Pruebe que un cono $C \subset E$ es convexo si y solo si cumple $C + C \subset C$.

2. Pruebe que el espacio normado E del ejercicio 3 de autocontrol del 6. 2, es un espacio de Banach.
3. Sean K_1 y K_2 compactos en el espacio normado E y λ un escalar. Demuestre que $K_1 + K_2$ y λK_1 son también compactos.
4. Demuestre que el espacio cociente de un espacio de Banach por un subespacio vectorial cerrado (ver ejercicio 2 de autocontrol 6. 2) es también un espacio de Banach.

6.4 SERIES EN ESPACIOS NORMADOS

Comencemos este epígrafe con el estudio de un tipo de sucesiones que aparecen en los espacios normados, las cuales son de gran importancia en las aplicaciones, especialmente a los métodos numéricos.

Dada una sucesión (x_k) de elementos de un espacio normado $(E, \|\cdot\|)$ y gracias a la estructura vectorial de E , se puede construir la sucesión de *sumas parciales*

$$S_n = x_1 + x_2 + \dots + x_n = \sum_{k=1}^n x_k \quad (45)$$

(suma parcial de orden n correspondiente a la sucesión (x_k)).

Definición 6.4.1. Se dice que la sucesión (x_k) de elementos del espacio normado $(E, \|\cdot\|)$ determina una *serie convergente*, si la correspondiente sucesión (S_n) de sus sumas parciales converge.

Al límite S de la sucesión (S_n) , cuando este existe, se le llama *suma de la serie* correspondiente a la sucesión (x_k) , y se denota por

$$S := \lim (S_n) = \sum_{k=1}^{\infty} x_k. \quad (46)$$

Se dice también que la serie $\sum_{k=1}^{\infty} x_k$ es *convergente* si el límite de la sucesión (S_n) existe, en cuyo caso se utiliza la notación (46). En caso contrario se dice que la serie $\sum_{k=1}^{\infty} x_k$ es *divergente*.

Si la serie $\sum_{k=1}^{\infty} x_k$ es convergente, al elemento $S - S_n$ se le llama *resto de orden n* de la *serie* y se denota por

$$r_n = S - S_n = \sum_{k=n+1}^{\infty} x_k. \quad (47)$$

Diremos que la serie $\sum_{k=1}^{\infty} x_k$ es *absolutamente convergente* si la serie de números

reales positivos $\sum_{k=1}^{\infty} \|x_k\|$ converge.

Proposición 6.4.2. La serie $\sum_{k=1}^{\infty} x_k$ es convergente si y solo si su resto tiende a cero.

La demostración se deduce obviamente de la definición anterior. Los problemas de convergencia de series en espacios funcionales de Banach, han dado lugar a importantes ramas de la matemática y en particular del Análisis funcional. Baste citar la teoría de las series de Fourier que da lugar al Análisis Armónico con sus innumerables aplicaciones en la Física Matemática. Por este motivo nos dedicaremos a continuación al estudio de las series en espacios de Banach.

Proposición 6.4.3 (criterio de Cauchy de convergencia de series en un espacio de Banach). La serie $\sum_{k=1}^{\infty} x_k$ converge en el espacio de Banach $(E, \|\cdot\|)$ si y solo si satisface la *condición de Cauchy*:

Para todo $\epsilon > 0$ existe $N = N(\epsilon) \in \mathbb{R}$ tal que si $m \geq n \geq N$, entonces

$$\left\| \sum_{k=n}^m x_k \right\| < \epsilon. \quad (48)$$

Demostración: El lector puede comprobar que la propiedad (48) es equivalente a la condición de Cauchy para la sucesión (S_n) , la cual es equivalente a la convergencia de (S_n) por la completitud de $(E, \|\cdot\|)$. ■

Proposición 6.4.4. Toda serie absolutamente convergente en un espacio de Banach, es convergente.

Demostración: En efecto, para cada par de números naturales $m \geq n$ se tiene

$$\left\| \sum_{k=n}^m x_k \right\| \leq \sum_{k=n}^m \|x_k\|. \quad (49)$$

Si la serie $\sum_{k=1}^{\infty} x_k$ converge absolutamente en $(E, \|\cdot\|)$, entonces la serie de números reales positivos $\sum_{k=1}^{\infty} \|x_k\|$ converge en $\mathbb{R}$.

Pero $\mathbb{R}$ es un espacio de Banach con respecto a la norma definida por el

módulo. Luego, aplicando el criterio (48) a la serie $\sum_{k=1}^{\infty} \|x_k\|$ y teniendo en cuenta (49), se obtiene el resultado deseado, gracias a 6.4.3. ■

Ejemplo 13 (series convergentes y no absolutamente convergentes). En los cursos de Análisis Clásico se demuestra que la serie de números reales $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ es convergente pero no es absolutamente convergente.

Definición 6.4.5. La serie $\sum_{k=1}^{\infty} x_k$ en el espacio normado $(E, \|\cdot\|)$ se llama *condicionalmente convergente* si es convergente pero no absolutamente convergente.

Ejemplo 14 (teorema de Riemann sobre el reordenamiento de series condicionalmente convergente). A continuación enunciamos sin demostración el siguiente resultado de Riemann que puede ser hallado en la literatura sobre Análisis Matemático:

Si la serie de números reales $\sum_{k=1}^{\infty} a_k$ es condicionalmente convergente, entonces para cada número real a existe un reordenamiento (b_k) de los términos de la sucesión (a_k) (es decir, para cada a_k hay un b_j tal que $a_k = b_j$ y viceversa), tal que $\sum_{k=1}^{\infty} b_k = a$.

El lector puede comprobar sin mucha dificultad que una serie absolutamente convergente $\sum_{k=1}^{\infty} x_k$ converge siempre al mismo valor para cualquier reordenamiento de la sucesión (a_k) .

La búsqueda de otros espacios normados diferentes de $\mathbb{R}$, en los cuales se cumpla un análogo del teorema de Riemann ha sido tema de investigación para matemáticos ilustres como Stefan Banach y sobre esta temática quedan aún en la actualidad varios problemas abiertos.

Veamos algunos criterios para la convergencia absoluta de series cuya demostración se deja de ejercicio al lector.

Proposición 6.4.6 (criterios de convergencia absoluta de series). Si la sucesión (a_k) del espacio normado $(E, \|\cdot\|)$ satisface una de las propiedades siguientes:

- i) $\sum_{k=1}^n \|a_k\|$ está uniformemente acotada con respecto a n ;
- ii) $\|a_k\| \leq \alpha_k$ donde α_k es el término general de una serie numérica convergente; entonces la serie $\sum_{k=1}^{\infty} a_k$ es absolutamente convergente.

La noción de base algebraica de un espacio vectorial introducida en el 6.1, puede ser considerada en un sentido más amplio en los espacios normados teniendo en cuenta la estructura topológica asociada a la norma. Para distinguir esta nueva noción del concepto de base algebraica de un espacio vectorial, hablaremos también de *base de un espacio normado*. Nosotros consideramos solamente aquellos espacios normados en que la base está compuesta por una cantidad de elementos a lo más numerable.

Definición 6.4.7 (base de un espacio normado). Una familia a lo más numerable de elementos $(x_n)_{n=1}^N$ ($N \leq \infty$) del espacio normado $(E, \|\cdot\|)$ se denomina *base de E* , si cada $x \in E$ puede ser representado en la forma única

$$x = \sum_{n=1}^N \lambda_n x_n. \quad (50)$$

donde (λ_n) es una sucesión numérica en K .

Si $N < \infty$ se dice que estamos en presencia de un espacio normado con base finita y si $N = \infty$, entonces a $(E, \|\cdot\|)$ se le llama *espacio normado con base numerable*.

Al número N se le llama dimensión topológica de $(E, \|\cdot\|)$.

La siguiente proposición nos muestra algunas propiedades de las bases de un espacio normado.

Proposición 6.4.8. Sea $(E, \|\cdot\|)$ un espacio normado, entonces se cumple:

- i) Los elementos de cualquier base de $(E, \|\cdot\|)$ son vectores linealmente independientes.
- ii) Toda base de $(E, \|\cdot\|)$ es una familia total (6.2.10) en $(E, \|\cdot\|)$.
- iii) Si $\{e_k\}_{k=1}^N$ es una base de $(E, \|\cdot\|)$, para cada M ($1 \leq M < N$) se tiene que

$$E = \overline{\{e_k\}_{k=1}^M} \oplus \overline{\{e_k\}_{k=M+1}^N} = E_1 \oplus E_2,$$

donde por $E_1 = \overline{\{e_k\}_{k=1}^M}$ ($E_2 = \overline{\{e_k\}_{k=M+1}^N}$) se denota la clausura del espacio vectorial generado por los vectores

$$\{e_k\}_{k=1}^M \text{ } (\{e_k\}_{k=M+1}^N).$$

Además, las familias de vectores $\{e_k\}_{k=1}^M$ y $\{e_k\}_{k=M+1}^N$ son bases de E_1 y E_2 respectivamente.

iv) Si el espacio normado $(E, \|\cdot\|)$ es de dimensión algebraica finita, toda base de E es también una base algebraica y viceversa.

v) Todas las bases de $(E, \|\cdot\|)$ tienen el mismo número de elementos.

Demostración:

i) Sean $\{e_k\}_{k=1}^N$ una base de $(E, \|\cdot\|)$ y

$$\theta = \sum_{k=1}^n \lambda_n e_n.$$

Como $\theta = \sum_{k=1}^N 0 e_n$, por la unicidad de la representación se tiene que $\lambda_k = 0$

para todo k , de donde se deduce que los vectores $\{e_k\}_{k=1}^N$ son linealmente independientes.

ii) Se deduce inmediatamente de la definición de serie y de familia total.

iii) En efecto, todo $x \in E$ se escribe en la forma única

$$x = \sum_{k=1}^N \lambda_k e_k = \sum_{k=1}^M \lambda_k e_k + \sum_{k=1}^N \lambda_k e_k = x_1 + x_2,$$

donde $1 \leq M \leq N$. (En realidad estamos considerando el caso $M < \infty$; el caso $M = N = \infty$ es similar aunque debe ser analizado un poco más cuidadosamente.)

Es claro que $x_1 \in E_1$, $x_2 \in E_2$. No resulta difícil probar que la suma de E_1 con E_2 es directa.

iv) Se obtiene directamente de la definición de base algebraica.

v) Sean $\{e_k\}_{k=1}^N$ y $\{b_j\}_{j=1}^M$ dos bases de $(E, \|\cdot\|)$. Supongamos por comodidad que ambas bases son a lo más numerables. Entonces si $M = N = \infty$ no hay contradicción con lo que queremos probar.

Supongamos por el absurdo que $M < N$, es decir $M \in \mathbb{N}$.

De aquí se deduce que E es un espacio vectorial de dimensión algebraica finita e igual a M . Por i) y la definición de base en E , habrían N vectores ($N > M$) linealmente independientes); contradicción. ■

Veamos que las propiedades expuestas en la proposición no pueden ser mejoradas. Para ello, analicemos los siguientes enunciados sin demostración. Un trabajo interesante para el lector es tratar de construir ejemplos que muestren la veracidad de estos enunciados:

- i) Existen conjuntos de vectores linealmente independientes en espacios normados y que no son bases.
- ii) Existen familias totales que no son bases.

En el próximo epígrafe veremos que en los espacios euclidianos existe una equivalencia entre las bases y las familias totales formadas por vectores ortogonales.

- iii) Si F es una familia total en $(E, \|\cdot\|)$ constituida por vectores linealmente independientes y si $\{F_1, F_2\}$ es una partición de F ;

$$F = F_1 \cup F_2, F_1 \cap F_2 = \emptyset,$$

entonces, en general, no tiene que cumplirse que

$$E = \overline{F_1} \oplus \overline{F_2}$$

(Explique ¿por qué?)

- iv) En algunos espacios normados de dimensión infinita, existen bases que no son bases algebraicas. Acerca de esto hablamos en el 6.1 cuando se introdujo la definición de base algebraica. Allí vimos que en el espacio normado $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ cualquier base algebraica tiene la potencia del continuo. Además, en el ejemplo 5 vimos que $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ tiene subfamilias totales numerables. Por el enunciado de ii) no podemos asegurar de lo anterior que $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ tiene base numerable como espacio normado. Este resultado, aunque cierto, es mucho más profundo y no entraremos aquí en su estudio.

Para culminar este epígrafe veamos la proposición siguiente:

Proposición 6.4.9. Todo espacio normado con base a lo más numerable es separable.

Demostración: Probemos que en un espacio normado $(E, \|\cdot\|)$ con base a lo más numerable $\{x_n\}_{n=1}^N$ ($N \leq \infty$) existe un subconjunto finito o numerable

denso. El caso $N < \infty$ es trivial ya que en ese caso $(E, \|\cdot\|)$ es un espacio normado de dimensión finita y por lo tanto es homeomorfo (6.2.13) a un espacio euclideo del tipo K^n ($K = \mathbb{R}$ o $\mathbb{C}$) el cual sabemos que es separable. Consideremos entonces que $N = \infty$ y definamos, a partir de la familia numerable $\{x_n\}_{n=1}^{\infty}$, el subconjunto A de E constituido por los elementos de la forma

$$\sum_{n=1}^p r_n x_n, \quad (51)$$

donde p es cualquier entero positivo y los coeficientes r_n son números racionales.

Mostremos que el conjunto A es numerable. En efecto, si denotamos por A_p al conjunto de todos los elementos de la forma (51) para p fijo, entonces, la potencia de A_p coincide con la de $\mathbb{Q}^p$, y por lo tanto, A_p es numerable. Por otra parte, no es difícil verificar que

$$A = \bigcup_{p=1}^{\infty} A_p,$$

de donde se deduce que A es numerable por ser la unión numerable de conjuntos numerables. Por la definición de base (6.4.7), todo $x \in E$ puede ser representado en la forma

$$x = \sum_{n=1}^{\infty} \lambda_n x_n = \lim_{p \rightarrow \infty} \sum_{n=1}^p \lambda_n x_n.$$

Para cada $p \in \mathbb{N}$ seleccionamos números racionales $(r_n^{(p)} (n = 1), \dots, p)$ tales que

$$\left| r_n^{(p)} - \lambda_n \right| < \frac{1}{p^2 \|x_n\|}$$

y supongamos

$$x^{(p)} = \sum_{n=1}^p r_n^{(p)} x_n,$$

entonces,

$$x^{(p)} \in A_p \subset A \text{ y}$$

$$\left\| x^{(p)} - \sum_{n=1}^p \lambda_n x_n \right\| = \left\| \sum_{n=1}^p (r_n^{(p)} - \lambda_n) x_n \right\| \leq p \cdot \frac{1}{p^2} = \frac{1}{p}.$$

Luego,

$$\begin{aligned} \|x^{(p)} - x\| &\leq \left\| x^{(p)} - \sum_{n=1}^p \lambda_n x_n \right\| + \left\| \sum_{n=1}^p \lambda_n x_n - x \right\| \leq \\ &\leq \frac{1}{p} + \left\| \sum_{n=1}^p \lambda_n x_n - x \right\| \xrightarrow{p \rightarrow \infty} 0, \end{aligned}$$

de donde se tiene que $x = \lim(x^{(p)})$, lo cual prueba que A es denso en E , y por lo tanto, E es separable. ■

Ejemplo 15 (espacios normados con base numerable). En los espacios $l^1(\mathbb{N})$ y $l^2(\mathbb{N})$ una base numerable está formada por una sucesión de vectores unitarios. En efecto si $x = (\xi_1, \xi_2, \dots, \xi_n, \dots) \in l^1(\mathbb{N})$ y si $e_n(0, \dots, 0, 1, 0, \dots)$, donde el 1 está ubicado en el lugar n -ésimo, son los vectores unitarios, entonces considerando

$$x^{(k)} = \sum_{n=1}^k \xi_n e_n = (\xi_1, \xi_2, \dots, \xi_k, 0, \dots, 0, \dots)$$

se tiene que

$$\|x^{(k)} - x\| = \sum_{n=k+1}^{\infty} |\xi_n| \longrightarrow 0.$$

Luego $x = \sum_{n=1}^{\infty} \xi_n e_n$. Inversamente, si para x se tiene una representación

del tipo $x = \sum_{n=1}^{\infty} \lambda_n e_n$, es fácil ver que $\lambda_n = \xi_n$ para todo n . De aquí concluimos

que la colección de vectores unitarios $\{e_n\}$ es una base numerable de $l^1(\mathbb{N})$. Un razonamiento similar es también válido para $l^2(\mathbb{N})$. Sin embargo, estos vectores unitarios no constituyen una base en el espacio $l^\infty(\mathbb{N})$. En efecto, si esto ocurriera, entonces cada $x \in l^\infty(\mathbb{N})$ podría escribirse en la forma

$$\sum_{n=1}^{\infty} \xi_n e_n. \quad (52)$$

Definiendo $x^{(k)}$ como en el caso anterior tendríamos

$$\|x^{(k)} - x\|_\infty = \sup_{n>k} |\xi_n|$$

pero esta última expresión no tiene necesariamente que tender a cero cuando $k \longrightarrow \infty$, es decir, la serie (52) no converge necesariamente con respecto a la norma de $l^\infty(\mathbb{N})$. En el capítulo 2 habíamos visto que $l^\infty(\mathbb{N})$ no es separable, luego por 6.4.9 no puede tener ninguna base numerable.

Nosotros no consideraremos en este epígrafe el problema de la construcción de bases en $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ pues este problema es mucho más complicado

aunque ya hemos indicado que este espacio normado tiene base numerable. No obstante, en los próximos epígrafes entraremos con más detalle en el estudio de las bases en el espacio $C([a, b], \mathbb{R})$ provisto de las normas $\|\cdot\|_1$ y $\|\cdot\|_2$ definidas en (32) y (33).

Ejercicios de Autocontrol

1. Demuestre que en un espacio normado toda serie absolutamente convergente converge siempre al mismo límite para cualquier reordenamiento de sus términos.
2. Demuestre que un espacio normado es de Banach si y solo si, cuando cada serie absolutamente convergente es también convergente.

6.5 ESPACIOS EUCLIDEANOS

En este epígrafe estudiaremos espacios normados en los cuales aparecen determinados conceptos geométricos que al conjugarse con la estructura vectorial y la estructura métrica, nos permiten obtener resultados más fuertes que los encontrados en la teoría expuesta hasta el momento. Por ejemplo, en los espacios euclidianos:

- a) se mejoran los resultados sobre mejor aproximación desarrollados en el epígrafe anterior;
- b) se establece una relación entre las bases y las familias totales mediante las llamadas series de Fourier.

El concepto adicional que introduciremos y que además es compatible, en un sentido que precisaremos más adelante, con la estructura de espacio normado es el de *ortogonalidad*.

Precisamente a partir de este nuevo concepto se puede generalizar a los espacios euclidianos el conocido resultado de los cursos elementales de Geometría que nos dice que la menor distancia de un punto a una recta o a un plano, está dada por la longitud del segmento que parte de dicho punto y corta perpendicularmente a la recta o al plano. Veremos que este principio básico de la teoría geométrica de los espacios euclidianos, es la herramienta más poderosa para el desarrollo de la teoría de optimización (mejor aproximación) en espacios euclidianos.

Definición 6.5.1. Un espacio euclidiano es un espacio vectorial H sobre el cuerpo $K(\mathbb{R} \text{ o } \mathbb{C})$ provisto de un *producto escalar* definido sobre $H \times H$. Para

cada par de vectores $x, y \in H$ el producto escalar de x por y es un escalar que denotaremos mediante $\langle x, y \rangle$:

$$\langle, \rangle : \begin{matrix} H \times H \\ (x, y) \rightsquigarrow \langle x, y \rangle \end{matrix} \longrightarrow K \quad (53)$$

El producto escalar satisface los axiomas siguientes:

$$(\mathbf{PE1}) \quad \langle x, y \rangle = \overline{\langle y, x \rangle}$$

$$(\mathbf{PE2}) \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$$

$$(\mathbf{PE3}) \quad \langle \lambda x, y \rangle = \lambda \langle x, y \rangle$$

$$(\mathbf{PE4}) \quad \langle x, x \rangle \geq 0 \text{ y } \langle x, x \rangle = 0 \text{ si y solo si } x = \theta.$$

La barra sobre el miembro derecho del axioma **(PE1)** denota el paso al complejo conjugado. El espacio euclideo se llama real si $K = \mathbb{R}$ y complejo si $K = \mathbb{C}$.

Observación: En la literatura matemática se llama también con frecuencia a los espacios euclideos espacios pre-hilbertianos. La justificación de este nombre quedará a-clarada más adelante.

En los espacios euclideos reales, el axioma **(PE1)** se reduce a:

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

ya que $\langle x, y \rangle$ es siempre un número real. En general para espacios euclideos complejos el axioma **(PE1)** nos garantiza que $\langle x, x \rangle$ es real para todo $x \in H$. De **(PE2)** y **(PE3)** se deduce la linealidad del producto escalar con respecto al primer argumento. Si conjugamos estos dos axiomas con **(PE1)**, tendremos para el segundo argumento:

$$\langle x, \lambda y + \mu z \rangle = \overline{\lambda} \langle x, y \rangle + \overline{\mu} \langle x, z \rangle. \quad (54)$$

A partir de ahora, al número $\sqrt{\langle x, x \rangle}$ lo denotaremos por $\|x\|$. Nuestro primer objetivo es verificar que $\|x\|$ define efectivamente una norma sobre H . Los axiomas **(PE1)** y **(PE3)** nos dan (N2) mientras que de **(PE4)** se deduce (N1).

La demostración de la desigualdad triangular (N3) es un poco más complicado. Para ello es necesario probar un importante lema que es fundamental a lo largo de este epígrafe.

Lema 6.5.2 (desigualdad de Cuchy-Buniakowsky). Para todo par de vectores x, y en un espacio euclideo se cumple

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \quad (55)$$

La igualdad ocurre si y solo si

$$x = \lambda y \text{ ó } y = \theta.$$

Demostración: Si $y = \theta$ la desigualdad ocurre trivialmente. Supongamos entonces que $y \neq 0$. Para todo escalar $\lambda \in K$, se tiene

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \lambda \langle y, x \rangle - \bar{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle$$

En particular, para $\lambda = \langle x, x \rangle / \langle y, y \rangle$, tendremos

$$0 \leq \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}$$

lo que equivale a

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle} = \|x\| \|y\|.$$

Se deja de ejercicio al lector comprobar la segunda parte del lema. ■

Proposición 6.5.3. En un espacio euclideo $(H, \langle, \rangle)$ la función:

$$\|x\| = \sqrt{\langle x, x \rangle} \text{ es una norma.}$$

Demostración: Queda solamente por establecer la desigualdad triangular (N_3) . Para cualesquiera $x, y \in H$, tenemos

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \\ &\leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2. \end{aligned}$$

Por la desigualdad de Cauchy-Buniakovsky se deduce que

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Extrayendo la raíz cuadrada a ambos lados se obtiene el resultado deseado. ■

Observación: De ahora en adelante llamremos también espacio euclideo al espacio normado $(E, \|\cdot\|)$ donde la norma puede ser obtenida a partir de un producto escalar $\langle, \rangle$ poniendo $\|x\| = \sqrt{\langle x, x \rangle}$

Si se introduce un producto escalar convenientemente puede mostrarse que muchos de los espacios normados considerados hasta el momento son espacios euclideos en el sentido de la observación anterior.

Ejemplo 16. El espacio $\mathbb{R}^n$ de los n -uplos de número reales, provisto de la norma $\|\cdot\|_2$ (29) es un espacio euclideo. Basta notar que para los vectores $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$ de $\mathbb{R}^n$, la aplicación de $\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ definida por:

$$\langle x, y \rangle = \sum_{i=1}^n \xi_i \eta_i, \tag{56}$$

es un producto escalar real (compruébelo), y además, $\|x_2\| = \sqrt{\langle x, x \rangle}$.

Sabemos que las normas $\|\cdot\|_2$, $\|\cdot\|_1$ y $\|\cdot\|_\infty$ definidas por (25)–(27), proveen a $\mathbb{R}^n$ de la misma estructura topológica (los mismos abiertos) y por tanto, de la misma convergencia y además por 6.2.13 ellas son también, en cierto sentido, equivalentes desde el punto de vista métrico. Sin embargo, veremos más adelante que de ellas solamente $\|\cdot\|_2$ provee a $\mathbb{R}^n$ de estructura de espacio euclideo.

Esta observación resulta fundamental para el estudio de la estructura geométrica de los espacios $\mathbb{R}^n$.

Ejemplo 17. El espacio real $l_2(\mathbb{N})$ se convierte en un espacio euclideo para el producto escalar de los vectores

$x = (\xi_1, \xi_2, \dots, \xi_n \dots)$, $y = (\eta_1, \eta_2, \dots, \eta_n \dots)$ definido por

$$\langle x, y \rangle_2 = \sum_{i=1}^{\infty} \xi_i \eta_i. \quad (57)$$

La desigualdad de Hölder para $l^2(\mathbb{N})$ (ver capítulo 2) que en este caso coincide con la desigualdad de Cauchy-Schwartz correspondiente a (57), garantiza que $|\langle x, y \rangle| \leq \|x\| \|y\|$, y por lo tanto, el producto escalar es siempre finito. La norma definida por $\sqrt{\langle x, x \rangle}$ es la norma usual de $l^2(\mathbb{N})$ (29).

¿Cómo se debe definir el producto escalar en $l^2(\mathbb{N})$ para el caso complejo?

Ejemplo 18. El espacio real $C([a, b], \mathbb{R})$ es un espacio euclideo para el producto escalar.

$$\langle x, y \rangle = \int_a^b f(t) g(t) dt. \quad (58)$$

De nuevo la desigualdad de Hölder nos garantiza que $\langle x, y \rangle$ es finito. Obviamente la norma asociada a este producto escalar es $\|\cdot\|_2$ (33).

Veamos que, en efecto, toda norma no proviene necesariamente de un producto escalar.

Proposición 6.5.4 (identidad del paralelogramo). Un espacio normado $(E, \|\cdot\|)$ es euclideo si y solo si se cumple la identidad del paralelogramo: (fig. 93).

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \quad (59)$$

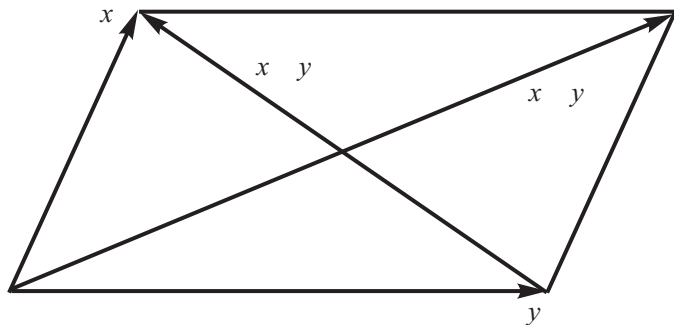


Figura 93

Demostración: Si E es euclideo, (59) se comprueba directamente desarrollando los axiomas del producto escalar. Recíprocamente, si se cumple (59) basta considerar

$$4\langle x, y \rangle = \|x + y\|^2 - \|x - y\|^2 +$$

$$+ i\|x + iy\|^2 - i\|x - iy\|^2$$

en el caso complejo, y

$$4\langle x, y \rangle = \|x + y\|^2 - \|x - y\|^2$$

en el caso real, y comprobar que $\langle, \rangle$ así definido es un producto escalar en E tal que $\|x\| = \sqrt{\langle x, x \rangle}$. Esto último se deja de ejercicio al lector. ■

Ejemplo 19 (espacio normado no euclideo). El espacio normado:

$C([a, b], \mathbb{R})$ provisto de la norma $\|\cdot\|_\infty$ no es euclideo. En efecto, tomando $f(x) = \sin^2 x$ y $g(x) = \cos^2 x$ se tiene que:

$$\|f + g\| = \|f - g\|_\infty = \|f\|_\infty = \|g\|_\infty = 1,$$

y por lo tanto, no se cumple la identidad del paralelogramo.

Veamos la siguiente propiedad del producto escalar.

Proposición 6.5.5 (continuidad del producto escalar). Supongamos que $x_n \rightarrow x$ y $y_n \rightarrow y$ en un espacio euclideo.

Entonces, $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$.

Observación: De ahora en adelante cuando hablemos de conceptos topológicos o métricos en un espacio euclideo nos estaremos refiriendo a los correspondientes a la norma asociada:

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

Demostración: Como la sucesión (x_n) es convergente, ella es acotada. Consideremos que $\|x_n\| \leq M$ para todo n . Por otra parte,

$$\begin{aligned}
 |\langle x_n, y_n \rangle - \langle x, y \rangle| &\leq |\langle x_n, y_n \rangle - \langle x_n, y \rangle| + |\langle x_n, y \rangle - \langle x, y \rangle| = \\
 &= |\langle x_n, y_n - y \rangle| + |\langle x_n - x, y \rangle|.
 \end{aligned}$$

Aplicando la desigualdad de Cauchy-Schwartz, obtenemos

$$|\langle x_n, y_n \rangle - \langle x, y \rangle| \leq \|x_n\| \|y_n - y\| + \|x_n - x\| \|y\|.$$

Como $\|x_n\|$ está acotado,

$$|\langle x_n, y_n \rangle - \langle x, y \rangle| \leq M \|y_n - y\| + \|x_n - x\| \|y\| \longrightarrow 0. \blacksquare$$

Definiremos a continuación la noción de ortogonalidad y veremos algunas de sus propiedades.

Definición 6.5.6. Dos vectores x y y de un espacio euclideo $(H, \langle, \rangle)$ se llaman *ortogonales* si $\langle x, y \rangle = 0$ (Notación: $x \perp y$). Dado un subconjunto S de H , el conjunto de todos los vectores x de H que son ortogonales a cada uno de los vectores de S se llama *complemento ortogonal* de S y se denota por $S^\perp$:

$$S^\perp = \{x \in H : y \in S \Rightarrow \langle x, y \rangle = 0\} \quad (60)$$

Evidentemente, el complemento ortogonal del conjunto unitario constituido por el vector nulo θ de H , es todo el espacio H . Además, para cualquier subconjunto S de H , $S^\perp$ es un subespacio vectorial cerrado. En efecto, $S^\perp$ es un subespacio vectorial ya que cualquier combinación lineal de vectores ortogonales a un conjunto es también ortogonal a dicho conjunto. Por otra parte, si (x_n) es un sucesión de $S^\perp$ convergente; $(x_n) \longrightarrow x$, entonces para todo $y \in S$ se tiene que $\langle x_n, y \rangle = 0$ y por la continuidad de producto escalar (6.5.5) $0 = \langle x_n, y \rangle \longrightarrow \langle x, y \rangle$, de donde se concluye que $x \in S^\perp$.

Proposición 6.5.7 (propiedades del complemento ortogonal). Sean S y T subconjuntos de un espacio euclideo $(H, \langle, \rangle)$; entonces:

- i) $S^\perp$ es un subespacio vectorial cerrado de H : $S^\perp = \overline{S^\perp}$.
- ii) $S^\perp = \overline{S}^\perp$
- iii) $S^\perp = [S]^\perp$.
- iv) $S \cap [S]^\perp$ es $\emptyset$ o igual al elemento nulo de H .
- v) $S \subset S^{\perp\perp}$.
- vi) Si $S \subset T$ entonces $T^\perp \subset S^\perp$.
- vii) $S^\perp = S^{\perp\perp\perp}$.

viii) $S^\perp = H \Leftrightarrow S = \{\theta\}$.

ix) Si S es una familia total en H , entonces $S^\perp = \{\theta\}$.

Demostración:

i) Fue probado en el comentario antes de enunciar la proposición: ii) y iii). Se deducen inmediatamente de i).

iv) Si $S \cap S^\perp$ es diferente del vacío y $x \in S \cap S^\perp$, entonces $\langle x, x \rangle = 0$; luego, por **(PE4)** $x = \theta$.

v) Si $x \in S$, entonces, $x \perp y$ para todo $y \in S^\perp$; luego $x \in S^{\perp\perp}$.

vi) Si $y \in T^\perp$, entonces, $y \perp x$ para todo $x \in S$ ya que $S \subset T$. Por lo tanto $y \in S^\perp$.

vii) De v) se tiene que $S^\perp \subset S^{\perp\perp\perp}$. Además, $S \subset S^{\perp\perp}$, lo cual por vi) implica que $S^{\perp\perp\perp} \subset S^\perp$. De aquí que $S^{\perp\perp\perp} = S^\perp$.

viii) $\Leftarrow$ es evidente.

$\Rightarrow$ Si $S^\perp = H$, entonces, $S^{\perp\perp} = H^\perp$.

Es fácil comprobar que $H^\perp = \{\theta\}$, luego de v) se deduce que $S = \{\theta\}$.

ix) Sea S una familia total en H . Por ii) y iii), $S^\perp = [S]^\perp = [\overline{S}]^\perp = H^\perp = \{\theta\}$. ■

Veamos a continuación dos problemas importantes que tienen relación con la noción de ortogonalidad en los espacios euclidianos:

I. Las propiedades de los conjuntos ortonormales numerables;

II. El problema de la mejor aproximación.

Para analizar el problema I, comencemos por el resultado siguiente:

Proposición 6.5.8 (teorema generalizado de Pitágoras). En un espacio euclideo $(H, \langle, \rangle)$ se cumplen los enunciados siguientes:

i) Si $x \perp y$, entonces $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

ii) Si $x = \sum_{n=1}^{\infty} x_n$ es una serie convergente de vectores ortogonales dos a dos, entonces, $\|x\|^2 = \sum_{n=1}^{\infty} \|x_n\|^2$.

(Luego una condición necesaria para la convergencia de una serie de vectores ortogonales es la convergencia de la serie numérica (61)).

Demostración: Basta probar ii). Por la distributividad del producto escalar tenemos:

$$\|x\|^2 = \langle x, y \rangle = \left\langle x, \sum_{n=1}^{\infty} x_n \right\rangle = \sum_{n=1}^{\infty} \langle x, x_n \rangle =$$

$$= \sum_{n=1}^{\infty} \left\langle \sum_{m=1}^{\infty} x_m, x_n \right\rangle = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle x_m, x_n \rangle.$$

Pero $\langle x_m, x_n \rangle = 0$ para $m \neq n$, y por lo tanto,

$$\|x\|^2 = \sum_{n=1}^{\infty} \langle x_n, x_n \rangle = \sum_{n=1}^{\infty} \|x_n\|^2. \blacksquare$$

Dada la importancia de los sistemas ortogonales numerables, es decir, aquellas familias numerables de puntos de un espacio euclideo cuyos elementos son ortogonales entre sí, realizaremos un estudio más detallado de ellos. Más adelante veremos aplicaciones de estos sistemas a las llamadas series de Fourier.

Veamos primeramente la definición siguiente:

Definición 6.5.9 (sistema ortogonal y sistema ortonormal). Un sistema de vectores $S = \{x_\alpha\}_{\alpha \in I}$ en un espacio euclideo se llama ortogonal si cada par de vectores diferentes de S son ortogonales:

$$\alpha \neq \beta \Rightarrow \langle x_\alpha, x_\beta \rangle = 0. \quad (62)$$

El sistema S se llama ortonormal si es ortogonal y además

$$\langle x_\alpha, x_\alpha \rangle = \|x_\alpha\|^2 = 1, \text{ para todo } \alpha \in I. \quad (63)$$

No es difícil comprobar que el sistema S es ortonormal si y solo si, para todo $\alpha, \beta \in I$:

$$\langle x_\alpha, x_\beta \rangle = \delta_{\alpha, \beta}, \quad (64)$$

donde

$$\delta_{\alpha,\beta} = \begin{cases} 0, & \text{si } \alpha \neq \beta \\ 1, & \text{si } \alpha = \beta \end{cases} \quad (65)$$

es la delta de Kronecker.

Proposición 6.5.10. Todo sistema ortogonal de puntos en un espacio euclideo separable es a lo sumo numerable.

Demostración: En efecto, sin perder generalidad, podemos admitir que el sistema $\{x_\alpha\}_{\alpha \in I}$ no solo es ortogonal sino también ortonormal en H de lo contrario, podríamos tomar $\left\{ \frac{x_\alpha}{\|x_\alpha\|} \right\}$ con

$$\|x_\alpha - x_\beta\| = \sqrt{2} \text{ para todo } \alpha \neq \beta.$$

Consideremos la familia de las bolas $B(x_\alpha, 1/2)$. Evidentemente estas bolas no se intersectan. Si H es separable, entonces existirá en H un conjunto $\{\psi_n\}$ numerable denso.

Luego, cada bola $B(x_\alpha, 1/2)$ debe contener al menos un elemento ψ_n , y por lo tanto, el número de estas bolas (y, por consiguiente, el número de elementos $\{x_\alpha\}$ es a lo sumo numerable. ■

Ejemplo 20 (espacio euclideo no separable). Hasta el momento hemos visto solamente ejemplos de espacios euclideos que son separables. Un ejemplo de espacio euclideo no separable, puede construirse de la manera siguiente:

Consideremos sobre el segmento $[0, 1]$ todas las funciones reales $x(t)$ que son diferentes de cero, en un conjunto a lo sumo numerable y la suma $\sum x^2(t_i)$, tomada respecto a estos puntos, es finita.

Definamos en este espacio las operaciones de adición y multiplicación por números como las habituales adición y multiplicación de funciones y definamos el producto escalar de x por y , de la manera siguiente:

$$\langle x, y \rangle = \sum x(t_i) y(t_i)$$

donde la suma se toma según el conjunto de puntos t_i tales que $x(t_i) y(t_i) \neq 0$.

No resulta difícil demostrar que en este espacio no existe ningún subconjunto denso numerable. Esto se deja de ejercicio al lector.

En nuestra exposición no nos hemos referido a bases ortogonales; demostremos el siguiente teorema general, análogo al teorema de existencia de base ortogonal en los espacios euclideos de dimensión finita. Para simplificar la demostración consideremos que el espacio H es separable.

Teorema 6.5.11 (proceso de ortonormalización de Gram-Shmidt).

Sea $(H, \langle, \rangle)$ un espacio euclideo complejo separable y $\{x_n\}$ un sistema de vectores linealmente independientes y a lo sumo numerable de vectores de H . Entonces existe en H un sistema de vectores $\{\varphi_n\}$ que satisface las condiciones siguientes:

- i) Los sistemas $\{\varphi_n\}$ y $\{x_n\}$ tienen la misma cantidad de elementos.
- ii) El sistema $\{\varphi_n\}$ es ortonormal.
- iii) Todo elemento φ_n es una combinación lineal de los elementos $\{x_n\}$ de la forma:

$$\varphi_n = a_{n1}x_1 + \dots + a_{nn}x_n, \text{ con } a_{nn} \neq 0.$$

- iv) Todo elemento x_n se representa en la forma

$$x_n = b_{n1}\varphi_1 + \dots + b_{nn}\varphi_n,$$

donde $b_{nn} \neq 0$

Además, todo elemento del sistema $\{\varphi_n\}$ queda determinado unívocamente por las condiciones ii), iii) y iv), salvo un factor de módulo 1.

Observación: Obviamente cualesquiera de las condiciones iii) o iv) implica que

$$[\{x_n\}] = [\{\varphi_n\}].$$

Demostración: Obviamente la propiedad i) se deduce de iii) y iv). Representemos el elemento φ_1 en la forma

$$\varphi_1 = a_{11}x_1;$$

entonces, a_{11} se determina por la condición

$$\langle \varphi_1, \varphi_1 \rangle = a_{11}^2 \langle x_1, x_1 \rangle = 1,$$

de donde

$$a_{11} = \frac{1}{b_{11}} \frac{\pm 1}{\sqrt{\langle x_1, x_1 \rangle}}.$$

Está claro que φ_1 está unívocamente determinado (salvo un factor de módulo 1). Supongamos que ya han sido construidos los elementos φ_k con $k < n$ que verifican las condiciones ii), iii) y iv). En este caso, se puede representar x_n en la forma

$$x_n = b_{n1}\varphi_1 + \dots + b_{n,n-1}\varphi_{n-1} + h_n \text{ donde}$$

$$\langle h_n, \varphi_n \rangle = 0, \quad k < n.$$

En efecto, los coeficientes correspondientes b_{nk} y, por consiguiente, el elemento h_n quedan unívocamente determinados por las condiciones

$$\begin{aligned}\langle h_n, \varphi_n \rangle &= \langle x_n - b_{n1}\varphi_1 - \dots - b_{n,n-1}\varphi_{n-1}, \varphi_k \rangle = \\ &= \langle x_n, \varphi_k \rangle - b_{nk} \langle \varphi_k, \varphi_k \rangle = 0.\end{aligned}$$

Es evidente que $\langle h_n, h_n \rangle \neq 0$, ya que la suposición $\langle h_n, h_n \rangle = 0$ estaría en contradicción con la independencia lineal del sistema $\{x_n\}$. Supongamos

$$\varphi_n = \frac{h_n}{\sqrt{\langle h_n, h_n \rangle}}.$$

De esta construcción inductiva se desprende que h_n y por consiguiente, también φ_n , se expresan mediante $x_1, \dots, x_n$, es decir,

$$\varphi_n = a_{n1}x_1 + \dots + a_{nn}x_n, \text{ donde}$$

$$a_{nn} = \frac{1}{\sqrt{\langle h_n, h_n \rangle}} \neq 0. \text{ Además,}$$

$$\langle \varphi_n, \varphi_n \rangle = 1, \langle \varphi_n, \varphi_k \rangle = 0 \quad (k < n),$$

$$\text{y } x_n = b_{n1} + \dots + b_{nn}\varphi_n, \text{ siendo}$$

$$b_{nn} = \sqrt{\langle h_n, h_n \rangle} \neq 0,$$

es decir, $\{\varphi_n\}$ verifica las condiciones planteadas en el teorema. ■

Ejemplo 21 (polinomios ortogonales). Consideremos el espacio euclideo

$$C([-1, 1], \mathbb{R}) \text{ con el producto escalar } \langle f, g \rangle_2 = \int_{-1}^1 f(t) g(t) dt.$$

Las funciones linealmente independientes $1, t, t^2, \dots, t^n, \dots$, generan el subespacio de los polinomios. El procedimiento de Gram-Schmidt puede ser aplicado a la sucesión $1, t, t^2, \dots$, y obtendremos una sucesión ortonormal $\{e_n\}_{n=0}^\infty$. Obviamente los e_n son polinomios por ser combinaciones lineales de polinomios. Puede demostrarse que $e_n(t) = \sqrt{\frac{2n+1}{2}} P_n(t)$

donde los $P_n(t)$ son los conocidos polinomios de Legendre

$$P_n(t) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dt^n} \{(1-t^2)^n\}.$$

Del teorema de Weierstrass sobre la aproximación uniforme de toda función continua por polinomios y de la desigualdad

$$\int_{-1}^1 |f(t)|^2 dt \leq 2 \max_{-1 \leq t \leq 1} |f(t)|^2,$$

la cual se satisface para toda función continua, se deduce que la familia $\{1, t, t^2, \dots, t^n, \dots\}$ es total en $(C([-1, 1], \mathbb{R}), \langle \cdot, \cdot \rangle_2)$. Por la observación que aparece después del teorema 6.5.11 concluimos que la sucesión $\{e_n\}$ constituye un sistema ortonormal total en el espacio antes mencionado. En el próximo epígrafe veremos otra demostración de este hecho.

Al ortonormalizar el sistema $\{1, t, t^2, \dots, t^n, \dots\}$ con respecto a otros productos escalares definidos sobre $C([-1, 1], \mathbb{R})$, por ejemplo

$$\langle f, g \rangle = \int_{-1}^1 f(t) g(t) (1-t)^\alpha (1+t)^\beta dt$$

con $\alpha, \beta \geq 1$, se obtienen nuevos sistemas de polinomios ortonormales clásicos conocidos con el nombre de polinomios de Jacobi de tipo (α, β) . En este caso, para definir el producto escalar se ha utilizado una *función de peso* del tipo $(1-t)^\alpha (1+t)^\beta$. Utilizando otras funciones de peso, como por ejemplo e^{-x} en el intervalo $[0, \infty]$ en lugar de $[-1, 1]$, se obtienen los llamados polinomios de Legendre–Hermite.

Los polinomios ortogonales han sido objeto de investigación de la teoría clásica de aproximación.

Tal y como se había expresado anteriormente, estudiaremos la relación entre las bases y las familias totales constituidas por vectores ortonormales.

Para ello, se hace necesario introducir la definición siguiente:

Definición 6.5.11. Sea $(H, \langle, \rangle)$ un espacio euclideo separable y $\{e_n\}$ un sistema ortonormal en H .

Para cada $x \in H$, se llama *n-ésimo coeficiente de Fourier* de x con respecto al sistema $\{e_n\}$, al número

$$a_n = \langle x, e_n \rangle, \quad n = 0, 1, \dots \quad (66)$$

y a la serie (por ahora formal)

$$\sum_n a_n e_n, \quad (67)$$

se llama *serie de Fourier* del elemento x según el sistema ortonormal $\{e_n\}$.

Ejemplo 22. Consideremos el espacio euclideo $C([- \pi, \pi], \mathbb{R})$ con el producto escalar definido en el ejemplo 21 (sustituyendo $[-1, 1]$ por $[- \pi, \pi]$).

Las funciones

$$e_k(t) = \begin{cases} \frac{1}{\sqrt{2\pi}}, & \text{si } k = 0 \\ \frac{1}{\sqrt{\pi}} \sin kt, & \text{si } k = 1, 2, \dots \\ \frac{1}{\sqrt{\pi}} \cos kt, & \text{si } k = 1, 2, \dots \end{cases} \quad (68)$$

constituyen un sistema ortonormal en el espacio antes mencionado, lo cual puede comprobar el lector mediante un cálculo directo.

De manera análoga a como fue probado para los polinomios de Legendre en el ejemplo 21, puede comprobarse que los polinomios trigonométricos $\{e_k(t)\}$ constituyen un sistema ortonormal total en $C([-\pi, \pi], \mathbb{R})$.

El sistema (68) aparece en la teoría clásica de las series de Fourier.

Como consecuencia del teorema 6.5.14, que veremos posteriormente, se tiene que cada $f \in C([-\pi, \pi], \mathbb{R})$ tiene un desarrollo en serie de Fourier del tipo

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \operatorname{sen} nt), \quad (69)$$

donde

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) dt, \quad (70)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos ntdt \quad (n = 1, 2, \dots) \text{ y} \quad (71)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \operatorname{sen} ntdt \quad (n = 1, 2, \dots). \quad (72)$$

La serie (69) converge en el sentido de la norma $\|\cdot\|_2$ definida en (33). El estudio de las propiedades de la serie de Fourier y de su convergencia a f en diferentes sentidos (puntual, uniforme, etcétera), es uno de los objetivos de la teoría clásica de las series de Fourier. Implícitamente en la demostración del teorema de Weierstrass en el capítulo 2, ya hemos utilizado las series de Fourier. Allí vimos que si llamamos $S_n(t)$ a la suma particular n -ésima de la serie (69) y definimos las sumas de Cesaro como el promedio

$$\sigma_n(t) = \frac{S_0(t) + S_1(t) + \dots + S_n(t)}{n+1} \quad (73)$$

entonces $\sigma_n(t)$ converge uniformemente a $f(t)$ si $[-\pi, \pi]$ es además periódica.

Lógicamente, después de la definición anterior, nos surgen las siguientes preguntas: ¿es convergente la serie (67)?, es decir, ¿tiende a algún límite la sucesión de sus sumas parciales en el sentido de la métrica del espacio H ? y si converge, ¿coincide su suma con el elemento inicial x ?

Para responder a estas preguntas veamos las dos proposiciones siguientes y el teorema 6.5.14.

Proposición 6.5.12. Sean $(H, \langle \cdot, \cdot \rangle)$ un espacio euclideo real separable, $\{e_k\}$ un sistema ortonormal en H y $f \in H$; entonces para cada $n \in \mathbb{N}$, el elemento de la forma

$$\tilde{S}_k = \sum_{k=1}^n a_k e_k : (a_k \in \mathbb{R})$$

que se encuentra a menor distancia de f , coincide con la suma parcial de orden n de la serie de Fourier de f (es decir, la distancia mínima de f a $\tilde{S}_k$ se obtiene tomando en calidad de a_k , los coeficientes de Fourier de f).

Observación: Nosotros no nos dedicaremos a estudiar aquí el problema general de la distancia mínima a un espacio vectorial generado por un número finito arbitrario (no necesariamente ortonormal) de vectores. A este problema se dedica la teoría de las llamadas matrices de Gram y su base se encuentra formulada en el teorema 6.5.16.

Demostración: En efecto, calculemos la distancia entre f y una suma del tipo S_n . Como el sistema $\{e_k\}$ es ortonormal, tendremos

$$\|f - \tilde{S}_n\|^2 = \left\langle f - \sum_{k=1}^n a_k e_k, f - \sum_{k=1}^n a_k e_k \right\rangle =$$

$$= \langle f, f \rangle - 2 \left\langle f, \sum_{k=1}^n a_k e_k \right\rangle +$$

$$+ \left\langle \sum_{k=1}^n a_k e_k, \sum_{j=1}^n a_j e_j \right\rangle =$$

$$= \|f\|^2 - 2 \sum_{k=1}^n a_k f_k + \sum_{k=1}^n a_k^2 =$$

$$= \|f\|^2 - \sum_{k=1}^n f_k^2 + \sum_{k=1}^n (a_k + f_k)^2$$

donde $f_k = \langle f, e_k \rangle$ son los coeficientes de Fourier del elemento f con respecto al sistema $\{e_k\}$.

Está claro que esta expresión alcanza su mínimo cuando el último sumando es igual a cero, es decir, para

$$a_k = f_k, \quad k = 1, \dots, n. \quad \blacksquare \quad (74)$$

De la demostración de la proposición anterior se deduce que si tomamos $a_k = f_k$, entonces,

$$\|f - S_n\|^2 = \|f\|^2 - \sum_{k=1}^n f_k^2.$$

Proposición 6.5.13 (desigualdad de Bessel). Sean $(H, \langle, \rangle)$ y $\{e_k\}$ como en la proposición 6.5.12; entonces para toda $f \in H$ se cumple

$$\sum_k f_k^2 \leq \|f\|^2 \quad (76)$$

donde $f_k = \langle f, e_k \rangle$, $k = 1, 2, \dots$, son los coeficientes de Fourier de f con respecto al sistema $\{e_k\}$.

Demostración: De (75) se deduce que para todo n

$$\sum_{k=1}^n f_k^2 \leq \|f\|^2,$$

luego (76) se satisface inmediatamente si $\{e_k\}$ es una familia finita y por paso al límite si es numerable. ■

Podemos ahora enunciar un teorema que nos explica la relación existente entre las bases y las familias totales compuestas por vectores ortonormales. (Compárese con 6.4.10).

Teorema 6.5.14. Sean $(H, \langle, \rangle)$ y $\{e_k\}$ como en la proposición 6.5.12.

Entonces son equivalentes:

i) Para todo $f \in H$ se cumple la identidad de Parseval

$$\sum_k f_k^2 = \|f\|^2 \quad (77)$$

donde f_k son los coeficientes de Fourier de f con respecto al sistema $\{e_k\}$.

ii) Toda $f \in H$ se puede representar mediante su serie de Fourier, es decir, para cada $f \in H$ la serie de Fourier correspondiente a f converge precisamente a f .

iii) El sistema $\{e_k\}$ es una base en H ;

iv) El sistema $\{e_k\}$ es total en H .

Demostración:

i) $\Rightarrow$ ii) se deduce inmediatamente de (75).

ii) $\Rightarrow$ iii) es evidente.

iii) $\Rightarrow$ iv) ya que toda base es siempre una familia total.

vi) $\Rightarrow$ i) Supongamos que el sistema $\{e_k\}$ es total, es decir, todo elemento $f \in H$ se puede aproximar con precisión arbitraria mediante combinaciones lineales del tipo

$$\sum_{k=1}^n a_k e_k.$$

De 6.5.12 sabemos que la suma parcial S_n de la serie de Fourier de f , nos da una aproximación mejor a f . Por consiguiente, la serie de Fourier de f converge a f y de (75) se deduce que tiene lugar la identidad de Parseval. ■

Como consecuencia del teorema anterior obtenemos el siguiente corolario:

Corolario 1 (existencia de bases ortonormales). En todo espacio euclideo separable existe siempre una base ortonormal.

Demostración: En efecto, si $\{\varphi_n\}$ es un conjunto numerable denso en H , de él se puede extraer un sistema total constituido por vectores linealmente independientes. Para ello, basta con omitir de la sucesión $\{\varphi_n\}$ todos aquellos elementos φ_k que pueden representarse como combinación lineal de φ_i con $i < k$.

Aplicando el proceso de ortonormalización de Gram-Schmidt (6.5.11), encontramos un sistema ortonormal total en H , que según 6.5.14 será una base ortonormal de H . ■

Observación: La existencia de bases ortonormales constituye un factor importante para la aplicación de los métodos numéricos a la solución de ecuaciones en espacios normales.

Para tener una idea de la importancia del resultado del corolario anterior, aunque no veremos aquí la demostración por ser un resultado bastante más profundo, debemos señalar que hay espacios de Banach en los cuales no se puede demostrar la existencia de bases.

Veamos otra propiedad importante de las bases ortonormales.

Proposición 6.5.15. En un espacio euclideo separable, toda base ortonormal es un conjunto ortonormal maximal.

Demostración: Sea $\{e_k\}$ una base ortonormal en H . Si suponemos que dicha

base no constituye un conjunto ortonormal maximal, entonces, existirá $x \in H$ no nulo ortogonal a todos los e_k . Por el teorema 6.5.14, x puede expresarse mediante su serie de Fourier con respecto al sistema $\{e_k\}$

$$x = \sum_k \langle x, e_k \rangle e_k;$$

pero como $\langle x, e_k \rangle = 0$ para todo k , entonces $x = 0$, ¡contradicción!. ■

Observación: En el próximo epígrafe veremos que en los espacios de Hilbert también se cumple el inverso de la proposición anterior, y por lo tanto, la maximalidad es también otra propiedad en los espacios de Hilbert separables, que junto a las expuestas en el teorema 6.5.14, caracteriza a los conjuntos ortonormales que son bases. Notemos que en la proposición anterior nos referimos al concepto usual de maximalidad; un sistema ortonormal se llama maximal en H si deja de ser ortonormal al añadirse cualquier otro elemento de H no perteneciente a dicho sistema.

Pasemos finalmente a estudiar el segundo problema relacionado con el concepto de ortogonalidad en los espacios euclidianos, en decir, el problema de la mejor aproximación, el cual ya introdujimos en el 6.3 y que hemos tratado de nuevo en la proposición 6.5.12.

Notemos que el resultado obtenido en 6.5.12 puede ser interpretado geoméricamente del modo siguiente: El elemento

$$f - \sum_{k=1}^n a_k e_k$$

es ortogonal a todas las combinaciones lineales del tipo

$$\sum_{k=1}^n \beta_k e_k$$

es decir, es ortogonal al subespacio generado por los elementos $\{e_1, e_2, \dots, e_n\}$ cuando, y solamente cuando, se cumple la condición (74) (compruebe esta afirmación). Por consiguiente, el resultado obtenido en 6.5.12 representa una generalización del conocido teorema de la Geometría elemental.

La longitud de la perpendicular bajada de un punto dado a una recta o a un plano es menor que la longitud de cualquier oblicua trazada por el mismo punto.

Precisamente, la expresión abstracta de este resultado es la primera versión del teorema general de optimización que expondremos a continuación en los espacios euclidianos. En el próximo epígrafe veremos otras versiones con conclusiones más fuertes para los espacios de Hilbert.

Teorema 6.5.16. Sean $(H, \langle \cdot, \cdot \rangle)$ un espacio euclideo, M un subespacio vectorial de H y x un vector arbitrario de H . Si existe un vector $m_0 \in M$ tal que $\|x - m_0\| \leq \|x - m\|$ para todo $m \in M$, entonces m_0 es único. Una condición necesaria y suficiente para que $m_0 \in M$ sea el vector minimizante (único) en M es que el vector error $x - m_0$ sea ortogonal a M ($x - m_0 \in M^\perp$).

Demostración: Probaremos primero que si m_0 es un vector minimizante, entonces $x - m_0$ es ortogonal a M . Supongamos por el contrario que existe un $m \in M$, que no es ortogonal a $x - m_0$. Sin pérdida de generalidad, podemos asumir que $\|m\| = 1$ y que $\langle x - m_0, m \rangle = \delta \neq 0$. Definamos el vector m_1 en M como $m_1 = m_0 + \delta m$.

Entonces,

$$\begin{aligned} \|x - m_1\|^2 &= \|x - m_0 - \delta m\|^2 = \|x - m_0\|^2 - \langle x - m_0, \delta m \rangle \\ &\quad - \langle \delta m, x - m_0 \rangle + |\delta|^2 = \\ &= \|x - m_0\|^2 - |\delta|^2 < \|x - m_0\|^2. \end{aligned}$$

Luego, si $x - m_0$ no es ortogonal a M , m_0 no es vector minimizante.

Probemos ahora que si $x - m_0 \in M^\perp$, entonces, m_0 es el único vector minimizante.

Para cualquier $m \in M$, por el teorema de Pitágoras generalizado (6.5.8) tenemos

$$\|x - m\|^2 = \|x - m_0 + m_0 - m\|^2 = \|x - m_0\|^2 + \|m_0 - m\|^2$$

y por lo tanto,

$$\|x - m\| > \|x - m_0\| \text{ para todo } m \neq m_0. \blacksquare$$

La versión tridimensional de este teorema se ilustra en la (fig. 94).

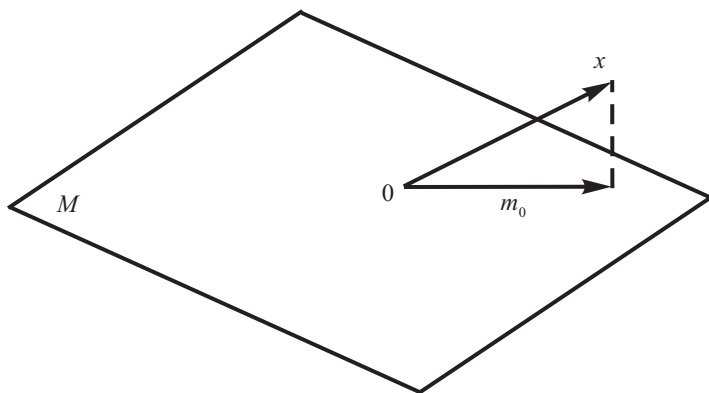


Figura 94

Observación: Notemos que en el teorema 6.5.16, no hemos establecido la existencia del vector minimizante, lo cual solo podemos asegurar hasta el momento, si M es de dimensión finita (6.3.7). Solo hemos probado hasta aquí que si dicho vector m_o existe, este es único y $x - m_o$ es ortogonal al subespacio M . En el próximo epígrafe introduciremos una clase de espacios euclidianos llamados espacios de Hilbert, donde se garantiza la existencia del vector minimizante.

Ejercicios de Autocontrol

1. Demuestre que el espacio euclideo del ejemplo 20 no es separable.
2. Calcule la serie de Fourier en $L_2[-\pi, \pi]$ de las funciones

a) $f(x) = \text{signo}(2x - 1)$;

b) $f(x) = e^{\lambda x}$

con respecto al sistema trigonométrico del ejemplo 22 (68).

3. Consideremos sobre el espacio $C([-1, 1], \mathbb{R})$ el producto escalar

$$\langle f, g \rangle = \int_{-1}^1 f(t) g(t) \frac{dt}{\sqrt{1-t^2}}.$$

Aplice el proceso de ortonormalización de Gram-Schmidt en este espacio a la sucesión de polinomios mónicos $\{1, x, x^2, \dots, x^n, \dots\}$.

4. Consideremos el espacio euclideo $C([-1, 1], \mathbb{R})$ provisto del producto escalar

$$\langle f, g \rangle = \int_{-1}^1 f(t) g(t) dt.$$

Halle en este espacio euclideo el complemento ortogonal a los conjuntos siguientes:

- a) funciones iguales a cero para $x \leq 0$
- b) funciones iguales a cero en el punto $x = 0$

- 5) Llamamos ángulo entre dos vectores x, y de un espacio euclideo $(H, \langle, \rangle)$ al ángulo φ tal que

$$\cos \varphi = \frac{\langle x, y \rangle}{\|x\| \|y\|}.$$

Calcule los ángulos del triángulo formado por las funciones $f_1(x) \equiv 0$, $f_2(x) \equiv 1$ y $f_3(x) \equiv x$, en el espacio euclideo del ejercicio 4.

6.6 ESPACIOS DE HILBERT

Introduzcamos una de las definiciones más importantes en el desarrollo de la matemática.

Definición 6.6.1. Se llama *espacio de Hilbert* a todo espacio euclideo completo.

Observación: Podemos dar las siguientes expresiones equivalentes a la definición de espacio de Hilbert:

- i) Un espacio de Hilbert es un espacio euclideo de Banach.
- ii) Un espacio de Hilbert es un espacio normado completo en el cual la norma está generada por un producto escalar.

Ejemplo 23. El espacio euclideo $\mathbb{R}^n$ con el producto escalar (56) y $l^2(\mathbb{N})$ con el producto escalar (57) son completos, luego, son espacios de Hilbert. Más adelante veremos que dichos espacios son, en cierto sentido, un modelo de espacios de Hilbert separables (vea el teorema 6.6.7).

El siguiente resultado nos muestra un método para la construcción de espacios de Hilbert (compárese con 6.3.9).

Proposición 6.6.2. (completamiento de un espacio euclideo) Sea $(H, \langle, \rangle)$ un espacio euclideo; entonces existe un completamiento $(\tilde{H}, \langle, \rangle_c)$

de H que es un espacio de Hilbert, tal que $\langle, \rangle_c$ es una prolongación continua de $\langle, \rangle$.

Demostración: De 6.3.9 se desprende que todo espacio euclideo $(H, \langle, \rangle)$ tiene un completamiento con respecto a la norma asociada a $\langle, \rangle$ que es un espacio de Banach, cuya norma es una extensión de la norma definida por el producto escalar $\langle, \rangle$. Por ser dicho producto escalar una aplicación uniformemente continua de $H \times H$ en $\mathbb{R}(\mathbb{C})$, puede prolongarse continuamente a $\tilde{H} \times \tilde{H}$, donde $\tilde{H}$ es el completamiento de H . Es claro que esta extensión es un producto escalar $\langle, \rangle_c$ en $\tilde{H}$ y por continuidad de la norma definida por $\langle, \rangle_c$ se obtiene precisamente la norma de H . ■

Ejemplo 24. Consideremos de nuevo el espacio $C([0, 1], \mathbb{R})$ provisto del producto escalar

$$\langle f, g \rangle_2 = \int_0^1 f(t) g(t) dt. \quad (78)$$

Es fácil comprobar que $(C([0, 1], \mathbb{R}), \langle, \rangle_2)$ es un espacio euclideo (ejemplo 18) el cual vimos además, en el capítulo 3, que no es completo. Este espacio admite un completamiento que denotaremos por $L_2[0, 1]$ que es un espacio de Hilbert.

Un resultado interesante es que el producto escalar en $L_2[0, 1]$ puede expresarse mediante una noción de integral más general llamada *integral de Lebesgue*, que coincide para las funciones continuas con la integral de Cauchy, a partir de la cual se define $\langle, \rangle_2$.

Veamos algunas propiedades de los espacios de Hilbert que “mejoran” las correspondientes en los espacios euclideos. La obtención de este primer grupo de propiedades en los espacios de Hilbert, se debe fundamentalmente a que hemos añadido la noción de completitud.

Proposición 6.6.3. En un espacio de Hilbert $(H, \langle, \rangle)$, para una sucesión (x_n) de elementos dos a dos ortogonales, son equivalentes:

- i) $\sum_{n=1}^{\infty} x_n$ es convergente.
- ii) $\sum_{n=1}^{\infty} \|x_n\|^2$ es convergente.

Demostración: La implicación i) $\Rightarrow$ ii) se cumple en cualquier espacio euclideo por el teorema de Pitágoras (6.5.8).

Inversamente, supongamos que la serie $\sum_{n=1}^{\infty} \|x_n\|^2$ converge y supongamos además que:

$$S_m = \sum_{n=1}^m x_n.$$

Para $m > p$ se tiene

$$S_m - S_p = \sum_{n=p+1}^m x_n,$$

y por el teorema de Pitágoras,

$$\|S_m - S_p\|^2 = \sum_{n=p+1}^m \|x_n\|^2.$$

De la convergencia de la serie $\sum_{n=1}^{\infty} \|x_n\|^2$ se deduce que la sucesión (S_n) es de Cauchy en H , luego, es convergente con lo cual probamos que la serie $\sum_{n=1}^{\infty} x_n$ converge. ■

De la desigualdad de Bessel (6.5.13), se deduce que para que los números $a_1, a_2, \dots, a_n, \dots$ representen los coeficientes de Fourier de un elemento en un espacio euclideo H , es necesario que la serie

$$\sum_n a_n^2$$

converja. Resulta que en un espacio de Hilbert esta condición no solo es necesaria, sino también suficiente.

Teorema 6.6.4 (teorema de Riesz-Fisher). Sea e_n un sistema ortonormal arbitrario en un espacio de Hilbert H , y sean los números

$$a_1, a_2, \dots, a_n, \dots$$

tales que la serie

$$\sum_{k=1}^{\infty} a_k^2$$

converge. Entonces, existe un elemento $x \in H$ tal que

$$a_k = \langle x, e_k \rangle$$

y

$$\sum_{k=1}^{\infty} a_k^2 = \langle x, x \rangle = \|x\|^2.$$

Demostración: Definamos

$$x_n = \sum_{k=1}^n a_k e_k.$$

Entonces, por el teorema de Pitágoras (6.5.8)

$$\|x_{n+p} - x_n\|^2 = \sum_{k=n+1}^{n+p} a_k^2.$$

Como la serie $\sum_{k=1}^{\infty} a_k^2$ converge, de aquí se deduce, en virtud de la completitud de H , la convergencia de la sucesión $\{x_n\}$ a un elemento $x \in H$. Además,

$$\langle x, e_i \rangle = \langle x_n, e_i \rangle + \langle x - x_n, e_i \rangle \quad (79)$$

donde el primer sumando del miembro derecho es igual a_i para $n \geq i$, mientras que el segundo sumando tiende a cero cuando $n \rightarrow \infty$ debido a la continuidad del producto escalar (6.5.5).

El miembro izquierdo de la igualdad (79) no depende de n , por eso, pasando al límite cuando $n \rightarrow \infty$ a ambos lados, obtenemos

$$\langle x, e_i \rangle = a_i.$$

De acuerdo con la definición de x , $\|x - x_n\| \rightarrow 0$ cuando $n \rightarrow \infty$ y por eso

$$\sum_{k=1}^{\infty} a_k^2 = \langle x, x \rangle.$$

En efecto,

$$\left\langle x - \sum_{k=1}^n a_k e_k, x - \sum_{k=1}^n a_k e_k \right\rangle = \langle x, x \rangle - \sum_{k=1}^n a_k^2 \xrightarrow{n \rightarrow \infty} 0.$$

Con esto el teorema queda demostrado. ■

Del teorema de Riesz-Fisher se deduce el siguiente resultado, el cual es de gran utilidad.

Teorema 6.6.5. Para un sistema ortonormal $\{e_n\}$ en un espacio de Hilbert separable $(H, \langle, \rangle)$ son equivalentes:

- i) $\{e_n\}$ es un sistema ortonormal maximal en H .
- ii) $\{e_n\}$ es una familia total en H .
- iii) No existe ningún elemento diferente de θ en H que sea ortogonal a todos los elementos del sistema $\{e_n\}$.

(Compárese con 6.5.9 ix) y 6.5.15)

Demostración: i) $\Rightarrow$ ii) Supongamos que la familia $\{e_n\}$ no es total; entonces por 6.5.14 ii), existe $x \in H$ que no puede ser representado por su serie de Fourier con respecto a $\{e_n\}$:

$$x \neq \sum_{n=1}^{\infty} a_n e_n,$$

donde $a_n = \langle x, e_n \rangle$.

Por el teorema de Riesz-Rischer (6.6.3) sabemos que la serie de Fourier de x converge a un cierto elemento de H , el cual en nuestro caso es diferente de x .

Pero entonces el elemento $\sum_{n=1}^{\infty} a_n e_n$ es un vector no nulo de H , ortogonal a todos los e_n luego, el sistema $\{e_n\}$ no es maximal.

ii) $\Rightarrow$ iii) Es consecuencia inmediata de 6.5.9 ix) y se cumple en todo espacio euclideo. Esto puede verse también de la siguiente forma: si $\{e_n\}$ es total, por la equivalencia i) iv) de 6.5.14, se cumple la identidad de Parseval para todo $x \in H$. Luego, si x es ortogonal a todos los elementos del sistema $\{e_n\}$, todos sus coeficientes de Fourier a_k se anulan, y por la identidad de Parseval obtenemos

$$\langle x, x \rangle = \sum_{k=1}^{\infty} a_k^2 = 0,$$

de donde $x = \theta$.

iii) $\Rightarrow$ i) Se deduce de la definición de maximidad. (Vea la observación que aparece después de la proposición 6.5.15). ■

Ejemplo 25. Consideremos de nuevo los polinomios ortogonales de Legendre $\{e_n\}$ introducidos en el ejemplo 21 y demostremos, aplicando 6.6.5, que este sistema es total en $L_2[-1, 1]$.

Por el teorema 6.5.11 es suficiente probar que no existe una función no idénticamente nula en $L_2[-1, 1]$ ortogonal al subespacio de los polinomios.

Supongamos que existe un $f \in L_2[-1, 1]$ ortogonal a cada polinomio mónico t^n , es decir,

$$\int_{-1}^1 t^n f(t) dt = 0, \text{ para todo } n = 0, 1, \dots \quad (80)$$

Por el hecho de que $C([-1, 1], \mathbb{R})$ es denso en $L_2[-1, 1]$ se puede demostrar que basta tomar f en (80) (compruébelo). Nuestro objetivo entonces es demostrar que de (80) se deduce que $f \equiv 0$.

En efecto, por ser f continua, el teorema de Weierstrass nos dice que dado $\epsilon > 0$, existe un polinomio

$$P(t) = \sum_{k=0}^N a_k t^k$$

tal que

$$|f(t) - P(t)| < \epsilon \text{ para todo } t \in [-1, 1].$$

Entonces, como f es ortogonal a los polinomios,

$$\int_{-1}^1 |f(t)|^2 dt = \int_{-1}^1 f(t)(f(t) - P(t))dt \leq \epsilon$$

$$\leq \epsilon \int_{-1}^1 |f(t)| dt \leq \epsilon \sqrt{2} \left\{ \int_{-1}^1 |f(t)|^2 dt \right\}^{1/2}$$

donde la última desigualdad es un caso particular de la desigualdad de Cauchy-Buniakowsky. Esto nos prueba que

$$\int_{-1}^1 |f(t)|^2 dt \leq 2\epsilon^2.$$

Como ϵ es arbitrario y $|f|$ es continua y positiva, deducimos que $f(t) = 0$. Con esto hemos probado de nuevo según 6.6.5, que el subespacio de los polinomios es denso en $L_2[-1, 1]$.

Veamos un corolario importante del teorema de Riesz-Fischer que nos dice que desde el punto de vista de la estructura introducida por el producto escalar, todos los espacios de Hilbert son equivalentes. El concepto *equivalencia* entre espacios de Hilbert debe conservar:

la estructura conjuntista,

la estructura vectorial,

la estructura métrica y

la estructura geométrica asociada al producto escalar, y además, debe mantener la compatibilidad entre estas estructuras.

A este efecto se introduce la definición siguiente:

Definición 6.6.6. Dos espacios de Hilbert $(H_1, \langle \cdot, \cdot \rangle_1)$ y $(H_2, \langle \cdot, \cdot \rangle_2)$ son isomorfos si existe una aplicación biyectiva

$$i : H_1 \longrightarrow H_2 \tag{81}$$

tal que

$$\langle x, y \rangle_1 = \langle i(x), i(y) \rangle_2 \tag{82}$$

para todos $x, y \in H_1$.

Notemos que de la propiedad (82) se deduce que para todos $x_1, x_2, x \in H_1$, $\lambda \in \mathbb{R}(\text{o } \mathbb{C})$ se tiene

$$i(x_1 + x_2) = i(x_1) + i(x_2); \tag{83}$$

$$i(\lambda x) = \lambda i(x). \tag{84}$$

En efecto, para cada $y \in H_1$

$$\langle i(x_1 + x_2), i(y) \rangle_2 = \langle x_1 + x_2, y \rangle_1 =$$

$$= \langle x_1, y \rangle_1 + \langle x_2, y \rangle_1 = \langle i(x_1), i(y) \rangle_2 +$$

$$+ \langle i(x_2), i(y) \rangle_2,$$

de donde, agrupando convenientemente

$$\langle i(x_1 + x_2) - i(x_1) - i(x_2), i(y) \rangle_2 = 0.$$

Por la sobreyectividad de la aplicación i , la arbitrariedad de y , y utilizando el hecho que $H_2^\perp = \{\theta\}$, concluimos que

$$i(x_1 + x_2) - i(x_1) - i(x_2) = \theta$$

o sea, se cumple (83).

No resulta difícil verificar (84).

Por otra parte de (82) se deduce que para las normas asociadas a los respectivos productos escalares, se tiene

$$\|x\|_1 = \|i(x)\|_2,$$

es decir i determina una isometría con respecto a las estructuras métricas subyacentes.

Teorema 6.6.7 (isomorfismo). Todos los espacios de Hilbert separables de igual dimensión son isomorfos entre sí.

Demostración: En el caso de espacios de dimensión finita la demostración es evidente y por ello se deja al lector. Por ello, demostraremos solo el caso de dimensión infinita. El lector puede verificar sin dificultad que la relación de isomorfía es una relación de equivalencia entre espacios de Hilbert. Luego, basta demostrar que todo espacio de Hilbert separable H es isomorfo a un espacio modelo.

Como espacio modelo tomaremos el espacio $l^2(\mathbb{N})$ provisto del producto escalar (57). Por ser H un espacio de Hilbert separable, del Corolario 1 del Teorema 6.5.14, concluimos que existe una base ortonormal numerable $\{e_n\}$ en H . Definamos la aplicación $i: H \longrightarrow l^2(\mathbb{N})$

de la manera siguiente:

$$i(x) = (a_1, a_2, \dots, a_n, \dots) \quad (85)$$

donde $a_i = \langle x, e_i \rangle$ es el i -ésimo coeficiente de Fourier de x con respecto a la base ortonormal $\{e_n\}$. La aplicación i está bien definida ya que por la

desigualdad de **Bessel** $\sum_{n=1}^{\infty} a_n^2 < \infty$.

Mostraremos que i , definida por (85), establece una correspondencia biunívoca entre H y $l^2(\mathbb{N})$. La inyectividad de i se deduce del teorema 6.6.5 y la sobreyectividad, del teorema de Riesz-Fischer (6.6.4). En efecto, si $x, y \in H$ con $x \neq y$, entonces $x - y \neq \theta$. Por 6.6.5 y por ser $\{e_n\}$ una base de H , $x - y$ no puede ser ortogonal a todos los elementos e_n . Luego, para algún $n \in \mathbb{N}$ se tendrá

$$\langle x - y, e_n \rangle \neq 0,$$

es decir,

$$\langle x - y, e_n \rangle \neq \langle y, e_n \rangle,$$

y por tanto,

$$i(x) \neq i(y)$$

con lo cual queda probada la inyectividad.

El lector puede verificar que la sobreyectividad de i es precisamente la tesis del teorema de Riesz-Fischer.

Nos queda por probar que

$$\langle x, y \rangle = \langle i(x), i(y) \rangle_2 = \sum_{n=2}^{\infty} a_n b_n, \quad (86)$$

donde $\{a_n\}$ y $\{b_n\}$ son los coeficientes de Fourier de x y y respectivamente y además, hemos denotado por $\langle, \rangle_2$ al producto escalar (57) definido en $l^2(\mathbb{N})$.

Notemos que la identidad de Parseval (la cual se cumple en este caso según 6.5.14), puede ser escrita también en la forma

$$\langle x, x \rangle = \langle i(x), i(x) \rangle_2. \quad (87)$$

Sean ahora $x, y \in H$. De (87) tenemos que se cumple

$$\langle x + \lambda y, x + \lambda y \rangle = \langle i(x) + \lambda i(y), i(x) + \lambda i(y) \rangle_2$$

para todo escalar λ . De aquí, por las propiedades del producto escalar se tiene que

$$\lambda \langle x, y \rangle + \lambda \langle y, x \rangle = \lambda \langle i(x), i(y) \rangle_2 + \lambda \langle i(y), i(x) \rangle_2$$

Tomemos $\lambda = 1$ y $\lambda = i$ en (88). Entonces de (88) se deduce que $\langle x, y \rangle$ y $\langle i(x), i(y) \rangle_2$ tiene la misma parte real e imaginaria, y por lo tanto, son iguales, con lo cual queda probado (86) y de hecho el teorema. ■

Pasemos ahora a ver otras propiedades de los espacios de Hilbert más vinculadas al concepto de ortogonalidad y los problemas de distancia mínima.

Comencemos por el siguiente teorema que mejora al 6.5.14 en el sentido de que garantiza la existencia del vector minimizante.

Teorema 6.6.8 (teorema clásico sobre la proyección). Sea M un subespacio vectorial cerrado de H ; entonces para cada vector $x \in H$, existe un único vector $m_0 \in M$ tal que $\|x - m_0\| \leq \|x - m\|$ para todo $m \in M$.

Una condición necesaria y suficiente para que $m_0 \in M$ sea el vector minimizante es que $x - m_0$ sea ortogonal a M .

Demostración: La unicidad y la ortogonalidad han sido ya establecidas en el teorema 6.5.16. Luego, se necesita solamente establecer la existencia del vector minimizante.

Si $x \in M$ entonces $m_0 = x$ y se obtiene el resultado deseado. Supongamos que $x \notin M$ y sea $\delta = \inf_{n \in M} \|x - m\|$. Debemos hallar un elemento $m_0 \in M$ tal que $\|x - m_0\| = \delta$. Con este propósito, tomemos una sucesión (m_k) de vectores de M tal que $\|x - m_k\|_{k \rightarrow \infty} \rightarrow \delta$.

Por la ley del paralelogramo

$$\begin{aligned} \|(m_j - x) + (x - m_i)\|^2 + \|(m_j - x) - (x - m_i)\|^2 &= \\ &= 2\|m_j - x\|^2 + \|x - m_i\|^2. \end{aligned}$$

Reordenando convenientemente obtenemos

$$\|m_j - m_i\|^2 = 2\|m_j - x\|^2 +$$

$$+ 2\|x - m_i\|^2 - 4\left\|x - \frac{m_i + m_j}{2}\right\|^2.$$

Para todo par i, j de índices, el vector $(m_i + m_j)/2$ pertenece a M ya que M es un *subespacio vectorial*. Entonces, por la definición de δ ,

$$\left\|x - \left(\frac{m_i + m_j}{2}\right)\right\| \geq \delta \text{ y obtenemos.}$$

$$\|m_j - m_i\|^2 \leq 2\|m_j - x\|^2 + 2\|x - m_i\|^2 - 4\delta^2.$$

Como $\|x - m_i\|^2 \rightarrow \delta^2$ cuando $i \rightarrow \infty$, concluimos que

$$\|m_j - m_i\|^2 \rightarrow 0 \text{ cuando } i, j \rightarrow \infty.$$

Luego, la sucesión (m_k) es de Cauchy en M y como M es un subespacio cerrado del espacio completo H , de aquí se obtendrá la existencia de un límite m_0 en M . Por continuidad de la norma, se prueba que $\|x - m_0\| = \delta$. ■

Observación: Notemos que durante la demostración de la existencia del vector minimizante en el teorema 6.6.8, no hemos hecho referencia explícita al producto escalar, sino solamente se utiliza la norma asociada. Sin embargo, implícitamente hemos utilizado el producto escalar al hacer uso de la ley del paralelogramo.

Existe una amplia clase de espacios de Banach donde también tiene sentido el teorema 6.6.8, aunque este no puede ser extendido a cualquier espacio de Banach.

Como consecuencia del teorema 6.6.8 veremos a continuación un resultado que justifica el nombre de *complemento ortogonal* para denotar al conjunto de los vectores ortogonales a un conjunto dado. Si el conjunto dado es un subespacio vectorial cerrado de un espacio de Hilbert, entonces su complemento contiene suficientes vectores adicionales como para generar a todo el espacio.

Teorema 6.6.9. Si M es un subespacio vectorial cerrado de un espacio de Hilbert H , entonces $H = M \oplus M^\perp$ y además, $M = M^{\perp\perp}$.

Demostración: Sea $x \in H$. Por el teorema de la proyección 6.6.8, existe un único vector $m_0 \in M$ tal que $\|x - m_0\| \leq \|x - m\|$ para todo $m \in M$ y $n_0 = x - m_0 \in M^\perp$. Luego, $x = m_0 + n_0$ con $m_0 \in M$ y $n_0 \in M^\perp$.

Para probar que esta representación es única basta probar que $M \cap M^\perp = \{\theta\}$ lo cual fue probado en 6.5.7 iv).

Para probar que $M = M^{\perp\perp}$ es solamente necesario probar que $M^{\perp\perp} \subset M$ ya que por 6.5.7 v) se tiene la inclusión en sentido contrario.

Sea $x \in M^{\perp\perp}$; por la primera parte de este teorema $x = m + n$ donde $m \in M$ y $n \in M^\perp$. Como ambos, x y m , pertenecen a $M^{\perp\perp}$, tendremos que $x - m \in M^{\perp\perp}$, y por lo tanto, $n \in M^{\perp\perp}$ y por ello, $n \perp n$ lo cual implica que $n = \theta$. Luego, $x = m \in M$ y $M^{\perp\perp} \subset M$. ■

Teniendo en cuenta el resultado anterior se introduce la definición siguiente:

Definición 6.6.10. Sea M un subespacio cerrado de un espacio de Hilbert H y $x \in H$. El vector único $m_0 \in M$ tal que $x - m_0 \in M^\perp$ se llama *proyección ortogonal* de x sobre M .

Del teorema 6.6.9 se desprende el siguiente corolario que completa la proposición 6.5.7 para los espacios de Hilbert.

Corolario 1. Sea S un subconjunto arbitrario de un espacio de Hilbert. Entonces se cumple:

- i) $S^{\perp\perp} = \overline{[S]}$ (donde por $\overline{[S]}$ se denota la clausura del subespacio vectorial generado por S);
- ii) $S^\perp = \{\theta\} \Leftrightarrow S$ es una familia total en H .

Demostración:

- i) Por 6.5.7 ii) y iii) se tiene que

$$S^{\perp\perp} = [S]^{\perp\perp} = \overline{[S]}^{\perp\perp}$$

y por 6.6.9 $\overline{[S]}^{\perp\perp} = \overline{[S]}$,

de donde $S^{\perp\perp} = \overline{[S]}$.

ii) la implicación $\Leftarrow$ fue ya obtenida en 6.5.7 ix) para un espacio euclideo cualquiera. Para probar la implicación en sentido contrario notemos que si $S^{\perp} = \{\theta\}$ entonces,

$$S^{\perp\perp} = \{\theta\}^{\perp} = H,$$

pero por 6.6.9 $S^{\perp\perp} = \overline{[S]}$, y por tanto, S es una familia total en H . ■

Observación: El punto i) del corolario anterior, nos dice que $S^{\perp\perp}$ coincide con el menor subespacio vectorial cerrado de H que contiene a S . El punto ii), no es mas que una generalización de 6.6.5 iii) a conjuntos cualesquiera.

Analicemos un ejemplo interesante de la teoría de ecuaciones diferenciales, donde se utilizan las series de Fourier con respecto a ciertos sistemas ortonormales para encontrar la solución.

Ejemplo 26 (Sistemas oscilatorios continuos. Problema de Bernoulli sobre las oscilaciones de una cuerda). Desarrollaremos en este ejemplo un problema, en la actualidad elemental, de los llamados *problemas en valores propios* para sistemas oscilatorios descritos por ecuaciones diferenciales. Este tipo de problema es el objeto de estudio de la teoría de operadores diferenciales y su análisis, nos brinda información de carácter fundamentalmente cualitativo en relación con los problemas que aparecen en la física-matemática.

A partir del siglo XVIII, cuando Daniel Bernoulli halló la solución al problema de oscilaciones de una cuerda, el progreso del análisis y de la matemática aplicada tuvieron una gran influencia en la formación de la teoría de las ecuaciones diferenciales e integrales, y en el desarrollo del cálculo de variaciones.

La teoría de oscilaciones, como se denomina comúnmente al estudio de los problemas en valores propios para sistemas oscilatorios, tiene enumerables aplicaciones, comenzando por problemas de la ingeniería y terminando por los problemas más complejos de la física teórica.

Nosotros expondremos de manera resumida las ideas de Daniel Bernoulli sobre el problema de oscilaciones de una cuerda. Este problema puede ser formulado de la manera siguiente:

Hallar una expresión funcional que describa el movimiento de una cuerda cuyos extremos se encuentren fijos, si se conocen la posición y la velocidad de cada uno de sus puntos en el instante inicial de tiempo $t = 0$.

Por comodidad supondremos que la longitud de la cuerda es igual a π y que sus extremos están fijos en los puntos $x = 0$ y $x = \pi$, para todo instante

de tiempo t y denotaremos mediante $w(x, t)$ la desviación del punto x de la cuerda a lo largo de la dirección transversal en el momento t .

Supongamos además, que la forma inicial de la cuerda está dada por una función $\varphi(x) = w(x, 0)$, y la velocidad inicial en cada punto viene dada por la función $\psi(x) = w_t(x, 0)$, donde el subíndice t indica la derivada parcial de w con respecto a t . Si nos encontramos en ausencia de fuerzas externas, toda solución del problema de Bernoulli debe satisfacer una ecuación en derivadas parciales de segundo orden, la cual para unidades de medida escogidas convenientemente, puede ser escrita en la forma usual

$$w_{tt} - w_{xx} = 0. \quad (89)$$

El método utilizado por Bernoulli para resolver este problema, consiste en buscar primeramente los tipos de movimiento (soluciones de (89)) más simples posibles; precisamente aquellos para los cuales cada punto de la cuerda realiza oscilaciones armónicas simples y que se diferencian de las oscilaciones de otros puntos solamente por la amplitud. Un movimiento tal, que se llama *oscilación propia* de la cuerda, se describe por una función $w(x, t)$ del tipo

$$w(x, t) = u(x) f(t), \quad (90)$$

donde $u(x)$ es una función que depende únicamente de la variable x y $f(t)$ de la variable t .

Veremos más tarde que para que se satisfagan las condiciones

$$w(0, t) = w(\pi, t) = 0, \quad (91)$$

la función $f(t)$ debe ser periódica.

Desde el punto de vista de la acústica, esto significa que cada pequeña porción de la cuerda produce un sonido de la misma intensidad que cualquier otra pequeña porción, de manera tal que toda la cuerda emite un tono musical puro; *el tono fundamental o uno de sus sobretonos*; entonces, el problema de hallar la oscilación de la cuerda se reduce a encontrar sus oscilaciones propias.

A diferencia de los sistemas descritos por ecuaciones diferenciales ordinarias, en los cuales el estado del sistema se describe mediante un vector $q(t)$ de n coordenadas, en el problema de oscilaciones de una cuerda es imposible dar su posición en cualquier momento de tiempo t con ayuda de un número finito de números. La posición de la cuerda se describe por una función $w(x, t)$ definida para cada t en el intervalo $0 \leq x \leq \pi$, y por este motivo, decimos que este problema es de *dimensión infinita*. Precisamente, es esta la causa por la cual surge la necesidad de la introducción de los espacios funcionales en el estudio de las ecuaciones diferenciales.

En el caso de sistemas cuyo estado venga descrito por una función vectorial $q(t)$ de n componentes, se dice que estamos en presencia de un sistema con n *grados de libertad*. En este caso puede demostrarse que todo posible movimiento de un tal sistema es la *superposición* de n movimientos particulares, llamados *oscilaciones propias*, cuyas frecuencias se determinan completamente de las propiedades físicas del sistema.

En el caso que estamos analizando se dice que tenemos un sistema con *infinitos grados de libertad*. Bernoulli encontró, que las oscilaciones propias de la cuerda, es decir, las oscilaciones del tipo $w(x, t) = u(x) f(t)$, son posibles solamente cuando las funciones $u(x)$ salvo un factor escalar, pertenecen a una cierta sucesión especial de funciones

$$u_1(x), u_2(x), \dots, u_n(x), \dots \quad (92)$$

llamadas *funciones propias del sistema*. Las frecuencias correspondientes a estas funciones propias se llaman *frecuencias propias del sistema* y están completamente determinadas por las características físicas del sistema.

Después de todo lo dicho anteriormente, podemos plantear el problema de Bernoulli matemáticamente de la forma siguiente:

Hallar la solución $w(x, t)$ de la ecuación (89) que satisface las condiciones iniciales:

$$w(x, 0) = \varphi(x), \quad (93)$$

$$w_t(x, 0) = \psi(x) \quad (94)$$

y las condiciones de contorno:

$$w(0, t) = w(\pi, t) = 0. \quad (95)$$

Para la solución, Bernoulli razonó de la siguiente forma. Supongamos que existe una solución de (93), que puede representarse como

$$w(x, t) = u(x) f(t).$$

Sustituyendo esta función w en (89), obtenemos

$$u(x) f''(t) - u''(x) f(t) = 0,$$

o equivalentemente

$$\frac{u''(x)}{u(x)} = \frac{f''(t)}{f(t)}.$$

Del hecho que la parte izquierda de esta igualdad no depende de t y la parte derecha no depende de x , ellas pueden ser iguales ente sí para todos x y t , solamente si ambas son iguales a una misma constante que denotaremos por $-\lambda$. Así pues, tenemos

$$\frac{u''}{u} = \frac{f''}{f} = -\lambda$$

o lo que es lo mismo,

$$-u'' = \lambda u$$

$$-f'' = \lambda f.$$

Estas ecuaciones son homogéneas, es decir, si u_1 y u_2 son soluciones de la primera de ellas, entonces $au_1 + bu_2$ también será solución para cualesquiera constantes a y b . La función idénticamente nula es solución trivial de cualquier ecuación homogénea. Obviamente para cada λ el espacio de las soluciones constituye un espacio vectorial. A cada solución no idénticamente nula se le llama solución no trivial.

Del hecho que los extremos de la cuerda están fijos, tendremos

$$w(0, t) = u(0) f(t) = 0,$$

$$w(\pi, t) = u(\pi) f(t) = 0,$$

para todo valor de t . Por consiguiente, si excluimos el caso trivial $f(t) \equiv 0$ se tiene que $u(0) = u(\pi) = 0$. Estas condiciones homogéneas deben ser también satisfechas por la solución de la ecuación $-u'' = \lambda u$.

Si no tenemos en cuenta las condiciones de contorno, entonces existirá solución no trivial para todo valor de λ . En efecto, como la ecuación $-u'' = \lambda u$ es lineal y homogénea con coeficientes constantes, cualquiera de sus soluciones tiene la forma

$$u = au_1 + bu_2,$$

donde a y b son constantes arbitrarias y u_1, u_2 son dos soluciones particulares:

$$u_1 = e^{ivx}$$

$$u_2 = e^{-ivx}$$

donde $v = \sqrt{\lambda}$.

Pero si tenemos en cuenta las condiciones de contorno $u(0) = u(\pi) = 0$, entonces la situación cambia sustancialmente. El problema

$$\begin{aligned} -u'' &= \lambda u \\ u(0) &= u(\pi) = 0, \end{aligned} \tag{96}$$

tendrá solución no trivial solamente para algunos valores especiales de λ .

La función $u = ae^{ivx} + be^{-ivx}$ satisface la condición $u(0) = 0$ solamente para $a = -b$ y la segunda $u(\pi) = 0$, solamente para $\lambda = n^2$, donde n es un número entero. De esta forma, llegamos a que todas las posibles soluciones son del tipo

$$u_n = d_n (e^{inx} - e^{-inx}) = c_n \operatorname{sen} nx,$$

donde $c_n \neq 0$, y $n = 1, 2, 3, \dots$

Pero entonces, para la función $f(t)$ llegamos a la ecuación $f''(t) + n^2 f(t) = 0$ cuya solución, para cada valor de n , es de la forma

$$f_n(t) = a_n \cos nt + b_n \operatorname{sen} nt,$$

donde a_n y b_n son constantes arbitrarias y de esta forma, vemos como aparece la periodicidad a que hacíamos referencia al inicio. En fin, la n -ésima oscilación propia se describe por la fórmula

$$w_n(x, t) = \operatorname{sen} nx(a_n \cos nt + b_n \operatorname{sen} nt). \quad (97)$$

La situación que hemos descrito es típica para los problemas en valores propios. Primeramente se tiene un problema mixto para una cierta ecuación diferencial, es decir, se busca la solución de una ecuación diferencial lineal que satisface condiciones iniciales y de contorno auxiliares. Nos planteamos el problema de hallar soluciones del tipo $w(x, t) = u(x)f(t)$, y de esta forma, llegamos a ecuaciones diferenciales para f y u dependientes de un cierto parámetro λ llamado *parámetro espectral*. Para cada λ fijo obtenemos problemas de contorno, para los cuales la función idénticamente nula es una solución. Sin embargo, para una cierta sucesión de valores (λ_n) que en nuestro caso coincide con la sucesión.

$\lambda_1 = 1^2, \lambda_2 = 2^2, \dots, \lambda_n = n^2, \dots$,
existen también otras soluciones $u_n(x)$.

En los problemas en valores propios para problemas de contorno en ecuaciones diferenciales los números λ_n se llaman *valores propios* del problema y las funciones u_n son las llamadas *funciones propias*. El conjunto de todos los valores propios, cada uno de los cuales se considera con cierta multiplicidad, se llama *espectro puntual* del problema.

Después de esta disgresión, regresamos al problema de Bernoulli. Bernoulli razonó de la forma siguiente: ¿se puede o no, con ayuda de estas oscilaciones propias particulares de la cuerda, obtener la solución que satisfaga las condiciones iniciales (94)?; es decir, ¿se pueden elegir constantes a_n y b_n tales que la solución

$$w(x, t) = \sum_{n=1}^{\infty} \operatorname{sen} nx(a_n \cos nt + b_n \operatorname{sen} nt) \quad (98)$$

satisfaga las condiciones iniciales $w(x, 0) = \varphi(x)$, $w_1(x, 0) = \psi(x)$?

Una vez llegado a este punto, notemos que es necesario precisar primeramente el sentido en que vamos a entender la convergencia de la serie (98).

Utilizando el hecho de que el sistema

$$\left\{ \sqrt{\frac{2}{\pi}} \operatorname{sen} kx \right\}_{k=1}^{\infty} \quad (99)$$

es una base en $L_2[0, \pi]$ (compruébelo), y por el teorema de Reisz-Fischer, se concluye que si

$$\sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) < \infty \quad (100)$$

entonces, la serie (98) converge con respecto a x en $L_2[0, \pi]$ para cada t fijo y en $L_2[-\pi, \pi]$ con respecto a t , para cada x fijo.

Si por algún método se pudiera probar que la serie (98) converge puntualmente en $t = 0$ y que además, la operación de derivación con respecto a t conmuta con el signo de sumatoria, obteniéndose otra serie que también converge puntualmente en $t = 0$, entonces las condiciones iniciales (94) implicarían que

$$\varphi(x) = \sum_{n=1}^{\infty} a_n \operatorname{sen} nx, \quad (101)$$

$$\psi(x) = \sum_{n=1}^{\infty} nb_n \operatorname{sen} nx. \quad (102)$$

Supongamos entonces que a_n y b_n son tales que se pueden realizar las operaciones antes mencionadas de manera que se obtenga (101) y (102).

Por la completitud del sistema (99) en $L_2[0, \pi]$, se tendría que si $\varphi(x)$ y $\psi(x)$ pertenece a $L_2[0, \pi]$ (por ejemplo, $\varphi(x), \psi(x) \in C([0, \pi], \mathbb{R})$) entonces los números a_n y b_n vienen determinados unívocamente por los coeficientes de Fourier de φ y ψ con respecto al sistema (99) y las series (101) y (102) convergen en $L_2[0, \pi]$;

$$a_n = \frac{2}{\pi} \int_0^{\pi} \varphi(x) \operatorname{sen} nxdx,$$

$$nb_n = \frac{2}{\pi} \int_0^{\pi} \psi(x) \operatorname{sen} nxdx.$$

Finalmente se debe verificar que para los a_n y b_n obtenidos, es posible realizar las operaciones con la serie (98) mediante las cuales obtuvimos las expresiones (101) y (102).

Si esto es posible entonces la serie (98) nos da la solución buscada del problema de Bernoulli.

Volvamos de nuevo al problema de la mejor aproximación en un espacio de Hilbert. Hasta ahora hemos considerado con algún detalle el problema de aproximar un vector arbitrario en un espacio de Hilbert por vectores de un subespacio vectorial de dimensión finita. Fue probado que si los vectores $\{e_1, e_2, \dots, e_n\}$ son ortonormales, entonces el vector de mejor aproximación al vector x en el espacio vectorial M generado por los $e_i, (i = 1, 2, \dots, n)$ es

$$\left. \begin{aligned} \langle x, e_1 \rangle &= a_1 \\ \langle x, e_2 \rangle &= a_2 \\ &\vdots \\ \langle x, e_n \rangle &= a_n \end{aligned} \right\} \quad (104)$$
$$\begin{aligned}\langle x, e_1 \rangle &= a_1 \\ \langle x, e_2 \rangle &= a_2 \\ &\vdots \\ \langle x, e_n \rangle &= a_n\end{aligned}$$
$$x_0 = \sum_{i=1}^n \beta_i e_i$$
$$\begin{array}{lcl} \langle e_1, e_1 \rangle \beta_1 + \langle e_2, e_1 \rangle \beta_2 + \dots + \langle e_n, e_1 \rangle \beta_n & = & a_1, \\ \langle e_1, e_2 \rangle \beta_1 + \langle e_2, e_2 \rangle \beta_2 + \dots + \langle e_n, e_2 \rangle \beta_n & = & a_2, \\ \vdots & & \vdots \\ \langle e_1, e_n \rangle \beta_1 + \dots & + & \langle e_n, e_n \rangle \beta_n = a_n \end{array}$$

Demostración: Sea M el subespacio vectorial n -dimensional generado por los vectores $\{e_1, \dots, e_n\}$. La variedad lineal definida por las restricciones del problema de minimización, es una traslación del subespacio $M^\perp$. Como $M^\perp$ es cerrado, la existencia y unicidad de una solución óptima es una consecuencia del teorema de la proyección. Más aún, sabemos que la solución óptima x_0 es

ortogonal a $M^\perp$, por lo cual $x_0 \in M^{\perp\perp}$. Pero como M es cerrado, entonces $M^{\perp\perp} = M$.

De todo lo anterior concluimos que $x_0 \in M$, o lo que es equivalente

$$x_0 = \sum_{i=1}^n \beta_i e_i.$$

Los n coeficientes β_i son seleccionados de manera tal que x_0 satisfaga las n restricciones. Esto conduce a las n ecuaciones expuestas en el enunciado del teorema. ■

Ejemplo 27 (aplicación a un problema de optimización). La velocidad angular w de las aspas de un ventilador, controlado por una fuente de corriente variable u , está determinada por la siguiente ecuación diferencial de primer orden:

$$\dot{w}(t) + w(t) = u(t); \quad (105)$$

donde $u(t)$ es la corriente suministrada en el instante t . La posición angular θ de una aspa es la integral de w con respecto al tiempo. Asumamos que el ventilador parte inicialmente del reposo, es decir, $\theta(0) = w(0) = 0$ y que se desea hallar la función u que determina la corriente aplicada, de manera tal que la energía gastada durante un segundo para rotar el aspa a una nueva posición de reposo $\theta = 1$, $w = 0$, sea mínima.

Supondremos que la energía es proporcional a $\int_0^1 u^2(t) dt$. Este es un problema simple de la llamada *teoría de control* en el cual el *funcional criterio* a optimizar, depende solamente de la función control u . El problema puede ser resuelto si es tratado como un problema de norma mínima en el espacio de Hilbert $L_2[0, 1]$.

Podemos integrar la ecuación diferencial de primer orden (105) y obtendremos la fórmula explícita

$$w(1) = \int_0^1 e^{(t-1)} u(t) dt \quad (106)$$

para la velocidad angular final correspondiente a cualquier control u . De la ecuación (105) se deduce además la ecuación

$$\theta(1) = \int_0^1 u(t) dt - w(1), \quad (107)$$

la cual combinada con la ecuación (106) nos da,

$$\theta(1) = \int_0^1 \{1 - e^{(t-1)}\} u(t) dt. \quad (108)$$

Si buscamos la función control $u(t)$ como un elemento del espacio de Hilbert $L_2[0, 1]$, las relaciones anteriores pueden ser expresadas como

$$\begin{aligned} w(1) &= \langle y_1, u \rangle, \\ \theta(1) &= \langle y_2, u \rangle, \end{aligned} \tag{109}$$

donde

$$y_1(t) = e^{(t-1)},$$

$$y_2(t) = 1 - e^{(t-1)}.$$

Con estas definiciones el problema original es equivalente a hallar $u \in L_2[0, 1]$ de norma mínima sujeta a las restricciones

$$\begin{aligned} 0 &= \langle y_1, u \rangle \\ 1 &= \langle y_2, u \rangle. \end{aligned} \tag{110}$$

De acuerdo al teorema 6.6.11 la solución óptima debe estar en el subespacio vectorial generado por y_1 y y_2 . Es decir,

$$u(t) = \alpha + \beta e^t,$$

donde las constantes α y β se seleccionan de manera que se satisfagan las restricciones (110). Si calculamos las constantes obtendremos la solución buscada:

$$u(t) = \frac{1}{3 - e} [1 + e - 2e^{-t}].$$

Como hemos podido observar, existen dos formas básicas en el planteamiento de los problemas de norma mínima en un espacio de Hilbert que se reducen a la solución de un número finito de ecuaciones lineales. En su forma general ambos problemas están relacionados con el cálculo de la menor distancia de un punto a una variedad lineal. En un caso la variedad lineal tiene dimensión finita y en el otro, tiene codimensión finita. Para comprender la relación entre estos dos problemas que representan uno de los problemas más simples de una relación de dualidad, consideremos la (fig. 95).

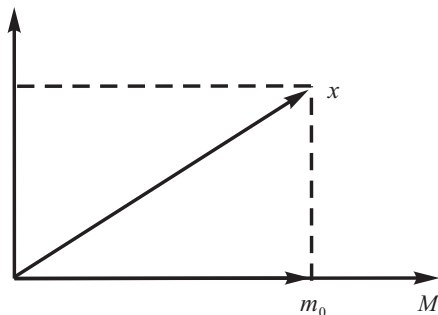


Figura 95

Sea M un espacio vectorial cerrado en un espacio de Hilbert H y sea x un vector arbitrario en H . Podemos formular dos problemas; el primero consiste en proyectar x sobre M y el segundo en proyectar x sobre $M^\perp$.

La situación es completamente simétrica ya que $M^{\perp\perp} = M$. Si resolvemos uno de estos problemas el otro es automáticamente resuelto ya que m_0 es la proyección de x sobre M , entonces $x - m_0$ es la proyección de x sobre $M^\perp$. Luego, si M o $M^\perp$ es de dimensión finita, ambos problemas pueden ser reducidos a resolver un sistema finito de ecuaciones lineales.

Veamos finalmente un problema de distancia mínima en los espacios de Hilbert de carácter un tanto diferente a los que hemos analizado hasta ahora. Más exáctamente, veremos como muchos de los resultados obtenidos pueden ser generalizados de las variedades lineales a conjuntos convexos. El siguientes teorema, en el cual se trata el problema de norma mínima ilustrado en la figura 96, es una generalización directa del teorema sobre la proyección.

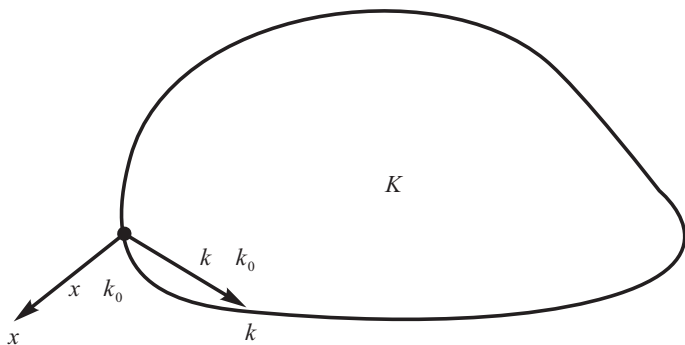


Figura 96

Teorema 6.6.12 (distancia mínima a un cerrado convexo). Sea x un vector en un espacio de Hilbert H y sea K un subconjunto cerrado convexo de H . Entonces existe un único vector $k_0 \in K$ tal que

$$\|x - k_0\| \leq \|x - k\|$$

para todo $k \in K$. Más aún, una condición necesaria y suficiente para que k_0 sea el único vector minimizante es que $\langle x - k_0, k - k_0 \rangle \leq 0$ para todo $k \in K$.

Demostración: Para probar la existencia, sea $\{k_i\}$ una sucesión en K tal que

$$\|x - k_i\| \longrightarrow \delta = \inf_{k \in K} \|x - k\|.$$

Por la ley del paralelogramo

$$\|x_i - k_j\|^2 = 2\|k_i - x\|^2 + 2\|k_j - x\|^2 - 4\left\|x - \frac{k_i + k_j}{2}\right\|^2.$$

Por la convexidad de K , $(k_i + k_j)/2$ está en K ; de aquí que

$$\left\| x - \frac{k_i + k_j}{2} \right\| \geq \delta,$$

y por lo tanto

$$\|x_i - k_j\|^2 \leq 2\|k_i - x\|^2 + 2\|k_j - x\|^2 - 4\delta^2 \longrightarrow 0.$$

Como la sucesión (k_i) es de Cauchy y K es cerrado, convergerá a un elemento $k_0 \in K$, y por continuidad de la norma $\|x - k_0\| = \delta$.

Para probar la unicidad, supongamos que existe otro $k_1 \in K$ tal que $\|k_1 - x\| = \delta$. La sucesión

$$k_n = \begin{cases} k_0, & \text{si } n \text{ par} \\ k_1, & \text{si } n \text{ impar} \end{cases}$$

cumple que $\|x - k_n\| \longrightarrow \delta$, luego por el argumento anterior $\{k_n\}$ es de Cauchy, y por lo tanto, converge. Obviamente, esto puede ocurrir en este caso tan solo si $k_1 = k_0$.

Probemos ahora que si k_0 es el único vector minimizante, entonces,

$$\langle x - k_0, k - k_0 \rangle \leq 0$$

para todo $k \in K$. Supongamos por el contrario que existe un vector $k_1 \in K$ tal que $\langle x - k_0, k_1 - k_0 \rangle = \epsilon > 0$. Consideremos los vectores $k_\alpha = (1 - \alpha)k_0 + \alpha k_1$; con $0 \leq \alpha \leq 1$.

Como K es convexo, cada $k_\alpha \in K$. Además,

$$\begin{aligned} \|x - k_\alpha\|^2 &= \|(1 - \alpha)(x - k_0) + \alpha(x - k_1)\|^2 = \\ &= (1 - \alpha)^2 \|x - k_0\|^2 + 2\alpha(1 - \alpha) \langle x - k_0, x - k_1 \rangle + \alpha^2 \|x - k_1\|^2. \end{aligned}$$

La expresión $\|x - k_\alpha\|^2$ es una función diferenciable de α con derivada en $\alpha = 0$ igual a

$$\begin{aligned} \frac{d}{d\alpha} \|x - k_\alpha\|^2_{\alpha=0} &= -2\|x - k_0\|^2 + \langle x - k_0, x - k_1 \rangle = \\ &= -2\langle x - k_0, k_1 - k_0 \rangle = -2\epsilon < 0. \end{aligned}$$

Luego, para algún α positivo suficientemente pequeño $\|x - k_\alpha\| < \|x - k_0\|$ lo cual contradice la propiedad minimizante de k_0 . De aquí concluimos que no puede existir un tal k_1 .

Inversamente, supongamos que $k_0 \in K$ es tal que $\langle x - k_0, k - k_0 \rangle \leq 0$ para todo $k \in K$. Entonces para cualquier $k \in K, k \neq k_0$, tendremos

$$\begin{aligned} \|x - k\|^2 &= \|x - k_0 + k_0 - k\|^2 = \|x - k_0\|^2 + \\ &+ 2\langle x - k_0, k_0 - k \rangle + \|k_0 - k\|^2 > \|x - k_0\|^2, \end{aligned}$$

y por lo tanto, k_0 es el único vector minimizante. ■

Ejercicios de Autocontrol

1. Halle en el espacio de Hilbert $L_2[0, 1]$ el complemento ortogonal a los conjuntos siguientes:

a) polinomios de x ;

- b) polinomios de x^2 ;
- c) polinomios con término independiente nulo;
- d) polinomios cuya suma de coeficientes es cero.

2. Demuestre que el conjunto de funciones $\left\{ \sqrt{\frac{2}{\pi}} \sin kx \right\}$ es una base ortonormal en $L_2 [0, \pi]$.

3. Sean H_1 y H_2 espacios de Hilbert y sea $i : H_1 \longrightarrow H_2$ sobreyectiva tal que $i(\lambda x + \mu y) = \lambda i(x) + \mu i(y)$.

Pruebe que i es un isomorfismo (6.6.6) si y solo si $\|i(x)\|_2 = \|x\|_1$ para todo $x \in H_1$.

4. Halle la función lineal $x(t) = at + b$ que minimiza la expresión

$$\int_{-1}^1 |t - x(t)| dt.$$

5. Dada una función x buscamos un polinomio p de grado n o menor tal que minimice $\int_0^1 |x(t) - p(t)|^2 dt$ satisfaciendo el requerimiento

$$\int_0^1 p(t) dt = 0.$$

a) Pruebe que este problema tiene una solución única.

b) Pruebe que este problema puede ser resuelto hallando primeramente el polinomio q de grado $\leq n$ que minimiza la expresión

$$\int_0^1 |x(t) - q(t)|^2 dt \text{ y después buscando } p \text{ de grado } n \text{ o menor que minimiza } \int_0^1 |q(t) - p(t)|^2 dt \text{ satisfaciendo el requerimiento}$$

$$\int_0^1 p(t) dt = 0.$$

c) Tome $n = 2$, $x(t) = t^2$ y calcule $p(t)$

6.7 OPERADORES LINEALES Y FUNCIONALES LINEALES

El estudio de los operadores lineales y funcionales lineales es fundamental para el tratamiento de una innumerable cantidad de problemas en los espacios vectoriales, muchos de ellos relacionados con las aplicaciones. Cuando se

combinan la estructura vectorial de los espacios normados con la estructura métrica, aparece la necesidad de introducir un tipo particular de aplicaciones continuas, que tienen la propiedad de “conservar” la estructura vectorial; tales aplicaciones son los *operadores lineales continuos* y los *funcionales lineales continuos*.

El objetivo fundamental del Análisis Funcional es precisamente el estudio de los operadores y funcionales lineales continuos en espacios de funciones.

Nosotros, sin embargo, no estudiaremos en este epígrafe los operadores y funcionales lineales desde el punto de vista del Análisis Funcional, sino nos dedicaremos a estudiar sus propiedades más simples y a analizar cómo se fortalece la propiedad de continuidad cuando se hace compatible con la estructura algebraica.

Durante este epígrafe veremos tres resultados muy importantes sobre operadores y funcionales lineales continuos, estos son:

- el teorema de Riesz sobre la representación de los funcionales lineales continuos en un espacio de Hilbert,
- el teorema sobre el límite puntual de una sucesión de operadores lineales continuos y
- el teorema de Hahn Banach.

A pesar de que estos son resultados que se enmarcan en el contexto del Análisis Funcional, hemos decidido tratarlos aquí porque tienen relación con otros temas importantes desarrollados a lo largo del texto como son: ortogonalidad, límite puntual y uniforme de una sucesión de funciones continuas y separación de conjuntos por funciones continuas.

Comencemos por algunas definiciones de carácter general.

Definición 6.7.1. Sean E y F espacios vectoriales sobre K ($K = \mathbb{R}$ o $\mathbb{C}$). Convendremos en llamar *operador* a toda aplicación

$$A : E \longrightarrow F.$$

En el caso en que $F = K$ diremos que A es un *funcional*.

El operador (funcional) A se llama *aditivo*, si

$$A(x + y) = A(x) + A(y) \tag{111}$$

para todo $x, y \in E$.

Se dice que A es *homogéneo*, si

$$A(\lambda x) = \lambda A(x) \tag{112}$$

para todo $\lambda \in K, x \in E$.

Diremos que A es un operador (funcional) *lineal* de E en F (de E en K) si es aditivo y homogéneo.

Observación: Para no complicar la notación en la definición 6.7.1 hemos denotado de la misma forma las operaciones de suma y producto por escalares en E y F . A partir de este momento adoptaremos esta convención salvo si se señala lo contrario.

Como una consecuencia simple de la definición 6.7.1 tenemos la proposición siguiente:

Proposición 6.7.2. Sean E y F espacios vectoriales y $A : E \longrightarrow F$ un operador aditivo; entonces:

- i) $A(\theta) = \theta$;
- ii) $A(-x) = -A(x)$ para todo $x \in E$;
- iii) $A(\lambda x) = \lambda A(x)$ se satisface para todo $\lambda \in \mathbb{Q}$.

Observación: De ahora en adelante, si no se indica lo contrario, utilizaremos la misma notación θ para el elemento nulo de ambos espacios vectoriales E y F .

Demostración:

- i) Para cualquier $x \in E$

$$A(x) = A(x + \theta) = A(x) + A(\theta),$$

y por lo tanto, $A(\theta) = \theta$.
- ii) Como $x - x = \theta$, entonces con la ayuda de i), tendremos

$$A(x) + A(-x) = A(x - x) = A(\theta) = \theta,$$

de donde $A(-x) = -A(x)$.
- iii) Si n es un entero positivo, entonces

$$A(nx) = \underbrace{A(x + x + \cdots + x)}_{n \text{ veces}} = nA(x).$$

Si $\lambda = 1/n$, donde n es un entero positivo, se tendrá

$$A(x) = A\left(n\frac{1}{n}x\right) = nA\left(\frac{1}{n}x\right).$$

y por lo tanto, $A\left(\frac{1}{n}x\right) = \frac{1}{n}A(x)$.

Si $\lambda = m/n$, con m y n enteros positivos, entonces por lo demostrado previamente

$$A(\lambda x) = A\left(m\frac{1}{n}x\right) = mA\left(\frac{1}{n}x\right) = \frac{m}{n}A(x) = \lambda A(x).$$

Finalmente, si λ es un número racional negativo, es decir, $\lambda = -r$, donde r tiene forma $\frac{m}{n}$ (m, n enteros positivos), entonces por la demostración para los racionales positivos y de ii), obtenemos

$$A(\lambda x) = A(-rx) = -A(rx) = -rA(x) = \lambda A(x).$$

Si $\alpha = 0$ la igualdad (112) se satisface automáticamente por i).■

Puede probarse, aún en el simple caso cuando E y F coinciden con la recta real $\mathbb{R}$, que la aditividad de A no implica su homogeneidad, es decir, la veracidad de (112) para todo λ . En el caso en que A sea lineal, puede comprobarse fácilmente que para cualesquiera elementos $x_1, x_2, \dots, x_n \in E$ y escalares $\lambda_1, \lambda_2, \dots, \lambda_n$.

$$A\left(\sum_{k=1}^n \lambda_k x_k\right) = \sum_{k=1}^n \lambda_k A(x_k). \quad (113)$$

Esta es la llamada propiedad de *distributividad* de un operador lineal. De aquí se deduce obviamente que si $A : E \rightarrow F$ es un operador lineal, entonces para cualquier subespacio vectorial M de E , $A[M]$ es un subespacio vectorial de F .

Por otra parte, la compuesta de $B \circ A$ de dos operadores lineales $A : E \rightarrow F$ y $B : F \rightarrow G$ es también un operador lineal. De esto se deduce que para todo operador lineal $A : E \rightarrow E$, cualquier potencia k -ésima A^k ($k \in \mathbb{N}$).

$$A^k = \underbrace{AA \cdots A}_{k \text{ veces}}$$

es un operador lineal.

Observación: Si $\{\varphi_\alpha\}_{\alpha \in I}$ es una base algebraica del espacio vectorial E y $(\psi_\beta)_{\beta \in j}$ una base algebraica del espacio vectorial F , entonces cualquier operador $A : E \rightarrow F$ está completamente definido si se da, para cada $\alpha \in I$ la expresión de $A(\varphi_\alpha)$ en términos de la base $(\psi_\beta)_{\beta \in j}$ de F . En efecto, cada $x \in E$ tiene una expresión única de la forma

$$x = \sum_{k=1}^n \lambda_k \varphi_{\alpha_k}, \quad \text{y por lo tanto,} \quad A(x) = \sum_{k=1}^n \lambda_k A(\varphi_{\alpha_k}).$$

Veamos a continuación algunos ejemplos de operadores y funcionales lineales.

Ejemplo 28 (matrices). Sea A una aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$. Una base de $\mathbb{R}^n$ está constituida por los vectores coordenados

$$\varphi_k = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n, \quad k = 1, 2, \dots, n$$

donde para cada k el número 1 ocupa el k -ésimo lugar en el n -uplo e_k .

Esta es la llamada base canónica de $\mathbb{R}^n$. De manera similar se construye la base canónica $(\psi_j)_{j=1}^m$ de $\mathbb{R}^m$.

Por la observación anterior, A está completamente determinado si se da la expresión de $A(\varphi_k)$ en términos de $(\psi_j)_{j=1}^m$, para cada $k = 1, \dots, n$.

$$\text{Si suponemos } A(\varphi_k) = \sum_{j=1}^m a_{jk} \psi_j,$$

entonces la matriz

$$(A) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (114)$$

describe completamente al operador A .

En efecto, el producto de una matriz de orden $n \times m$ por un vector $1 \times n$ se define de manera tal que si para los elementos de $\mathbb{R}^n$, $x = x_1\varphi_1 + \dots + x_n\varphi_n$, ($\mathbb{R}^m$, $y = y_1\psi_1 + \dots + y_m\psi_m$) utilizamos la notación

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y := \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

entonces se obtenga el mismo resultado al multiplicar (A) por el vector x que al evaluar el operador A en el vector x .

La teoría de matrices, es por consiguiente, la teoría de operadores en espacios vectoriales de dimensión finita.

Ejemplo 29 (operadores de proyección ortogonal). A partir de la definición 6.6.10 de proyección ortogonal de un vector sobre un subespacio cerrado de un espacio de Hilbert, se introducen los llamados operadores de proyección ortogonal. Si M es un subespacio vectorial cerrado del espacio de Hilbert H , se llama proyector ortogonal asociado a M , y se denota por P_M , al operador

$$P_M : H \longrightarrow H,$$

donde $P_M(x)$ es la proyección ortogonal de x sobre M . No es difícil comprobar que P_M es lineal y por ello, se deja de ejercicio al lector.

Además,

$$P_M^2 = P_M \cdot P_M = P_M, \quad (115)$$

es decir,

$$P_M(x) = x \quad (116)$$

para todo $x \in M$.

Ejemplo 30 (operadores simétricos en espacios de Hilbert). Un operador A de un espacio de Hilbert $(H, \langle, \rangle)$ en sí mismo se llama simétrico si se cumple

$$\langle Ax, y \rangle = \langle x, Ay \rangle \quad (117)$$

para todo par x, y de elementos de H . No es difícil comprobar que todo operador simétrico es lineal (se deja de ejercicio al lector).

Un resultado mucho más profundo y que se deriva del llamado teorema del grafo cerrado a que hicimos referencia en la introducción de este capítulo, dice que todo operador simétrico definido sobre un espacio de Hilbert es continuo. Este resultado no será estudiado en este libro pues se sale del marco de nuestros intereses, pero lo enunciamos pues nos muestra una vez más la relación intrínseca que existe entre conceptos de origen diferente como son la continuidad (de carácter topológico) y la simetría (de carácter geométrico).

Si se considera en calidad de espacio de Hilbert, los espacios euclidianos de dimensión finita $\mathbb{R}^n$, no resulta difícil verificar que a los operadores simétricos de $\mathbb{R}^n$ en sí mismo, corresponden las matrices simétricas de orden $n \times n$, es decir, aquellas tales que $a_{ij} = a_{ji}$, para todo par i, j con $1 \leq i, j \leq n$.

Ejemplo 31 (operadores unitarios). Son aquellos operadores definidos de un espacio normado $(E, \|\cdot\|)$ sobre sí mismo y tales que

$$\|Ax\| = \|x\| \quad (118)$$

para todo $x \in E$. Si E es además euclidiano y A lineal, no es difícil comprobar que

$$\langle Ax, Ay \rangle = \langle x, y \rangle \quad (119)$$

para todos $x, y \in E$ (verifíquelo).

El lector puede comprobar que en el caso en que $E = \mathbb{R}^n$, a los operadores lineales unitarios corresponden las matrices de orden $n \times n$ que tiene determinante a igual a 1.

Ejemplo 32 (operadores integrales). Sea $E = C([0, 1], \mathbb{R})$ y definamos el operador $A : E \longrightarrow E$ mediante

$$Ax = \int_0^1 K(s, t) x(t) dt, \quad (120)$$

donde la función K es continua en el cuadrado unidad $0 \leq s \leq 1, 0 \leq t \leq 1$.

El lector puede verificar que el operador A es lineal a partir de las propiedades de la integral de Cauchy para funciones continuas.

Ejemplo 33 (funcional lineal de evaluación puntual). Sea E el espacio vectorial del ejemplo 32; un simple cálculo nos muestra que para cada $t_0 \in [0, 1]$ fijo la aplicación

$$\delta_{t_0} : E \longrightarrow \mathbb{R}$$

definida por

$$\delta_{t_0}(x) = x(t_0), \quad (121)$$

es un funcional lineal, llamado *evaluación puntual en el punto t_0* .

Ejemplo 34 (funcional lineal integral). Sea E el espacio vectorial del ejemplo 32 y sea f una función en E . El lector puede verificar que la aplicación

$$l_f : E \longrightarrow \mathbb{R}$$

definida por

$$l_f(x) = \int_0^1 f(t) x(t) dt, \quad (122)$$

es un funcional lineal sobre E .

En lo adelante supondremos que los espacios vectoriales E y F están provistos de sendas estructuras de espacios normados, y estudiaremos algunas propiedades y caracterizaciones de los operadores lineales continuos de E y F .

Comencemos por la propiedad elemental siguiente:

Proposición 6.7.3. Todo operador aditivo y continuo entre dos espacios normados reales es necesariamente homogéneo, y por lo tanto, lineal.

Demostración: En 6.7.2 iii) probamos que cada operador aditivo satisface (112) para todo número racional λ . Si λ es irracional, entonces por la densidad de $\mathbb{Q}$ en $\mathbb{R}$, podemos hallar una sucesión (r_n) de números racionales que $r_n \longrightarrow \lambda$. Pero entonces, para cualquier $x \in E$, $r_n x \longrightarrow \lambda x$, y por la continuidad de A junto con la veracidad de (112) para los números racionales, se tiene

$$A(\lambda x) = \lim A(r_n x) = \lim r_n A(x) = \lambda A(x)$$

lo cual prueba la homogeneidad del operador A . ■

Para los operadores lineales en espacios de dimensión finita se tiene el resultado importante siguiente.

Teorema 6.7.4. Todo operador lineal que actúa entre dos espacios normados de dimensión finita es continuo.

Demostración: Sea $A : E \longrightarrow F$ un operador lineal entre los espacios vectoriales E y F de dimensión finita m y n respectivamente. Sea $\{e_1, \dots, e_n\}$ una base de E . Entonces cada $x \in E$ tiene una expresión única del tipo

$$x = \sum_{k=1}^m \lambda_k e_k,$$

y por lo tanto,

$$A(x) = \sum_{k=1}^m \lambda_k A(e_k).$$

Si ahora la sucesión $(x^{(p)}) \xrightarrow{p \rightarrow \infty} x$ en E , donde

$$x^{(p)} = \sum_{k=1}^m \lambda_k^{(p)} A(e_k),$$

entonces, siguiendo la idea de la demostración de la proposición 6.2.9, se tiene que $\lambda_k(p) \longrightarrow \lambda_k$ cuando $p \longrightarrow \infty$, para todo $k = 1, \dots, m$.

Por lo tanto,

$$A(x^{(p)}) = \sum_{k=1}^m \lambda_k^{(p)} A(e_k) \longrightarrow \sum_{k=1}^m \lambda_k A(e_k) = A(x),$$

donde la flecha en este caso denota la convergencia en F , con lo cual se prueba la continuidad de A . ■

Cuando introducimos la noción de continuidad en espacios métricos cualesquiera, comenzamos por dar la definición de continuidad puntual y posteriormente la de continuidad global. Para los operadores lineales en espacios normados veremos que es suficiente la continuidad en un punto para que se tenga la continuidad global.

Proposición 6.7.5. Un operador lineal $A : E \longrightarrow F$ es continuo (globalmente) si y solo si es continuo al menos en un punto $x_0 \in E$.

Demostración: La implicación $\Rightarrow$ se cumple obviamente. Para la demostración de la implicación en sentido contrario supongamos que A es continuo en el punto $x_0 \in E$.

Sea x otro punto arbitrario de E y $(x^{(n)})$ una sucesión en E que converge a x por la continuidad de la suma se tiene que $x_n - x + x_0 \longrightarrow x_0$. Por otra

parte, como hemos supuesto la continuidad de A en x_0 , tendremos que

$$A(x_n - x + x_0) = A(x_n) - A(x) + A(x_0) \longrightarrow A(x_0). \quad (123)$$

Pero basándonos de nuevo en la continuidad de la suma, de (123) se deduce que

$$A(x_n) - A(x) \longrightarrow 0,$$

es decir, $A(x_n) \longrightarrow A(x)$,

de donde concluimos la continuidad de A en un punto arbitrario $x \in E$. ■

Pasemos a estudiar la relación entre los conceptos *continuidad* y *acotación* para operadores y funcionales lineales.

Observación: A partir de este momento, en algunos casos, denotaremos indistintamente por $\|\cdot\|$ las normas en todos los espacios normados a que haremos referencia, salvo si se indica lo contrario.

Sabemos que un conjunto B de un espacio normado $(E, \|\cdot\|)$ está acotado si existe una constante $M > 0$ tal que $\|x\| \leq M$ para todo $x \in B$.

No es difícil demostrar que todo operador lineal continuo transforma conjuntos acotados en conjuntos acotados. En efecto, supongamos que $A : E \longrightarrow F$ es un operador lineal para el cual existe un conjunto acotado B en E tal que $A[B]$ no es un subconjunto acotado de F . Entonces para cada $n \in \mathbb{N}$ se puede encontrar un elemento $x_n \in B$ tal que

$$\|A(x_n)\| \geq n. \quad (124)$$

Como B es acotado, existe una constante $M > 0$ tal que

$$\|x_n\| \leq M \text{ para todo } n \in \mathbb{N}. \quad (125)$$

De (125) se deduce obviamente que la sucesión $\frac{x_n}{n}$ converge a cero en E . Pero

de (124) se tiene que $\left\|A\left(\frac{x_n}{n}\right)\right\| \geq 1$, y por lo tanto,

$$A\left(\frac{x_n}{n}\right) \not\rightarrow A(\theta) = \theta, \quad (126)$$

luego A no sería continuo, con lo cual llegamos a una contradicción.

Más aún, se tiene el teorema siguiente:

Teorema 6.7.6 (relación entre continuidad y acotación). Para un operador lineal $A : (E, \|\cdot\|_1) \longrightarrow (F, \|\cdot\|_2)$ son equivalentes:

- i) A es continuo.
- ii) A transforma conjuntos acotados de E en conjuntos acotados de F .
- iii) $A[B_E]$ es un conjunto acotado de F , donde $B_E = \{x \in E : \|x\|_1 \leq 1\}$ es la bola unidad de E .
- iv) Existe una constante $M > 0$ tal que $\|Ax\|_2 \leq M\|x\|_1$ para todo $x \in E$.

Observación: La propiedad de transformar conjuntos acotados en acotados, caracteriza solamente la continuidad en la clase de los operadores lineales. Pueden darse ejemplos triviales de operadores no lineales que transforman conjuntos acotados en acotados y que no son continuos en ningún punto. Basta considerar la llamada función de Dirichlet

$$D : \mathbb{R} \longrightarrow \mathbb{R}$$

que le asocia el elemento 0 a todo número racional y el 1 a todo irracional. Por este motivo también conviene en llamar a los *operadores lineales continuos*, *operadores lineales acotados*.

Demostración: i) $\Rightarrow$ ii) Ya ha sido demostrado anteriormente.

ii) $\Rightarrow$ iii) Es evidente.

iii) $\Rightarrow$ iv) Supongamos que no existe una constante $M > 0$ tal que se cumpla (125); entonces para cualquier entero positivo n existe un elemento $x_n \in E$ tal que

$$\|A(x_n)\|_2 > n\|x_n\|_1. \quad (127)$$

Si consideramos

$$x'_n = \frac{x_n}{\|x_n\|_1}, \quad (128)$$

entonces $x'_n \in B_E$ y (127) contradice iii).

iv) $\Rightarrow$ i) De (126) se concluye inmediatamente la continuidad de A en $x_0 = \theta$ y por la proposición 6.7.5 se concluye i). ■

Observación: La caracterización iv) del teorema 6.7.6 para el caso de los funcionales lineales continuos, se expresa sustituyendo el módulo usual (en $\mathbb{R}$ o $\mathbb{C}$) en lugar de $\|\cdot\|_2$ en (126).

Pasemos a ver algunos ejemplos de operadores y funcionales lineales continuos.

Ejemplo 35 (matrices). Del teorema 6.7.4 se deduce que toda matriz real (compleja) de orden $n \times m$ (114) define un operador lineal continuo de $\mathbb{R}^n$ en $\mathbb{R}^m$ (de $\mathbb{C}^n$ en $\mathbb{C}^m$).

Ejemplo 36 (operadores de proyección ortogonal). Si M es un subespacio vectorial cerrado de un espacio de Hilbert H , entonces el operador de proyección ortogonal correspondiente P_M es continuo. En efecto, para todo $x \in H$

$$x = P_M(x) + (x - P_M(x)),$$

donde $P_M(x) \perp (x - P_M(x))$. Por el teorema de Pitágoras generalizado

$$\|x\|^2 = \|P_M(x)\|^2 + \|x - P_M(x)\|^2,$$

de donde

$$\|P_M(x)\| \leq \|x\|. \quad (129)$$

De (129) y del punto iv) del teorema 6.7.6 se deduce la continuidad de P_M .

Ejemplo 37 (operadores unitarios). De (118) y del punto iv) del teorema 6.7.6 se concluye inmediatamente que todo operador unitario es continuo.

Ejemplo 38 (operadores integrales). Sea $E = C([0, 1], \mathbb{R})$ y $A : E \rightarrow E$ un operador integral definido como en el ejemplo 32 mediante (120). Se deja al lector comprobar que, con la hipótesis de continuidad impuesta en el ejemplo 32 al núcleo $K(s, t)$ del operador integral A , $A(x) \in E$ para todo $x \in E$.

Tomemos $x \in E$ y computemos la expresión $\|A(x)\|_\infty$.

$$\begin{aligned} \|A(x)\|_\infty &= \max_{0 \leq s \leq 1} \left| \int_0^1 K(s, t) x(t) dt \right| \leq \\ &\leq \max_{0 \leq s \leq 1} \left\{ \int_0^1 |K(s, t)| dt \right\} \max_{0 \leq t \leq 1} |x(t)| = \\ &= \max_{0 \leq s \leq 1} \left\{ \int_0^1 |K(s, t)| dt \right\} \|x\|_\infty. \end{aligned} \quad (130)$$

Luego por iv) de 6.7.6 hemos demostrado que el operador lineal

$$A : (E, \|\cdot\|_\infty) \rightarrow (E, \|\cdot\|_\infty)$$

es continuo.

Ejemplo 39 (funcional lineal de evaluación puntual). Sea $(E, \|\cdot\|_\infty)$ el espacio normado del ejemplo 38. Entonces, para cada $t_0 \in [0, 1]$ fijo, la evaluación puntual δ_{t_0} , definida en (121), es un funcional lineal continuo sobre $(E, \|\cdot\|_\infty)$. En efecto, para cada $x \in E$

$$|\delta_{t_0}(x)| = |x(t_0)| \leq \|x\|_\infty. \quad (131)$$

Ejemplo 40 (funcional lineal integral). Consideremos de nuevo el espacio normado $(E, \|\cdot\|_\infty)$ del ejemplo 38. Para cada $f \in E$ fija, el funcional lineal l_f definido mediante (122) es continuo, ya que

$$l_f(x) \leq \int_0^1 |f(t)| dt \|x\|_\infty \quad (132)$$

Notemos que del teorema 6.7.6 iv) se deduce que para todo operador lineal y continuo

$$A : (E, \|\cdot\|_1) \longrightarrow (F, \|\cdot\|_2),$$

la cantidad

$$\|A\| := \sup_{x \in E, x \neq \theta} \frac{\|A(x)\|_2}{\|x\|_1} \quad (133)$$

es un número positivo.

Precisamente $\|A\|$ coincide con el ínfimo de todas las constantes $M > 0$ para las cuales tiene lugar el punto iv) del teorema 6.7.6.

El lector puede comprobar sin dificultad que

$$\|A\| := \sup_{x \in E, \|x\|_1 \leq 1} \|A(x)\|_2 = \sup_{x \in E, \|x\|_1 = 1} \|A(x)\|_2 \quad (134)$$

Consideremos ahora el espacio $L(E, F)$ de todos los operadores lineales y continuos (acotados) de E en F . Este espacio puede proveerse de una estructura natural de espacio vectorial mediante la suma y el producto por escalares usual de operadores.

Si

$$A, B \in L(E, F) \text{ y } \lambda \in K \text{ (} K = \mathbb{R} \text{ o } \mathbb{C}\text{),}$$

entonces,

$$(A + B)(x) := A(x) + B(x)$$

$$(\lambda A)(x) := \lambda A(x).$$

Teorema 6.7.7. Para todo par de espacios normados $(E, \|\cdot\|_1)$ y $(F, \|\cdot\|_2)$ el espacio vectorial de los operadores lineales y continuos de E en F provisto de la aplicación

$$\|\cdot\| : L(E, F) \xrightarrow{A \mapsto \|A\|} \mathbb{R}_+ \quad (135)$$

definida por (133), es un espacio normado.

Demostración: Se deja al lector verificar, en calidad de ejercicio, que la aplicación (135) satisface los requerimientos de una norma. ■

Ejemplo 41 (norma de un operador de proyección ortogonal). Sea M un operador vectorial cerrado en un espacio de Hilbert H y P_M el correspondiente operador de proyección ortogonal sobre H . De (129) se deduce que

$$\|P_M\| \leq 1,$$

además si $x \in M$ entonces

$$P_M(x) = x,$$

y por lo tanto,

$$\|P_M\| = 1. \quad (136)$$

Ejemplo 42 (norma de un operador unitario). De (118) se concluye inmediatamente que todo operador unitario tiene norma 1.

Ejemplo 43 (norma de un operador integral). Consideremos el operador integral

$$A : (E, \|\cdot\|_\infty) \longrightarrow (E, \|\cdot\|_\infty),$$

donde $E = C([0, 1], \mathbb{R})$ el cual fue analizado en el ejemplo 38. De (130) se tiene que

$$\|A\| \leq \max_{0 \leq s \leq 1} \int_0^1 |K(s, t)| dt.$$

Podemos probar que la cantidad que se encuentra en el miembro derecho de esta desigualdad es precisamente la norma de A . En efecto, sea s_0 el punto del intervalo $[0, 1]$ en el cual la función continua

$$\int_0^1 |K(s, t)| dt$$

alcanza su máximo.

Dado $\epsilon > 0$ sea p un polinomio que se aproxima uniformemente a la función $K(s_0, t)$ con un orden menor que ϵ

$$\max_{0 \leq s \leq 1} |K(s_0, t) - p(t)| < \epsilon$$

y sea x una función en $C([0, 1], \mathbb{R})$ con $\|x\|_\infty \leq 1$ y que se aproxima a la función discontinua

$$\operatorname{sgn} p(t) = \begin{cases} 1, & \text{si } p(t) \geq 0 \\ -1, & \text{si } p(t) < 0 \end{cases}$$

en el sentido que

$$\left| \int_0^1 p(t) x(t) dt - \int_0^1 |p(t)| dt \right| < \epsilon.$$

Esta última aproximación se construye fácilmente ya que p tiene solamente un número finito de cambios de signo (los detalles se dejan al lector).

Para este x tenemos

$$\begin{aligned}
 & \left| \int_0^1 K(s_0, t) x(t) dt \right| \geq \left| \int_0^1 p(t) x(t) dt \right| - \\
 & - \left| \int_0^1 [K(s_0, t) - p(t)] x(t) dt \right| \geq \\
 & \geq \left| \int_0^1 p(t) x(t) dt \right| - \epsilon \geq \int_0^1 |p(t)| dt - 2\epsilon \geq \\
 & \geq \int_0^1 |K(s_0, t)| dt - \left| \int_0^1 [|K(s_0, t)| - |p(t)|] dt \right| - 2\epsilon \geq \\
 & \geq \int_0^1 |K(s_0, t)| dt - 3\epsilon.
 \end{aligned}$$

De aquí, como $\|x\|_\infty \leq 1$, tenemos que

$$\|A\| \geq \int_0^1 |K(s_0, t)| dt - 3\epsilon.$$

Pero como ϵ es arbitrario y la desigualdad en sentido inverso fue establecida anteriormente, tenemos que

$$\|A\| = \max_{0 \leq s \leq 1} \int_0^1 |K(s, t)| dt. \quad (137)$$

En los ejercicios de autocontrol el lector podrá ver cómo varía la expresión de la norma de A si se sustituye $\|\cdot\|_\infty$ por otras normas en $C([0, 1], \mathbb{R})$ como espacio de partida o de llegada.

Ejemplo 44 (norma de un funcional lineal de evaluación puntual).

Sea $(E, \|\cdot\|_\infty)$ el espacio normado del ejemplo anterior y sea $t_0 \in [0, 1]$. De (131) se deduce que la evaluación puntual en t_0 es un funcional lineal continuo cuya norma es

$$|\delta_{t_0}(x)| \leq 1.$$

Si tomamos la función constante $x(t) = \langle 1 \rangle$, entonces

$$|\delta_{t_0}(x)| = \|x\|_\infty,$$

y por lo tanto,

$$\|\delta_{t_0}\| = 1.$$

Ejemplo 45 (norma de un funcional lineal integral).

Sea $(E, \|\cdot\|_\infty)$ el espacio normado del ejemplo 43. De (132) se concluye que, para cada $f \in E$,

l_f es un funcional lineal continuo sobre $(E, \|\cdot\|_\infty)$ cuya norma

$$\|l_f\| \leq \int_0^1 |f(t)| dt. \quad (138)$$

Utilizando argumentos más finos del Análisis funcional y de la teoría de integración, que no entraremos a analizar aquí, es posible demostrar que existen funciones reales continuas f definidas sobre $[0, 1]$ para las cuales no se cumple la igualdad en (138). Sin embargo, puede demostrarse que si consideramos el espacio vectorial $E = C([0, 1], \mathbb{R})$ provisto de la norma

$$\|x\|_1 = \int_0^1 |x(t)| dt,$$

entonces $l_f : (E, \|\cdot\|_1) \rightarrow \mathbb{R}$ sigue siendo un funcional lineal continuo, y además,

$$\|l_f\| = \|f\|_\infty. \quad (139)$$

En efecto, l_f es continuo ya que

$$|l_f(x)| \leq \|f\|_\infty \int_0^1 |x(t)| dt = \|f\|_\infty \|x\|_1.$$

De aquí se deduce que

$$\|l_f\| \leq \|f\|_\infty. \quad (140)$$

La demostración de la igualdad en (140) se deja de ejercicio al lector.

Después de haber analizado los ejemplos anteriores, estudiaremos algunas propiedades del espacio normado $(L(E, F), \|\cdot\|)$.

Teorema 6.7.8 (límite en norma de operadores lineales continuos).

El espacio normado $(L(E, F), \|\cdot\|)$ es cerrado con respecto al límite definido por la norma $\|\cdot\|$. Es decir, si (A_n) es una sucesión de operadores lineales continuos de E en F , y A es un operador de E en F tal que

$$\|A_n - A\| \xrightarrow{n \rightarrow \infty} 0, \quad (141)$$

entonces $A \in L(E, F)$.

Observación: La expresión $\|A_n - A\|$ tiene sentido para cualquier operador $A : E \rightarrow F$, aún no siendo lineal.

Demostración: La verificación de la linealidad de A se deja de ejercicio al lector.

Pasemos a probar la continuidad de A y para ello, utilicemos el teorema 6.7.6 iv)

Por ser $|||\cdot|||$ una norma, tendremos que para todo entero positivo n

$$|||A||| \leq |||A - A_n||| + |||A_n|||. \quad (142)$$

Sea $N \in \mathbb{N}$ fijo suficientemente grande tal que $|||A - A_N|||$ (la existencia de N está garantizada por (141)). Por ser A_N un operador lineal continuo se tendrá que $|||A_N||| \leq M$ para alguna constante $M > 0$. Pero entonces, de (142) se deduce que

$$|||A||| \leq M + 1$$

y por 6.7.6 iv) concluimos que A es continuo. ■

Antes de pasar a la próxima propiedad del espacio $L(E, F)$ veamos una propiedad muy importante de los operadores lineales continuos conocida con el nombre de teorema de S. Banach-H. Steinhauss (teorema 6.7.10), y que como sabemos del capítulo 2, no se cumple para aplicaciones continuas cualesquiera. Comencemos por el lema auxiliar siguiente:

Lema 6.7.9. Si la sucesión de operadores lineales continuos (A_n) converge puntualmente a algún operador A y las normas $|||A_n|||$ están uniformemente acotadas con respecto a n , entonces el operador límite A también es lineal y continuo.

Demostración: Sean A_n y A operadores de $(E, \|\cdot\|_1)$ en $(F, \|\cdot\|_2)$. La linealidad de A se deduce de las siguientes relaciones, válidas para cualesquiera $x, y \in E$

$$\begin{aligned} A(\lambda x + \mu y) &= \lim A_n(\lambda x + \mu y) = \\ &= \lim [\lambda A_n(x) + \mu A_n(y)] = \\ &= \lambda \lim A_n(x) + \mu \lim A_n(y) = \lambda A(x) + \mu A(y). \end{aligned}$$

Teniendo en cuenta que $|||A_n||| \leq M$ para todo n , donde M es una cierta constante positiva, y pasando al límite en la desigualdad

$$\|A_n(x)\|_2 \leq M \|x\|_1,$$

llegamos a que

$$\|A(x)\|_2 \leq M \|x\|_1$$

para todo $x \in E$. De esta última desigualdad se concluye la continuidad de A , debido a 6.7.6 iv). ■

En los espacio de Banach la acotación de las normas de una sucesión de operadores lineales se deduce de su convergencia puntual y el resultado probado en el lema se fortalece sustancialmente.

Teorema 6.7.10 (S. Banach-H. Steinhauss). Si el espacio normado:

$(E, \|\cdot\|_1)$ es de Banach y la sucesión de operadores lineales continuos (A_n) en $L(E, F)$ converge puntualmente al operador A , entonces las normas $\|A_n\|$ están uniformemente acotadas con respecto a n y el operador límite A es también lineal continuo.

Demostración: Como consecuencia del lema 6.7.9 es suficiente demostrar la acotación uniforme de las normas $\|A_n\|$. Supongamos por el absurdo que estas normas no están uniformemente acotadas y probemos primeramente que para cualquier bola cerrada $B \in E$, el conjunto $\bigcup_{n=1}^{\infty} [B]$ no es acotado en F . En efecto, si para alguna esfera $B = B(x_0, r)$ este conjunto fuera acotado, entonces se cumpliría la desigualdad $\|A_n(x)\|_2 \leq C$ para todo $x \in B$ y $n = 1, 2, \dots$, donde C es una cierta constante positiva. Pero, para cualquier $x \in E$ con $\|x\|_1 \leq 1$, podemos considerar $x' = x_0 + rx$ y sabemos que $x' \in B$. Luego, de la relación

$$A_n(x) = A_n\left(\frac{x' - x_0}{r}\right) = \frac{1}{r}[A_n(x') - A_n(x_0)]$$

se tendría que, para cualquier $x \in E$ con $\|x\|_1 \leq 1$

$$\|A_n(x)\|_2 \leq \frac{1}{r}(\|A(x')\|_2 + \|A_n(x_0)\|_2) \leq \frac{2C}{r},$$

lo cual significa que $\|A_n\| \leq \frac{2C}{r}$ para todo $n \in \mathbb{N}$ y esto contradice la suposición de que las normas $\|A_n\|$ no estén uniformemente acotadas.

Una vez probado que el conjunto $\bigcup_{n=1}^{\infty} A_n[B]$ no está acotado, tomemos una bola cerrada fija $B = B_0 = B(x_0, r)$ de radio r en E .

Como una consecuencia de la no acotación de $\bigcup_{n=1}^{\infty} A_n[B_0]$ podemos hallar un número n_1 y un elemento $x_1 \in B_0$, tales que $\|A_{n_1}(x_1)\| > 1$. No existe limitación en considerar a x_1 como un punto inferior de la bola B_0 , ya que habríamos podido considerar otra bola con radio menor que r y seleccionar el punto x_1 en esa otra esfera. Como consecuencia de la continuidad del operador lineal A_{n_1} , existe otra bola cerrada $B_1 = B_1(x_1, r_1)$, en cuyos puntos se cumple que

$$\|A_n(x)\| \geq 1.$$

Como r puede ser arbitrariamente pequeño, podemos asumir que $r_1 \leq r/2$. Además, como x_1 es un punto interior de la bola B_0 , podemos suponer que r es tomado tan pequeño que

$$B_1 \subset B_0.$$

La sucesión de las normas $\|A_n\|$ con $n > n_1$ es también no acotada, y por lo tanto, repitiendo el razonamiento inicial, el conjunto

$$\bigcup_{n=1}^{\infty} A_n[B_1]$$

será de nuevo un conjunto no acotado.

Si repetimos el proceso anterior, podremos hallar un número natural $n_2 > n_1$ y un punto x_2 interior a la bola cerrada B_1 para los cuales

$$\|A_{n_1}(x_2)\|_2 > 2.$$

Posteriormente, podremos hallar una bola cerrada $B_2 \subset B_1$ para la cual se cumple que

$$\|A_{n_1}(x)\|_2 > 2 \text{ para todo } x \in B_2$$

y tal que su radio r satisfaga

$$r_2 \leq r_1/2 \leq r/2^2.$$

Continuando este proceso indefinidamente, obtenemos una sucesión (B_m) de bolas cerradas encajadas de radio r_m donde $r_m \rightarrow 0$ y la sucesión de índices $n_1 < n_2 < \dots < n_m < \dots$ es tal que

$$\|A_{n_m}(x)\|_2 \geq m,$$

para todo $x \in B_m$.

Como E es completo, entonces de acuerdo al principio de bolas encajadas estudiado en el capítulo 3, existe un punto x^* común a todas las bolas B_m . En este punto x^* se cumple que $\|A_{n_m}(x^*)\|_2 \geq m$ para cualquier m , es decir, la sucesión $(A_n(x^*))$ no está acotada en F y esto contradice la existencia de $\lim A_n(x^*)$. Con esta contradicción obtenemos la demostración del teorema.

■

La convergencia en norma para operadores lineales continuos de E en F coincide precisamente con la convergencia uniforme sobre la bola unidad de $E(\{x \in E : \|x\|_1 < 1\})$.

Este tipo de convergencia no se diferencia sustancialmente de la convergencia uniforme definida para las funciones continuas que actúan sobre un espacio métrico compacto, y por ello, no debe resultarnos extraño el haber obtenido un resultado como el enunciado en el teorema 6.7.8.

Es bueno señalar que no tendría sentido definir la convergencia uniforme sobre todo E para operadores lineales continuos; este concepto de convergencia no sería bueno. Por ejemplo, para operadores lineales de $\mathbb{R}$ en $\mathbb{R}$, los cuales se expresan en la forma

$$A_n(x) = \lambda_n x, \lambda_n \in \mathbb{R}$$

(rectas que pasan por el origen) se tiene que la sucesión A_n converge uniformemente al operador idénticamente nulo, si y solo si a partir de un cierto n , todos los λ_n son iguales a cero.

Sin embargo, si bien el resultado 6.7.8 era de esperar, no podemos decir lo mismo del teorema de S. Banach-H. Steinhass que es verdaderamente sorprendente. Recordemos que el límite puntual de una sucesión de aplicaciones continuas no es necesariamente continua. El teorema 6.7.10 nos dice que si el

espacio de partida es completo y las aplicaciones son además lineales, entonces el límite es también una aplicación lineal y continua, aunque no necesariamente se tiene la convergencia en norma.

Después de esta digresión pasemos de nuevo a ver otra propiedad importante del espacio $(L(E, F), \|\cdot\|)$.

Teorema 6.7.11. Sean $(E, \|\cdot\|_1)$ y $(F, \|\cdot\|_2)$ espacios normados. Si $(F, \|\cdot\|_2)$ es de Banach, entonces también lo es $(L(E, F), \|\cdot\|)$.

Demostración: Sea (A_n) una sucesión de Cauchy en $L(E, F)$.

Como para cualquier $x \in E$ se tiene

$$\|A_n(x) - A_m(x)\|_2 = \|(A_n - A_m)(x)\|_2 \leq$$

$$\leq \|A_n - A_m\| \cdot \|x\|,$$

y como además $\|A_n - A_m\| \rightarrow 0$ cuando $n, m \rightarrow \infty$ entonces la sucesión $(A_n(x))$ es una sucesión de Cauchy en F . Por ser F un espacio de Banach, existe el límite

$$\lim A_n(x),$$

para todo $x \in E$.

Como toda sucesión de Cauchy es acotada, entonces $\|A_n\|$ está acotado uniformemente por una constante $M > 0$. De acuerdo al Lema 6.7.9, el operador A definido por la fórmula

$$A(x) = \lim A_n(x),$$

pertenece a $L(E, F)$.

Probemos que (A_n) converge a A con respecto a la norma $\|\cdot\|$. Para un $\epsilon > 0$ arbitrario, podemos seleccionar $N \in \mathbb{N}$ tal que $\|A_n - A_m\| < \epsilon$ para $n, m > N$. Esta desigualdad significa que

$$\|A_n(x) - A_m(x)\|_2 < \epsilon \|x\|_1, \quad (143)$$

para todo $x \in E$. Pasando al límite en (143) cuando $m \rightarrow \infty$, tendremos que

$$\|A_n(x) - A(x)\|_2 \leq \epsilon \|x\|_1,$$

de donde se deduce que

$$\|A_n - A\| \leq \epsilon$$

o lo que es lo mismo: A_n converge a A en $(L(E, F), \|\cdot\|)$. Con esto queda probado que $(L(E, F), \|\cdot\|)$ es completo. ■

Utilizando el teorema de convergencia 6.7.8 se obtiene el siguiente resultado que es muy utilizado en la resolución de ecuaciones operacionales.

Teorema 6.7.12. Sea E un espacio de Banach, I el operador identidad de E y A un operador lineal continuo que aplica E en sí mismo y tal que $\|A\| < 1$. Entonces existe el operador inverso $(I - A)^{-1}$ el cual es lineal y continuo y se representa en la forma

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k, \quad (144)$$

donde mediante A^k se denota la composición de A consigo mismo k veces y la serie converge en la norma $\|\cdot\|$ del espacio de Banach $L(E, E)$.

Demostración: Demostraremos la existencia y acotación del operador $(I - A)^{-1}$ con lo cual quedará probada su linealidad ya que el inverso de cualquier operador lineal es también un operador lineal (compruébelo).

Como $\|A\| < 1$, tenemos que $\sum_{k=0}^{\infty} \|A^k\| \leq \sum_{k=0}^{\infty} \|A\|^k < \infty$.

Por la completitud de E , deducimos de 6.7.11 y 6.4.4 que la serie $\sum_{k=0}^{\infty} A^k$ converge en $(L(E, E), \|\cdot\|)$, y por lo tanto, su suma representa un operador lineal acotado. Para $n \in \mathbb{N}$ cualquiera tenemos

$$(I - A) \sum_{k=0}^n A^k = \sum_{k=0}^n A^k (I - A) = I - A^{n+1}$$

Pasando al límite cuando $n \rightarrow \infty$ y tomando en consideración que: $\|A^{n+1}\| \leq \|A\|^{n+1} \rightarrow 0$, encontramos

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

de donde,

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

que es lo que queríamos demostrar. ■

Hasta el momento hemos estudiado la teoría general de operadores en espacios normados. Dedicemos la última parte de este epígrafe a hablar con más detalle del caso particular de los funcionales lineales continuos que corresponden en los espacios métricos al estudio de las funciones reales continuas. Haremos énfasis fundamentalmente en tres temas:

1. Representación de los funcionales lineales continuos en un espacio de Hilbert (teorema de Riesz).

2. Relación entre los operadores y funcionales lineales en un espacio de Hilbert mediante los funcionales bilineales.
3. Teorema de Hahn-Banach.

Mediante este tercer tema veremos la relación entre los problemas siguientes:

- a) Extensión de funcionales lineales continuos en espacios normados.
- b) Extencia de funcionales lineales continuos en espacio normados.
- c) Separación de conjuntos convexos en espacios normados mediante hiperplanos cerrados.

Si E es un espacio vectorial sobre el cuerpo K ($K = \mathbb{R}$ o $\mathbb{C}$), entonces se llama *dual algebraico* de E y se denota por $A[E]$ al conjunto de todos los funcionales lineales definidos sobre E . Es obvio que $A[E]$ puede también proveerse de una estructura de espacio vectorial con respecto a la suma de funcionales y al producto usual por escalares; en efecto, si f y g son elementos de $A[E]$ y $\lambda, \mu \in K$, entonces, $\lambda f + \mu g \in A[E]$,

donde

$$(\lambda f + \mu g)(x) = \lambda f(x) + \mu g(x)$$

para todo $x \in E$.

Si suponemos que sobre E tenemos una estructura adicional de espacio normado, entonces los funcionales lineales y continuos definidos sobre E constituyen un subespacio vectorial de $A[E]$

Definición 6.7.13. Se llama *dual topológico* del espacio normado $(E, \|\cdot\|)$ al conjunto de todos los funcionales lineales continuos definidos sobre E y se denota por E^* .

Observación: Como ya hemos planteado, E^* es un subespacio vectorial de $A[E]$ y que no necesariamente coincide con $A[E]$ como veremos en el ejemplo 47. Obviamente E^* coincide con el espacio vectorial $L(E, K)$ ($K = \mathbb{R}$ o $\mathbb{C}$)

$$E^* = L(E, K)$$

Como K es completo, del teorema 6.7.11 concluimos que E^* es siempre completo con respecto a la norma $\|\cdot\|$:

$$\|f\| = \sup_{x \in E, x \neq \theta} \frac{|f(x)|}{\|x\|}.$$

Ejemplo 46. De 6.7.4 se deduce que para todo espacio normado de dimensión finita $A[E] = E^*$.

Ejemplo 47. Consideremos el espacio vectorial real $E = C([0, 1], \mathbb{R})$, provisto de la norma

$$\|x\|_1 = \int_0^1 |x(t)| dt.$$

Entonces la evaluación puntual $\delta_0 (t_0 \in [0, 1])$ es un funcional lineal sobre $(E, \|\cdot\|_1)$ que no es continuo. En efecto, tomemos por ejemplo $t_0 = 0$. La sucesión de funciones continuas

$$x_n(t) = \begin{cases} nt - 1, & \text{si } t \in [0, 1/n] \\ 0 & \text{si } t \in [1/n, 1] \end{cases}$$

converge a cero en $(E, \|\cdot\|_1)$. Sin embargo, $\delta_0(x_n) \not\rightarrow 0$, y por lo tanto, $\delta_0 : (E, \|\cdot\|_1) \rightarrow \mathbb{R}$ no es un funcional lineal continuo.

Pasemos ahora al estudio de la representación de los funcionales lineales continuos en un espacio de Hilbert. Por la continuidad del producto escalar en un espacio euclideo cualquiera H , sabemos que para todo $\alpha \in H$ el funcional lineal $f_\alpha : H \rightarrow K (K = \mathbb{R} \text{ o } \mathbb{C})$

$$f_\alpha : x \rightsquigarrow \langle x, \alpha \rangle \quad (145)$$

es continuo. Por otra parte, de la desigualdad de Cauchy-Buniakowsky

$$|\langle x, \alpha \rangle| \leq \|x\| \cdot \|\alpha\|,$$

de donde se deduce que

$$\|f_\alpha\| \leq \|\alpha\|.$$

Además, la relación $f_a(a) = \|a\|^2$ nos prueba que

$$\|f_a\| = \|a\|. \quad (146)$$

Recíprocamente se tiene:

Teorema 6.7.14 (teorema de Riesz). Sea $(H, \langle \cdot, \cdot \rangle)$ un espacio de Hilbert; entonces para todo funcional lineal continuo $f \in H^*$ existe un único vector $a \in H$ tal que

$$f(x) = f_a(x) = \langle x, a \rangle, (x \in H). \quad (147)$$

Más aún $\|f\| = \|a\|$ y todo $a \in H$ determina un único funcional lineal continuo mediante (147). En otras palabras, para todo espacio de Hilbert H , la aplicación

$$H \underset{a \rightsquigarrow f_a}{\rightsquigarrow} H^* \quad (148)$$

define una isometría de H sobre H^* .

Demostración: La segunda parte del teorema se deduce del comentario previo a dicho teorema, si logramos demostrar la existencia de $a \in H$ tal que se

verifique (147). Pasemos entonces a demostrar (147). Sea $f \in H^*$ y denotemos por H_0 al conjunto de todos los $y \in H$ tales que $f(y) = \theta$

Por la linealidad de f , H_0 será un subespacio vectorial de H . Además, por ser f continuo H_0 es un subespacio vectorial cerrado, en efecto,

$$H_0 = f^{-1}[\{0\}].$$

Si $H = H_0$, entonces $f \equiv \theta$ y el teorema estaría probado tomando $a = \theta$. Si $H \neq H_0$ entonces, de acuerdo al teorema 6.6.9, podemos escribir $H = H_0 \oplus H_0^\perp$. Como $H \neq H_0$ existe un vector no nulo $z \in H_0^\perp$. Como $z \notin H_0$, necesariamente $f(z) \neq 0$. Del hecho que $H_0^\perp$ es un subespacio vectorial de H , podemos elegir z de forma tal que $f(z) = 1$.

Probemos que z es un múltiplo escalar del vector a buscado. Dado cualquier $x \in H$ tenemos que $x - f(x)z \in H_0$ ya que $f(x - f(x)z) = f(x) - f(x)f(z) = 0$.

Como $z \in H_0^\perp$ tendremos que

$$\langle x - f(x)z, z \rangle = 0, \text{ es decir,}$$

$$\langle x, z \rangle = f(x) \|z\|^2, \text{ de donde}$$

$$f(x) = \left\langle x, z, \|z\|^{-2} \right\rangle. \text{ Luego, definiendo}$$

$$a = z \|z\|^{-2}, \text{ se deduce que } f(x) = \langle x, a \rangle.$$

El vector a es claramente único, ya que si b es otro vector tal que $f(x) = \langle x, b \rangle$ para todo $x \in H$, entonces

$$\langle x, b \rangle = f(x) = \langle x, a \rangle,$$

y por lo tanto,

$$\langle x, b - a \rangle = 0 \text{ para todo } x \in H.$$

De aquí que $b - a \in H^\perp$ de donde $b - a = \theta$ lo cual implica que $b = a$.

Por el comentario previo al teorema se concluye la demostración. ■

Utilizando el teorema 6.7.14 se puede establecer una relación entre operadores y funcionales lineales continuos en un espacio de Hilbert. Comencemos por la definición siguiente:

Definición 6.7.15. Sea E un espacio vectorial sobre el cuerpo K ($K = \mathbb{R}$ o $\mathbb{C}$). Un *funcional bilineal* sobre E es una aplicación $H : E \times E \rightarrow K$ que cumple:

$$\text{FB1 } h(x + y, z) = h(x, z) + h(y, z)$$

$$\text{FB2 } h(\lambda x, y) = \lambda h(x, y)$$

$$\text{FB3 } h(x, y) = \overline{h(y, x)}$$

para todos $x, y, z \in E$ y $\lambda \in K$, donde la raya sobre h denota el paso al número complejo conjugado en el caso en que $K = \mathbb{C}$.

Se deja al lector demostrar la siguiente proposición análoga a 6.7.5 y 6.7.6.

Proposición 6.7.16. Sea $(E, \|\cdot\|)$ un espacio vectorial normado y h un funcional bilineal sobre E , entonces son equivalentes:

i) h es continuo (globalmente).

ii) h es continuo en un punto $(a, b) \in E \times E$.

iii) Para todo conjunto acotado C en $E \times E$ existe una constante $M > 0$ tal que si $(x, y) \in C$,

$$|h(x, y)| \leq M.$$

iv) Si B es bola unidad en E ($x \in E : \|x\| \leq 1$) entonces h está acotada sobre $B \times B$.

v) Existe una constante $M > 0$ tal que para todo $x, y \in E$

$$|h(x, y)| \leq M \|x\| \|y\|.$$

La demostración se deja de ejercicio al lector.

Indicación: seguir la idea de las demostraciones de 6.7.5 y 6.7.6.

Del punto v) de la proposición anterior se deduce que se puede definir la *norma de un funcional bilineal* h mediante

$$\|h\| = \sup_{x, y \in E, x \neq \theta, y \neq \theta} \frac{|h(x, y)|}{\|x\| \|y\|} \quad (149)$$

Proposición 6.7.17. Sea $(H, \langle \cdot, \cdot \rangle)$ un espacio euclideo sobre el cuerpo K ($K = \mathbb{R} \text{ o } \mathbb{C}$) y A un operador lineal de H en sí mismo. Entonces la aplicación

$$h : H \times H \longrightarrow K,$$

definida por

$$h(x, y) = \langle Ax, y \rangle \quad (150)$$

es un funcional bilineal.

Además h es continuo si y solo si A es continuo y

$$\|h\| = \|A\|. \quad (151)$$

Demostración: La bilinealidad de h es una consecuencia directa de la linealidad de A y de las propiedades del producto escalar. De (150) y la continuidad del producto escalar se prueba inmediatamente que h es continuo si A lo es. Recíprocamente, si h es continuo, por 6.7.16 iv) existe una constante $M > 0$ tal que para todos $x, y \in H$

$$|h(x, y)| = |\langle Ax, y \rangle| \leq M \|x\| \|y\|. \quad \text{Tomando } y = Ax, \text{ tendremos}$$

$$\|A\|^2 \leq M \|x\| \|Ax\|, \quad \text{de donde, } \|Ax\| \leq M \|x\|,$$

y de aquí concluimos la continuidad de A .

La demostración de (151) se deduce inmediatamente de la definición y se deja de ejercicio al lector. ■

El siguiente teorema de representación, es muy importante para la solución de ecuaciones lineales en espacios de Hilbert.

Teorema 6.7.18 (Lax-Milgram). Para cada funcional bilineal continuo h definido sobre un espacio de Hilbert, existe un único operador lineal y continuo $A : H \longrightarrow H$ que *representa* a h en el sentido de (150) y además se cumple (151).

Demostración: Para cada $y \in H$ fijo, la aplicación

$$x \rightsquigarrow h(x, y)$$

es un funcional lineal continuo sobre H .

Por el teorema de representación de Riesz (6.7.14), existe un vector $a \in H$ tal que para todo $x \in H$,

$$h(x, y) = \langle x, a \rangle.$$

Como el vector a está unívocamente determinado por y , podemos establecer la correspondencia

$$y \rightsquigarrow a.$$

Denotaremos al vector a correspondiente a y , por $A(y)$.

Mostraremos que la aplicación así definida

$$A : H \longrightarrow H$$

es un operador lineal. En efecto,

$$h(x, y) = \langle x, A(y) \rangle$$

para todos $x, y \in H$. Luego,

$$h(x, \lambda y + \mu z) = \langle x, A(\lambda y + \mu z) \rangle \quad (152)$$

para todos $x, y, z \in H; \lambda, \mu \in K$ ($K = \mathbb{R}$ o $\mathbb{C}$). Por otra parte, de la bilinealidad de h se tiene que

$$\begin{aligned} h(x, \lambda y + \mu z) &= \overline{\lambda} h(x, y) + \overline{\mu} h(x, z) = \\ &= \overline{\lambda} \langle x, A(y) \rangle + \overline{\mu} \langle x, A(z) \rangle = \\ &= \langle x, \lambda A(y) + \mu A(z) \rangle. \end{aligned} \quad (153)$$

De (152) y (153) se deduce que

$$\langle x, A(\lambda y + \mu z) - \lambda A(y) - \mu A(z) \rangle = 0$$

para todo $x \in H$, de donde se concluye que

$$A(\lambda y + \mu z) = \lambda A(y) + \mu A(z)$$

o sea A es lineal.

La tesis del teorema se obtiene ahora aplicando el resultado de la proposición 6.7.17. ■

Veamos cómo se aplican los resultados sobre representación de funcionales lineales continuos a la solución de ecuaciones en espacios de Hilbert.

Ejemplo 48 (soluciones débiles de ecuaciones operacionales en espacios de Hilbert). Sea A un operador acotado en el espacio de Hilbert $(H, \langle, \rangle)$ y sea H un subespacio vectorial cerrado de H . Diremos que $x \in H$ es una *solución débil* de la ecuación

$$x + A^2x = y \quad (y \in H) \quad (154)$$

con respecto a H , si para todo $z \in H$ se cumple que

$$\langle x, z \rangle + \langle Ax, Az \rangle = \langle y, z \rangle \quad (155)$$

Demostraremos la existencia de soluciones débiles para (154). Para ello consideremos H provisto del nuevo producto escalar

$$[x, y] = \langle x, y \rangle + \langle Ax, Ay \rangle. \quad (156)$$

La comprobación de que (156) define un producto escalar en H se deja de ejercicio al lector. La norma asociada al producto escalar (156) es

$$\|x\|_A = \sqrt{[x, x]} = \sqrt{\|x\|^2 + \|A(x)\|^2},$$

luego, para cada $y \in H$ fijo y todo $z \in H$ se tiene:

$$|\langle y, z \rangle| \leq \|y\| \|z\| \leq \|y\| \|z\|_A.$$

De esta desigualdad se tiene que para cada y fijo, la aplicación

$$z \rightsquigarrow \langle y, z \rangle$$

es un funcional lineal y continuo en el espacio euclideo $(H, [,])$. El lector puede comprobar por la continuidad de A y por ser H cerrado en H , que $(H, [,])$ es también un espacio de Hilbert.

Luego, por el teorema de representación de Riesz 6.7.14, para cada $y \in H$ existe $x \in H$ único tal que

$$[x, z] = \langle y, z \rangle$$

para todo $z \in H$. Esto último es precisamente lo que queríamos demostrar: x es la solución débil de la ecuación (154) con respecto a H .

En la práctica, en ocasiones se trata de elegir el subespacio convenientemente de manera que la solución débil (la cual siempre existe) coincida con la solución de (154) en sentido usual.

Esta técnica de encontrar soluciones de ecuaciones a partir de sus soluciones débiles, es muy utilizada en las teorías de los problemas de contorno para las ecuaciones diferenciales elípticas en derivadas parciales. Para más información al respecto consulte el libro de V.P. Mijailov: *Ecuaciones en Derivadas Parciales*.

Pasemos finalmente al estudio del teorema de Hahn-Banach en sus diferentes versiones: *extensión y geométrica*, con lo cual podremos ver la relación entre los problemas de prolongación lineal continua y separación de convexos por funcionales lineales continuos. Estos problemas corresponden a los ya estudiados en el capítulo 2 para funciones reales continuas en espacios métricos arbitrarios.

El teorema de Hahn-Banach es una de las herramientas fundamentales del Análisis Funcional, y juega un papel muy importante en el estudio de los problemas de optimización en espacios vectoriales.

Comencemos por la llamada forma algebraica del teorema de Hahn-Banach, del cual se deducen los teoremas de extensión de funcionales lineales continuos, pero antes veamos la definición siguiente:

Definición 6.7.19. Sea f un funcional lineal definido sobre un subespacio vectorial M de un espacio vectorial E . Un funcional lineal g se llama una *extensión lineal* de f , si g está definido sobre un subespacio vectorial N de E que contiene a M y además, la restricción de g a M (g/M) coincide con f . En este caso se dice que g es una extensión de f desde M hasta N .

Dicho en forma simple, el teorema de Hahn-Banach nos asegura que todo funcional lineal continuo f definido sobre un subespacio vectorial M de un espacio normado E , puede ser extendido a un funcional lineal continuo g definido sobre todo el espacio E y con la misma norma que la norma de f sobre M es decir,

$$\|g\| = \|f\|_M = \sup_{m \in M} \frac{|f(m)|}{\|m\|}.$$

Nosotros veremos una versión más general de este resultado en el cual se reemplazan las normas por funcionales sublineales. Esta generalización nos servirá más adelante para probar la versión geométrica del teorema de Hahn-Banach.

Definición 6.7.20. Una función real p definida sobre un espacio vectorial real E se llama *funcional sublineal* si satisface:

$$\text{SL1)} \quad p(x + y) \leq p(x) + p(y) \text{ para todos } x, y \in E.$$

$$\text{SL2)} \quad p\lambda x = \lambda p(x), \text{ para todo } \lambda \geq 0 \text{ y } x \in E.$$

Evidentemente, cualquier norma es un funcional sublineal.

Teorema 6.7.21 (teorema de Hahn-Banach en la forma de extensión). Sea E un espacio vectorial normado real y p un funcional sublineal

continuo sobre E . Sea f un funcional lineal definido sobre un subespacio vectorial M de E y que satisfice

$$f(m) \leq p(m)$$

para todo $m \in M$. Entonces existe una extensión g de f a todo E tal que

$$g(x) \leq p(x)$$

sobre todo E .

Demostración: Aunque el teorema es válido en cualquier espacio normado, en nuestra demostración asumiremos que E es separable. El resultado general es obtenido exactamente por el mismo método mediante una simple aplicación del lema de Zorn. El lector familiarizado con el lema de Zorn (vea capítulo 1) no debe tener mucha dificultad para generalizar la prueba. La idea básica es extender f en una dimensión cada vez y aplicar la hipótesis de inducción.

Supongamos que y es un vector de E que no está en M . Consideremos todos los elementos del subespacio vectorial generado por M y y , es decir, $[M + y]$. Un elemento x en $[M + y]$ tiene una representación única de la forma $x = m + \lambda y$, donde $m \in M$ y λ es un número real. Toda extensión f_0 de f desde M hasta $[M + y]$ tiene la forma

$$f_0(x) = f(m) + \lambda f_0(y),$$

y por lo tanto f_0 está totalmente determinada por la constante $f_0(y)$.

Nosotros debemos probar que esta constante puede ser seleccionada de manera tal que $f_0(x) \leq p(x)$ sobre $[M + y]$.

Pero para dos elementos cualesquiera m_1 y m_2 en M , tenemos

$$\begin{aligned} f(m_1) + f(m_2) &= f(m_1 + m_2) \leq p(m_1 + m_2) \leq \\ &\leq p(m_1 - y) + p(m_2 + y), \end{aligned}$$

de donde

$$f(m_1) - p(m_1 - y) \leq p(m_2 + y) - f(m_2).$$

Luego,

$$\sup_{m \in M} [f(m) - p(m - y)] \leq \inf_{m \in M} [p(m + y) - f(m)].$$

Por lo tanto, existe una constante c tal que

$$\sup_{m \in M} [f(m) - p(m - y)] \leq c \leq \inf_{m \in M} [p(m + y) - f(m)].$$

Entonces, para el vector $x = m + \lambda y \in [M + y]$, definiremos

$$f_0(x) = f(m) + \lambda c.$$

Ahora debemos probar que

$$f_0(m + \lambda y) \leq p(m + \lambda y).$$

Si $\lambda > 0$, entonces

$$\lambda c + f(m) = \lambda \left[c + f\left(\frac{m}{\lambda}\right) \right] \leq$$

$$\leq \lambda p\left(\frac{m}{\lambda} + y\right) \leq \mu p\left(\frac{m}{\mu} - y\right) = p(m + \lambda y).$$

Si $\lambda = -\mu < 0$, entonces,

$$-\mu c + f(m) = \mu \left[-c + f\left(\frac{m}{\mu}\right) \right] \leq$$

$$\leq \mu p\left(\frac{m}{\mu} - y\right) = p(m - \mu y).$$

Luego, $f_0(m + \lambda y) \leq p(m + \lambda y)$ para todo $\lambda \in \mathbb{R}$ y f_0 es una extensión de f desde M hasta $[M + y]$.

Sea $\{x_1, x_2, \dots, x_n \dots\}$ un conjunto numerable denso en E . De este conjunto de vectores seleccionamos un subconjunto $\{y_1, y_2, \dots, y_n, \dots\}$ formado por vectores linealmente independientes y que no pertenezcan a M . El conjunto $\{y_1, y_2, \dots, y_n, \dots\}$ junto con el subespacio M generan un subespacio vectorial S denso en E . El funcional f puede ser extendido a un funcional g_0 sobre el subespacio S de manera que

$$g_0(x) \leq p(x)$$

para todo $x \in S$. Para ello basta extender f desde M hasta $[M + y_1]$ después hasta $[[M + y_1] + y_2]$ y así sucesivamente aplicando la hipótesis de inducción.

Finalmente, el funcional lineal g_0 (el cual es continuo ya que p lo es) puede ser extendido por continuidad desde el subespacio denso S a todo el espacio E . Para cada $x \in E$ existe una sucesión (s_n) en S que converge a x .

La extensión g buscada viene dada por

$$g(x) = \lim g_0(s_n).$$

Obviamente g es lineal y por la continuidad de p se tendrá que

$$g(x) \leq p(x)$$

para todo $x \in E$, con lo cual queda demostrado el teorema. ■

Corolario 1 (extensión de funcionales lineales continuos en espacios normados). Sea f un funcional lineal continuo definido sobre un subespacio vectorial M de un espacio normado real E . Entonces, existe un funcional lineal continuo g definido sobre todo E que es una extensión de f y cuya norma es igual a la norma de f sobre M .

Para la demostración basta tomar $\|f\| \|x\|$ y aplicar el teorema de Hahn-Banach.

Corolario 2 (existencia de elementos no nulos en E^*). Sea x un elemento de un espacio normado $(E, \|\cdot\|)$. Entonces existe un funcional lineal

continuo g sobre E tal que

$$g(x) = \|x\|$$

Demostración: Supongamos que $x \neq \theta$. Sobre el subespacio unidimensional generado por el vector x , definamos el funcional lineal f mediante

$$f(\lambda x) = \lambda \|x\|.$$

Entonces f es un funcional lineal acotado con norma igual a uno. Por el corolario 1 del teorema 6.7.21, f puede ser extendido a un funcional lineal continuo g sobre E con norma unidad. Este funcional satisface los requerimientos del corolario. Si $x = \theta$ nos sirve cualquier funcional continuo sobre E (cuya existencia acabamos de demostrar). ■

Observación: El corolario 2 del teorema 6.7.21 nos muestra la riqueza en elementos del espacio dual correspondiente a cualquier espacio normado; este resultado es importante ya que muchos problemas en los espacios normados se pueden plantear de manera equivalente en los correspondientes duales. Precisamente, gracias a la riqueza del espacio dual, a veces resulta más fácil resolver el problema dual y después de obtenida la solución pasar de nuevo al espacio original.

Para finalizar este capítulo, veremos el problema de separación de conjuntos convexos en espacios normados mediante funcionales lineales continuos. La base de este tipo de resultado se encuentra también en el teorema de Hahn-Banach pero en su forma geométrica.

Existe una diferencia conceptual entre los funcionales lineales vistos como elementos de un espacio dual o vistos como hiperplanos generados en el espacio primario. Estos puntos de vista diferentes, nos permiten relacionar los aspectos más relevantes de un espacio y su dual mediante una interpretación geométrica de dichos espacios.

Definición 6.7.22. Un hiperplano H en un espacio vectorial E es un conjunto de la forma $H = x_0 + M$, donde M es un subespacio vectorial propio maximal en E , es decir, M es un subespacio vectorial diferente de H y tal que si N es cualquier otro subespacio vectorial que contiene a M , entonces $M = N$ o $N = E$.

Esta definición de hiperplano est hecha sin alguna referencia explícita a los funcionales, ya que esta refleja la interpretación geométrica de un hiperplano. Sin embargo, los hiperplanos están íntimamente relacionados con los funcionales lineales como lo muestran las dos propiedades siguientes:

Proposición 6.7.23. Sea H un hiperplano en el espacio vectorial E ; entonces existe un funcional lineal f sobre E y una constante c tales que

$$H = \{x \in E : f(x) = c\}.$$

Inversamente, si f es un funcional lineal no nulo sobre E , el conjunto $\{x \in E : f(x) = c\}$ es un hiperplano en E .

Demostración: Sea H un hiperplano en E . Entonces H es la traslación de un subespacio vectorial M de E

$$H = x_0 + M.$$

Si $x_0 \notin M$, entonces $[x_0 + M] = E$ y para $x = \lambda x_0 + m$ con $m \in M$, definimos $f(x) = \lambda$.

Entonces $H = \{x \in E : f(x) = 1\}$.

Si $x_0 \in M$ tomamos $x_1 \notin M$, $E = [x_1 + M]$ y $H = M$ y definimos para $x = \lambda x_1 + m$, $f(x) = \lambda$.

Entonces

$$H = \{x \in E : f(x) = 0\}.$$

Inversamente sea f un funcional no nulo sobre E y sea:

$M = \{x \in E : f(x) = 0\}$. Obviamente M es un subespacio vectorial de E . Sea $x_0 \in E$ tal que $f(x_0) = 1$. Entonces para cualquier $x \in E$, $[x - f(x)x_0] = 0$, y por lo tanto, $x - f(x)x_0 \in M$. Luego, $E = [x_0 + M]$, de donde se deduce que M es un subespacio vectorial propio maximal en E . Para cualquier número real c , sea x_1 cualquier elemento para el cual $f(x_1) = c$. Entonces, $\{x \in E : f(x) = c\} = \{x \in E : f(x - x_1) = 0\} = M + x_1$, que es un hiperplano. ■

La proposición 6.7.23 es de carácter general y es aplicable a hiperplanos en un espacio vectorial arbitrario. Un hiperplano H en un espacio normado E , debe ser cerrado o denso en E ya que como H es una variedad lineal maximal y por ser H también una variedad lineal, entonces $\overline{H} = H$ o $\overline{H} = E$. Para nuestro propósito nos interesan fundamentalmente los hiperplanos cerrados en un espacio normado E . Estos hiperplanos corresponden a los funcionales lineales continuos sobre E .

Proposición 6.7.24. Sea f un funcional lineal no nulo sobre el espacio normado E . Entonces el hiperplano $H = \{x \in E : f(x) = c\}$ es cerrado para cada constante c si y solo si f es continuo.

Demostración: Supongamos primeramente que f es continuo; entonces H es cerrado por ser la preimagen mediante f del subconjunto cerrado unitario $\{c\}$.

Inversamente, supongamos que $M = \{x \in E : f(x) = 0\}$ es cerrado.

Sea $E = [x_0 + M]$ y supongamos que $(x_n) \rightarrow x$ en E ; entonces $x_n = \lambda_n x_0 + x_m$ y si denotamos por d la distancia de x_0 a M (la cual es positiva ya que M es cerrado), tendremos

$$|\lambda_n - \lambda| d \leq \|x_n - x\| \rightarrow 0$$

y por lo tanto, $\lambda_n \rightarrow \lambda$. Por otra parte, $f(x_n) = \lambda_n f(x_0) + f(m_n) = \lambda_n f(x_0) \rightarrow \lambda f(x_0) = f(x)$. Luego, f es continuo sobre E . ■

Si f es un funcional lineal no nulo sobre un espacio vectorial E , se puede asociar al hiperplano $H = \{x \in E : f(x) = c\}$ los cuatro conjuntos siguientes:

$$\begin{aligned} &\{x \in E : f(x) \leq c\}, \{x \in E : f(x) < c\} \\ &\{x \in E : f(x) \geq c\}, \{x \in E : f(x) > c\} \end{aligned}$$

llamados *semiespacios* determinados por H . Los dos primeros se llaman *semiespacios negativos* determinados por f y los otros dos *semiespacios positivos*. Si f es continuo, entonces los semiespacios $\{x \in E : f(x) < c\}$ y $\{x \in E : f(x) > c\}$ son abiertos mientras que $\{x \in E : f(x) \leq c\}$ y $\{x \in E : f(x) \geq c\}$ son cerrados.

Como hemos visto hasta el momento, existe una correspondencia entre los hiperplanos cerrados en un espacio normado E y los elementos de su espacio dual E^* . Luego, es de esperar que aquellos resultados en que juegue un papel fundamental el espacio dual E^* puedan ser visualizados en términos de hiperplanos cerrados.

Por este motivo se puede demostrar una versión geométrica del teorema de Hahn-Banach la cual, en su forma más simple, nos dice que dado un conjunto convexo C con interior no vacío y un punto x_0 que no pertenece a $\overset{\circ}{C}$, existe un hiperplano cerrado que contiene a x_0 pero disjunto con $\overset{\circ}{C}$.

Antes de demostrar este importante teorema de separación veamos la definición auxiliar siguiente:

Definición 6.7.25. Sea C un conjunto convexo en un espacio vectorial normado E y supongamos que el elemento nulo θ es un punto interior de C . Entonces el *funcional* de Minkowsky p de C es una función real definido sobre E mediante

$$p(x) = \inf \left\{ r \in \mathbb{R} : \frac{x}{r} \in C, r > 0 \right\}. \quad (157)$$

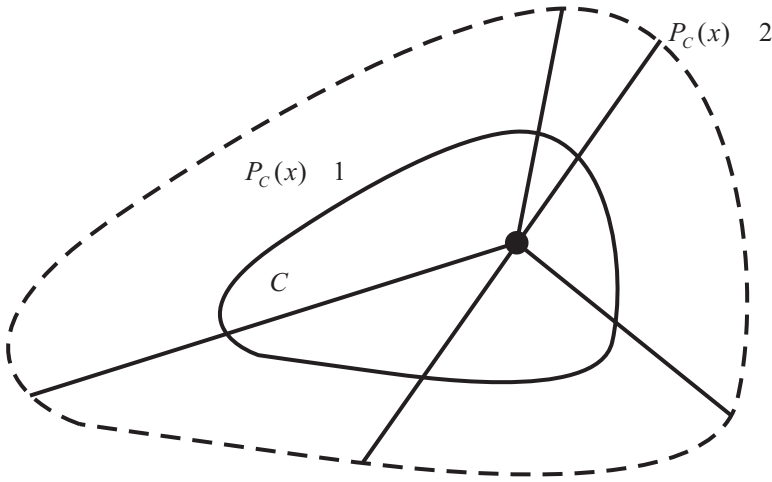


Figura 97

Notemos que para C igual a la bola de E , el correspondiente funcional de Minkowsky es $p(x) = \|x\|$. En el caso general, $p(x)$ define una especie de distancia desde el origen hasta x medida con respecto a C ; $p(x)$ es el factor por el cual debe ser multiplicado C hasta que contenga a x (fig. 97).

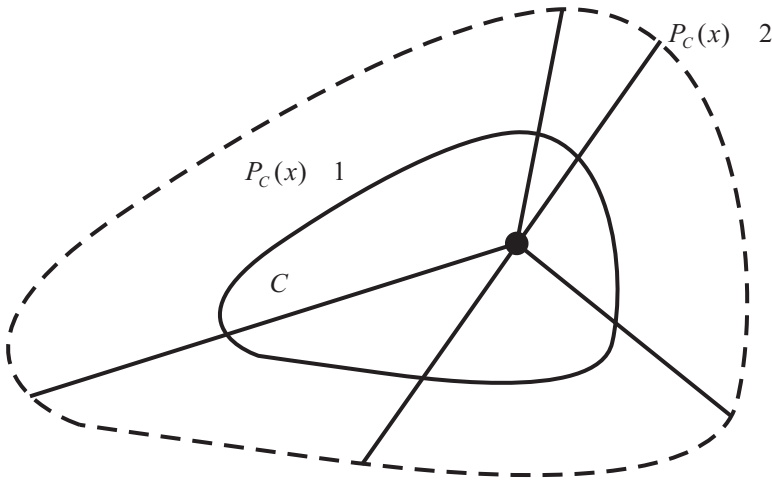


Figura 97

Lema 6.7.26. Sea C un conjunto convexo que contiene el elemento nulo θ en su interior; entonces el funcional de Minkowsky p de C satisface:

- i) $\infty > p(x) \geq 0$ para todo $x \in E$.

ii) $p(\lambda x) = \lambda p(x)$ para $\lambda > 0$.

iii) $p(x + y) \leq p(x) + p(y)$.

iv) p es continuo.

v) $C = \{x \in E : p(x) \leq 1\}$,

$$\overset{\circ}{C} = \{x \in E : p(x) < 1\}.$$

Demostración: i) Como C contiene a una bola de centro en θ , entonces dado $x \in E$ existe un $r > 0$ tal que $x/r \in C$. Luego $p(x)$ es finito para todo x . Obviamente $p(x) \geq 0$.

ii) Para $\lambda > 0$.

$$\begin{aligned} p(\lambda x) &= \inf \left\{ r \in \mathbb{R} : \frac{\lambda x}{r} \in C, r > 0 \right\} \\ &= \inf \left\{ \lambda r' \in \mathbb{R} : \frac{x}{r'} \in C, r' > 0 \right\} \\ &= \inf \left\{ r' \in \mathbb{R} : \frac{x}{r'} \in C, r' > 0 \right\} \\ &= \lambda p(x). \end{aligned}$$

iii) Dados $x, y \in E$ y $\epsilon > 0$, seleccionemos $t, s \in \mathbb{R}$ tales que $p(x) < t < p(x) + \epsilon$,

$$p(y) < s < p(y) + \epsilon.$$

Por ii) $p\left(\frac{x}{t}\right) < 1$ y $p\left(\frac{y}{s}\right) < 1$, y por lo tanto, $\frac{x}{t}, \frac{y}{s} \in C$. Sea $r = t + s$.

Por la convexidad de C .

$$\left(\frac{t}{r}\right)\left(\frac{x}{t}\right) + \left(\frac{s}{r}\right)\left(\frac{y}{s}\right) = \frac{(x+y)}{r} \in C.$$

Entonces, $p(x+y)/r < 1$. Pero por ii) tendremos que $p(x+y) \leq r \leq p(x) + p(y) + 2\epsilon$. Por la arbitrariedad de ϵ se deduce la subaditividad de p .

iv) Sea ϵ el radio de una bola cerrada centrada en θ y contenida en C ; entonces para cualquier $x \in E$, $\epsilon x / \|x\| \in C$ y por lo tanto, $p(\epsilon x / \|x\|) \leq 1$. Por ii) se tiene que $p(x) \leq (1/\epsilon) \|x\|$. Esto prueba que p es continuo en θ . Sin embargo, por iii) tenemos que

$$p(x) = p(x - y + y) \leq p(x - y) + p(y)$$

y

$$p(y) = (y - x + x) \leq p(x - y) + p(x).$$

De aquí concluimos que

$$-p(y-x) \leq p(x) - p(y) \leq p(x-y).$$

de donde se deduce la continuidad sobre E a partir de la continuidad en θ .

v) Se obtiene inmediatamente de iv). ■

Teorema 6.7.27 (teorema de Mazur. Forma geométrica del teorema de Hahn Banach). Sea E un espacio vectorial normado real y C un conjunto convexo en E con interior no vacío. Supongamos que V es una variedad lineal en E que no contiene puntos interiores de C . Entonces existe un hiperplano cerrado H en E que contiene a V pero que no contiene puntos interiores de C ; es decir, existe un elemento $f \in E^*$ y una constante c tales que

$$f(v) = c \text{ para todo } v \in V,$$

$$f(x) < c \text{ para todo } x \in \overset{\circ}{C}.$$

Demostración: Haciendo una traslación apropiada podemos suponer que $\theta \in \overset{\circ}{C}$. Sea M el subespacio vectorial de E generado por V ; entonces V es un hiperplano en M que no contiene el θ , luego, existe un funcional lineal continuo f_0 sobre M tal que

$$V = \{x \in M : f_0(x) = 1\}.$$

Sea p el funcional de Minkowsky de C . Como V no contiene puntos interiores de C , tendremos que:

$$f_0(x) = 1 \leq p(x) \text{ para todo } x \in V.$$

Por homogeneidad $f_0(\lambda x) = \lambda \leq p(\lambda x)$ para $x \in V$ y $\lambda > 0$. Además, para $\lambda < 0$, $f_0(\lambda x) \leq 0 \leq p(\lambda x)$. Entonces

$$f_0(x) \leq p(x) \text{ para todo } x \in M.$$

Por el teorema de Hahn-Banach, existe una extensión f desde M hasta E que cumple

$$f(x) \leq p(x) \text{ para todo } x \in E.$$

Como $f(x) \leq p(x)$ sobre E y p es continuo, por el lema 6.7.26, tendremos que f es continuo y además $f(x) < 1$ para todo $x \in \overset{\circ}{C}$.

Luego, H es el hiperplano cerrado buscado. ■

Existen varios corolarios y modificaciones de este teorema, de los cuales veremos el siguiente:

Teorema 6.7.28 (teorema de separación de Eidelheit). Sean C_1 y C_2 conjuntos convexos en el espacio vectorial normado E tales que C_1 tiene interior no vacío y C_2 no contiene puntos interiores de C_1 . Entonces existe un hiperplano cerrado H que separa C_1 y C_2 , es decir, existe $f \in E^*$ tal que

$$\sup_{x \in C_1} f(x) \leq \inf_{x \in C_2} f(x).$$

En otras palabras C_1 y C_2 están en semiespacios opuestos determinados por el hiperplano H (fig. 98).

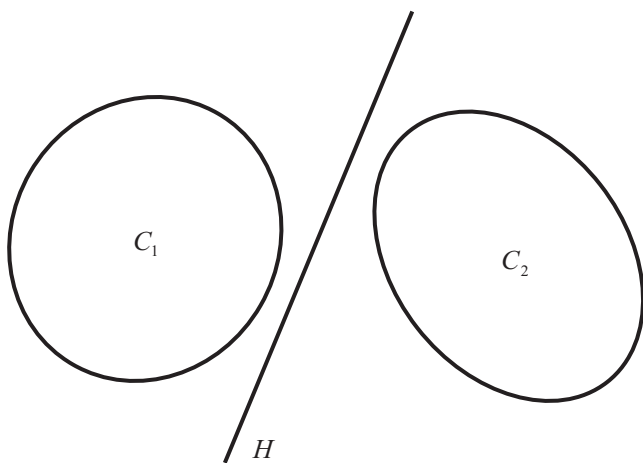


Figura 98

Demostración: Sea $C = C_1 - C_2$; entonces C contiene un punto interior y $\theta \notin \overset{\circ}{C}$. Del teorema 6.7.27 no resulta difícil concluir que existe un $f \in E^*$ no nulo, tal que $f(x) \leq 0$ para todo $x \in C$ (verifíquelo). Entonces para $x_1 \in C_1, x_2 \in C_2$

$$f(x_1) \leq f(x_2).$$

Consecuentemente, existe un número real c tal que

$$\sup_{x_1 \in C_1} f(x_1) \leq c \leq \sup_{x_2 \in C_2} f(x_2).$$

Entonces el hiperplano deseado será.

$$H = \{x \in E : f(x) = c\}. \blacksquare$$

Aunque no lo veremos en este texto, la formulación geométrica del teorema de Hahn-Banach así como los teoremas de separación que de él se deducen, son una herramienta muy potente en el estudio de los problemas generales de la teoría de optimización.

Ejercicios de Autocontrol

1. Pruebe que un funcional f definido sobre un espacio normado E es continuo si y solo si $\text{Ker } f = \{x \in E : f(x) = 0\}$ es cerrado en E .

2. Encuentre un ejemplo de un operador A entre dos espacios normados E y F tal que $\text{Ker } f = \{x \in E : f(x) = \theta\}$ es cerrado en E y A no es continuo.
3. Demuestre la proposición 6.7.16.
4. Demuestre que la norma de un funcional lineal continuo f en un espacio normado, es el inverso de la distancia desde el hiperplano $\{x : f(x) = 1\}$ hasta el elemento nulo θ .
5. Pruebe que $|||\cdot|||$ es una norma en $L(E, F)$ y que además, si $A \in L(E, F)$, $B \in L(F, G)$ y entonces $B \circ A \in L(E, G)$ y $|||B \circ A||| \leq |||B||| |||A|||$.
6. Demuestre la igualdad (139) del ejemplo 45.
7. Demuestre que si $A \in L(E, E)$ es no nulo y λ es un escalar $|\lambda| < \frac{1}{|||A|||}$, entonces $I - \lambda A$ es invertible. Dé una expresión para $(I - \lambda A)^{-1}$ en forma de una serie.
8. Demuestre que para un operador A simétrico y continuo en un espacio de Hilbert H , se cumple

$$|||A||| = \sup_{\|x\| \leq 1} \langle A(x), x \rangle.$$
9. Sea H un hiperplano en un espacio normado E . Demuestre que si H no contiene al elemento nulo, entonces existe un único funcional lineal f sobre E tal que $H = \{x \in E : f(x) = 1\}$.
10. Demuestre que si C es un cerrado convexo en un espacio normado y $x \notin C$, entonces existe un semiespacio cerrado que contiene a C y no contiene a x .
11. Deduzca del ejercicio 10 que todo conjunto cerrado y convexo en un espacio normado, es igual a la intersección de todos los semiespacios cerrados que lo contienen.

CAPÍTULO 7

DIFERENCIACIÓN EN ESPACIOS NORMADOS

En el capítulo 6 analizamos el papel fundamental que juegan los operadores y funcionales lineales continuos en la teoría de los espacios normados, debido a la compatibilidad entre la estructura métrica y la estructura vectorial en estos espacios. Podemos decir que los conceptos *operador* y *funcional* lineal corresponden a los espacios normados de la misma manera que el concepto general de función continua corresponde a la estructura métrica tal y como fue estudiado en el capítulo 2.

En este capítulo veremos que, *localmente*, determinadas funciones continuas definidas sobre espacios normados pueden *aproximarse* (en un sentido que precisaremos más adelante), por operadores o funciones lineales. Este hecho resulta verdaderamente muy útil sobre todo en el estudio de muchos problemas de optimización entre los que se encuentra el problema clásico del cálculo de variaciones y en general en el llamado Análisis Funcional no lineal.

El concepto adicional que es necesario introducir para lograr la *aproximación* lineal local a que hacíamos referencia, es el de *diferenciabilidad* y a su estudio dedicaremos este capítulo. Con este objetivo, en las cinco primeras secciones estudiaremos los conceptos *diferenciación fuerte* (diferenciación según Frechet) y *diferenciación débil* (diferenciación según Gateaux).

Para evitarnos complicaciones de carácter técnico a lo largo de todo el capítulo, supondremos que los espacios normados a que haremos referencia son reales.

En este contexto analizaremos las situaciones siguientes:

$$f : U \subset \mathbb{R} \longrightarrow E, \quad (1)$$

$$f : U \subset E \longrightarrow F, \quad (2)$$

$$f : U \subset E \longrightarrow \mathbb{R}, \quad (3)$$

donde E y F son espacios normados y U es un abierto de $\mathbb{R}$ en (1) y un abierto de E en (2) y (3).

En el 7.1 veremos la generalización inmediata de la diferenciación para funciones reales definidas sobre un intervalo (caso (1)). Además, en este epígrafe introduciremos la noción de diferencial en sentido fuerte o diferencial según Frechet, estudiando sus propiedades fundamentales.

El 7.2 está dedicado a la relación entre diferenciación e integración abstracta. Los resultados del 7.2 son utilizados en el 7.3, donde se introduce el concepto *derivadas de orden superior*, para obtener una generalización de la conocida fórmula de Taylor del Análisis Clásico y de algunos estimados como el teorema del valor medio, útiles en la teoría de aproximación.

En 7.4 se introduce la noción de diferenciación en sentido débil o diferenciación según Gateaux donde se estudian sus propiedades así como su relación con el concepto de *diferenciación en sentido fuerte*.

En el 7.5 se aplican los resultados de los epígrafes anteriores al estudio de los problemas de extremos de funcionales del tipo (3) obteniéndose como caso particular, algunas conclusiones para el problema clásico del cálculo de variaciones. Con este epígrafe culmina el capítulo; sin embargo, debemos señalar que para tener una visión más amplia sobre la posibilidad de generalizar la noción de diferenciabilidad, el lector debe dedicarse en algún momento al estudio de la diferenciación de funciones en sentido generalizado también conocida con el nombre de diferenciación según Soboliev para funciones

$$f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}, \quad (4)$$

donde U es un abierto de $\mathbb{R}^n$.

Esta noción de diferenciabilidad tiene innumerables aplicaciones, en particular en la *teoría de los problemas de contorno* para las *ecuaciones diferenciales en derivadas parciales*, etcétera, las cuales no veremos en este libro ya que su estudio requiere el desarrollo de temas que salen del alcance y objetivo de este texto.

7.1 DIFERENCIACION EN SENTIDO FUERTE. DIFERENCIAL Y DERIVADA SEGÚN FRECHET

Como es conocido de los cursos de Análisis Clásico, decimos que una función real $f : [a, b] \longrightarrow \mathbb{R}$ es diferenciable en el punto $x_0 \in]a, b[$ si existe el límite del cociente incremental de f en ese punto:

$$f \text{ derivable en } x_0 \leftrightarrow \text{ existe } \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad (5)$$

A este límite, si existe, se le denota por $f'(x_0)$ y se le llama derivada de la función f en el punto x_0 .

Evidentemente la derivada de la función f en el punto x_0 es un número real cuya interpretación geométrica es la pendiente de la recta tangente a la gráfica de f (en coordenadas cartesianas) y que corta a dicha gráfica en el punto $(x_0, f(x_0))$.

Supongamos que ahora nos encontramos en la situación de (1), es decir, U es un abierto de $\mathbb{R}$ (que para comodidad supondremos igual a un intervalo $]a, b[$), E es un espacio normado y

$$f :]a, b[\longrightarrow E. \quad (6)$$

Si $x_0 \in]a, b[$, podemos definir la diferenciabilidad de f en el punto x_0 de la misma forma que en (5).

Al límite del cociente incremental, si existe, lo continuaremos llamando derivada de f en el punto x_0 y lo seguiremos denotando mediante $f'(x_0)$.

Notemos que en este caso la operación de resta en el numerador del cociente incremental es posible gracias a la estructura algebraica de E . Lo mismo ocurre con respecto a la división por $h \in \mathbb{R}$ que interpretamos como la multiplicación en E por el escalar $1/h$. Además, en este caso el límite en (5) se refiere al límite en el espacio normado E .

Obviamente, $f'(x_0)$ es ahora un elemento del espacio normado E y no es sencillo en ese caso una interpretación geométrica simple como en el caso en que $E = \mathbb{R}$.

Sin embargo, notemos que la definición dada anteriormente en el caso (6) es equivalente a decir:

existe un elemento $f'(x_0) \in E$ tal que

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \alpha(x_0, h), \quad (7)$$

donde $\alpha(x_0, h)$ es una función de las variables x y h definidas para h en un entorno de $0 \in \mathbb{R}$ y que cumple

$$\frac{\|\alpha(x_0, h)\|_E}{|h|} \xrightarrow{h \rightarrow 0} 0. \quad (8)$$

La expresión (7) nos dice que $f'(x_0)$ determina la parte lineal del incremento de f en una proximidad del punto x_0 . En efecto, notemos que la aplicación (que denotaremos también por $f'(x_0)$)

$$f'(x_0) : \mathbb{R}_{h \rightarrow f'(x_0)h} \longrightarrow E \quad (9)$$

es un operador lineal continuo de $\mathbb{R}$ en E . Precisamente este operador nos brinda una *aproximación lineal* de los valores de la función f en un entorno del punto x_0 .

Esta nueva versión de la definición de diferenciabilidad puede ser generalizada sin dificultad al caso (2).

Definición 7.1.1. Sean E y F espacios normados, U un subconjunto abierto de E y $f : U \longrightarrow F$ una aplicación.

Diremos que f es diferenciable en sentido fuerte en un punto dado $x_0 \in U$, si existe un operador lineal continuo $A_{x_0} \in L(E, F)$ tal que

$$f(x_0 + h) - f(x_0) = A_{x_0}(h) + \alpha(x_0; h), \quad (10)$$

donde

$$\frac{\|\alpha(x_0; h)\|_F}{\|h\|_E} \longrightarrow 0, \text{ cuando } \|h\|_E \longrightarrow 0. \quad (11)$$

La expresión $A_{x_0}(h)$ (que para cada $h \in E$ representa, evidentemente, un elemento del espacio normado F) se llama *diferencial fuerte* (o *diferencial de Frechet*) de la aplicación f en el punto x_0 . El propio operador lineal A_{x_0} se llama derivada (más precisamente *derivada fuerte*) de la aplicación f en el punto x_0 .

Por analogía con (7) denotaremos al operador derivada mediante $f'(x_0)$:

$$A_{x_0} = f'(x_0). \quad (12)$$

Diremos que $f: U \longrightarrow F$ es diferenciable en sentido fuerte sobre U si lo es en cada punto $x \in U$ y denotaremos mediante $D(U, F)$ al conjunto de todas las funciones derivables en sentido fuerte de U en F .

Observación: A diferencia del caso (6) en que cuando existe la derivada siempre define un operador lineal y continuo, en el caso general en que $\mathbb{R}$ se sustituye por E , puede ocurrir que el incremento $f(x_0 + h) - f(x_0)$ sea lineal pero no continuo con respecto a h y también que dicho incremento no sea lineal en h .

Veamos algunos ejemplos de funciones derivables según Frechet.

Ejemplo 1. Sea $U = E = \mathbb{R}^n$ y $F = \mathbb{R}$ y tomemos $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ con derivadas parciales continuas con respecto a cada variable. Como es sabido de los cursos de Análisis Clásico, la existencia de las derivadas parciales de una función real definida sobre $\mathbb{R}^n$ no es una condición suficiente para la existencia de su diferencial. Veremos ahora que si suponemos la continuidad de las derivadas parciales, entonces f es diferenciable según Frechet (lo cual resulta ser en este caso la noción de diferenciability usual), y además, para cada $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, se tiene

$$f'(x)(h) = \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} h_i, \quad (13)$$

es decir, $f'(x)$ coincide, para cada $x \in \mathbb{R}^n$, con el funcional lineal y continuo definido sobre $\mathbb{R}^n$ mediante el producto escalar por el vector

$$(\text{grad } f) = \left\{ \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right\}. \quad (14)$$

Probemos que la expresión (13) es la diferencial de Frechet de f en el punto $x \in \mathbb{R}^n$.

Obviamente $f'(x_0) = (\text{grad } f)(x_0)$ es lineal y continuo con respecto a h . Luego, basta con verificar las propiedades (10) y (11) de la diferencial de Frechet.

Dado $\epsilon > 0$, la continuidad de las derivadas parciales implica que existe una vecindad $B(x; \delta)$ tal que

$$\left| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(y) \right| < \epsilon/n$$

para todo $y \in B(x; \delta)$, $i = 1, 2, \dots, n$.

Definamos los vectores unitarios en la forma usual $e_i = (0, \dots, 0, 1, 0, \dots, 0)$

y para $h = \sum_{i=1}^n h_i e_i$, supongamos que $g_0 = \theta$, $g_k = \sum_{i=1}^k h_i e_i$, $k = 1, 2, \dots, n$.

Notemos que $|g_k| \leq |h|$ para todo k , entonces,

$$\begin{aligned} & \left| f(x+h) - f(x) - \sum_{i=1}^n \frac{\partial f}{\partial x_i} h_i \right| = \\ & = \left| \sum_{k=1}^n \left\{ f(x+g_k) - f(x+g_{k-1}) - \frac{\partial f}{\partial x_k} h_k \right\} \right| \leq \\ & \leq \sum_{k=1}^n \left| f(x+g_k) - f(x+g_{k-1}) - \frac{\partial f}{\partial x_k} h_k \right|. \end{aligned}$$

Examinemos ahora el k -ésimo término en la suma anterior. El vector $x+g_k$ difiere de $x+g_{k-1}$ solamente en la k -ésima componente. En efecto, $x+g_k = x+g_{k-1} + h_k e_k$. Entonces, por el teorema del valor medio para funciones de una variable real, tendremos

$$f(x+g_k) - f(x+g_{k-1}) = \frac{\partial f}{\partial x_k}(x+g_{k-1} + \alpha e_k) h_k$$

para algún $0 \leq \alpha \leq 1$.

Por otra parte, $x+g_{k-1} + \alpha e_k \in B(x; \delta)$ si $|h| < \delta$, luego

$$\left| f(x+g_k) - f(x+g_{k-1}) - \frac{\partial f}{\partial x_k} h_k \right| < \frac{\epsilon}{n} |h|.$$

Finalmente concluimos que

$$\left| f(x+h) - f(x) - \sum_{i=1}^n \frac{\partial f}{\partial x_k} h_i \right| < \epsilon |h|$$

para todo h con $|h| < \delta$, con lo cual se verifica (10) y (11). Veamos ahora el caso de un funcional definido sobre un espacio de dimensión infinita.

Ejemplo 2. Sea $U = E = C([0, 1], \mathbb{R})$ y sea $f : E \rightarrow \mathbb{R}$ definido mediante

$$f(x) = \int_0^1 g(x(t), t) dt, \quad (15)$$

donde se supone que existe la derivada parcial $\frac{\partial g}{\partial x} = g_x$, siendo continua con respecto a x y t . Entonces, se puede probar que existe la derivada de Frechet de f en todo E y

$$f'(x)(h) = \int_0^1 g_x(x(t), t) h(t) dt, \quad (16)$$

para todo $h \in E$. En efecto,

$$\left| f(x+h) - f(x) - \int_0^1 g_x(x(t), t) h(t) dt \right| = \left| \int_0^1 \{g(x+h, t) - g(x, t) - g_x(x(t), t) h(t)\} dt \right|.$$

Para cada t fijo tenemos, por el teorema del valor medio aplicado a funciones de una variable:

$$g(x(t) + h(t), t) - g(x(t), t) = g_x(\bar{x}(t), t) h(t)$$

donde $|x(t) - \bar{x}(t)| \leq |h(t)|$. Dado $\epsilon > 0$, la continuidad uniforme de g_x en x y t implica que existe un $\delta > 0$ tal que para $\|h\|_\infty < \delta$, se tiene

$$|g_x(x+h, t) - g_x(x, t)| < \epsilon.$$

Entonces, tenemos

$$\left| f(x+h) - f(x) - \int_0^1 g_x(x, t) h(t) dt \right| =$$

$$= \left| \int_0^1 (g_x(\bar{x}, t) - g_x(x, t)) h(t) dt \right| \leq \epsilon \|h\|_\infty,$$

siempre que $\|h\|_\infty < \delta$; de donde se obtiene el resultado deseado.

Ejemplo 3. Consideremos en el espacio de Hilbert real H el funcional $f(x) = \|x\|^2$. Entonces $f(x+h) - f(x) = \|x+h\|^2 - \|x\|^2 = 2\langle x, h \rangle + \|h\|^2$.

Luego, la expresión $2\langle x, h \rangle$ constituye la parte principal de esta expresión y, por consiguiente

$$f'(x) = 2x.$$

Notemos que basándonos en el teorema sobre la representación de los funcionales lineales continuos en un espacio de Hilbert, hemos identificado aquí los elementos de H^* (dual topológico de H) con los correspondientes en H .

Veamos ahora algunas propiedades elementales que se desprenden directamente de la definición de derivada según Frechet.

Proposición 7.1.2 (unicidad de la derivada). Si la aplicación $f : U \longrightarrow F$, donde U es abierto en E , es diferenciable en sentido fuerte en el punto $x \in U$, entonces la derivada correspondiente se determina de manera única.

Demostración: En efecto, supongamos que existen dos operadores $A, B \in L(E, F)$ tales que

$$\begin{aligned} f(x+h) - f(x) &= A(h) + \alpha_1(x; h) = \\ &= B(h) + \alpha_2(x; h). \end{aligned}$$

Entonces,

$$A(h) - B(h) = \alpha_2(x; h) - \alpha_1(x; h)$$

y en virtud de (11),

$$\frac{\|A(h) - B(h)\|_F}{\|h\|_E} \longrightarrow 0$$

cuando $\|h\|_E \longrightarrow 0$.

De aquí se concluye que

$$\|A - B\| = 0,$$

y por lo tanto, $A = B$ con lo cual queda probada la unicidad enunciada. ■

Proposición 7.1.3. Sea $f : E \longrightarrow F$; entonces se cumplen las propiedades siguientes:

- i) Si $f(x) = y_0 = \text{constante}$, se tiene $f'(x) = \theta$ (es decir $f'(x)$ coincide con el operador idénticamente nulo).
- ii) Si f es lineal y continuo, entonces $f'(x) = f$ para todo $x \in E$.

Demostración:

i) Se deja de ejercicio al lector.

ii) En efecto, por definición

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

luego $\alpha(x; h) \equiv \theta$. ■

La siguiente proposición nos muestra que efectivamente la diferenciabilidad según Frechet es una exigencia más fuerte que la continuidad.

Proposición 7.1.4. Sean $U \subset E$ y $f : U \longrightarrow F$ diferenciable en el punto $x_0 \in U$; entonces f es continua en x_0 .

Demostración:. En efecto, por (10) $\|f(x_0 + h) - f(x_0)\|_F = \|f'(x_0)(h) + \alpha(x; h)\|_F$. Por (11) y la continuidad de $f'(x_0) \in L(E, F)$, dado $\epsilon > 0$, existe $\delta > 0$ tal que si $\|h\|_E < \delta$, entonces

$$\begin{aligned} & \|f'(x_0)(h) + \alpha(x; h)\| \leq \\ & \leq \|f'(x_0)\| \|h\|_E + \epsilon \|h\|_E = \end{aligned}$$

$$(\|f'(x_0)\| + \epsilon) \|h\|_E,$$

de donde se concluye la continuidad de f en x_0 . ■

Proposición 7.1.5. Sean U abierto en E y f, g aplicaciones de U en F . Si f y g son diferenciables en sentido fuerte en el punto $x_0 \in U$, entonces:

- i) $f + g$ es diferenciable en el punto x_0 y $(f + g)'(x_0) = f'(x_0) + g'(x_0)$;
- ii) para todo escalar $\lambda \in \mathbb{R}$. λf diferenciable en el punto x_0 y $(\lambda f)'(x_0) = \lambda f'(x_0)$.

Demostración: De las definiciones de suma de operadores y de producto de un operador por un escalar, tenemos que

$$\begin{aligned} (f + g)(x_0 + h) &= f(x_0 + h) + g(x_0 + h) = \\ &= f(x_0) + g(x_0) + f'(x_0)(h) + g'(x_0)(h) + \end{aligned}$$

$$\alpha_1(x_0; h) + \alpha_2(x_0; h),$$

y además,

$$\lambda f(x_0 + h) = \lambda f(x_0) + \lambda f'(x_0)(h) + \lambda \alpha_1(x_0; h)$$

de donde se deducen las igualdades de i) y ii). ■

Corolario 1. El espacio $D(E, F)$ de todas las aplicaciones diferenciables según Frechet de E en F , es un espacio vectorial provisto de las operaciones de suma y producto usual de funciones por escalares. Además, la aplicación que a cada $f \in D(E, F)$ le hace corresponder su derivada f' , es un operador

lineal de $D(E, F)$ en el espacio de las aplicaciones de E en $L(E, F)$

$$D(E, F) \xrightarrow{f \rightsquigarrow f'} L(E, F)^E \quad (17)$$

La demostración es una consecuencia inmediata de la proposición anterior.

Entre las propiedades de la derivada de funciones numéricas que se generalizan a la derivada según Frechet, la menos obvia es la siguiente:

Teorema 7.1.6 (derivada de una función compuesta). Sean E , F y G tres espacios normados.

Sean además, $x_0 \in E$, f una aplicación definida sobre una vecindad $U(x_0)$ con valores en F y g una aplicación definida sobre una vecindad $V(y_0)$ en F con valores en G , donde $y_0 = f(x_0)$.

Si f es diferenciable en sentido fuerte en el punto $x_0 \in E$ y g es diferenciable en sentido fuerte en el punto $y_0 \in F$, entonces la función compuesta $h = g \circ f$ también será diferenciable en el punto x_0 y se cumple

$$h'(x_0) = g'(y_0)f'(x_0). \quad (18)$$

Demostración: En efecto, de acuerdo con las suposiciones hechas se tiene que

$$f(x_0 + \epsilon) = f(x_0) + f'(x_0)(\epsilon) + \alpha_1(x_0; \epsilon).$$

$$g(y_0 + \eta) = g(y_0) + g'(y_0)(\eta) + \alpha_2(y_0; \eta).$$

$f'(x_0)$ y $g'(y_0)$ son operadores lineales continuos y por eso

$$h(x_0 + \epsilon) = g(y_0 + f'(x_0)(\epsilon) + \alpha_1(x_0; \epsilon)) =$$

$$= g(y_0) + g'(y_0)(f'(x_0)(\epsilon) + \alpha_1(x_0; \epsilon)) + \alpha_2(f'(x_0)(\epsilon) +$$

$$+ \alpha_1(x_0; \epsilon) = g(y_0) + g'(y_0)f'(x_0)(\epsilon) + \alpha_3(x_0; \epsilon).$$

con lo cual queda demostrada la formula (18). ■

En el caso en que f y g son funciones numéricas, la fórmula (18) se convierte en la conocida regla de diferenciación de una función compuesta.

Como consecuencia del resultado anterior veremos la proposición siguiente:

Proposición 7.1.7. Sea f diferenciable según Frechet sobre un intervalo $U \subset E$ y con valores en F . Sea $x \in U$ y supongamos que $x + \alpha h \in U$ para todo $\alpha, 0 \leq \alpha \leq 1$. Entonces:

$$\|f(x + h) - f(x)\|_F \leq \|h\|_E \sup_{0 \leq \alpha \leq 1} \|f'(x + \alpha h)\|. \quad (19)$$

Demostración: Como consecuencia del teorema de Hahn-Banach sabemos que para cada $y \in F$ se puede encontrar un elemento $y^* \in F^*$ tal que

$$y^*(y) = \|y\|_F. \quad (20)$$

Tomemos $y = f(x+h) - f(x)$; entonces la función

$$\varphi(\alpha) = y^*[f'(x + \alpha h)]$$

está definida sobre el intervalo $[0, 1]$ y por la regla de derivada de la compuesta, su derivada será igual a

$$\varphi'(\alpha) = y^*[f'(x + \alpha h)].$$

Por el teorema del valor medio aplicado a funciones de una variable real tenemos que

$$\varphi(1) - \varphi(0) = \varphi'(\alpha_0), 0 < \alpha_0 < 1,$$

y por lo tanto,

$$|y^*[f(x+h) - f(x)]| \leq \|y^*\| \sup_{0 < \alpha < 1} \|f'(x + \alpha h)\| \|h\|_E.$$

Como se cumple la condición de alineamiento (20), entonces

$$\|f(x+h) - f(x)\|_F \leq \|h\|_E \sup_{0 < \alpha < 1} \|f'(x + \alpha h)\|$$

con lo cual se obtiene el resultado deseado. ■

De (19) se deduce también la desigualdad siguiente:

Corolario 1. Supongamos que se satisfacen las mismas condiciones de la proposición (7.1.7); entonces se cumple la desigualdad:

$$\begin{aligned} & \|f(x+h) - f(x) - f'(x)(h)\|_F \leq \\ & \|h\|_E \sup_{0 < \alpha < 1} \|f'(x + \alpha h) - f'(x)\| \end{aligned} \quad (21)$$

Demostración: Basta considerar para cada x fijo la aplicación

$$f(h) = f(x+h) - f'(x)(h)$$

y aplicar (19) para mayorar

$$\|g(h) - g(\theta)\|_E. \quad \blacksquare$$

Ejercicios de Autocontrol

1. Calcule la derivada y el diferencial del funcional $f(x) = \|x\|$ sobre un espacio normado arbitrario.
2. Sobre el espacio $E = (C([0, 1], \mathbb{R}), \|\cdot\|_\infty)$ definamos el funcional

$$f(x) = \left[x \left(\frac{1}{2} \right) \right]^2.$$

Calcule su diferencial y derivada en sentido fuerte.

3. Utilizando el corolario 1 de la proposición 7.1.7 obtenga una aproximación para el incremento $f(x) - f(0)$ del funcional f definido en el ejemplo 2.

7.2 RELACIÓN ENTRE DERIVACIÓN E INTEGRACIÓN ABSTRACTA

En esta sección introduciremos el concepto de *integración abstracta en espacios de Banach*, y desarrollaremos brevemente algunas de sus propiedades que serán utilizadas en 7.3 para la obtención de la fórmula de Taylor.

Supongamos nuevamente que nos encontramos en el caso de funciones definidas sobre un intervalo $[a, b]$ y con valores en un espacio normado E . Sabemos que la derivada de una tal función $f : [a, b] \rightarrow E$ en cualquier punto $x_0 \in [a, b]$, si es que existe, está representada por un elemento del espacio E . Para tales funciones, que llamaremos funciones abstractas, se puede definir la integral (de Riemann) de la forma siguiente:

Definición 7.2.1. Sea f una función abstracta definida sobre el intervalo $[a, b]$ y con valores en el espacio de Banach E .

Consideremos la colección de las sumas integrales

$$\sum_{k=0}^{n-1} f(\epsilon_k) (t_{k+1} - t_k) \quad (22)$$

correspondientes a las particiones

$$\alpha = t_0 < t_1 < \dots < t_n = b \quad (23)$$

con $\epsilon_k \in [t_k, t_{k+1}]$

Si existe el límite de las sumas integrales (22) cuando $\max |t_{k+1} - t_k| \rightarrow 0$ entonces se dice que la función f es integrable sobre el intervalo $[a, b]$.

Al límite, cuando existe, se le llama *integral de la función abstracta f sobre el intervalo $[a, b]$* y se denota mediante

$$\int_a^b f(t) dt \quad (24)$$

Notemos que en la definición anterior nos estamos refiriendo la límite de una red y no al de una sucesión de la misma forma que en el caso de la integral de Riemann clásica para funciones con valores reales.

Debido a que la integral es un límite en el espacio de Banach E , la integral es precisamente un elemento de E .

Razonamientos análogos a los empleados para funciones que toman valores numéricos, permiten demostrar que la integral de una función abstracta continua sobre un intervalo compacto $[a, b]$ siempre existe.

Además, ella tiene propiedades análogas a las propiedades de la integral usual de Riemann. El lector puede, sin dificultad, verificar las propiedades que aparecen a continuación:

i) Si A es un operador lineal continuo del espacio E en otro espacio F , entonces

$$\int_a^b A f(t) dt = A \int_a^b f(t) dt. \quad (25)$$

ii) Si f es de la forma $g(t)x_0$, donde f es una función numérica y x_0 un elemento fijo de E , se tiene

$$\int_a^b f(t) dt = x_0 \int_a^b g(t) dt, \quad (26)$$

donde $\int_a^b g(t) dt$ es una integral de Riemann en sentido usual.

iii) Si f y g son funciones abstractas integrables sobre el intervalo $[a, b]$ y si λ, μ son números reales entonces $\lambda f + \mu g$ también es integrable y

$$\int_a^b (\lambda f(t) + \mu g(t)) dt = \lambda \int_a^b f(t) dt + \mu \int_a^b g(t) dt \quad (27)$$

iv) Si la función abstracta $f(t)$ es integrable, entonces también lo es la función numérica $\|f(t)\|$ y se cumple

$$\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt. \quad (28)$$

Veamos ahora cómo se puede aplicar esta noción de integral al problema de la reconstrucción de una aplicación definida entre dos espacios normados, a partir de su derivada.

Consideremos una aplicación f definida sobre un segmento $J = [x_0, x_0 + \Delta x]$ del espacio normado E :

$$J = \{x \in E : x = x_0 + t\Delta x, t \in [0, 1]\}$$

y con valores en el espacio $\mathbf{BC}(E, F)$ de todas las aplicaciones continuas y acotadas de E en otro espacio de Banach F :

$$f : J \longrightarrow \mathbf{BC}(E, F) \quad (29)$$

donde entendemos por función acotada a toda aquella función de E en F que transforme todo conjunto acotado de E en un conjunto acotado de F . Evidentemente:

$$L(E, F) \subset \mathbf{BC}(E, F). \quad (30)$$

En el espacio $\mathbf{BC}(E, F)$ se puede introducir una métrica llamada *métrica de la convergencia uniforme sobre los subconjunto acotados de E* mediante

$$d(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(f, g)}{1 + d_n(f, g)} \quad (31)$$

donde

$$d_n(f, g) = \sup_{\|x\|_E \leq n} \|f(x) - g(x)\|_F. \quad (32)$$

Obviamente d es la métrica generada por la familia de pseudométricas (d_n) y para una sucesión (f_k) en $\mathbf{BC}(E, F)$ se tiene:

$$(f_k) \xrightarrow{d} f \Leftrightarrow f_k \text{ converge a } f \text{ uniformemente sobre cualquier subconjunto acotado de } E. \quad (33)$$

Luego, la métrica inducida por d sobre el subespacio $L(E, F)$ de $\mathbf{BC}(E, F)$ es topológicamente equivalente a la métrica asociada a la norma usual de operadores lineales continuos, es decir, para una sucesión (A_n) de operadores lineales continuos de E en F se cumple

$$(A_n) \xrightarrow{d} A \Leftrightarrow \|A_n - A\| \longrightarrow 0. \quad (34)$$

Si una función f del tipo (29) es continua, teniendo sobre $\mathbf{BC}(E, F)$ la métrica d , entonces se puede definir la integral de f sobre el segmento J de la manera siguiente.

Definición 7.2.2. Sea $f : J \longrightarrow \mathbf{BC}(E, F)$ continua, donde $J = [x_0, x_0 + \Delta x] \subset E$; entonces se llama integral de f sobre el segmento J , a la integral de la función abstracta

$$f(x_0 + t\Delta x) \Delta x \quad (35)$$

sobre el intervalo $[0, 1]$ (aquí $f(x_0 + t\Delta x)\Delta x$ denota, para cada $t \in [0, 1]$, la evaluación de la aplicación $f(x_0 + t\Delta x) \in \mathbf{BC}(E, F)$ en el elemento $\Delta x \in E$).

Como notación pondremos

$$\int_{x_0}^{x_0+\Delta x} f(x) dx = \int_0^1 f(x_0 + t\Delta x) \Delta x dt. \quad (36)$$

Observación: El lector puede comprobar fácilmente que la continuidad de f , implica la continuidad de la función abstracta $f(x_0 + t\Delta x)\Delta x$, lo cual asegura su integrabilidad según la definición 7.2.1.

Apliquemos estas ideas al problema de la reconstrucción de una función a partir de su derivada.

Teorema 7.2.3 (fórmula de Newton-Leibnitz generalizada). Sea f una aplicación que actúa de E en F y que tiene derivada en sentido fuerte sobre el segmento $[x_0, x_0 + \Delta x]$ de E la cual, además, es continua considerada como aplicación de $[x_0, x_0 + \Delta x]$ en $L(E, F)$. Entonces existe la integral

$$\int_{x_0}^{x_0+\Delta x} f'(x) dx, \quad (37)$$

y se cumple la igualdad

$$\int_{x_0}^{x_0+\Delta x} f'(x) dx = f(x_0 + \Delta x) - f(x_0). \quad (38)$$

Demostración: En efecto, por definición

$$\int_{x_0}^{x_0+\Delta x} f'(x) dx = \lim_{\lambda \rightarrow 0} \sum_{k=0}^{n-1} f'(x_0 + t_k \Delta x) (\Delta x) (t_{k+1} - t_k)$$

$$= \lim_{\lambda \rightarrow 0} \sum_{k=0}^{n-1} f'(\bar{x}_k) (\Delta x_k).$$

donde $\bar{x}_k = x_0 + t_k \Delta x$, $\Delta x_k = (t_{k+1} - t_k) \Delta x$ y $\lambda = \max_k |t_{k+1} - t_k|$.

Al mismo tiempo, para cualquier partición del segmento $0 \leq t \leq 1$ tenemos:

$$f(x_0 + \Delta x) - f(x_0) = \sum_{k=0}^{n-1} f(x_0 + t_{k+1} \Delta x) - f(x_0 + t_k \Delta x)$$

$$= \sum_{k=0}^{n-1} [f(x_{k+1}) - f(x_k)].$$

De la desigualdad (21), encontramos

$$\left\| \sum_{k=0}^{n-1} f(\overline{x}_{k+1}) - f(\overline{x}_k) - f'(\overline{x}_k) \Delta x_k \right\|_F \leq \|\Delta x\|_E$$

$$\sum_{k=0}^{n-1} (t_{k+1} - t_k) \sup_{0 < \alpha < 1} \|f'(\overline{x}_k + \alpha \Delta x) - f'(\overline{x}_k)\|.$$
(39)

Como la derivada f' es continua y por consiguiente, también uniformemente continua sobre el segmento $[x_0, x_0 + \Delta x]$, el miembro derecho de la desigualdad (39) tiende a cero cuando disminuyen indefinidamente las longitudes de los elementos de la partición del segmento $[x_0, x_0 + \Delta x]$. Luego, de aquí se concluye la igualdad (38). ■

Ejercicios de Autocontrol

1. Demuestre las propiedades (25)-(28) de la integral abstracta.
2. Aplique la fórmula de Newton-Leibnitz y obtenga su expresión explícita en el caso del funcional del ejemplo 2.
3. Demuestre que mediante (31) y (32) se define una métrica d en el espacio $BC(E, F)$ y que cumple (33).
4. Demuestre la equivalencia (34).

7.3 DERIVADAS DE ORDEN SUPERIOR Y FÓRMULA DE TAYLOR

Sea f una aplicación diferenciable en sentido fuerte y que actúa de E en F . Su derivada $f'(x)$ es, para cada $x \in E$, un elemento de $L(E, F)$, es decir, f' define una aplicación del espacio E en el espacio de los operadores lineales y continuos $L(E, F)$. Si esta aplicación también es diferenciable entonces a su derivada se le llama *segunda derivada* de f y se denota mediante el símbolo f'' . Luego, para cada $x \in E$ donde exista la segunda derivada de f , se tiene que

$$f''(x) \in L(E, L(E, F)). \quad (40)$$

Probemos que los elementos del espacio $L(E, L(E, F))$, de los operadores lineales y continuos que actúan de E en $L(E, F)$, admiten una interpretación

más cómoda a partir de las llamadas *aplicaciones bilineales continuas*.

Definición 7.3.1. Una *aplicación bilineal* del espacio normado E en el espacio normado F es una aplicación

$$B : E \times E \longrightarrow F$$

tal que para cualesquiera elementos x_1, x_2, y_1, y_2 de E y cualesquiera números reales λ y μ satisface

$$B(\lambda x_1 + \mu x_2, y) = \lambda B(x_1, y) + \mu B(x_2, y),$$

$$B(x, \lambda y_1 + \mu y_2) = \lambda B(x, y_1) + \mu B(x, y_2).$$

Como podemos notar, las condiciones exigidas a B nos dicen que es lineal con respecto a cada uno de sus argumentos.

Se deja al lector verificar la siguiente proposición cuyo análogo fue enunciado para los funcionales bilineales en el capítulo 6.

Proposición 7.3.2. Una aplicación bilineal B de E en F es continua si y solo si existe un número positivo M tal que

$$\|B(x, y)\|_F \leq M \|x\|_E \|y\|_E, \quad (41)$$

para todos los $x, y \in E$.

La demostración se deja al lector.

El menor de los números M que satisface la condición (41) se llama *norma de la aplicación bilineal* B y se denota por $\|B\|$.

De la manera usual se definen las operaciones bilineales para las aplicaciones bilineales continuas del espacio E en el espacio F , constituyendo un espacio vectorial que denotaremos mediante $B(E^2, F)$.

Proposición 7.3.3. Los espacios normados $L(E, L(E, F))$ y $B(E^2, F)$ son isométricamente isomorfos.

Demostración: Recordemos que una isometría ente dos espacios normados X y Y es una aplicación lineal biyectiva $i : X \longrightarrow Y$ tal que

$$\|x\|_X = \|i(x)\|_Y,$$

para todo $x \in X$.

Construyamos una isometría de $L(E, L(E, F))$ sobre $B(E^2, F)$. Para ello, a cada elemento A del espacio $L(E, L(E, F))$ le haremos corresponder un elemento B de $B(E^2, F)$ mediante

$$B(x, y) = (A(x))(y), \quad (42)$$

es decir, $i(A) = B$.

Es obvio que esta correspondencia es lineal. Probemos que es, además, isométrica y transforma el espacio $L(E, L(E, F))$ en todo el espacio $B(E^2, F)$. En efecto, si $z = B(x, y) = (A(x))(y)$, tenemos

$$\|z\|_F \leq \|A(x)\|_{L(E, F)} \|y\|_E \leq \|A\| \|x\|_E \|y\|_E,$$

de donde,

$$\|B\| \leq \|A\|. \quad (43)$$

Por otro lado, dada una aplicación bilineal B , la aplicación $y \rightsquigarrow (A(x))(y) = B(x, y)$ es, para cada x fijo, una aplicación lineal del espacio E en F . Por consiguiente, a cada $x \in E$ se le hace corresponder un elemento $A(x)$ del espacio $L(E, F)$. Es obvio que $A(x)$ depende linealmente de x , es decir que la aplicación bilineal B determina a un elemento A del espacio $L(E, L(E, F))$, lo cual demuestra la biyección de i .

Además, está claro que la aplicación B se construye a partir de A mediante la fórmula (42) y que

$$\|A(x)\|_{L(E, F)} = \sup_{\|y\|_E < 1} \|(A(x))(y)\|_F =$$

$$= \sup_{\|y\|_E < 1} \|B(x, y)\|_F \leq \|B\| \|x\|_E$$

de donde

$$\|A\| \leq \|B\|. \quad (44)$$

Comparando (43) y (44), obtenemos

$$\|A\| = \|B\| = \|i(A)\|,$$

lo cual demuestra que i es una isometría. ■

Hemos visto que la segunda derivada $f''(x)$ de una aplicación $f: E \rightarrow F$ es un elemento del espacio $L(E, L(E, F))$. De acuerdo con la proposición 7.3.3 podemos considerar que $f''(x)$ es un elemento del espacio $B(E^2, F)$.

Ejemplo 4. Sean $E = \mathbb{R}^m$ y $F = \mathbb{R}^n$; entonces toda aplicación lineal de E en F se puede representar mediante una matriz. Luego, la derivada $f'(x)$ de una aplicación que actúa de $\mathbb{R}^m$ en $\mathbb{R}^n$, es una matriz (dependiente de $x \in \mathbb{R}^m$). Si en $\mathbb{R}^m$ y $\mathbb{R}^n$ consideramos las bases canónicas $e_1, \dots, e_m$ y $v_1, \dots, v_n$, respectivamente (constituidas por los vectores unitarios en la dirección de los ejes coordenados), tendremos para $x \in \mathbb{R}^m$, $y = f(x) \in \mathbb{R}^n$

$$x = x_1 e_1 + x_2 e_2 + \dots + x_m e_m,$$

$$y = f_1 v_1 + f_2 v_2 + \dots + f_n v_n$$

y en este caso,

que es lineal con respecto a cada uno de sus argumetos, en el mismo sentido de la definición 7.3.1.

Al igual que en el caso de las aplicaciones bilineales, puede comprobarse que una aplicación n -lineal N de E en F es continua si y solo si existe un número positivo M tal que

$$\|N(x_1, \dots, x_m)\|_F \leq M \|x_1\|_E \dots \|x_n\|_F, \quad (47)$$

para todos los $x_i \in E$, $i = 1, \dots, n$.

De manera trivial se definen las operaciones de suma y producto por escalares para las aplicaciones n -lineales. El espacio $N(E^n, F)$ se convierte entonces en un espacio vectorial normado si además, se define la norma del elemento $N \in N(E^n, F)$ como el ínfimo de las constantes M para las cuales se satisface la desigualdad (47).

El lector puede verificar que el espacio normado $N(E^n, F)$ es isométricamente isomorfo al espacio al cual, por definición pertenece la n -ésima derivada $f^{(n)}(x)$ de $f : E \rightarrow F$ cuando ella existe.

Supongamos ahora que la aplicación f tiene derivada n -ésima en sentido fuerte en un punto x del espacio normado E . De acuerdo al isomorfismo de que hablamos anteriormente, podemos dar una interpretación más clara a la diferencial de n -ésimo orden de la función f en el punto x . Hemos definido la diferencial en sentido fuerte de la aplicación f en el punto x como el resultado de aplicar al elemento $h \in E$ el operador lineal $f'(x)$:

$$df := f'(x)(h). \quad (48)$$

La diferencial de segundo orden se define como

$$d^2f := f''(x)(h, h), \quad (49)$$

es decir, como la expresión cuadrática asociada a la aplicación bilinial $f''(x) \in B(E^2, F)$.

De manera análoga, la diferencial de orden n se define mediante

$$d^n f := f^{(n)}(x)(h, \dots, h), \quad (50)$$

es decir, como aquel elemento del espacio F en el cual se transforma mediante la aplicación n -lineal $f^{(n)}(x)$ el elemento $(h, \dots, h) \in E^n$.

Utilizaremos ahora la noción introducida de diferencial de orden n para obtener una importante generalización de la conocida fórmula de Taylor del Análisis Clásico.

Sabemos que la diferenciabilidad fuerte de una aplicación $f : E \longrightarrow F$ significa que la diferencia $f(x+h) - f(x)$ se puede representar como la suma de un término lineal y otro sumando de orden superior al primero respecto a $\|h\|_E$. Con el objetivo de obtener una aproximación más exacta del incremento $f(x+h) - f(x)$ demostraremos la llamada fórmula de Taylor generalizada.

Teorema 7.3.5 (fórmula de Taylor generalizada). Sea f una aplicación definida sobre el abierto $U \subset E$ con valores en F y tal que $f^{(n)}(x)$ es uniformemente continua en U . Entonces se cumple la igualdad:

$$\begin{aligned} f(x+h) - f(x) &= f'(x)(h) + \frac{1}{2!}f''(x)(h, h) + \dots \\ &\dots + \frac{1}{n!}f^{(n)}(x)(h, \dots, h) + w(x; h). \end{aligned} \quad (51)$$

donde

$$\frac{\|w(x, h)\|_F}{\|h\|_E^n} \longrightarrow 0 \text{ cuando } \|h\|_E \longrightarrow 0. \quad (52)$$

Demostración: La demostración se realiza por inducción. Para $n = 1$ la igualdad (51) es trivial ya que coincide con la definición de diferencial según Frechet.

Supongamos que esta es válida para $n - 1$, cualquiera que sea la aplicación que satisfaga las condiciones del teorema. Entonces para la aplicación f' tendríamos:

$$\begin{aligned} f'(x+h) &= f'(x) + f''(x)(h) + \frac{1}{2!}f'''(x)(h, h) + \dots \\ &+ \dots + \frac{1}{(n-1)!}f^{(n)}(x)(h, \dots, h) + w_1(x; h), \end{aligned} \quad (53)$$

donde

$$\frac{\|w_1(x; h)\|_F}{\|h\|_E^n} \longrightarrow 0 \text{ cuando } \|h\|_E \longrightarrow 0.$$

Integrando en el segmento $[x, x+h]$ ambos miembros de la igualdad (53) y empleando la fórmula (38) de Newton-Leibnitz generalizada, obtenemos

$$\begin{aligned} f(x+h) - f(x) &= \int_0^1 f'(x+th) h dt = \\ &= \int_0^1 \left\{ f'(x) + t f''(x)(h) + \frac{1}{2!} t^2 f'''(x)(h, h) + \right. \\ &\quad \left. + \frac{1}{(n-1)!} t^{n-1} f^{(n)}(x)(h, \dots, h) \right\} h dt + R_n. \end{aligned} \quad (54)$$

donde $R_n = \int_0^1 w_1(x; th) h dt.$

De (54) se tiene que

$$f(x+h) - f(x) = f'(x)(h) + \frac{1}{2!}f''(x)(h, h) + \dots$$

$$\dots + \frac{1}{n!}f^{(n)}(x)(h, \dots, h) + R_n,$$

siendo

$$\|R_n\|_F \leq \int_0^1 \|w_1(x; h)\| \cdot \|h\| dt,$$

y obviamente, el miembro derecho de esta desigualdad satisface la condición (52), por lo cual

$$R_n = w(x; h).$$

Con esto queda probado el teorema. ■

En 7.5 veremos varias aplicaciones de la fórmula de Taylor al estudio de los problemas de extremos de funcionales.

Ejercicios de Autocontrol

1. Demuestre la proposición 7.3.2.
2. Aplique la fórmula de Taylor hasta el orden dos, para obtener un desarrollo del funcional f definido sobre $C([a, b], \mathbb{R})$, mediante:

$$f(x) = \int_a^b \int_a^b K(t, \tau) x(t) x(\tau) dt d\tau,$$

donde $K(t, \tau)$ es una función real continua en el cuadrado $[a, b] \times [a, b]$ que satisface la condición

$$K(t, \tau) = K(\tau, t).$$

3. Utilice la fórmula de Taylor para obtener un polinomio que aproxime a la función

$$f(x, y) = \sin xy$$

en el punto $(\pi + 1/10, \pi + 1/10)$ con un orden menor que 10^{-3} .

7.4 DIFERENCIACIÓN EN SENTIDO DÉBIL (DIFERENCIACIÓN SEGÚN GATEAUX). RELACIÓN ENTRE LAS DIFERENCIALES SEGÚN FRECHET Y GATEAUX

Hasta el momento hemos estudiado un concepto fuerte de diferenciabilidad conocido como diferencial de Frechet. Ahora introduciremos una nueva noción

más débil que la anterior y que es el análogo en el Análisis Clásico de la llamada derivada direccional. Este concepto es a veces suficiente para la obtención de condiciones suficientes en los problemas de optimización de funcionales que veremos en el próximo epígrafe.

Definición 7.4.1. Sea f una aplicación que actúa de E en F . Se llama *diferencial en sentido débil o diferencial según Gateaux* de la aplicación f en el punto x , al límite

$$Df(x; h) = \frac{d}{dt} f(x + th)|_{t=0} = \lim_{t \rightarrow 0} \frac{f(x + th) - f(x)}{t}, \quad (55)$$

donde la convergencia se entiende como la convergencia según la norma del espacio F .

Observación:. La diferencial débil puede no existir y aún cuando exista, a diferencia de la diferencial fuerte, puede no ser lineal con respecto a h . Si esta linealidad tiene lugar, es decir, si

$$Df(x; h) = f'_d(x)(h), \quad (56)$$

donde $f'_d(x)$ es un operador lineal de E en F , entonces a este operador se le llama derivada débil o *derivada según Gateaux* de la aplicación f .

Señalemos además que para las derivadas débiles no se cumple, como regla general, el teorema sobre la diferenciación de la función compuesta. Se sugiere al lector buscar un ejemplo de este hecho.

La siguiente proposición nos muestra que, en efecto, la diferenciación según Frechet es más fuerte que la diferenciación según Gateaux.

Proposición 7.4.2. Sea f una aplicación de E en F . Si f es diferenciable en sentido fuerte en el punto $x \in E$, entonces también lo es en sentido débil y además sus diferenciales (derivadas) fuerte y débil coinciden.

Demostración: En efecto, para una aplicación fuertemente diferenciable, tenemos:

$$f(x + th) - f(x) = f'(x)(th) + \alpha(x; th),$$

y por lo tanto,

$$\frac{f(x + th) - f(x)}{t} = f'(x)(h) + \frac{\alpha(x; th)}{t} \longrightarrow f'(x)(h)$$

cuando $t \longrightarrow 0$. ■

De la proposición anterior obtenemos que todos los ejemplos, dados en 7.1 de funciones Frechet diferenciables, nos brindan ejemplos de funciones

diferenciables según Gateaux; sin embargo, podemos debilitar las suposiciones hechas sobre f y obtener funciones débilmente diferenciables que no lo son en sentido fuerte, tal y como veremos en el proximo ejemplo. No obstante, debemos señalar que existen resultados del tipo obtenidos en 7.1 para funciones fuertemente diferenciables y que se cumplen también bajo la condición más débil de diferenciability según Gateaux. Tal es el caso de la proposición 7.1.7 y su corolario, cuya verificación para el caso de funciones débilmente diferenciables se deja al lector.

Ejemplo 5. El presente ejemplo nos muestra que la diferenciability en sentido débil y fuerte son conceptos diferentes aún en el caso de espacios de dimensión finita. Efectivamente, se conoce de los cursos de Análisis Clásico para una función numérica $f(x) = f(x_1, \dots, x_n)$, que la existencia de la derivada direccional $\frac{d}{dt}f(x + th)$ para cualquier h fijo no implica necesariamente en el caso $n \geq 2$, la diferenciability (fuerte) de esta función, es decir, la posibilidad de representar su incremento $f(x + th) - f(x)$ como suma de una parte lineal respecto a h y un término que tiende a 0 cuando $h \rightarrow 0$, más rápidamente que $|h|$.

Como ejemplo trivial de esta situación podemos citar la función de dos variables

$$f(x_1, x_2) = \begin{cases} \frac{x_1^3 x_2}{x_1^4 + x_2^2} + x_1 + x_2, & \text{si } x_1^2 + x_2^2 \neq 0, \\ 0, & \text{si } x_1 = x_2 = 0. \end{cases}$$

No es difícil verificar que esta función es continua en todo el plano incluyendo el origen $(0, 0)$. Además, ésta tiene diferencial débil en el punto $(0, 0)$, ya que

$$\lim_{t \rightarrow 0} \frac{f(0 + th) - f(0)}{t} = \lim_{t \rightarrow 0} \left(h_1 + h_2 + \frac{t^3 h_1^3 h_2}{t^4 h_1^4 + t^2 h_2^2} \right) = h_1 + h_2$$

Sin embargo, esta diferencial no constituye la parte lineal principal del incremento de la función f en el punto $(0, 0)$. En efecto, sea

$$\alpha(0; h) = f(0 + h) - f(0) - (h_1 + h_2) = \frac{h_1^3 h_2}{h_1^4 + h_2^2}.$$

Entonces, tomando $h_2 = h_1^2$, tendremos

$$\lim_{|h| \rightarrow 0} \frac{\alpha(0, h)}{|h|} = \lim_{h_1 \rightarrow 0} \frac{h_1^5}{2h_1^4 \sqrt{h_1^2 + h_2^2}} = \frac{1}{2} \neq 0.$$

Recordemos, sin embargo, como vimos en el ejemplo 1 que la existencia y continuidad de las n derivadas parciales de f es suficiente para la existencia de su diferencial fuerte. Notemos además, que la derivada parcial $\frac{\partial f}{\partial x_k}$ coincide

con el límite (55) cuando h es el k -ésimo vector coordenado unitario.

Ejemplo 6. Para toda función abstracta $f : \mathbb{R} \longrightarrow E$, la diferenciabilidad en sentido fuerte y débil coinciden.

El siguiente teorema nos muestra cuando se cumple el inverso de la proposición 7.4.2.

Teorema 7.4.3. Sea f una aplicación de E en F . Si la derivada débil $f'_d(x)$ de la aplicación f existe en una vecindad $V(x_0)$ del punto x_0 y representa en esta vecindad una función continua de la variable x con valores en $L(E, F)$, entonces la derivada fuerte $f'(x_0)$ existe en el punto x_0 y coincide con la débil.

Demostración: Por hipótesis, la aplicación f tiene derivada débil, es decir $df(x_0; h) = f'_d(x_0)(h)$. Tomemos h de manera tal que $x_0 + h \in V(x_0)$ y consideremos la expresión

$$\alpha(x_0; h) = f(x_0 + h) - f(x_0) - f'_d(x_0)(h). \quad (57)$$

Sea ahora φ un elemento arbitrario del espacio F^* (dual de F); entonces se tiene:

$$\varphi(\alpha(x_0; h)) = \varphi(f(x_0 + h) - f(x_0)) - \varphi(f'_d(x_0)(h)). \quad (58)$$

Si consideramos la aplicación

$$g(t) = f(x_0 + th)$$

de argumento numérico t , esta función es diferenciable respecto a t y para ésta

$$\frac{dg}{dt} = \lim_{t \rightarrow 0} \varphi \left(\frac{f(x_0 + th + \Delta th) - f(x_0 + th)}{\Delta t} \right).$$

Luego, aplicando a g la fórmula del valor medio del Análisis Clásico, podemos escribir la igualdad (58) en la forma

$$\varphi(\alpha(x_0; h)) = \varphi[f'_d(x_0 + \tau h) - f'_d(x_0)](h) \quad (59)$$

donde $0 < \tau < 1$. Para h fijo, el elemento $\varphi \in F^*$ se puede escoger de manera tal, que $\|\varphi\| = 1$ y que cumpla la desigualdad

$$|\varphi(\alpha(x_0; h))| \leq \frac{1}{2} \|\alpha(x_0; h)\|_F \|\varphi\| = \frac{1}{2} \|\alpha(x_0; h)\|_F.$$

Por lo tanto, de este resultado y de la igualdad (59) encontramos que

$$\|\alpha(x_0; h)\|_F \leq 2 \|f'_d(x_0 + \tau h) - f'_d(x_0)\| \|h\|_E.$$

Pero como hemos supuesto que $f'_d(x)$ es una función continua de x , entonces

$$\lim_{h \rightarrow 0} \|f'_d(x_0 + \tau h) - f'_d(x_0)\| = 0$$

por lo cual

$$\lim_{h \rightarrow 0} \frac{\|\alpha(x_0; h)\|_F}{\|h\|_E} = 0,$$

y con ello, queda probado que $f'_d(x_0)(h)$ coincide con la diferencial fuerte de f en el punto x_0 . ■

Ejercicios de Autocontrol

1. Dé un ejemplo de dos aplicaciones para las cuales existe la derivada en sentido débil pero no en sentido fuerte y que para su compuesta no se cumpla la regla de diferenciación dada en el teorema 7.1.6.
2. Sobre el espacio normado $(C([0, 1], \mathbb{R}), \|\cdot\|_\infty)$ definamos el funcional $f(x) = \max_{0 \leq t \leq 1} x(t)$. Determine para cuales funciones x existe la diferencial en sentido débil de f y es además lineal en h .
3. Pruebe que la función

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4}, & \text{si } x \neq 0 \\ 0, & \text{si } x = 0 \end{cases}$$

es Gateaux diferenciable pero no continua en $x = y = 0$.

7.5 EXTREMOS LOCALES DE FUNCIONALES

Aplicaremos en este epígrafe los conceptos de diferenciabilidad según Gateaux y Frechet al problema de la minimización o maximización de un funcional definido sobre un espacio vectorial normado. Este ha sido históricamente uno de los temas más estudiados del Análisis Funcional no lineal, encontrado su primera formalización en el llamado Cálculo de Variaciones. Sin embargo, existe una teoría moderna mucho más amplia de los llamados problemas extremales y que incluye como caso particular al Cálculo de Variaciones clásico. En toda esta teoría, generalmente se supone que dado un funcional $f : E \rightarrow \mathbb{R}$, este alcanza un máximo local o un mínimo en un punto $x_0 \in E$ (x_0 perteneciendo a una cierta parte E), y se hallan condiciones necesarias para que x_0 sea un tal punto de extremo local. La obtención de condiciones suficientes de extremo es, por lo general, un problema mucho más difícil y requiere de la imposición de condiciones más fuertes al comportamiento local o global del funcional f .

En este epígrafe nos limitaremos a utilizar los resultados antes obtenidos, para exponer algunas de la técnicas más usuales en el estudio de los problemas de extremos locales de funcionales.

A lo largo de todo el epígrafe, entenderemos por *funcional* toda aplicación $f : E \longrightarrow \mathbb{R}$, donde E es un espacio normado real.

Definición 7.5.1. Sea f un funcional real definido sobre un subconjunto Ω de un espacio normado E . Un punto $x_0 \in \Omega$ se llama *mínimo local* de f sobre Ω si existe una bola abierta $B(x_0, r)$ tal que $f(x_0) \leq f(x)$; para todo $x \in \Omega \cap B(x_0, r)$.

El punto x_0 se llama *mínimo local estricto* de f sobre Ω si la desigualdad estricta $f(x_0) < f(x)$ se satisface para todo $x \in \Omega \cap B(x_0, r)$ con $x \neq x_0$ y *máximo local estricto* se define análogamente. Diremos que x_0 es un *extremo local* de f sobre Ω si es un máximo o un mínimo local.

El conjunto Ω sobre el cual se plantea el problema de extremos se llama *conjunto admisible*.

Si Ω es compacto y f continua sabemos, por los resultados del capítulo 4, que existe un máximo absoluto y un mínimo absoluto de f sobre Ω . En muchas ocasiones, en particular en muchos problemas de la física, resulta a veces más importante el concepto de extremo local que el de extremo absoluto. Además, en los problemas de la práctica no se tiene, por lo general, la compacidad de Ω y aún cuando se tenga, los resultados obtenidos en el capítulo 4 no nos brindan ningún criterio para *detectar* los posibles extremos locales de f .

Para los funcionales definidos sobre $\mathbb{R}^n$, es bien conocida la siguiente condición necesaria de extremo: si la función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ es diferenciable en el punto $x_0 = (x_1^0, \dots, x_n^0)$ y tiene un extremo local en este punto, entonces $df = 0$ en dicho punto, lo cual es equivalente a que

$$\frac{\partial f}{\partial x_1}(x_0) = \frac{\partial f}{\partial x_2}(x_0) = \dots = \frac{\partial f}{\partial x_n}(x_0) = 0.$$

Esta condición se extiende inmediatamente a funcionales definidas sobre un espacio normado arbitrario.

Teorema 7.5.2 (condición necesaria de extremo local). Sea f un funcional real definido sobre el espacio normado E y diferenciable en sentido débil en el punto $x_0 \in E$. Una condición necesaria para que f tenga un extremo local $x_0 \in E$ es que

$$Df(x_0; h) = 0 \text{ para todo } h \in E. \quad (60)$$

Demostración: Para todo $h \in E$, la función real $f(x_0 + th)$ de la derivada real t debe alcanzar un extremo en $t = 0$; entonces, por un resultado conocido del cálculo clásico

$$Df(x_0; h) = \left. \frac{d}{dt} f(x_0 + th) \right|_{t=0} = 0. \blacksquare$$

Un punto $x_0 \in E$ en el cual $Df(x_0; h) = 0$ para todo $h \in E$ se llama *punto estacionario* de f . El teorema 7.5.2 nos dice que los extremos locales de f sobre todo E deben buscarse en el conjunto de sus puntos estacionarios. El resultado anterior también es una condición necesaria de extremo local cuando f está definida sobre un subconjunto abierto Ω de E en lugar de sobre todo E .

A pesar de la simpleza del teorema anterior esto no opaca su gran utilidad. La mayor parte de los resultados del Cálculo de Variaciones pueden obtenerse como una consecuencia del teorema 7.5.2.

Muchos problemas interesantes sobre optimización de funcionales pueden ser tratados en un espacio normado apropiado, reduciéndose al cálculo de una cierta diferencial. Existen muchas generalizaciones importantes cuando el conjunto admisible en lugar de ser todo el espacio E se sustituye por un subconjunto Ω con determinadas propiedades. Nosotros no entraremos en ese detalle (vea el problema 1 de autocontrol), en su lugar nos dedicaremos a estudiar un poco el problema clásico del Cálculo de Variaciones aplicando el resultado del teorema 7.5.2.

En la Mecánica Clásica, el comportamiento de un sistema físico de partículas está completamente determinado por las propiedades de una cierta función que depende de las posiciones $x(t)$ de las partículas, sus velocidades $\dot{x}(t)$ y el tiempo t .

A tal función se le llama *Lagrangiana* del sistema y se denota por $L(x(t), \dot{x}(t), t)$. Se llama acción de la Lagrangiana L a lo largo de la trayectoria $x(t)$ durante el intervalo $[t_1, t_2]$ a la integral

$$J = \int_{t_1}^{t_2} L(x(t), \dot{x}(t), t) dt. \quad (61)$$

Uno de los principios básicos de la Mecánica Clásica es el llamado *principio de mínima acción* de Halmilton que nos dice que la trayectoria $x_0(t)$ descrita por un sistema físico de partículas es precisamente aquella a lo largo de la cual la acción (61) alcanza un valor mínimo. Si de alguna manera logramos obtener caracterizaciones bien sean necesarias o suficientes para los extremos de J , a tales caracterizaciones le llamaremos *ecuaciones de movimiento* del sistema descrito por la Lagrangiana L .

El problema Clásico del Cálculo de Variaciones consiste precisamente en hallar una función $x_0(t)$ en el intervalo $[t_1, t_2]$ que minimice a una integral del tipo (61).

Para especificar el problema completamente, debemos aclarar cual es la clase de funciones $x(t)$ sobre la cual se busca el extremo, es decir cuál es el

conjunto admisible. Nosotros supondremos que la función L es continua en $x, \dot{x}$ y t y que además, tiene derivadas parciales continuas L_x y $L_{\dot{x}}$ con respecto a x y $\dot{x}$.

En calidad de espacio normado consideraremos el espacio

$$C^1([t_1, t_2], \mathbb{R}) = \left\{ \begin{array}{l} x \in C([t_1, t_2], \mathbb{R}) : x \\ \text{es continuamente derivable} \end{array} \right\} \quad (62)$$

provisto de la norma

$$\|x\|_{\infty}^{(1)} = \max_{t_1 \leq t \leq t_2} |x(t)| + \max_{t_1 \leq t \leq t_2} |\dot{x}(t)|. \quad (63)$$

El lector puede verificar sin dificultad que $\|\cdot\|_{\infty}^{(1)}$ es una norma sobre:

$C^1([t_1, t_2], \mathbb{R})$ y que el espacio normado obtenido es de Banach.

Analizaremos la versión más simple del problema donde se supone que los extremos $x(t_1)$ y $x(t_2)$ son fijos. Con esta última suposición delimitamos totalmente el conjunto admisible y finalmente obtenemos el siguiente planteamiento del *problema clásico del Cálculo de Variaciones*:

Hallar una condición necesaria de extremo para el funcional

$$J : \Omega \longrightarrow \mathbb{R}, \quad \Omega \subset E,$$

donde J viene dado por

$$J(x) = \int_0^{t_2} f(x(t), \dot{x}(t), t) dt \quad (64)$$

y f satisface las propiedades exigidas a L . Además, aquí

$$E = (C^1([t_1, t_2], \mathbb{R}), \|\cdot\|_{\infty}^{(1)}) \quad (65)$$

y

$$\Omega = \{x \in E : x(t_1) = a, x(t_2) = b; a \text{ y } b \text{ fijos}\}. \quad (66)$$

Teorema 7.5.3 (problema clásico del Cálculo de Variciones. Ecuaciones de Euler-Lagrange). Supongamos que $f(x, \dot{x}, t)$ donde

$$f : \mathbb{R}^2 \times [t_1, t_2] \longrightarrow \mathbb{R}$$

es una función continua con respecto a las variables $x, \dot{x}, t$ y continuamente derivable con respecto a x y $\dot{x}$.

Entonces una condición necesaria para que la función $x(t)$ sea un extremo del funcional J (ver (64)) sobre el conjunto Ω (vea (66)) es que $x(t)$ satisfaga la *ecuación de Euler-Lagrange*.

$$\frac{d}{dt} f_x(x, \dot{x}, t) = f_x(x, \dot{x}, t). \quad (67)$$

Demostración: Tomemos una función admisible x ($x \in \Omega$) y consideremos vectores de la forma $x + h$ que también sean admisibles. La clase de tales vectores es llamada *clase de variaciones admisibles*. En nuestro caso, es claro que la clase de variaciones admisibles es el subconjunto de $C^{(1)}([t_1, t_2], \mathbb{R})$, constituido por todas las funciones continuamente derivables que se anulan en t_1 y t_2 . Según el teorema 7.5.2, la condición de extremo es que para todos estos h se cumpla:

$$DJ(x; h) = 0.$$

Pero la diferencial débil de J es:

$$DJ(x; h) = \frac{d}{d\alpha} \int_{t_1}^{t_2} f(x + \alpha h, \dot{x} + \alpha \dot{h}, t) dt \Big|_{\alpha=0}$$

o, de forma equivalente

$$DJ(x; h) = \int_{t_1}^{t_2} f_x(x, \dot{x}, t) h(t) dt + \int_{t_1}^{t_2} f_{\dot{x}}(x, \dot{x}, t) h(t) dt. \quad (68)$$

Del resultado del ejemplo 2 se deduce que esta diferencial es precisamente la diferencial de J según Frechet. Igulemos esta diferencial a cero y asumamos adicionalmente por comodidad que $f_{\dot{x}}$ tiene derivada continua con respecto a t cuando $x(t)$ coincide con la solución óptima. En este caso podemos integrar por partes en (68) obteniendo

$$DJ(x; h) = \int_{t_1}^{t_2} [f_x(x, \dot{x}, t) - \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t)] h(t) dt +$$

$$+ f_{\dot{x}}(x, \dot{x}, t) h(t) \Big|_{t_1}^{t_2} = 0.$$

El segundo sumando del miembro derecho se anula ya que $h(t_1) = h(t_2) = 0$; luego la condición necesaria para que $x(t)$ sea extremo del funcional J es

$$\int_{t_1}^{t_2} [f_x(x, \dot{x}, t) - \frac{d}{dt} f_{\dot{x}}(x, \dot{x}, t)] h(t) dt = 0 \quad (69)$$

para todo h de $C^{(1)}([t_1, t_2], \mathbb{R})$ que se anula en t_1 y t_2 .

Como el término que multiplica a $h(t)$ en el integrando es continuo, se prueba inmediatamente (vea lema 7.5.4) que se debe anular idénticamente sobre $[t_1, t_2]$.

Por lo tanto, concluimos que la función extremal $x(t)$ debe satisfacer la ecuación de Euler-Lagrange (67). ■

Observación: A priori no tendría sentido la derivación con respecto a t en la ecuación de Euler-Lagrange, si no hubiéramos supuesto en la demostración

del teorema anterior la continuidad de $\frac{d}{dt}f_x$ en la solución óptima. Hasta el momento no tenemos nada que nos asegure que en efecto, la función extremal $x(t)$, es suficientemente suave como para que ocurra la suposición adicional hecha en el teorema. Los tres lemas siguientes nos muestran que es posible hacer esta suposición.

Lema 7.5.4. Si $\alpha(t)$ es continua sobre $[t_1, t_2]$ y $\int_{t_1}^{t_2} \alpha(t) h(t) dt = 0$ para toda $h \in C^{(1)}([t_1, t_2], \mathbb{R})$, con $h(t_1) = h(t_2) = 0$, entonces $\alpha(t) = 0$ sobre $[t_1, t_2]$.

Demostración: Supongamos que $\alpha(t)$ no es idénticamente nula, digamos positiva, para algún $t \in [t_1, t_2]$.

Entonces, $\alpha(t)$ es positiva sobre algún intervalo $[t', t''] \subset [t_1, t_2]$.

Sea

$$h(t) = \begin{cases} (t - t')^2(t'' - t)^2, & \text{si } t \in [t', t''] \\ 0, & \text{en otro caso.} \end{cases}$$

Obviamente la función h satisface las hipótesis del lema y

$$\int_{t_1}^{t_2} \alpha(t) h(t) dt > 0 \text{ lo cual conduce a una contradicción con lo supuesto. } \blacksquare$$

Lema 7.5.5. Si $\alpha(t)$ es continua en $[t_1, t_2]$ y $\int_{t_1}^{t_2} \alpha(t) h(t) dt = 0$ para toda $h \in C^{(1)}([t_1, t_2], \mathbb{R})$, con $h(t_1) = h(t_2) = 0$ entonces $\alpha(t) = c$ sobre $[t_1, t_2]$, donde c es una constante.

Demostración: Sea $\int_{t_1}^{t_2} \alpha(t) dt = K$ y consideremos $c = K/(t_2 - t_1)$.

Obviamente

$$\int_{t_1}^{t_2} [\alpha(t) - c] dt = 0.$$

Definamos

$$h(t) = \int_{t_1}^t [\alpha(\tau) - c] d\tau.$$

Entonces $h(t)$ satisface las hipótesis del lema y

$$\begin{aligned} \int_{t_1}^{t_2} [\alpha(t) - c]^2 dt &= \int_{t_1}^{t_2} [\alpha(t) - c] h(t) dt = \\ &= \int_{t_1}^{t_2} \alpha(t) h(t) dt - c[h(t_2) - h(t_1)] = 0, \end{aligned}$$

de donde concluimos que $\alpha(t) \equiv c$. ■

Lema 7.5.6. Si $\alpha(t)$ y $\beta(t)$ son continuas sobre $[t_1, t_2]$ y

$$\int_{t_1}^{t_2} [\alpha(t) h(t) + \beta(t) \dot{h}(t)] dt = 0 \tag{70}$$

para todo $h \in C^{(1)}([t_1, t_2], \mathbb{R})$ con $h(t_1) = h(t_2) = 0$ entonces β es derivable y $\dot{\beta}(t) = \alpha(t)$ en $[t_1, t_2]$.

Demostración: Definamos

$$A(t) = \int_{t_1}^t \alpha(\tau) d\tau.$$

Entonces integrando por partes tenemos

$$\int_{t_1}^{t_2} \alpha(t) h(t) dt = - \int_{t_1}^{t_2} A(t) \dot{h}(t) dt.$$

Luego, (70) se convierte en

$$\int_{t_1}^{t_2} [-A(t) + \beta(t)] \dot{h}(t) dt = 0.$$

lo cual, por el lema 7.5.5 implica

$$\beta(t) = A(t) + c,$$

para alguna constante c . De aquí concluimos, por la definición de A , que $\dot{\beta}(t) = \alpha(t)$. ■

Observación: Como consecuencia del lema 7.5.6, es claro que la ecuación de Euler-Lagrange se deduce directamente de la ecuación (68) sin asumir a priori la derivabilidad de f_x con respecto a t .

Ejemplo 7 (arcos de longitud mínima). Dados a y b en $\mathbb{R}^n$, emplearemos las ecuaciones de Euler-Lagrange para determinar la curva continuamente diferenciable de longitud mínima que conecta los puntos a y b . Es decir, se trata de minimizar el funcional

$$J(x) = \int_{t_1}^{t_2} \sqrt{1 + (\dot{x}(t))^2} dt,$$

el cual representa la longitud de la curva x en $\mathbb{R}^n$, parametrizada por la función continuamente diferenciable $x(t)$ en el intervalo $[t_1, t_2]$. Es decir,

$$\min \{ J(x) : x \in C^{(1)}([t_1, t_2], \mathbb{R}^n), x(t_1) = a, x(t_2) = b \}.$$

De (67) se obtiene inmediatamente

$$\frac{d}{dt} \frac{\partial}{\partial \dot{x}} \sqrt{1 + (\dot{x})^2} = 0,$$

o de forma equivalente $\dot{x} = \text{constante}$.

Luego, hemos probado que la curva de longitud mínima (curva geodésica) que conecta dos puntos a y b de $\mathbb{R}^n$ es el segmento de recta que une estos dos puntos.

En general podemos plantearnos el problema de calcular la curva geodésica que une dos puntos cualesquiera de una superficie dada manteniéndose siempre sobre dicha superficie. Este es un problema más complejo del Cálculo de Variaciones y que no veremos en este texto.

Ejemplo 8 (un problema de planificación). Supongamos que nos encontramos en la situación de una persona que debe hacer un plan de *inversiones y gastos*, para un intervalo de tiempo T , de manera tal, que se maximice el disfrute total y teniendo una cantidad de recursos fijos S .

Asumiremos que no se tiene ninguna otra entrada de recurso aparte de la que proviene de las inversiones hechas con la cantidad S . El disfrute obtenido por unidad de recurso invertido (*a lo que llamaremos utilidad de la inversión*) en un instante de tiempo t dado ($0 \leq t \leq T$) es una cierta función U que depende del nivel de los gastos al cual denotaremos por $r(t)$.

Luego deseamos maximizar una expresión del tiempo

$$\int_0^T e^{-\beta t} U(r(t)) dt.$$

Aquí el factor $e^{-\beta t}$ refleja el hecho de que el disfrute en un futuro interesa menos que en un tiempo cercano.

Si $x(t)$ es el capital total en el instante t , entonces la velocidad $\dot{x}(t)$ con que varían los recursos de que disponemos es proporcional a $x(t)$ si le restamos la razón de gastos $r(t)$. Es decir,

$$\dot{x}(t) = \alpha x(t) - r(t).$$

donde la constante α denota el nivel de las inversiones por unidad de recursos invertidos. Luego, nuestro problema consiste en maximizar la expresión

$$\int_0^T e^{-\beta t} U(\alpha x(t) - \dot{x}(t)) dt$$

sujeto a las restricciones $x(0) = S$ y $x(T) = 0$ (o algún otro valor fijo). De hecho, hay una restricción adicional que es $x(t) \geq 0$, pero en el caso que consideraremos se satisface automáticamente.

La ecuación de Euler-Lagrange para este problema es

$$\alpha e^{-\beta t} U'(\alpha x(t) - \dot{x}(t)) \frac{d}{dt} e^{-\beta t} U'(\alpha x(t) - \dot{x}(t)) = 0$$

donde U' es la derivada de la función de utilidad U . Esto nos conduce a

$$\frac{d}{dt} U'(\alpha x(t) - \dot{x}(t)) = (\beta - \alpha) U'(\alpha x(t) - \dot{x}(t)). \quad (71)$$

Luego, integrando (71), hallamos que

$$U' [r (t)] = U' (r (0)) e^{(\beta-\alpha)t}. \tag{72}$$

Por lo tanto, la forma en que r depende del tiempo t , puede ser obtenida explícitamente si se especifica la expresión de U .

Una función de utilidad simple y que refleja la noción intuitiva de disminución marginal del disfrute (es decir, U' es decreciente), así como el disfrute marginal infinito cuando no se hacen gastos (es decir, $U' [0] = \infty$), es la función

$$U (r) = 2r^{1/2}.$$

Sustituyendo esta expresión de U en (72), obtenemos

$$\alpha x (t) - \dot{x} (t) = r (t) = r (0) e^{2(\alpha-\beta)t}.$$

Integrando esta última ecuación, llegamos a que

$$\begin{aligned} x (t) &= e^{\alpha t} x (0) + \frac{r (0)}{\alpha - 2\beta} [e^{\alpha t} - e^{2(\alpha-\beta)t}] = \\ &= \left\{ x (0) - \frac{r (0)}{2\beta - \alpha} \right\} e^{\alpha t} + \frac{r (0)}{2\beta - \alpha} e^{2(\alpha-\beta)t}. \end{aligned} \tag{73}$$

Si suponemos que $\alpha > \beta > \alpha/2$, podemos hallar $r (0)$ de (73) si exigimos que $x (T) = 0$.

Entonces,

$$r (0) = \frac{(2\beta - \alpha)x (0)}{1 - e^{(\alpha-2\beta)T}},$$

es decir, durante el disfrute óptimo de los recursos, estas deben crecer inicialmente y a partir de un momento comienza a decrecer hasta cero.

Pasemos finalmente a estudiar una condición suficiente de extremo para funcionales definidas sobre un espacio de Banach. Hasta el momento, solo hemos visto una condición necesaria de extremo para funcionales. Esta condición nos dice (7.5.2) que si el funcional f es Gateaux diferenciable sobre E , entonces sus extremos locales deben buscarse entre aquellos puntos x de E donde se anula su diferencial en sentido débil.

Si f está definido sobre un espacio de dimensión finita $\mathbb{R}^n$, entonces resolver el problema de si uno de estos puntos es o no efectivamente un extremo, debe considerarse la segunda diferencial. Más exactamente se cumple:

1. Si la función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ tiene diferencial de segundo orden y alcanza un mínimo local en el punto $(x_1^0, \dots, x_n^0)$, entonces en ese punto la matriz $d^2 f$ es definida positiva. (Análogamente, si en el punto $(x_1^0, \dots, x_n^0)$ hay un máximo local, en ese punto $d^2 f$ es definida negativa.)
2. Si en el punto $x_0 = (x_1^0, \dots, x_n^0)$ se satisfacen las condiciones

$$df(x_0) = 0 \text{ y } d^2f = \sum_{i,k=1}^n \frac{\partial^2 f(x_0)}{\partial x_i \partial x_k} x_i x_k > 0$$

para todo $x = (x_1, \dots, x_n)$ no idénticamente nulo, entonces la función $f(x)$ tiene en ese punto un mínimo local (máximo local si $d^2f < 0$).

Analicemos en qué medida se conservan estos resultados para funcionales definidas sobre un espacio de Banach.

Teorema 7.5.7 (condición suficiente de mínimo). Sea f un funcional definido en el espacio de Banach E y que verifica las condiciones:

- i) $f'(x_0) = 0$.
- ii) $f''(x_0)$ es un funcional cuadrático estrictamente positivo.

Entonces f tiene un mínimo local en el punto x_0 .

Demostración: Recordemos que un funcional cuadrático B definido en un espacio normado E se llama estrictamente positivo, cuando se puede encontrar una constante $c > 0$, tal que se cumple $B(x, x) \geq c \|x\|^2$, para todo $x \in E$.

Elijamos $\epsilon > 0$ suficientemente pequeño de manera que para $\|h\| < \epsilon$ la magnitud $w(x_0, h)$ en la magnitud (74) satisfaga la condición

$$|w(x_0; h)| \leq \frac{c}{4} \|h\|^2,$$

donde la constante $c > 0$ se elige de manera tal que $f''(x_0)(h, h) \geq c \|h\|^2$.

Entonces,

$$f(x_0 + h) - f(x_0) = \frac{1}{2} f''(x_0)(h, h) +$$

$$+ w(x_0, h) > \frac{c}{4} \|h\|^2 > 0$$

para $\|h\| < \epsilon$, lo cual contradice la propiedad de mínimo local de x_0 . ■

Teorema 7.5.8 (condición necesaria de mínimo local). Sea f un funcional real definido sobre el espacio de Banach E con la segunda derivada continua en una vecindad del punto x_0 . Si este funcional alcanza un mínimo en el punto x_0 , entonces $d^2f(x_0) \geq 0$, es decir,

$$d^2f(x_0)(h, h) \geq 0, \text{ para todo } h \in E.$$

Demostración: Por la fórmula de Taylor, tenemos

$$f(x_0 + h) - f(x_0) = f'(x_0)(h) + \frac{1}{2} f''(x_0)(h, h) + w(x_0, h),$$

donde $\frac{|w(x_0, h)|}{\|h\|^2} \rightarrow 0$ cuando $\|h\| \rightarrow 0$.

Si el funcional f tiene un mínimo en x_0 entonces $f'(x_0) = 0$ y nos queda

$$f(x_0 + h) - f(x_0) = \frac{1}{2}f''(x_0)(h, h) + w(x_0, h). \quad (74)$$

Si para algún h admisible tiene lugar la desigualdad

$$f''(x_0)(h, h) < 0 \quad (75)$$

entonces, teniendo en cuenta que

$$f''(x_0)(\epsilon h, \epsilon h) = \epsilon^2 f''(x_0)(h, h),$$

se tendrá que existen elementos de h de norma tan pequeña como se quiera para los cuales también se cumple (75). Pero el signo de la expresión (74) depende, para $\|h\|$ suficientemente pequeño, del signo de su término principal $\frac{1}{2}f''(x_0)(h, h)$ y obtendríamos que

$$f(x_0 + h) - f(x_0) = \frac{1}{2}f''(x_0)(h, h) + w(x_0, h) < 0,$$

es decir, para ese valor de h

$$f(x_0 + h) < f(x_0)$$

lo cual sería una contradicción con el hecho de que x_0 es un mínimo local de f . Análogamente se analiza el caso cuando x_0 es un máximo local. ■

El teorema demostrado es una generalización inmediata del teorema correspondiente para funciones de un número finito de variables. Sin embargo, la situación cambia en el caso de la condición suficiente según nos muestra el ejemplo siguiente:

Ejemplo 9. En este ejemplo mostraremos que la condición mencionada anteriormente $f''(x_0)(h, h) > 0$, que es suficiente para la existencia de mínimo entre aquellos puntos en los cuales $f'(x_0) = 0$ en el caso de funciones de n variables reales, no resulta suficiente para funciones definidas en un espacio de Banach de dimensión infinita. En efecto consideremos, en el espacio de Hilbert $l^2(\mathbb{N})$, el funcional

$$f(x) = \sum_{n=1}^{\infty} \frac{x_n^2}{n^3} - \sum_{n=1}^{\infty} x_n^4,$$

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

Obviamente, en el punto $x = 0$ la primera diferencial de este funcional se anula y la segunda es igual a $d^2f(0)(h, h) = \sum_{n=1}^{\infty} \frac{h_n^2}{n^3}$, es decir, representa un funcional cuadrático definido positivo. Sin embargo, en el punto 0 no hay un mínimo local, ya que $f(0) = 0$ y $f(0, \dots, 0, \frac{1}{n}, 0, \dots) = \frac{1}{n^5} - \frac{1}{n^4} < 0$.

Por consiguiente, en cualquier vecindad del punto $x = 0$ existen puntos $x \in l^2(\mathbb{N})$ en los cuales $f(x) < f(0)$.

Introduzcamos el concepto siguiente:

Definición 9. Un funcional cuadrático B definido sobre el espacio normado E se llama estrictamente positivo, cuando existe una constante $c > 0$ tal que $B(x, x) \geq c \|x\|^2$, para todo $x \in E$.

Observación: El lector puede verificar sin dificultad, que en un espacio de dimensión finita la positividad en sentido estricto de un funcional bilineal cuadrático es equivalente a que sea definido positivo. Sin embargo, como nos muestra el ejemplo 9 ambas nociones no son equivalentes en espacios normados de dimensión infinita.

En el cálculo de variaciones se establecen condiciones suficientes de extremo más finas que dependen de la estructura concreta del funcional analizado y por lo tanto, son más sencillas de verificar en casos particulares. Estas condiciones, sin embargo, se salen del objetivo de este texto y por ello no las expondremos en él.

Ejercicios de Autocontrol

1. Pruebe el siguiente teorema que es una generalización de 7.5.2:

Sea f un funcional real definido sobre un espacio vectorial normado E . Supongamos que x_0 es un mínimo local de f sobre el conjunto convexo Ω de E y que f es Gateaux diferenciable en x_0 . Entonces,

$$Df(x_0; x - x_0) \geq 0$$

para todo $x \in \Omega$.

2. Demuestre que el espacio

$$(C^{(1)}([t_1, t_2], \mathbb{R}), \|\cdot\|_{\infty}^{(1)})$$

definido en (62)–(63) es un espacio de Banach, pero que no es de Banach si se sustituye $\|\cdot\|_{\infty}^{(1)}$ por $\|\cdot\|_{\infty}$.

3. Demuestre el enunciado de la observación que aparece después de la definición 7.5.9.
4. Sea f el funcional definido sobre $(C([t_1, t_2], \mathbb{R}), \|\cdot\|_{\infty})$ mediante

$$f(x) = \int_a^b K(t, x(t)) dt$$

donde K es una función continuamente diferenciable.

Utilice el teorema 7.5.2 para demostrar que la condición

$$\frac{d}{dt}K_x(t, x(t)) = 0,$$

determina una condición necesaria de extremo para el funcional f .

EJERCICIOS

CAPÍTULO 1

1. Demuestre que las siguientes propiedades son equivalentes:

- a) $A \subset B$
- b) $A \cap B = A$
- c) $A \cup B = B$

2. Demuestre las inclusiones siguientes:

- a) $(A \cap C) \cup (B \cap D) \subset (A \cup B) \cap (C \cup D)$
- b) $(B \setminus C) \setminus (B \setminus A) \subset A \setminus C$
- c) $A \setminus C \subset (A \setminus B) \cup (B \setminus C)$

3. Demuestre las igualdades siguientes:

- a) $A \setminus (B \setminus C) = (A \setminus B) \setminus (A \cap C)$
- b) $(A \setminus B) \setminus C = (A \setminus C) \setminus (B \setminus C)$
- c) $(A \setminus B) \cup (B \setminus C) \cup (C \setminus A) \cup (A \cap B \cap C) = A \cup B \cup C$
- d) $(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C) = (A \cap C) \setminus B$
- e) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$
- f) $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$

4. Diga si es cierta la implicación siguiente: $A \setminus B = C \Rightarrow A = B \cup C$

5. Diga si son ciertas las igualdades siguiente:

- a) $A \setminus (B \cup C) = (A \setminus B) \setminus C$
- b) $A \cup (B \setminus C) = (A \cup B) \setminus C$
- c) $(A \setminus B) \cup C = (A \cup C) \setminus B$

Si alguna de ellas no se cumple diga entonces en que sentido se satisface la inclusión.

6. Demuestre que la igualdad $A \setminus (B \setminus C) = (A \setminus B) \cup C$ es cierta si $A \supset C$ y no se cumple si $C \setminus A \neq \emptyset$

7. Demuestre que para conjuntos E, F y G cualesquiera se cumplen las igualdades siguientes:

- a) $E \times (F \cup G) = (E \times F) \cup (E \times G)$
- b) $(F \cup G) \times E = (F \times E) \cup (G \times E)$
- c) $E \times (F \cap G) = (E \times F) \cap (E \times G)$
- d) $(F \cap G) \times E = (F \times E) \cap (G \times E)$

8. Diga si son ciertas las igualdades siguientes:

- a) $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$
- b) $(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)$

9. Demuestre que: $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$

10. Demuestre que: $(P \times Q) \setminus (A \times B) = ((P \setminus A) \times Q) \cup P \times (Q \setminus B)$

11. Establezca una correspondencia biunívoca entre el conjunto $\mathbb{N}$ de los números naturales y el conjunto S de los números enteros pares positivos.

12. Halle explícitamente una correspondencia biunívoca entre los intervalos $[0, 1]$ y $[a, b]$.

13. Halle explícitamente una correspondencia biunívoca entre los intervalos $[0, 1[$ y $[0, \infty[$.

14. Halle explícitamente una correspondencia biunívoca entre los intervalos $[0, 1]$ y $[0, \infty[$.

15. Construya una correspondencia biunívoca entre una circunferencia de radio 1 y el segmento $[0, 1]$.

16. Halle explícitamente una correspondencia biunívoca entre un círculo de radio 1 y el complemento de dicho círculo abierto, en el plano complejo.

17. Diga cuál es la potencia del conjunto formado por todos los triángulos del plano cuyos vértices tienen coordenadas racionales.

18. Diga cuál es la potencia del conjunto formado por todas las fracciones racionales con coeficientes enteros en el numerador y en el denominador.

19. Demuestre que el conjunto de todas las circunferencias del plano cuyos radios son racionales al igual que las coordenadas de sus centros, es numerable.

20. Demuestre que si la distancia entre dos puntos cualesquiera de un conjunto E sobre la recta real $\mathbb{R}$ es mayor que un número real fijo $r > 0$, entonces el conjunto E es finito o numerable.
21. Demuestre que el conjunto formado por todos los subconjuntos finitos de un conjunto numerable es también numerable.
22. Demuestre, utilizando el teorema de reordenamiento de Riemann para series (ver capítulo 6), que el conjunto de todas las sucesiones que se obtienen permutando el orden de los elementos de $\mathbb{N}$, tiene la potencia del continuo.
23. Sean $f : X \longrightarrow Y$; $A, A_1, A_2 \subset X$; $B \subset Y$. Pruebe que:

a) $f[A_1 \cap A_2] \subset f[A_1] \cap f[A_2]$

b) $f[A_1 \setminus A_2] \supset f[A_1] \setminus f[A_2]$

c) $A \subset f^{-1}[f[A]]$

d) $f[f^{-1}[B]] \subset B$

e) Dé ejemplos donde la inclusión no pueda ser sustituida por la igualdad en los incisos anteriores.

f) Demuestre que la igualdad en d) ocurre si y solo si $B \subset f[X]$.

g) $f[A] \cap B = f[A \cap f^{-1}[B]]$

h) Si $f^{-1}[f[A_1]] = A_1$, entonces

$$f[A_1 \cap A_2] = f[A_1] \cap f[A_2]$$

i) Si $g = f/A$, entonces $g^{-1}[B] = A \cap f^{-1}[B]$

j) Si $X = \bigcup_{\alpha \in I} X_\alpha$ y $f_\alpha = f|_{X_\alpha}$ entonces $f^{-1}[B] \bigcup_{\alpha \in I} f^{-1}[B]$

CAPÍTULO 2

1. Diga si la aplicación $d(x, y) = \sqrt{|x - y|}$ define una métrica sobre el conjunto $\mathbb{R}$ de los número reales
2. Sea C el conjunto de puntos de una circunferencia del plano $\mathbb{R}^2$. Definamos la distancia entre dos puntos $x, y \in C$ como la longitud del arco de C más corto que une a ambos puntos. ¿Satisface o no esta distancia los axiomas de métrica?

3. Sea X el conjunto de todas las rectas del plano $\mathbb{R}^2$ que no pasan por el origen de coordenadas. Definamos la distancia entre dos rectas:

$$l_1 : x \cos \alpha_1 + y \operatorname{sen} \alpha_1 - p_1 = 0$$

$$l_2 : x \cos \alpha_2 + y \operatorname{sen} \alpha_2 - p_2 = 0$$

(donde $0 \leq \alpha_1 < 2\pi$; $0 \leq \alpha_2 < 2\pi$; $p_1 > 0$; $p_2 > 0$), mediante las fórmulas:

$$\text{a) } \rho_1(l_1, l_2) = \sqrt{(p_1 - p_2)^2 + (\operatorname{sen} \alpha_2 - \operatorname{sen} \alpha_1)^2 + (\cos \alpha_2 - \cos \alpha_1)^2}$$

$$\text{b) } \rho_2(l_1, l_2) = |p_2 - p_1| + |\operatorname{sen} \alpha_2 - \operatorname{sen} \alpha_1| + |\cos \alpha_2 - \cos \alpha_1|$$

$$\text{c) } \rho_3(l_1, l_2) = |p_2 - p_1| + |\operatorname{sen} \alpha_2 - \operatorname{sen} \alpha_1|$$

Interprete geoméricamente las expresiones de ρ_1 , ρ_2 y ρ_3 y diga si son métricas sobre X .

4. Diga si la aplicación $\rho(f, g) = \sup_{x \in [a, b]} |f'(x) - g'(x)|$

es una métrica sobre el espacio de las funciones reales continuamente derivables en el intervalo $[a, b]$.

5. Una métrica d sobre un conjunto X se llama ultramétrica si verifica la propiedad:

$$x \in X \wedge y \in X \wedge z \in X \Rightarrow d(x, z) \leq \max \{d(x, y), d(y, z)\}.$$

Muestre que si d es una ultramétrica entonces:

$$\text{a) } d(x, y) \neq d(y, z) \Rightarrow d(x, z) = \max \{d(x, y), d(y, z)\}$$

b) Toda bola en (X, d) es a la vez abierta y cerrada.

c) Todo punto de una bola es centro de esta.

d) Si dos bolas son no disjuntas, una de ellas está incluida en la otra.

e) La métrica discreta sobre X es una ultramétrica.

6. Sea X el espacio de las funciones reales definidas sobre $I = [0, 1]$ y Lipschitzianas, es decir, que cumplen

$$\sup \left\{ \frac{|f(x) - f(y)|}{|x - y|} : (x, y) \in I \times I : x \neq y \right\} = K_f < \infty$$

Pruebe que:

- a) X es un subespacio vectorial de $C([0, 1], \mathbb{R})$;
- b) $N(f, g) = d_\infty(f, g) + K_{f-g}$ define una métrica sobre X .
- c) Las métricas d_∞ y N , no son topológicamente equivalentes.

(Indicación: Construya una sucesión de funciones f_n tal que K_{f_n} sea constante para todo n y $(f_n) \xrightarrow{d_\infty} \langle 0 \rangle$).

7. Demuestre que el subespacio de todas las funciones del tipo $f(x) + g(y)$ donde $f, g \in C([0, 1], \mathbb{R})$, es cerrado en $(C([0, 1] \times [0, 1], \mathbb{R}), d_\infty)$.

8. Separe el alfabeto en clases homeomorfas:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

9. Sean (X, d) un espacio métrico, $A \subset X$, $\chi_A : X \rightarrow \mathbb{R}$ la función característica de A . Pruebe:

χ_A es continua en $x \in X \Leftrightarrow$ existe $V \in \mathcal{V}_d(x)$ tal que $V \subset A$ y $V \subset X \setminus A$.

10. Sea $C = \{1/m + 1/n : m, n \in \mathbb{N}^*\}$.

Halle el conjunto C' de los puntos de acumulación de C y el conjunto C'' de los puntos de acumulación de C' en $\mathbb{R}$.

11. Sea (X, d) un espacio métrico y $U \in \mathcal{O}_d$. Pruebe que $U \cap \overline{B} \subset \overline{U \cap B}$ para todo $B \subset X$.

12. Sean (X, d) y (Y, ρ) espacios métricos. Definamos en el espacio $C(X, Y)$ de las funciones continuas de X en Y , la pseudométrica

$$d_\infty(f, g) := \sup \{ \rho(f(x), g(x)) : x \in X \}.$$

Pruebe que:

- a) El subconjunto $C_\alpha(X, Y)$ de todas las funciones continuas y acotadas de X en Y es a la vez abierto y cerrado en $C(X, Y)$.
- b) d_∞ induce una métrica sobre $C_\alpha(X, Y)$.

13. Dé un ejemplo en $\mathbb{R}$ con la métrica usual de dos abiertos A y B tales que los conjuntos:

$$A \cap \overline{B}, B \cap \overline{A}, \overline{A \cap B} \text{ y } \overline{A} \cap \overline{B}$$

sean todos distintos.

14. Sean (X, d) un espacio métrico, y A y B subconjuntos de X . Pruebe que:

a) $\partial \overline{A} \subset \partial \overset{\circ}{A}$; $\partial \alpha \subset \partial A$.

Dé un ejemplo en $\mathbb{R}$ con la métrica usual donde estos tres conjuntos sean diferentes.

b) $\partial (A \cup B) \subset \partial A \cup \partial B$

Dé un ejemplo en $\mathbb{R}$ con la métrica usual donde estos tres conjuntos sean diferentes.

c) $\overline{A} \cap \overline{B} = \emptyset \Rightarrow \partial (A \cup B) = \partial A \cup \partial B$

15. Pruebe que en el espacio métrico (X, d) son equivalentes las proposiciones siguientes:

a) X es separable.

b) Todo $A \subset X$ que contiene solamente puntos aislados es numerable.

c) Toda colección de abiertos dos a dos disjuntos es numerable.

16. Sean (X, d_1) y (Y, d_2) , (Z, d_3) espacios métricos, $f : X \longrightarrow Y$ y $g : Y \longrightarrow Z$ aplicaciones. Pruebe que:

a) Si $g \circ f$ es abierta (cerrada) y f es una biyección continua, entonces g es abierta (cerrada).

b) Si $g \circ f$ es abierta (cerrada) y g es una inyección continua, entonces f es abierta (cerrada).

17. Pruebe que en $\mathbb{R}$ con la métrica usual toda biyección continua es abierta.

18. Sea $f : \mathbb{R} \longrightarrow \mathbb{C}$ con sus métricas usuales. Supongamos

$$\varphi(x, \delta) := \sup \{|f(s) - f(t)| : s, t \in]x - \delta, x + \delta[\}$$

$$\varphi(x) := \inf \{\varphi(x, \delta) : \delta > 0\}.$$

Pruebe que:

a) φ es superiormente continua.

b) f es continua en $x \Leftrightarrow \varphi(x) = 0$.

c) El conjunto de discontinuidad de una función de $\mathbb{R}$ en $\mathbb{C}$ es un G_δ .

- 19 Supongamos que nuestro planeta Tierra es una esfera idealmente suave, y consideremos el conjunto E de todos los puntos M de la superficie de la Tierra que satisface la propiedad siguiente:

Si de M caminamos 7 km al norte, después 7 km al este, espúes 7 km al sur, y finalmente regresamos de nuevo al punto M .

- a) ¿Es E un conjunto cerrado? Si no es cerrado, ¿Cuál es la clausura?
 b) ¿Cuál es el conjunto de los puntos de acumulación de E ?

20. Sea f_0 una función real continua fija definida sobre el segmento $[0, 1]$. Demuestre que el subconjunto E de $X = (C([0, 1], \mathbb{R}), d_\infty)$ constituido por todas las funciones f que satisfacen la desigualdad $f(x) \leq f_0(x)$ para todo $x \in [0, 1]$ es cerrado en X .

21. Sea A un subconjunto arbitrario de (X, d) .

Diga si es cierto que para cualquier punto $x_0 \in X$ se cumple siempre

$$d(x_0, A) = d(x_0, \overline{A})$$

$$\text{y } d(x_0, A) = d(x_0, \overset{\circ}{A}).$$

22. Demuestre que para dos subconjuntos arbitrarios A y B de (X, d) se cumple

$$d(A, B) = d(A, \overline{B}) = d(\overline{A}, B) = d(\overline{A}, \overline{B})$$

¿Se cumple también que $d(A, B) = d(\overset{\circ}{A}, \overset{\circ}{B})$?

23. Halle la clausura, en $\mathbb{R}$ con la métrica usual, de los subconjuntos siguientes:

a) $\left\{ x = \frac{p^2}{q^2} : p, q \in \mathbb{Z} \wedge q \neq 0 \right\}$

b) $\left\{ x = 2^{p/q} : p, q \in \mathbb{Z} \wedge q \neq 0 \right\}$

c) $\left\{ x = \frac{p^2}{4p^2 + q^2} : p, q \in \mathbb{Z} \wedge p \neq 0 \wedge q \neq 0 \right\}$

- d) Demuestre que el conjunto $\{x = m + n\epsilon : m, n \in \mathbb{Z} \wedge \epsilon \text{ irracional}\}$ es denso en $\mathbb{R}$.

¿Se obtiene el mismo resultado si se sustituye $\mathbb{Z}$ por los enteros pares?

24. Muestre que un espacio métrico producto $\prod_{n \in \mathbb{N}} X_n$ es discreto si y solo si cada espacio coordenado es discreto.
25. Pruebe que la unión de cualquier familia localmente finita de cerrados en un espacio métrico es cerrada.
26. Demuestre que el conjunto de puntos de la forma $l_n(r^2 + 1)$ donde r recorre el conjunto $\mathbb{Q}$ de número racionales es denso en $[0, \infty[$.
27. Sea $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ acotada y continua en todo punto excepto en $x = 0$. Sean $\sum_n a_n$ una serie numérica convergente de términos positivos y $\{x_1, x_2, \dots, x_n, \dots\}$ un conjunto denso numerable en $\mathbb{R}$. Halle el conjunto de puntos de continuidad y el conjunto de puntos de discontinuidad de la función
- $$f(x) = \sum_{k=1}^{\infty} a_n \varphi(x - x_k).$$
28. Demuestre que una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ no puede ser continua sobre un conjunto denso numerable y discontinua en el resto de los puntos.
29. Construya una función $f : \mathbb{R} \rightsquigarrow \mathbb{R}$ discontinua en todo punto, excepto en $x = 1$ y $x = -1$.
30. Construya una función $f : \mathbb{R} \rightsquigarrow \mathbb{R}$ discontinua en todo punto de $\mathbb{R}/\mathbb{Z}$ y continua en $\mathbb{Z}$.
31. Diga si existe una función $f : [0, 1] \rightsquigarrow \mathbb{R}$ continua en todos los puntos racionales de $[0, 1]$ y discontinua en los irracionales.
32. Investigue la continuidad de la función $f : \mathbb{R} \rightsquigarrow \mathbb{R}$, igual a x^2 en los puntos racionales e igual a $-x^2$ en los irracionales.
33. Dé un ejemplo de una función $f : [0, 1] \rightsquigarrow \mathbb{R}$ que sea discontinua en todo $[0, 1]$, pero que $|f|$ sea continua.

CAPÍTULO 3

1. Demuestre la no completitud de los siguientes espacios métricos y halle su completamiento:

a) la recta $\mathbb{R}$ con la distancia , $d(x, y) := |\arctan x - \arctan y|$,

b) la recta $\mathbb{R}$ con la distancia, $d(x, y) := |e^x - e^y|$.

2. En el conjunto $\{\Delta\}$ de todos los segmentos de una recta, definamos una distancia como la longitud de la diferencia simétrica

$$d(\Delta_1, \Delta_2) := |\Delta_1| |\Delta_2| - 2 |\Delta_1 \cap \Delta_2|.$$

Demuestre que d es una métrica y que el espacio métrico obtenido no es completo. Halle el completamiento.

3. Demuestre la no completitud del espacio de los polinomios con respecto a las distancia siguientes:

a) $d(P, Q) = \max_{x \in [0, 1]} |P(x) - Q(x)|$

b) $d(P, Q) = \int_0^1 |P(x) - Q(x)| dx$

c) $d(P, Q) = \sum_i |c_i|$, donde $P(x) - Q(x) = \sum_i c_i x^i$

4. Sea (X, d) un espacio métrico completo acotado y G el conjunto de todos los homeomorfismos de X en si. Definamos en G una métrica mediante la fórmula

$$D(f, g) = \sup \{d(f(x), g(x)) + d(f^{-1}(x), g^{-1}(x))\}$$

Demuestre la completitud del espacio métrico (G, D) .

5. Demuestre que el espacio métrico (E, N) definido en el ejercicio 2.6 del capítulo 2, es completo.

6. a) Sea (X, d) un espacio ultramétrico (vea ejercicio 5 del capítulo 2). Pruebe que para que una sucesión (x_n) sea de Cauchy en (X, d) es necesario y suficiente que $\lim_{n \rightarrow \infty} d(x_n, x_{n+1}) = 0$.

- b) Sea A un conjunto arbitrario y X el conjunto de todas las sucesiones infinitas $x = (x_n)$ de elementos de A . Para cada par de elementos distintos $x = (x_n)$, y $y = (y_n)$ de X sea $k(x, y)$ el menor entero n tal que $x_n \neq y_n$.

Consideremos

$$d(x, y) := \frac{1}{k(x, y)} \text{ si } x \neq y.$$

Demuestre que d es una métrica ultramétrica sobre X y que el espacio métrico (X, d) es completo.

7. Sea $f(x)$ una función numérica definida sobre $\mathbb{R}$ y que satisface $|f'(x)| \leq K$, donde K es una constante; $K > 1$.

Demuestre que la ecuación $f(x) = x$ tiene solución única.

8. Consideremos el sistema infinito de ecuaciones

$$x_i = \sum_{k \sim 1}^{\infty} c_{ik} x_k + b_i; i = 1, 2, \dots,$$

donde $\sum_{i,k=1} c_{ik}^2 < 1$; $\sum_{i=1} b_i^2 < \infty$.

Demuestre que este sistema tiene una y solo una solución $(x_1, x_2, \dots)$ en el espacio $l^2(\mathbb{N})$.

9. Consideremos la función $f(x) = \frac{1}{2} l_n x$ definida sobre $[1, \infty[$. Para cualesquiera $x, y \in [1, \infty[$ tenemos:

$$|f(x) - f(y)| = |f'(c) - (x - y)| \leq \frac{1}{2} |x - y|, \text{ donde } c \in]x, y[.$$

Luego, la aplicación f es constante; sin embargo, esta no posee puntos fijos (en efecto, la ecuación $x = \frac{1}{2} l_n x$ no posee soluciones reales: ¿por qué?)

¿Hay alguna contradicción con el teorema del punto fijo de Banach?

10. La función $f(x) = \frac{x^2}{2|x|}$ no tiene puntos fijos, aunque para cualesquiera

$x, y \in \mathbb{R}^*$ se cumple la desigualdad $|f(x) - f(y)| \leq \frac{1}{2} |x - y|$. ¿Hay alguna contradicción con el teorema del punto fijo de Banach?

- 11.* El siguiente ejercicio expone otro método de completamiento de un espacio métrico:

a) Sea (X, d) un espacio métrico. Dos sucesiones de Cauchy (x_n) y (y_n) en (X, d) se dice que son equivalentes si $\lim d(x_n, y_n) = 0$. Demuestre que esta relación es una relación de equivalencia.

b) Pruebe que el conjunto cociente $\overline{X}$ asociado a dicha relación y provisto de la métrica

$$d((x_n), (y_n)) = \lim d(x_n, y_n)$$

es un espacio métrico completo.

c) Pruebe que la aplicación

$$i : X \xrightarrow{x \rightsquigarrow (x_n)} i(X) \subset X,$$

donde $x_n = (n \in \mathbb{N})$, es una isometría y que el par $(i, \overline{i(x)})$ es, por tanto, un completamiento de (X, d) .

12. La función $f(x, y) = \frac{2xy}{x^2 + y^2}$ es continua sobre el conjunto abierto $0 < x^2 + y^2 < 4$. ¿Es o no esta función uniformemente continua sobre este conjunto? ¿Es esta uniformemente continua en el anillo abierto $1 < x^2 + y^2 < 4$?

CAPÍTULO 4

1. Demuestre que un espacio métrico (X, d) es compacto si y solo si (X, φ) es compacto para cualquier métrica φ topológicamente equivalente a d .
2. Sea (K_n) una sucesión de compactos tal que la intersección de cualquier subfamilia finita es no vacía. Demuestre que la intersección $\bigcap_{n \in \mathbb{N}} K_n$ es también no vacía.
3. Dado el conjunto compacto numerable:

$K = \left\{ 0, 1, \frac{1}{2}, \frac{1}{4}, \dots, \frac{1}{2^n}, \dots \right\}$, consideremos un cubrimiento de K mediante la familia de intervalos abiertos:

$]1 - \epsilon, 1 + \epsilon[, \left] \frac{1 - \epsilon}{2}, \frac{1 + \epsilon}{2} \right[, \left] \frac{1 - \epsilon}{4}, \frac{1 + \epsilon}{4} \right[, \dots, \left] \frac{1 - \epsilon}{2^n}, \frac{1 + \epsilon}{2^n} \right[, \dots$, y $] -\epsilon, \epsilon[$, donde ϵ es un número positivo fijo menor que $1/2$. Extraiga de esta familia de intervalos una subfamilia finita que siga siendo un cubrimiento de K .

4. Dado el conjunto numerable

$$E = \left\{ 1, \frac{1}{2}, \frac{1}{4}, \dots, \frac{1}{2^2}, \dots \right\},$$

consideremos un cubrimiento de E mediante la familia de intervalos abiertos:

$$]1 - \epsilon, 1 + \epsilon[, \left] \frac{1 - \epsilon}{2}, \frac{1 + \epsilon}{2} \right[, \dots, \left] \frac{1 - \epsilon}{2^n}, \frac{1 + \epsilon}{2^n} \right[, \dots$$

donde $0 < \epsilon < 1/3$. ¿Es posible extraer de este cubrimiento un subcubrimiento finito?

5. Consideremos el disco abierto D de centro O y radio 1 en el plano $\mathbb{R}^2$. Dibujemos la circunferencia concéntrica C de radio $1/3$ construyamos la familia de todos los posibles discos abiertos de radio $2/3$ cuyo centro se encuentra sobre la circunferencia C . Estos discos abiertos constituyen un cubrimiento infinito del conjunto D .

a) Demuestre que de este cubrimiento no se puede extraer un subcubrimiento finito.

b) ¿Puede extraerse algún subcubrimiento numerable?

6. Diga si es cierto o no el enunciado siguiente y explique la respuesta:

De cualquier cubrimiento de un compacto por clausura de conjuntos abiertos, se puede extraer un subcubrimiento finito.

7. Sabemos que en general las proyecciones sobre un espacio producto son abiertas pero no cerradas. Veremos ahora que las proyecciones paralelas a espacios coordinados compactos son cerradas. Pruebe que si (X, d) es compacto y (Y, d) métrico, entonces la proyección $(x, y) \rightsquigarrow y$ de $X \times Y$ sobre Y es cerrada.

8. Sea f una aplicación del espacio métrico (X, d) en (Y, φ) .

Pruebe que:

a) Si f es continua entonces el grafo de f es cerrado en $X \times Y$ pero el inverso no se cumple necesariamente (dé un contraejemplo para el caso $X = Y = \mathbb{R}$).

b) Si (Y, φ) es compacto, entonces $f : X \longrightarrow Y$ es continua si y solo si su grafo es cerrado en $X \times Y$.

(Indicación: utilice el ejercicio 7 del capítulo 4.)

9. Sean (X, d) un espacio métrico (Y, φ) compacto tal que $f : X \times Y \rightsquigarrow \mathbb{R}$ es continua. Pruebe que las aplicaciones

$$M(\cdot) = \sup(f(\cdot, y) : y \in Y)$$

$$m(\cdot) = \inf(f(\cdot, y) : y \in Y)$$

son continuas.

(Indicación: Para todo número $r \in \mathbb{R}$ defina $S_r = \{(x, y) : f(x, y) \geq r\}$ y utilice el ejercicio 7 del capítulo 4.)

10. Sean (X, d) un espacio métrico y $(E, \|\cdot\|)$ normado sobre $\mathbb{R}$ o $\mathbb{C}$. Sea $H \in F(X, E)$ equicontinuo en $x_0 \in X$ y $k > 0$.

Demuestre que el conjunto

$$H_k = \left\{ \sum_i c_i f_i : f_i \in H, \sum_i |c_i| \leq k \right\}$$

es equicontinuo.

11. Sea $I = [-1, 1]$ y sea para todo $n \in \mathbb{N}^*$

$$f_n : \mathbb{R}_+ \longrightarrow I$$

definido por:

$$f_n(x) = \sin \sqrt{(x + 4n^2\pi^2)}.$$

Pruebe que $\{f_n : n \in \mathbb{N}^*\}$ es una sucesión equicontinua de funciones que no es relativamente compacta en $(C(\mathbb{R}_+, I), d_\infty)$.

12. Sean $I = [0, 1]$ y $f : I \longrightarrow \mathbb{R}$ continua no idénticamente nula y tal que $f(0) = f(1) = 0$. Pruebe que para $g_n : x \longrightarrow f(x^n)$, $(x \in I)$ se cumple:

- a) $(g_n) \longrightarrow \langle 0 \rangle$ puntualmente,
- b) (g_n) no converge a $\langle 0 \rangle$ uniformemente.

13. Sea I un intervalo compacto en $\mathbb{R}$ y (f_n) una sucesión creciente de funciones $f_n \in F(I, \mathbb{R})$ que converge puntualmente sobre I a una función continua g .

Pruebe que g es monótona y que (f_n) converge uniformemente a g .

14. Sean (X, d) , (Y, φ) espacios compactos; entonces toda función $f : X \longrightarrow Y$ induce una aplicación

$$f^* = C(Y, \mathbb{R}) \longrightarrow C(X, \mathbb{R})$$

Pruebe que:

- a) f^* es un homeomorfismo (algebraico), del álgebra $C(Y, \mathbb{R})$ en $C(X, \mathbb{R})$.
- b) f es sobreyectiva si y solo si f^* es un isomorfismo de $C(Y, \mathbb{R})$ sobre una subálgebra de $C(X, \mathbb{R})$ que contiene a $\langle 1 \rangle$.
- c) f es inyectiva si y solo si f^* es sobreyectiva.

d) Si f es un homeomorfismo (topológico) entonces $(f^{-1}) = (f^*)^{-1}$.

CAPÍTULO 5

1. Sea E un conjunto conexo que contiene más de un punto. Demuestre que E no posee puntos aislados.
2. Demuestre que para que dos conjuntos no vacíos A y B sean separados, es necesario y suficiente que se cumpla la igualdad

$$(\overline{A} \cap B) \cup (A \cap \overline{B}) = \emptyset.$$
3. Demuestre que si A y B son ambos cerrados y ambos abiertos no vacíos, y además, $A \cap B = \emptyset$ entonces son separados.
4. Sea E un conjunto no conexo diferente del vacío. Pruebe que:
 - a) Si E es cerrado, entonces puede representarse como la unión de dos conjuntos cerrados disjuntos no vacíos.
 - b) Si E es abierto, entonces puede representarse como la unión de dos conjuntos abiertos disjuntos no vacíos.
5. Sean A y B subconjuntos del espacio métrico (X, d) .
 - a) Demuestre que si A y B son cerrados tales que $A \cup B$ y $A \cap B$ son conexos entonces A y B son conexos.
 - b) Muestre mediante un ejemplo que se obtiene el mismo resultado en a) si uno de los dos conjuntos no es cerrado.
 - c) Supongamos que A y B son conexos y que $\overline{A} \cap B \neq \emptyset$. Demuestre entonces que $A \cup B$ es conexo.
6. Demuestre que el conjunto de los puntos del plano con ambas coordenadas irracionales no es conexo.
7. Demuestre que un disco cerrado del plano $\mathbb{R}^2$ y una bola cerrada en $\mathbb{R}^3$ son conexos.
8. Sean $(X, d), (Y, \varphi)$ espacios métricos conexos, y A y B subconjuntos propios de X y Y respectivamente.
 Muestre que en el espacio producto $X \times Y$ el complemento de $A \times B$ es un conjunto conexo.

9. Pruebe que toda aplicación continua f de un espacio conexo $\mathbb{C}$ tal que e^f sea constante, es constante.
10. Sean (X, d) localmente conexo y $A \subset X$. Si δA es localmente conexo, pruebe que A también es localmente conexo.
11. Sean (X, d) localmente conexo, $X = A \cup B$, donde A y B son cerrados y $A \cap B$ es localmente conexo. Pruebe que entonces ambos conjuntos A y B son localmente conexos.
12. En el plano $\mathbb{R}^2$ sea el subconjunto

$$E = \left\{ (x, y) : (x \in \prod \wedge y \in [0, 1]) \vee (x \in Q \wedge y \in [-1, 0]) \right\}.$$

- a) Pruebe que E es conexo pero no localmente conexo.
- b) Si $t \rightsquigarrow (f(t), g(t))$ es una aplicación continua de $[0, 1]$ en E , pruebe que f es constante.

(Indicación: Si existen puntos $t_0 \in [0, 1]$ tales que $g(t_0) < 0$, considere el conjunto abierto $U \subset [0, 1]$ de todos los t tales que $g(t) < 0$ y aplique el hecho que todo abierto de $\mathbb{R}$ se puede plantear como una unión a lo más numerable de intervalos abiertos disjuntos.)

13. Sean (X, d_1) y (Y, d_2) conexos y f una aplicación continua de $X \times Y$ en (Z, d_3) tal que cada una de las aplicaciones parciales $f_y : X \rightarrow Z$ y $f_x : Y \rightarrow Z$ son continuas.

Demuestre que $f[X \times Y]$ es conexo.

14. Sea (X, d) compacto y conexo y sea $C \subset X$ cerrado. Pruebe que existe un cerrado conexo mínimo B tal que $C \subset B \subset X$.

CAPÍTULO 6

1. Demuestre que si A es cerrado y B compacto en el espacio normado $(E, \|\cdot\|)$, entonces $A + B$ es cerrado. Dé un ejemplo en que A y B son ambos cerrados pero $A + B$ no lo es.
2. En el espacio $C(\mathbb{R}, \mathbb{R})$ introduzca una familia numerable de seminormas mediante

$$p_N(f) = \max_{|x| \leq N} |f(x)|$$

y supongamos

$$d(f, g) = \sum_{N=1}^{\infty} \frac{1}{2^N} \frac{p_N(f - g)}{1 + p_N(f - g)}.$$

Diga para cuáles valores de $r > 0$, la bola $B_d(f, r)$ es un conjunto convexo en $C(\mathbb{R}, \mathbb{R})$.

3. Definamos una distancia en $C(\mathbb{R}, \mathbb{R})$ mediante la fórmula

$$d(f, g) = \sup_{x \in \mathbb{R}} \frac{|f(x) - g(x)|}{1 + |f(x) - g(x)|}.$$

Diga si esta distancia es compatible con la estructura vectorial de $C(\mathbb{R}, \mathbb{R})$.

4. Sea $\mathbb{R}^{\mathbb{N}}$ el espacio de todas las sucesiones de números reales, provisto de la métrica:

$$d(\{x_n\}, \{y_n\}) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|x_n - y_n|}{1 + |x_n - y_n|}.$$

Pruebe que:

- La métrica d es compatible con la estructura vectorial de $\mathbb{R}^{\mathbb{N}}$.
- Cada sucesión $x = \{x_n\}$ en $\mathbb{R}^{\mathbb{N}}$ tiene un sistema fundamental de vecindades conexas.
- No existe ninguna norma sobre $\mathbb{R}^{\mathbb{N}}$ topológicamente equivalente a la métrica d .

5. Demuestre que $\|x\|_p = \left(\sum_{k=1}^n (|x_k|^p) \right)^{1/p}$ no es una norma sobre $\mathbb{R}^n$ para $p < 1$ y $n \geq 2$.

6. Sea $(E, \|\cdot\|)$ un espacio de Banach y M un subespacio cerrado de E .

Demuestre que la fórmula

$$\|x\|_1 = \inf_{y \rightsquigarrow x \in M} \|y\|$$

define una norma sobre el espacio cociente E/M , y que además, el espacio normado $(E/M, \|\cdot\|_1)$ es de Banach.

7. Muestre que un espacio normado E es no separable si y solo si, cuando en él, hay un conjunto no numerable de bolas de radio 1 disjuntas dos a dos.

8. Se E un espacio de Banach y M un subespacio cerrado de E . Un vector $x \in E$ se llama ϵ -perpendicular a M , si para todo $y \in M$ se cumple la desigualdad $\|x + y\| \geq (1 - \epsilon) \|x\|$.

- Demuestre que para $\epsilon > 0$ cualquier subespacio propio M posee elementos 0-perpendiculares.
- Demuestre que la existencia de elementos 0-perpendiculares al subespacio M , determinado por la ecuación $f(x) = 0$ ($f \in E^*$), es equivalente a la existencia de $\max_{\|x\|} |f_1(x)|$.

9. Consideremos el espacio normado $E = ((C[0, 1], \mathbb{R}), \|\cdot\|_\infty)$ y sea M el subconjunto de E constituido por aquellas funciones f tales que

$$\int_0^{1/2} f(t) dt - \int_{1/2}^1 f(t) dt = 1.$$

Pruebe que M es cerrado conexo en C pero no contiene elementos de norma mínima.

10. Sea E un espacio normado y E^* su dual topológico:

- Pruebe que para cada $x \in E$ fijo, la aplicación

$$E^* \xrightarrow{f \mapsto f(x)} \mathbb{R}$$

define un funcional lineal y continuo sobre E^* cuya norma es igual a $\|x\|$. (Esto nos proporciona una inmersión natural de E en su *segunda dual* E^{**} , que a su vez el espacio dual E^* .)

- Pruebe que $\|x_n\|$ es acotada si x_n es una sucesión en E tal que $f(x_n)$ es acotada para todo $f \in E^*$.

11. Si la serie $\sum_{i=1}^{\infty} a_i x_i$ converge para toda sucesión (x_i) tal que $\sum_{i=1}^{\infty} x_i^2 < \infty$, pruebe entonces $\sum_{i=1}^{\infty} a_i^2 < \infty$.

12. Sea $\{e_n\}_{n=1}^{\infty}$ una familia ortonormal en el espacio de Hilbert H .

- Pruebe que esta familia es un ejemplo de un conjunto cerrado y acotado en H que no es compacto.

b) Sea Q el subconjunto de H de todos los elementos de la forma

$$x = \sum_{n=1}^{\infty} c_n e_n, \text{ donde } |c_n| \leq 1/n.$$

Pruebe que Q es compacto en H .

c) Más generalmente, sea $\{\delta_n\}$ una sucesión de números positivos, y sea S el conjunto de todos los $x \in H$ de la forma

$$x = \sum_{n=1}^{\infty} c_n e_n, \text{ donde } |c_n| \leq \delta_n.$$

Pruebe que S es compacto en H si y solo si $\sum_{n=1}^{\infty} \delta_n^2 < \infty$.

13. Calcule:

$$\min_{a,b,c} \int_{-1}^1 |x^2 - a - bx - cx^2|^2 dx$$

y halle

$$\max_g \int_{-1}^1 x^2 g(x) dx$$

donde $g \in L_2[-1, 1]$ está sujeta a las restricciones

$$\int_{-1}^1 g(x) dx = \int_{-1}^1 xg(x) dx = \int_{-1}^1 x^2 g(x) dx = 0;$$

$$\int_{-1}^1 |g(x)|^2 dx = 1.$$

14. Calcule:

$$\min_{a,b,c} \int_0^{\infty} |x^2 - a - bx - cx^2|^2 e^{-x} dx$$

15. Sea H un espacio de Hilbert. Si $x_0 \in H$ y M es subespacio vectorial cerrado de H , pruebe que

$$\min \{|x - x_0| : x \in M\} = \max \{|\langle x_0, y \rangle| : y \in M^{\perp}, \|y\| = 1\}.$$

16. Definamos el funcional lineal f sobre $L_2[0, 1]$ mediante

$$f(x) = \int_0^1 a(t) \int_0^1 b(s) x(s) ds dt$$

donde a y b son funciones reales continuas sobre $[0, 1]$. Pruebe que f es un funcional lineal continuo y halle un elemento $y(t) \in L_2[0, 1]$ tal que $f(x) = \langle x, y \rangle$.

17. Sea E^* el dual del espacio normado E . Pruebe que si E^* es separable entonces también lo es E .
18. ¿Qué se puede decir de los coeficientes de Fourier de $f \in C([-\pi, \pi], \mathbb{R})$ con respecto al sistema trigonométrico ortonormalizado, en cada uno de los casos siguientes:?
- f es par,
 - f es impar.
19. Exprese en términos de los coeficientes de Fourier de $f \in C([-\pi, \pi], \mathbb{R})$ con respecto al sistema trigonométrico ortonormalizado, las siguientes propiedades de la función f :
- $f(t + \pi/2) = f(t)$,
 - $f(t + \pi/k) = \lambda f(t)$. ¿Para cuales valores de $k \in \mathbb{Z}$ existen funciones no nulas que poseen esta propiedad?
20. Definamos el operador $P_k : l^2(\mathbb{N}) \rightsquigarrow l^2(\mathbb{N})$ mediante la fórmula:

$$P_k(\{x_n\}) = \{\epsilon_{nk}x_n\},$$

donde

$$\epsilon_{nk} = \begin{cases} 1, & \text{si } k \leq n, \\ 0, & \text{si } k > n. \end{cases}$$

¿Cuál de las siguientes convergencias tiene lugar cuando $k \rightarrow \infty$?

- $\|P_k\| \rightarrow 0$;
 - $P_k(\{x_n\}) \rightarrow 0$ en $l^2(\mathbb{N})$ para toda sucesión $\{x_n\} \in l^2(\mathbb{N})$.
 - $\langle P_k(\{x_n\}), \{y\} \rangle \rightarrow 0$ para toda sucesión $\{y_n\} \in l^2(\mathbb{N})$, donde $\langle, \rangle$ denota el producto escalar en $l^2(\mathbb{N})$.
21. Sea $\{e_1, e_2, \dots, e_k, \dots\}$ la base canónica en $l^2(\mathbb{N})$:

$$e_k = (0, \dots, 0, 1, 0, \dots).$$

Definamos el operador A_n por la fórmula

$$A_n e_k = \begin{cases} e_1, & \text{si } k = n \\ 0, & \text{si } k \neq n \end{cases}$$

Demuestre que $\|A_n\| = 1$ y que sin embargo,

$$\langle A_n(x), y \rangle \rightarrow 0$$

para todos $x, y \in l^2(\mathbb{N})$, donde $\langle, \rangle$ denota el producto escalar en $l^2(\mathbb{N})$.

22. Demuestre la continuidad de la operación de composición de operadores. Es decir, si $A_n \longrightarrow A$ en $L(F, G)$ y $B_n \longrightarrow B$ en $L(E, F)$, entonces $A_n B_n \longrightarrow AB$ en $L(E, G)$.

23. Sea A el operador de la multiplicación por la función continua $a(x)$ en el espacio $C([a, b], \mathbb{R}), \|\cdot\|$.

a) Demuestre que A es continuo si en calidad de $\|\cdot\|$ se toma $\|\cdot\|_\infty$

$$\text{ó } \|f\| = \|f\|_2 = \left(\int_a^b |f(t)|^2 dt \right)^{1/2}.$$

b) Calcule la norma de A respecto a $\|\cdot\|_\infty$ y $\|\cdot\|_2$.

24. Calcule la norma del operador identidad de $C([a, b], \mathbb{R}), \|\cdot\|_2$ en $C([a, b], \mathbb{R}), \|\cdot\|_1$, donde $\|\cdot\|_2$ viene definida en el ejercicio 23 y

$$\|f\|_1 = \int_a^b |f(t)| dt.$$

25. Sea $T(t)$ el operador de traslación $T(t)f(x) = f(x+t)$ definido en el espacio

$$E : \left\{ f : \mathbb{R} \longrightarrow \mathbb{R} : f \text{ continua} \wedge \int_{-\infty}^{\infty} |f(t)| dt < \infty \right\}$$

provisto de la norma

$$\|f\| = \int_{-\infty}^{\infty} |f(t)| dt.$$

a) Demuestre que

$$T(t)f \longrightarrow T(t_0)f \text{ para toda } f \in E, \text{ cuando } t \longrightarrow t_0.$$

b) Diga si se cumple que

$$T(t)f \longrightarrow T(t_0)f \text{ para toda } f \in E, \text{ cuando } t \longrightarrow t_0 \text{ en } L(E, E).$$

26. Sea A un operador lineal definido sobre un espacio de Hilbert H . El operador A^* se llama adjunto de A si cumple

$$\langle Ax, y \rangle = \langle x, A^*y \rangle, \text{ para todos } x, y \in H.$$

Pruebe que:

a) El adjunto de A siempre existe.

b) A^* es lineal.

c) $A^{**} := (A^*)^* = A$

d) A es continuo $\Leftrightarrow A^*$ es continuo. En este caso

$$|||A||| = |||A^*|||$$

e) Si A es continuo entonces

$$|||A^*A||| = |||AA^*||| = |||A|||^2 = |||A^*|||^2.$$

27. Sea $K(x, y)$ una función real continua definida sobre el cuadrado $[a, b] \times [a, b]$.

Definamos el operador integral A sobre $C([a, b], \mathbb{R}), \|\cdot\|_2$) mediante

$$Af(x) = \int_a^b K(x, t) f(t) dt.$$

a) Demuestre que A es acotado.

b) Calcule su norma.

c) Calcule el adjunto de A .

(Indicación: Vea ejercicio 23 del capítulo 6.)

28. Demuestre la continuidad y calcule la norma de los funcionales lineales

$$f_1(x) = ax(0) + bx(1),$$

$$f_2(x) = \sum_{i=1}^{\infty} \frac{x_n}{n}, \quad x = (x_n)$$

en los espacios $(C([0, 1], \mathbb{R}), \|\cdot\|_{\infty})$ y $l^2(\mathbb{N})$.

29. Se dice que $\lambda \in \mathbb{C}$ es un valor propio del operador lineal A definido en el espacio euclideo, H si existe $x \in H$ no nulo tal que $Ax = \lambda x$.

Al elemento x se le llama vector propio asociado al valor propio λ .

a) Demuestre que el conjunto de los vectores propios correspondientes a un mismo valor propio es un subespacio vectorial de H , el cual cerrado si A es continuo y se llama subespacio propio asociado al valor propio λ .

b) Demuestre que si A es simétrico entonces los subespacios propios asociados a valores propios diferentes, son ortogonales.

c) Calcule los vectores y valores propios del operador integral A del ejercicio 27 del capítulo 6 cuando $a = 0$, $b = 1$ y $K(x, y) = \cos 2\pi(x - y)$, y también cuando $K(x, y) = \min(x, y)$.

30. Sea A un operador positivo en el espacio euclideo H , es decir $\langle Ax, x \rangle \geq 0$ para todo $x \in H$.

a) Demuestre que para cada x fijo, la sucesión numérica

$h(k) = \ln(A^k x, x)$ es convexa, es decir,

$$h\left(\frac{k+1}{2}\right) \leq \frac{h(k) + h(l)}{2}.$$

b) Demuestre la desigualdad

$$\|A(x)\|^2 = \langle Ax, x \rangle \|A\|.$$

31. Sea P el espacio vectorial de todos los polinomios x con coeficientes reales. Sea P_+ y P_- los subconjuntos de P formados por los polinomios cuyos términos de mayor grado tienen coeficientes positivos y negativos respectivamente. Demuestre que P_+ y P_- son conjuntos convexos que no pueden ser separados por ningún hiperplano.

32. Demuestre que un subconjunto convexo B de un espacio normado E es denso en E si y solo si, cuando todo funcional lineal $f \in E^*$ que se anula sobre B es idénticamente nulo.

33. a) Demuestre que todo subconjunto cerrado convexo en un espacio vectorial normado E es la intersección de una cierta familia de semiespacios cerrados del tipo $\{x \in E : f(x) \leq c\}$, donde $f \in E^*$ y $c \in \mathbb{R}$.

b) Represente la bola unidad en $(C([0, 1], \mathbb{R}), \|\cdot\|_2)$ como la intersección de una familia numerable de semiespacios cerrados.

CAPÍTULO 7

1. Calcule los extremos locales de la función

$$\varphi(x, y) = \sin x \sin y.$$

2. Halle el supremo y el ínfimo de la función

$$\varphi(x, y) = (x + y)e^{-(x^2 + y^2)}$$

sobre el plano $\mathbb{R}^2$.

3. Pruebe la desigualdad $xe^{x(1+y^2)^2} \geq -e^{-1}$

para todo par de números reales x, y .

4. Sea $\varphi : E \longrightarrow \mathbb{R}$ donde E es un espacio métrico separable. Pruebe que el conjunto de puntos $p \in E$ donde φ tiene máximo local (estricto) es finito ó numerable.
5. Calcule las derivadas parciales, la diferencial y el módulo de la diferencial para las aplicaciones:
 - a) $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ definida por

$$\varphi(x, y) = (e^{xy}, \cos x, \log(1 + x^2 + y^2), x + \operatorname{arctg} y).$$
 - b) $\varphi : \mathbb{R} \longrightarrow \mathbb{R}^3$ definida por

$$\varphi(x) = (\cos x, \operatorname{sen} x, ax). \text{ (Esta curva se llama hélice.)}$$
6. Calcule las segundas derivadas de las aplicaciones definidas en los ejercicios 1, 2 y 5.
7. Sean f y g funcionales reales definidos sobre el subconjunto abierto U del espacio de Banach E y diferenciables según Frechet en $x_0 \in G$.

- a) Demuestre que $h(x) = f(x) \cdot g(x)$ es diferenciable según Frechet en x_0 y calcule su derivada.
- b) Si $g(x_0) \neq 0$, demuestre que $h(x) = f(x)/g(x)$ es diferenciable según Frechet en x_0 y calcule su derivada.

8. Sea $\{e_n\}$ la base canónica de $l^2(\mathbb{N})$ y consideremos la aplicación

$$f : l(\mathbb{N}) \longrightarrow l^2(\mathbb{N})$$

dada por la igualdad

$$f(\lambda e_n) = \lambda^{1+1/n} e_n \quad (n = 1, 2, \dots) \text{ y}$$

$$f(x) = 0, \text{ si } x \neq \lambda e_n.$$

Muestre que f es diferenciable en $x = \theta$ según Gateaux, pero no según Frechet.

9. Calcule $d^2 f(x)(h)$, si $f(x) = \|x\|$ para todo $x \in H$, $x \neq \theta$, donde H es un espacio de Hilbert real.

10. Sobre el espacio $(C[0, 1], \mathbb{R}, \|\cdot\|_\infty)$ definamos el funcional

$$f(x) = \int_0^1 |x(t)| dt.$$

Determine, para cuales x existe la diferencial de Gateaux $Df(x; h)$ y es lineal en h .

11. Sea $\varphi : \mathbb{R} \rightsquigarrow \mathbb{R}$ continuamente derivable y que satisface

$$|\varphi(x)| \leq K|x|$$

donde K es una constante positiva.

Halle la diferencial según Gateaux del funcional

$$f(x) = \sum_{i=1}^{\infty} \varphi(x_n),$$

donde

$$x = (x_n) \in l^2(\mathbb{N}),$$

Diga si f es también diferenciable según Frechet.

12. Aplique las ecuaciones de Euler-Lagrange para encontrar condiciones necesarias de extremo para los siguientes funcionales definidos en $((C^{(1)}([0, 1], \mathbb{R}), \|\cdot\|_{\infty}^{(1)}))$:

$$\text{a) } J(x) = \int_0^1 \left\{ x(t) / \dot{x}(t)^2 \right\} dt \text{ y}$$

$$\text{b) } J(x) = \int_0^1 \left\{ tx(t) + t^2 \dot{x}(t) + 1 \right\} dt$$

sujetos a las condiciones

$$x(0) = 0; x(1) = 1.$$

13. En algunos problemas de cálculo de variaciones se hace necesario considerar una clase de funciones más amplia que $(C^{(1)}([0, 1], \mathbb{R}))$.

Supongamos que deseamos hallar una condición necesaria de extremo para el funcional

$$J(x) = \int_a^b f(x, \dot{x}, t) dt,$$

entre todas las funciones x para las cuales $x(a) = A, x(b) = B$ y que tienen derivadas continuas sobre $[a, b]$ excepto posiblemente en un único punto de $[a, b]$. Supongamos que el extremo tiene derivada continua excepto en $c \in [a, b]$.

- a) Pruebe que la condición de Euler $f_x = \frac{\partial}{\partial t} f_{\dot{x}}$ sobre los intervalos $[a, c]$ y $[c, b]$ es una condición necesaria de extremo.

- b) Aplique estas condiciones al funcional

$$J(x) = \int_{-1}^1 x^2(1 - \dot{x}^2) dt, x(-1) = 0, x(1) = 1.$$

BIBLIOGRAFÍA

- AKHIEZER N.T., I.M. GLAZMAN: *Theory of Line Operators in Hilbert Space*. Frederick Ungar Publishing Co., New York, 1961 (traducido del ruso) (Tomos I y II).
- CHINN W.G., N.E. STEENROD: *First Concepts of Topology* New Mathematics Library, num.18, Random House, 1966.
- DIEUDINNÉ: J. *Foundations of Modern Analysis*. Academic Press, 1960.
- HINRICHSSEN D. y otros: *Topología general*, Editorial Pueblo y Educación. Ciudad de la Habana, 1977.
- KIRILOV, A.A., A.D. GVICHIANI: *Teoremas y problemas de análisis funcional*. Nauka, Moscú. 1979.
- KOLMOGOROV A.N., S.V. FOMIN : *Elementos de la teoría de funciones y del análisis funcional*, Editorial MIR, Moscú, 1972.
- LANG S.: *Differentiable Manifolds*, Addison-Wesley, 1972.
- LUENBERGUER D.G: *Optimization by Vector Space Methods*. J. Wiley, 1969
- LUISTERNK L., V. SOBOLEV: *Elements of the theory of Functional Analysis*. Frederick Ungar Publishing Co, New York, 1961 (traducido del ruso.)
- MIJAILOV V.P.: *Ecuaciones diferenciales en derivadas parciales*. 2da. ed; Editorial MIR, Moscú, 1982 (traducido al español).
- RUDIN W.: *Functional Analysis*. McGraw-Hill, 1973.
- SPIVAK M : *Cálculus on Mainfolds*. Benjamin, New York, 1965.
- TAYLOR A.E.: *Introduction to Functional Analysis*. Jhon Wiley and Sons Inc. New York, 1958.

- YOSIDA K.: *Functional Analysis*. Springer-Verlag, New York, 1968.

Índice Temático

- o, conjunto 46
- da, aplicación 76,138
- ica (seudo) 97
- do, conjunto 138
- ulación, punto de 161
- encia, conjunto 158
- ente, punto 157,160,163
- un conjunto 158
- o, punto 161
- ción (vea función) 63
- real 482
- real 484
- ciones acotadas, subespacio de 97
- nuas, subespacio de 97
- 46
- teorema de 289
- a, de elección 77,79
- erabilidad (primero) 152,267
- segundo) 152
- paración 115
- teorema de 235, 249
- cio de 239
- h, espacio de 378
- n-Steinhaus, teorema de 445
- de Möbius 113
- opológica 152
- ográfica 355
- valentes 386, 411
- (abiertas y cerradas) 134
- a de Klein 115
- a 40
- o de variaciones 494
- Canónica, descomposición 69
 - inyección 69
 - sobrección 69
- Cantor, teorema de 235
- Característica, función 69
- Cardinal, número 51
- Cartesiano, producto 45
- Cauchy, sucesión de 222
 - Integral de 193
- Cauchy-Buniakovsky,
 - desigualdad de 98,100
- Cerrado, conjunto 175
- Cerrada, aplicación 186
- Cociente, espacio 112
- Colección 59
- Combinación, lineal 354
- Compacidad, criterios de 254
- Compatible, métrica
 - con la estructura algebraica 358
 - con la estructura geométrica 413
 - con la estructura topológica 314
- Compactificación 393
- Compacto (espacio, conjunto) 254,257
 - localmente 301
 - relativamente 261
- Complemento, conjunto 45
 - ortogonal 394
- Completamiento, de un espacio métrico 240
 - de un espacio normado 378
 - de un espacio euclideo 408
- Completo, espacio métrico 227
 - subespacio 231
- Composición, de aplicaciones 68
 - de funciones continuas 68
- Conexa, componente 338
- Conexa por arcos, componente 347
- Conexo, (espacio, conjunto) 320
 - localmente 337

Conexo por arcos, espacio, conjunto 347
 Conjunto 36
 – cociente 62,112
 – producto 105
 – universal 39
 Constante, aplicación 68
 Continua en un punto, aplicación 124
 Contractante, aplicación 235
 Convergencia 90,94
 – de una red 90
 – de una sucesión 90
 – en media 90
 – puntual 90
 – uniforme 90
 Convexo, conjunto 373
 Coordenada, función 128
 Cota (superior, inferior) 73
 Cubo de Hilbert 265
 Cubrimiento 60
 – localmente infinito 204
 – puntualmente finito 204
 Curva 295, 297

D

D' Alembert, teorema de 310
 Denso, subconjunto 164
 Derivada, en sentido débil 488
 – en sentido fuerte, 470
 Derivado, conjunto 161
 Diagonal 309
 Diámetro 138
 Diferencia, conjunto 44
 Dígrafos evaluados 103
 Dirigido, conjunto 73
 Discreta, distancia 102
 Distancia 90,97
 – de costos 103
 – del bosque 107,126,133
 – discreta 119
 – entre conjuntos 140
 – inducida 108
 Distancias semejantes 116
 – topológicamente equivalentes 152

E

ϵ –cadena 334
 ϵ –red 255

Ecuación de Euler-Lagrange 494
 Ecuación diferencial 242
 Equicontinuo 284
 Equivalencia, relación de 61
 – clases de 61
 Esfera 134,135
 Espacio dual topológico, 450
 Espacio, euclideo 389
 – vectorial 353
 – generado 355
 Estrellado, conjunto 373
 Estructura, algebraica 104,358
 – bornológica 261
 – diferenciable 207
 – geométrica 104, 87
 – topológica 86
 – uniforme 224
 Euclideo, espacio 390
 Extensión 240
 Exterior, punto 178,179
 Extremo, local 492

F

Familia ortonormal 396
 Familias, de conjuntos 77
 – normales 302
 Fórmula, de Newton-Leibnitz
 generalizada 480
 – de Taylor generalizada 480
 Frontera, conjunto 179
 – punto 179
 F_σ -conjunto 177
 – Función 63
 – de Urysohn 197
 – propia 420, 422
 – separante 196
 Funcional 432
 – lineal 432
 – continuo 440

G

G_δ -conjunto 208
 Goursat, teorema de 248
 Grafo de 63

H

Hahn-Banach, teorema de (forma
 geométrica) 456

ma de (forma algebraica) 456
 t, cubo de 282
 cio de 408
 plano 459
 morfismos 152
 morfos, espacios 152

dad, aplicación 68
 n 65
 rsa 67
 sión 240
 al, de Lebesgue 251
 iemann 211
 r, de un conjunto 179
 o 179
 cción, de conjuntos 43
 olecciones 59
 milias 43
 o de 339
 a, función 67
 va, aplicación 65
 ría 108
 e espacios métricos 108

, curva de 270

botella de 115

continua, función 456
 que, número de 256
 ral de 251
 l paralelogramo 392

de Gram 424
 al, elemento 74
 no 273
 o de Newton 246
 a 97
 ucto 105
 as compatibles con la estructura

algebraica 358
 — con la estructura geométrica 413
 — con la estructura topológica 359
 Métrico espacio 97
 Minimal, elemento 74
 Mínimo 273
 Minkowsky, desigualdad de 99
 — funcional de 461
 Möbius, banda de 461
 Monótona, aplicación 128

N

Norma 362
 Normado, espacio 362
 Normas equivalentes 370
 Numerabilidad, primer axioma de 152
 — segundo axioma de 152
 Numerable, conjunto 54

O

Operador 431
 — integral 436, 437
 — isométrico 482
 — lineal 432
 — — continuo 431
 — proyección 69
 — simétrico 503
 Orden, relación de 71
 — filtrante 73
 — producto 61

P

Paralelogramo, ley del 392
 Partición 60
 — de unidad 206
 — de unidad subordinada 207
 Producto, conjunto 105
 — escalar 350
 — relación 61
 Prolongación, continua 68
 — de funciones reales continuas 192
 Producto, cartesiano 45
 — de espacios métricos 105
 Propiedad (global, local) 203
 Proyección, estereográfica 153
 — operador 69

– teorema de la 415

R

Red 211

Relación, de equivalencia 59

– funcional 63

Restricción, aplicación 68

Riesz, teorema de 451

Riesz-Fischer, teorema de 410

S

Saturado 62

Segmento 348

Semicontinua, aplicación 210

Semiespacio, cerrado 466

– positivo (negativo) 461

Seminorma 371

Separable, espacio 165

Separación, axiomas de 190

Separado, conjunto 190

Serie, absolutamente convergente 381

– condicionalmente convergente 383

– de elementos ortogonales 396

– de Fourier 169

Seudométrica 110

Seudométricas, semejantes 155

– topológicamente equivalentes 155

Sobreyectiva, aplicación 65

Soporte de una función 206

Stone-Weierstrass, teorema de 168

Subconjunto 38

Subespacio métrico 108

Sucesión convergente 118

Sumas (superiores e inferiores) 192

Supremo 110

T

Teoría, de aproximación 94

– de optimización 365

– espectral de operadores 351

Totalmente ordenado, conjunto 205

Tijonov, teorema de 280

U

Uniformemente continua, aplicación 222

Unión, de conjuntos 41

– de colecciones 59

– de familias 77

Urysohn, función de 197

– teorema de 199

V

Valor propio 418,422

Vecindad, en un espacio métrico 143

Vector propio 459

Vectores, linealmente dependientes 354

– linealmente independientes 354

W

Weierstrass, teorema de 168,171

Z

Zorn, Teorema de 83