Álgebra Lineal y Geometría

Manuel Castellet \ Irene Llerena

EDITORIAL REVERTÉ

Álgebra lineal y Geometría

Manuel Castellet
Catedrático de la Universidad Autónoma de Barcelona

Irene Llerena
Profesora Titular de la Universidad de Barcelona

Con la colaboración de

Carlos Casacuberta
Profesor Titular de la Universidad de Barcelona

ER

EDITORIAL
REVERTÉ

Barcelona · Bogotá · Buenos Aires · Caracas · México

Título de la obra original:
Àlgebra Lineal i Geometria

Edición en lengua catalana publicada por:
Publicacions de la Universitat Autònoma de Barcelona

Revisado por los autores

Edición en papel
ISBN: 978-84-291-5009-4

Edición e-book (PDF)
ISBN: 978-84-291-9285-8

Propiedad de:
EDITORIAL REVERTE, S.A.
Loreto 13-15 Local B
08029 Barcelona. España
Tel.: 93 419 33 36
Fax: 93 419 51 89
reverte@reverte.com

www.reverte.com

A Albert, Josep y Marc

Et surtout leurs adeptes y trouvent des jouissances analogues à celles que donnent la peinture et la musique. Ils admirent la délicate harmonie des nombres et des formes; ils s'émerveillent quand une découverte nouvelle leur ouvre une perspective inattendue; et la joie qu'ils éprouvent ainsi n'a-t-elle pas le caractère esthétique, bien que les sens n'y prennent aucune part?...

C'est pourquoi je n'hésite pas à dire que les mathématiques méritent d'être cultivées pour elles-mêmes et que les théories qui ne peuvent être appliquées à la physique doivent l'être comme les autres.

...Mais, le mathématicien pur qui oublierait l'existence du monde extérieur serait semblable à un peintre qui saurait harmonieusement combiner les couleurs et les formes, mais à qui les modèles, feraient défaut. Sa puissance créatrice serait bientôt tarie.

Henri Poincaré

Índice

I Divisibilidad en los números enteros

I.1 División entera. Ideales 9
I.2 Mínimo común múltiplo y máximo común divisor 10
I.3 Números primos entre sí y números primos 13
I.4 Congruencias 15
I.5 Los anillos $\mathbf{Z}/(m)$ 17
I.6 Ecuaciones diofánticas lineales 18
I.7 Nota histórica 19
I.8 Ejercicios 20
I.9 Ejercicios para programar 22

II Divisibilidad en el anillo de polinomios

II.1 Definición del anillo de polinomios 23
II.2 División entera e ideales en $K[x]$ 25
II.3 Mínimo común múltiplo y máximo común divisor 27
II.4 Polinomios irreducibles y polinomios primos entre sí . . . 30
II.5 Ceros de un polinomio 32
II.6 Polinomios irreducibles de $\mathbf{R}[x]$ 34
II.7 Los anillos $K[x]/(m(x))$ 35
II.8 Nota histórica 38
II.9 Ejercicios 39
II.10 Ejercicios para programar 40

III Grupos

III.1 Definición y ejemplos 41
III.2 Permutaciones 43
III.3 Subgrupos 47
III.4 Homomorfismos 49
III.5 Grupo cociente. Subgrupos normales 51
III.6 Producto directo de grupos 55

III.7 Grupos cíclicos 57
III.8 Grupos finitos 58
III.9 Nota histórica 62
III.10 Ejercicios 63
III.11 Ejercicios para programar 66

IV Espacios vectoriales
IV.1 Definición y ejemplos 67
IV.2 Subespacios vectoriales 70
IV.3 Bases de un espacio vectorial 72
IV.4 Fórmula de Grassmann. Suma directa de subespacios. . . 77
IV.5 Suma directa de espacios vectoriales 79
IV.6 Espacio vectorial cociente. 80
IV.7 Coordenadas 82
IV.8 Nota histórica 84
IV.9 Ejercicios 85
IV.10 Ejercicios para programar 87

V Aplicaciones lineales
V.1 Definición y ejemplos 89
V.2 Matriz asociada a una aplicación lineal 94
V.3 Teorema de isomorfismo 99
V.4 El espacio de las aplicaciones lineales 102
V.5 El álgebra de endomorfismos 103
V.6 El espacio dual 105
V.7 Subespacios ortogonales 109
V.8 Nota histórica 111
V.9 Ejercicios 111
V.10 Ejercicios para programar 114

VI Determinantes
VI.1 Determinante de n vectores 115
VI.2 Determinante de una matriz 121
VI.3 Determinante de un endomorfismo 122
VI.4 Regla de Laplace 124
VI.5 Cálculo del rango de una matriz 128
VI.6 Nota histórica 132
VI.7 Ejercicios 132
VI.8 Ejercicios para programar 134

VII Sistemas de ecuaciones lineales

VII.1 Planteo del problema 135
VII.2 Existencia de soluciones 136
VII.3 Regla de Cramer 137
VII.4 Resolución de un sistema de ecuaciones lineales 137
VII.5 Método de Gauss 140
VII.6 Cálculo de la matriz inversa 143
VII.7 Nota histórica 144
VII.8 Ejercicios 145
VII.9 Ejercicios para programar 146

VIII Estructura de los endomorfismos

VIII.1 Vectores propios y valores propios. Polinomio característico 149
VIII.2 Diagonalización de matrices 152
VIII.3 Polinomio mínimo 157
VIII.4 Subespacios invariantes 159
VIII.5 Grado del polinomio mínimo 166
VIII.6 El teorema de Cayley-Hamilton 166
VIII.7 Matriz canónica (general) de un endomorfismo 168
VIII.8 Matriz canónica de Jordan 173
VIII.9 Nota histórica 177
VIII.10 Ejercicios 177
VIII.11 Ejercicios para programar 181

IX Espacios afines

IX.1 Definición de espacio afín 184
IX.2 Traslaciones. Otra definición de espacio afín 186
IX.3 Variedades lineales 187
IX.4 Intersección y suma de variedades lineales 189
IX.5 Dependencia lineal de puntos 192
IX.6 Coordenadas baricéntricas 194
IX.7 Ecuaciones de una variedad en coordenadas baricéntricas 200
IX.8 Coordenadas cartesianas 201
IX.9 Ecuaciones de una variedad en coordenadas cartesianas 203
IX.10 Razón simple 205
IX.11 Orientación de un espacio afín real 209
IX.12 Semiespacios 210
IX.13 Nota histórica 211
IX.14 Ejercicios 212
IX.15 Ejercicios para programar 215

X Afinidades

X.1 Definición y primeras propiedades 217
X.2 Unos ejemplos 221
X.3 Más propiedades de las afinidades 225
X.4 Ecuaciones de una afinidad en una referencia cartesiana . 229
X.5 El grupo afín 233
X.6 Variedades invariantes 236
X.7 Clasificación de las afinidades de un espacio afín A en sí mismo 238
X.8 Afinidades de la recta afín 240
X.9 Afinidades del plano afín 241
X.10 Nota histórica 244
X.11 Ejercicios 245
X.12 Ejercicios para programar 247

XI Espacios vectoriales euclídeos y unitarios

XI.1 Formas bilineales y sesquilineales 249
XI.2 Producto escalar 251
XI.3 Norma 256
XI.4 Producto escalar y espacio dual 258
XI.5 Subespacios ortogonales 259
XI.6 Aplicaciones adjuntas y autoadjuntas 260
XI.7 Diagonalización de matrices simétricas y hermíticas . . . 262
XI.8 Producto vectorial 263
XI.9 Nota histórica 266
XI.10 Ejercicios 266
XI.11 Ejercicios para programar 268

XII Aplicaciones ortogonales. Aplicaciones unitarias

XII.1 Definiciones 271
XII.2 Diagonalización de matrices unitarias 274
XII.3 Forma canónica de una matriz ortogonal 274
XII.4 Los grupos $O(2)$ y $SO(2)$ 277
XII.5 Ángulos 280
XII.6 El grupo $O(3)$ 286
XII.7 Otra determinación de las rotaciones 289
XII.8 Composición de rotaciones 289
XII.9 Nota histórica 293
XII.10 Ejercicios 293
XII.11 Ejercicios para programar 295

XIII Espacios afines euclídeos

XIII.1 Espacios afines euclídeos 299
XIII.2 Distancia entre dos variedades lineales 301
XIII.3 Isometrías 304
XIII.4 Clasificación de los desplazamientos 306
XIII.5 Desplazamientos de la recta euclídea 307
XIII.6 Desplazamientos del plano euclídeo 307
XIII.7 Desplazamientos del espacio euclídeo tridimensional . . . 309
XIII.8 Semejanzas 313
XIII.9 Semejanzas del espacio afín euclídeo tridimensional . . . 315
XIII.10 Semejanzas del plano afín euclídeo 316
XIII.11 Algunos ejemplos y aplicaciones 318
XIII.12 Nota histórica 325
XIII.13 Ejercicios 326
XIII.14 Ejercicios para programar 330

Introducción

La imagen de un gran roble que tiene por raíces el álgebra, la geometría plana, la trigonometría, la geometría analítica y los números irracionales, por tronco el análisis, y diversas ramas, ya no es aceptada actualmente. Hoy en día, a finales del siglo 20, la imagen adecuada para representar las matemáticas es, tal como dice H. Eves, la de un baniano, un árbol con varios troncos, que desarrolla siempre troncos nuevos: cada rama del baniano, por un crecimiento fibroso, se extiende hacia abajo hasta llegar al suelo. Entonces arraiga y con el tiempo ese filamento se va volviendo grueso y fuerte hasta convertirse en un nuevo tronco con muchas ramas, cada una de las cuales lanza sus filamentos hacia el suelo.

Al igual que el gran roble, esos banianos son hermosos y tienen una larga vida. Se dice que el baniano de la India bajo el cual meditaba Buda todavía vive y sigue creciendo.

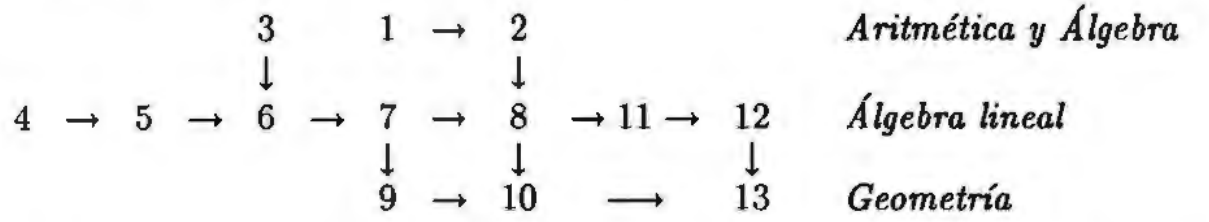
Se puede ascender al árbol por diferentes troncos, empezando por los fundamentos, que representan las raíces del tronco elegido. Todos los troncos están, evidentemente, interconectados por el complicado sistema de ramaje del árbol.

Nosotros hemos escogido tres troncos del baniano: la aritmética, el álgebra lineal y la geometría. De cada uno de estos troncos hemos presentado algunas raíces que han de permitir al estudiante ir subiendo por el árbol y, junto con los conocimientos adquiridos en otros troncos (análisis, álgebra, topología, etc.), poder moverse seguro por una parcela del gran baniano.

El presente texto es el fruto de la experiencia de varios años de los autores impartiendo las asignaturas "Geometría I" en la Universidad de Barcelona y "Álgebra I" en la Universidad Autónoma de Barcelona, y está fuertemente influenciado por el constante intercambio de ideas con Josep Vaquer.

El libro se puede dividir en tres partes: Aritmética y Álgebra (capítulos 1, 2 y 3), Álgebra lineal (capítulos 4, 5, 6, 7, 8, 11 y 12) y Geometría (capítulos 9, 10 y 13). Aunque cada parte tiene interés propio, todas ellas están íntimamente relacionadas dando unidad al texto. Si $A \to B$ significa

que para estudiar el capítulo B se necesita una parte esencial del A, los capítulos del texto pueden ordenarse así:



Hemos procurado que el lenguaje sea llano y que el texto pueda servir tanto al profesor como al alumno que lo lea por su cuenta. Al final de cada capítulo hemos incluido una breve nota histórica, que solamente pretende ser lo que su nombre indica.

Junto con una lista cuidadosamente elaborada de ejercicios clásicos, hemos incluido en cada capítulo una serie de ejercicios para programar. Para la resolución de esos ejercicios, suponemos que el estudiante es capaz de programar en algún lenguaje y pretendemos que con ellos profundice en la teoría y al mismo tiempo trabaje personalmente métodos de cálculo.

El libro está especialmente pensado para estudiantes de primer curso de las Facultades de Matemáticas, Física e Informática y de las Escuelas de Ingeniería y Arquitectura. Una selección de los capítulos puede ser adecuada también para los restantes estudios científicos, técnicos y socioeconómicos.

La colaboración de Carles Casacuberta ha sido decisiva para la confección de este libro. La lectura atenta que ha hecho del texto, sus sugerencias y la elaboración de la mayoría de los ejercicios para programar han significado una contribución inestimable que agradecemos de todo corazón.

La realización material del libro se ha visto facilitada por la aportación de Jordi Saludes, quien, gracias a su dominio del programa de edición TEX, ha hecho posible que el texto tenga este aspecto tan agradable.

M. CASTELLET – I. LLERENA

Capítulo I

Divisibilidad en los números enteros

I.1 División entera. Ideales

Designaremos por $\mathbf{Z}$ el conjunto de los números enteros. La teoría de la divisibilidad en $\mathbf{Z}$ es consecuencia de la siguiente importante propiedad.

Teorema 1.1 (de la división entera) *Dados $a, b \in \mathbf{Z}$, $b \neq 0$, existen dos únicos números enteros q y r que cumplen $a = bq + r$, $0 \leq r < |b|$. Estos números q y r se llaman el* cociente *y el* resto *de la división entera de a por b.*

Ejemplo:

$$-8 = 3 \cdot (-3) + 1, \qquad 3 = (-8) \cdot 0 + 3.$$

Si el resto de la división entera de a por b es 0, se dice que a es un *múltiplo* de b (escribiremos $a = \dot{b}$), que b es un *divisor* de a (escribiremos $b \mid a$), o que *a es divisible por b.* Indicaremos por (b) el conjunto de los múltiplos de b. Observemos que (b) cumple las dos propiedades siguientes:

- es cerrado por la suma; es decir, $a, c \in (b) \Rightarrow a + c \in (b)$;
- si $a \in (b)$ y c es cualquier entero, entonces $ac \in (b)$.

Proposición 1.2 *Si el subconjunto $I \subset \mathbf{Z}$ cumple*

1. $a, b \in I \Rightarrow a + b \in I$,
2. $a \in I$, $c \in \mathbf{Z} \Rightarrow ac \in I$,

entonces existe un $b \in I$ tal que $I = (b)$.

DEMOSTRACIÓN: Si $I = \{0\}$, entonces $I = (0)$. Si I contiene un elemento no nulo a, también contiene $-a = a \cdot (-1)$, y o bien a o bien $-a$ es positivo. Por tanto, I contiene enteros positivos. Sea b el menor de los positivos contenidos en I. Por 2, I contiene todos los múltiplos de b: $(b) \subset I$. Vamos a ver que $I \subset (b)$ y, por tanto, $I = (b)$. En efecto, dado $a \in I$ cualquiera, por (1.1),

$$a = bq + r.$$

Por 1 y 2, $r = a - bq = a + b(-q) \in I$; pero $0 \leq r < |b| = b$, y b es el menor de los positivos de I; así pues, $r = 0$ y, por tanto, $a = bq \in (b)$. □

Un subconjunto I que cumple las condiciones 1 y 2 de (1.2) se llama un *ideal* de $\mathbf{Z}$. El elemento b tal que $I = (b)$ se denomina *base* del ideal.

Ejercicio:

$(b) = (c)$ si y sólo si $c = \pm b$.

Observación:

$(a) \subset (b)$ si y sólo si $b \mid a$. Las cuestiones de divisibilidad equivalen, por tanto, a cuestiones sobre inclusiones entre ideales.

I.2 Mínimo común múltiplo y máximo común divisor

Dados números enteros $a_1, \ldots, a_n$, la intersección $(a_1) \cap \ldots \cap (a_n)$ es el conjunto de los números enteros múltiplos comunes de todos ellos. Este conjunto cumple las dos condiciones de (1.2) y, por tanto, $(a_1) \cap \ldots \cap (a_n) = (m)$ para un m conveniente. Este m está caracterizado por las dos propiedades siguientes:

- m es múltiplo común de $a_1, \ldots, a_n$;
- cualquier otro múltiplo común de $a_1, \ldots, a_n$ es múltiplo de m.

Diremos que m es el *mínimo común múltiplo* de $a_1, \ldots, a_n$ y escribiremos

$$m = \text{m.c.m.}(a_1, \ldots, a_n).$$

¡Atención!:

Observemos que también $-m$ es mínimo común múltiplo de $a_1, \ldots, a_n$.

Consideremos ahora la unión $(a_1) \cup \ldots \cup (a_n)$. Este conjunto, en general, no cumple las condiciones de (1.2). Por ejemplo, $(2) \cup (3)$ no contiene el $5 = 2 + 3$. Formemos a partir de $(a_1) \cup \ldots \cup (a_n)$ un subconjunto I de $\mathbf{Z}$ que cumpla las condiciones de (1.2). Por la condición 1, I debe contener todas las sumas de múltiplos de $a_1, \ldots, a_n$: $a_1c_1 + \ldots + a_nc_n$. No hace falta ampliar más; el conjunto

$$I = \{a_1c_1 + \ldots + a_nc_n \mid c_1, \ldots, c_n \in \mathbf{Z}\}$$

cumple ya las condiciones de (1.2) y, por tanto, existe un entero d tal que $I = (d)$. Denotaremos I por $(a_1, \ldots, a_n)$. Así pues, $I = (a_1, \ldots, a_n) = (d)$. Este número d está caracterizado por las dos propiedades siguientes:

- d es divisor común de $a_1, \ldots, a_n$, ya que ello equivale a afirmar que $a_i \in (d)$ para $i = 1, \ldots, n$. $(a_i = a_1 \cdot 0 + \ldots + a_i \cdot 1 + \ldots + a_n \cdot 0 \in I)$.
- Cualquier otro divisor d' común a $a_1, \ldots, a_n$ divide a d. En efecto, que d' sea divisor de $a_1, \ldots, a_n$ significa que $a_i \in (d')$, $i = 1, \ldots, n$. Por tanto, $\{a_1c_1 + \ldots + a_nc_n \mid c_i \in \mathbf{Z}\} \subset (d')$, es decir, $(d) \subset (d')$, lo cual implica que d' es un divisor de d.

Diremos que d es el *máximo común divisor* de $a_1, \ldots, a_n$ y escribiremos

$$d = \text{m.c.d.}(a_1, \ldots, a_n).$$

¡Atención!:

También $-d$ es máximo común divisor de $a_1, \ldots, a_n$.

Observemos que el máximo común divisor d es una suma de múltiplos de $a_1, \ldots, a_n$,

$$d = a_1r_1 + \ldots + a_nr_n.$$

Esta expresión es conocida como *identidad de Bézout.*

Acabaremos este apartado con un método práctico de cálculo del máximo común divisor y de la identidad de Bézout. El método se basa en el siguiente resultado:

Proposición 2.1 *Sea $a = bq + r$ la división entera de a por b. Entonces*

$$\text{m.c.d.}(a, b) = \text{m.c.d.}(b, r).$$

DEMOSTRACIÓN: El resultado es consecuencia de que $(a,b) = (b,r)$. En efecto, todo elemento $ac_1 + bc_2 \in (a,b)$ satisface $ac_1 + bc_2 = b(qc_1 + c_2) + rc_1 \in (b,r)$ y, recíprocamente, todo elemento $bn_1 + rn_2 \in (b,r)$ satisface $bn_1 + rn_2 = an_2 + b(n_1 - qn_2) \in (a,b)$. □

Si aplicamos reiteradamente esta proposición, obtenemos

$$\begin{array}{lll} a = bq + r, & (a,b) = (b,r), & r < |b|, \\ b = rq_1 + r_1, & (b,r) = (r,r_1), & r_1 < r, \\ r = r_1q_2 + r_2, & (r,r_1) = (r_1,r_2), & r_2 < r_1. \end{array}$$

Los sucesivos restos van disminuyendo y obtendremos, por tanto, en un momento dado resto cero:

$$\begin{array}{lll} r_{k-2} = r_{k-1}q_k + r_k, & (r_{k-2},r_{k-1}) = (r_{k-1},r_k), & r_k < r_{k-1}, \\ r_{k-1} = r_kq_{k+1} + 0, & (r_{k-1},r_k) = (r_k,0) = (r_k). & \end{array}$$

Así pues, $(a,b) = (r_k)$; es decir, $r_k = \text{m.c.d.}(a,b)$.

Este método para hallar el máximo común divisor se llama *algoritmo de Euclides.*

Para calcular el máximo común divisor de más de dos enteros, aplicamos:

Ejercicio:

$\text{m.c.d.}(a_1,a_2,a_3) = \text{m.c.d.}[\text{m.c.d.}(a_1,a_2),a_3]$ y, en general,
$\text{m.c.d.}(a_1,\ldots,a_n) = \text{m.c.d.}[\text{m.c.d.}(a_1,\ldots,a_{n-1}),a_n]$.

Las divisiones enteras efectuadas en el algoritmo de Euclides nos permiten expresar $d = r_k = \text{m.c.d.}(a,b)$ como suma de un múltiplo de a y un múltiplo de b. En efecto, en

$$d = r_k = r_{k-2} - r_{k-1}q_k$$

d se expresa como suma de un múltiplo de r_{k-2} y un múltiplo de r_{k-1}. Ahora bien, $r_{k-1} = r_{k-3} - r_{k-2}q_{k-1}$, y sustituyendo en la igualdad anterior obtenemos una expresión de d como suma de un múltiplo de r_{k-3} y un múltiplo de r_{k-2}. Volviendo a sustituir convenientemente, podemos expresar d como suma de múltiplos de r_{k-4} y r_{k-3}; y así sucesivamente hasta obtener la identidad de Bézout

$$d = ar + bs.$$

En el próximo apartado (3.2) demostraremos que si $m = \text{m.c.m.}(a,b)$ y $d = \text{m.c.d.}(a,b)$, entonces $md = \pm ab$. Esto nos permite calcular m si conocemos d. Para el cálculo del mínimo común múltiplo de más de dos números utilizamos:

Ejercicio:

$\text{m.c.m.}(a_1, a_2, a_3) = \text{m.c.m.}[\text{m.c.m.}(a_1, a_2), a_3]$ y, en general,
$\text{m.c.m.}(a_1, \ldots, a_n) = \text{m.c.m.}[\text{m.c.m.}(a_1, \ldots, a_{n-1}), a_n]$.

I.3 Números primos entre sí y números primos

Se dice que a y b son *primos entre sí* si $\text{m.c.d.}(a, b) = 1$.

Ejemplos:

1. $\text{m.c.d.}(3, 8) = 1$. Observemos que $1 = 3 \cdot 3 + 8 \cdot (-1)$.
2. Si $d = \text{m.c.d.}(a, b)$ y $a = da'$, $b = db'$, entonces $\text{m.c.d.}(a', b') = 1$. En efecto, si d' fuera un divisor común de a' y b', entonces dd' sería divisor común de a y b y, por tanto, un divisor de d. Esto sólo es posible si $d' = \pm 1$.

Teorema 3.1 (de Euclides) *Si $a \mid bc$ y $\text{m.c.d.}(a, b) = 1$, entonces $a \mid c$.*

DEMOSTRACIÓN: Si $1 = \text{m.c.d.}(a, b)$, podemos expresar el 1 como $1 = ar + bs$. Multiplicando por c obtenemos $c = acr + bcs$. Pero a divide a los dos sumandos y, por tanto, $a \mid c$. □

Proposición 3.2 *Si $m = \text{m.c.m.}(a, b)$ y $d = \text{m.c.d.}(a, b)$, entonces se cumple $md = \pm ab$.*

DEMOSTRACIÓN: Pongamos $a = da'$ y $b = db'$. Se trata de ver que $m = \pm da'b'$ es un mínimo común múltiplo de a y b. Es evidente que $da'b'$ es múltiplo común de a y b. Sea n otro múltiplo común de a y b; es decir, $n = ar = bs$. Entonces $a'dr = b'ds$, de donde $a'r = b's$ con a', b' primos entre sí. Entonces, por (3.1), a' divide a s, es decir, $s = a'h$ y $n = bs = db'a'h$. Así resulta que n es múltiplo de $db'a'$. □

Cualquier número entero p es divisible por ± 1 y por $\pm p$. Diremos que p es *primo* si estos son sus únicos divisores. El 1 y el -1 no se consideran números primos.

Proposición 3.3 *El conjunto de los números primos es infinito.*

DEMOSTRACIÓN: Lo demostraremos viendo que, dado un conjunto finito de números primos $N = \{p_1, \ldots, p_m\}$, siempre hay un número primo fuera de N. En efecto, consideremos $a = p_1 \cdots p_m + 1$. Si $b \mid a$, también $-b \mid a$; por tanto, a tiene siempre divisores positivos. Sea p el menor de los divisores positivos de a diferentes de 1. Claramente, p es primo. Si p fuera uno de los p_i, dividiría a $p_1 \cdots p_m$ y, por tanto, dividiría a $a - p_1 \cdots p_m = 1$. Esto es imposible, ya que $p \neq 1$. De ahí que $p \notin N$. □

Proposición 3.4 *Todo número entero a no nulo, $a \neq \pm 1$, es producto de números primos.*

DEMOSTRACIÓN: Tal como hemos visto en la demostración de (3.3), a tiene sempre un divisor primo $p_1 \neq \pm 1$. Así pues, tenemos $a = p_1 a_1$. Si $a_1 \neq \pm 1$, elijamos un divisor primo de a_1, $p_2 \neq \pm 1$, y tendremos $a_1 = p_2 a_2$. Luego $a = p_1 p_2 a_2$. Repitamos el mismo proceso si $a_2 \neq \pm 1$, y así sucesivamente. Ahora bien, $|a| > |a_1| > |a_2| > \ldots$. Llegará pues un momento en que tendremos $a = p_1 \cdots (p_n a_n)$ con $a_n = \pm 1$. Esto es una descomposición de a en números primos. □

La descomposición de un entero en producto de primos no es exactamente única. Por ejemplo,

$$12 = 2 \cdot 2 \cdot 3 = 2 \cdot (-2) \cdot (-3) = (-2) \cdot 3 \cdot (-2).$$

Hay, sin embargo, una cierta unicidad. Concretamente,

Proposición 3.5 *Si $p_1 \cdots p_n = q_1 \cdots q_m$ y todos los factores p_i, q_j son números primos, $i = 1, \ldots, n$, $j = 1, \ldots, m$, entonces $n = m$ y los números $\{p_1, \ldots, p_n\}$ son los mismos que los $\{q_1, \ldots, q_m\}$, salvo el signo (y el orden).*

DEMOSTRACIÓN: Observemos que si p, q son números primos, entonces o bien m.c.d.$(p, q) = 1$ o bien $p = \pm q$. Pero p_1 divide a $p_1 \cdots p_n = q_1(q_2 \cdots q_m)$. Por el teorema de Euclides (3.1), o bien $p_1 \mid q_2 \cdots q_m$, cuando m.c.d.$(p_1, q_1) = 1$, o bien $p_1 = \pm q_1$. En el primer caso, $p_1 \mid q_2(q_3 \cdots q_m)$; aplicando nuevamente el teorema de Euclides, obtenemos que $p_1 \mid q_3 \cdots q_m$, o $p_1 = \pm q_2$. Repitamos el proceso tantas veces como sea necesario. O bien hallaremos que p_1 es uno de los $q_j, j = 1, \ldots, m-2$, salvo el signo, o bien concluiremos que $p_1 \mid q_{m-1} q_m$, de donde $p_1 = \pm q_{m-1}$ o $p_1 = \pm q_m$.

Así pues, p_1 coincide, salvo el signo, con uno de los q_j. Cambiando el orden si es necesario, podemos suponer que $p_1 = \pm q_1$. Entonces $p_2 \cdots p_n = (\pm q_2) q_3 \cdots q_m$. El mismo razonamiento prueba que p_2 es igual, salvo el signo, a uno de los q_j, $j = 2, \ldots, m$, y así sucesivamente. Si $n < m$, llegaremos a la situación $1 = \pm q_{n+1} \cdots q_m$, y esto no es posible porque todos los q_j son diferentes de ± 1. Si $m > n$, llegaremos a $\pm p_{m+1} \cdots p_n = 1$, igualmente imposible. Por tanto, $n = m$. □

I.4 Congruencias

Fijemos $0 \neq m \in \mathbf{Z}$. Diremos que dos números enteros a y b son *congruentes módulo* m si $a - b \in (m)$. Esto equivale a decir que las divisiones enteras de a y b por m tienen el mismo resto. En efecto,

$$\left.\begin{array}{l} a = mq + r \\ b = mq_1 + r_1 \end{array}\right\} \Rightarrow a - b = m(q - q_1) + (r - r_1) \text{ con } |r - r_1| < |m|.$$

Por tanto, $a - b \in (m)$ si y sólo si $r = r_1$. Si a y b son congruentes módulo m, escribiremos $a \equiv b\,(m)$.

Es muy fácil ver que se cumplen las siguientes condiciones:

1. Para todo $a \in \mathbf{Z}$, $a \equiv a\,(m)$.
2. $a \equiv b\,(m) \Rightarrow b \equiv a\,(m)$.
3. $a \equiv b\,(m)$, $b \equiv c\,(m) \Rightarrow a \equiv c\,(m)$.

Formemos ahora subconjuntos de $\mathbf{Z}$ de la siguiente manera: cada subconjunto está formado por todos los números enteros que dan el mismo resto al efectuar la división entera por m. Obtenemos m subconjuntos:

- $(m) =$ conjunto de enteros que dan resto 0,
- $\{\dot{m} + 1\} =$ conjunto de enteros que dan resto 1,
- .....
- $\{\dot{m} + (|m| - 1)\} =$ conjunto de enteros que dan resto $|m| - 1$.

Estos conjuntos se llaman *clases de restos módulo* m. Designaremos por $\mathbf{Z}/(m)$ el conjunto de las clases de restos módulo m. Cada entero está en una de estas clases y sólo en una. Una clase queda, por tanto, bien determinada al dar uno cualquiera de sus elementos. Diremos que ese elemento es un *representante* de la clase.

Nota:

El proceso que acabamos de llevar a cabo es un caso particular de un proceso general muy usual en matemáticas. Se trata de lo siguiente: sea A un conjunto; una *relación* en A es un criterio que nos permite decidir si dos elementos cualesquiera de A, a y b, "satisfacen la relación" o no. Más exactamente: dar una relación en A es dar una colección de pares ordenados de elementos de A (que serán los elementos que "satisfacen la relación"); es decir, dar un subconjunto del producto cartesiano $A \times A$. Indicaremos por $a \sim b$ el hecho de que a esté relacionado con b. Ejemplos de relaciones son

- $a \sim b \Leftrightarrow a \mid b$.
- $a \sim b \Leftrightarrow a < b$.
- $a \sim b \Leftrightarrow a - b \in (m)$.

Una relación es *relación de equivalencia* si cumple

- Propiedad reflexiva: para todo $a \in A$, $a \sim a$.
- Propiedad simétrica: $a \sim b \Rightarrow b \sim a$.
- Propiedad transitiva: $a \sim b$, $b \sim c \Rightarrow a \sim c$.

De los ejemplos anteriores, sólo la congruencia módulo m es una relación de equivalencia. Toda relación de equivalencia nos permite dividir el conjunto A en subconjuntos disjuntos (*clases de equivalencia*) de la siguiente manera: cada clase está formada por todos los elementos relacionados entre sí. Las tres propiedades anteriores aseguran que todo elemento está en una y sólo en una clase. En efecto, designemos por $[a]$ la clase de todos los elementos relacionados con a. Claramente, $a \in [a]$. Supongamos que a está también en otra clase: $a \in [c]$. Entonces $a \sim c$ y las propiedades transitiva y simétrica nos dicen que todo elemento relacionado con a está también relacionado con c y viceversa. Es decir, $[a] = [c]$.

Una *partición* de A es una serie de subconjuntos de A tales que todo $a \in A$ está en uno y sólo en uno de esos subconjuntos. Una *clasificación* de los elementos de A no es otra cosa que una partición de A. Por ejemplo, clasificamos $\mathbf{Z}$ en pares e impares, o clasificamos las personas por su nacionalidad. Las clases de equivalencia forman una partición de A. Recíprocamente, una partición de A determina una relación de equivalencia: $a, b \in A$ son "equivalentes" si están en el mismo subconjunto de la partición. Por tanto, clasificar es lo mismo que formar clases por una relación de equivalencia. Esta relación viene a ser el criterio según el cual clasificamos. Por ejemplo, si queremos clasificar los números enteros, tendremos que fijar con qué criterio lo hacemos. Si lo hacemos según su paridad, situaremos dos enteros en la misma clase si ambos son pares o ambos son impares. Lo que hemos hecho no es sino dar una relación de equivalencia.

El conjunto de las clases de equivalencia se llama *conjunto cociente* y se denota por $A/\sim$.

I.5 Los anillos $\mathbf{Z}/(m)$

Queremos ahora definir unas operaciones en $\mathbf{Z}/(m)$ que desempeñen el papel de la suma y el producto en $\mathbf{Z}$. La manera más natural de hacerlo es definir

$$[a]+[b]=[a+b], \qquad [a]\cdot[b]=[ab].$$

Sin embargo, hay un problema. Consideremos unos representantes distintos de las clases $[a]$ y $[b]$: sean $[a_1]=[a], [b_1]=[b]$. Las mismas definiciones dan $[a_1]+[b_1]=[a_1+b_1]$, $[a_1]\cdot[b_1]=[a_1b_1]$. Las clases $[a_1+b_1]$, $[a_1b_1]$ que ahora obtenemos, ¿coinciden con las clases $[a+b]$, $[ab]$ antes obtenidas? En otras palabras, la suma y el producto definidos, ¿dependen de los representantes elegidos? La respuesta es no; en efecto,

$$\left.\begin{array}{l}[a_1]=[a] \Rightarrow a_1=a+\dot{m}\\ [b_1]=[b] \Rightarrow b_1=b+\dot{m}\end{array}\right\} \Rightarrow \left\{\begin{array}{l} a_1+b_1=a+b+\dot{m} \Rightarrow [a_1+b_1]=[a+b]\\ a_1b_1=ab+\dot{m} \Rightarrow [a_1b_1]=[ab].\end{array}\right.$$

Un conjunto A con dos operaciones $(a+b,\ a\cdot b)$ es un *anillo* si cumple:

- Propiedades de $+$:
  - Asociativa: $(a+b)+c=a+(b+c) \quad \forall a,b,c\in A.$
  - Conmutativa: $a+b=b+a \quad \forall a,b\in A.$
  - Existe un elemento, que denominaremos *cero* y designaremos por 0, tal que
    $$a+0=0+a=a \quad \forall a\in A.$$
  - Para cada $a\in A$ hay un elemento, que denominaremos el *opuesto* de a y denotaremos por $-a$, tal que $a+(-a)=0$.
- Propiedad de $\cdot$:
  - Asociativa: $(a\cdot b)\cdot c=a\cdot(b\cdot c) \quad \forall a,b,c\in A.$
- Propiedades que relacionan $+$ y $\cdot$:
  - Distributivas:
    $$a\cdot(b+c)=a\cdot b+a\cdot c,$$
    $$(a+b)\cdot c=a\cdot c+b\cdot c \quad \forall a,b,c\in A.$$

Si, además, se cumple que la operación $\cdot$ es conmutativa ($a\cdot b=b\cdot a$ para todo $a,b\in A$), se dice que A es un *anillo conmutativo*. Si existe un elemento $e\in A$ tal que $a\cdot e=e\cdot a=a$ para todo $a\in A$, se dice que A *tiene unidad*. El elemento e se llama la *unidad* de A y generalmente se designa por 1. Un elemento $a^{-1}\in A$ tal que $a\cdot a^{-1}=a^{-1}\cdot a=1$ se llama un *inverso* de a.

Observemos que en un anillo se cumple $a \cdot 0 = 0 \cdot a = 0$ para todo a. En efecto, $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Por tanto, sumando $-(a \cdot 0)$ a ambos lados, obtenemos $0 = a \cdot 0$. Resulta, pues, que en un anillo A el 0 no puede tener inverso. Un anillo conmutativo con unidad en el cual todo elemento distinto de cero posee inverso se llama un *cuerpo*. $\mathbf{Z}$ es un anillo conmutativo con unidad. El conjunto de los racionales $\mathbf{Q}$, el conjunto de los reales $\mathbf{R}$ y el conjunto de los complejos $\mathbf{C}$ son cuerpos.

$\mathbf{Z}/(m)$ es un anillo conmutativo con unidad, [1]. $\mathbf{Z}/(m)$ tiene, sin embargo, propiedades que no tenía $\mathbf{Z}$. Por ejemplo, el producto de dos elementos diferentes de [0] puede ser [0]. Así, en $\mathbf{Z}/(6)$, $[2] \cdot [3] = [0]$. A estos elementos se les llama *divisores de cero*. Por otro lado, hay elementos que tenen inverso. Por ejemplo, en $\mathbf{Z}/(8)$, $[3] \cdot [3] = [1]$. Observemos que, en un anillo, si un elemento es divisor de cero no puede tener inverso. En efecto, sea $a \cdot b = 0$ con $a \neq 0$ y $b \neq 0$. Si existe el inverso de a, resulta que $b = 1 \cdot b = (a^{-1} \cdot a) \cdot b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0$, en contra de lo que hemos supuesto.

Proposición 5.1 *Si* m.c.d.$(a, m) = 1$, $[a]$ *tiene un inverso en* $\mathbf{Z}/(m)$. *Si* m.c.d.$(a, m) = d \neq \pm 1, \pm m$, *entonces* $[a]$ *es un divisor de cero en* $\mathbf{Z}/(m)$.

DEMOSTRACIÓN: Si m.c.d.$(a, m) = 1$ podemos poner $1 = ar + ms$, de donde $[1] = [ar] = [a][r]$ y $[r]$ es inverso de $[a]$. Si $d =$ m.c.d.(a, m), pongamos $a = da'$, $m = dm'$. Entonces $am' = a'm \in [0]$, de donde $[a][m'] = [0]$ y $[m'] \neq 0$, ya que $0 < m' < |m|$. □

Corolario 5.2 *El anillo* $\mathbf{Z}/(p)$ *es un cuerpo si y sólo si* p *es primo.*

DEMOSTRACIÓN: Si p es primo, (5.1) nos dice que $\mathbf{Z}/(p)$ es un cuerpo. Si $\mathbf{Z}/(p)$ es un cuerpo, no puede tener divisores de cero (véase la observación hecha antes de (5.1)). Entonces (5.1) nos dice que p debe ser primo. □

I.6 Ecuaciones diofánticas lineales

Nnestro objetivo en este apartado es estudiar las soluciones enteras de la ecuación

$$ax + by = c,$$

donde $a, b, c \in \mathbf{Z}$. La primera proposición se refiere a la existencia de soluciones.

Proposición 6.1 *La ecuación diofántica* $ax + by = c$, $a, b, c \in \mathbf{Z}$, *tiene solución si y sólo si el máximo común divisor de* a *y* b *divide a* c.

Ejercicio:

Demostrar esta proposición.

Supongamos, pues, que $ax + by = c$ tiene solución. Dividiendo por $d =$ m.c.d.(a, b), obtenemos una ecuación con las mismas soluciones, $a'x + b'y = c'$, en la cual m.c.d.$(a', b') = 1$. Multipliquemos la identidad de Bézout $1 = a'r + b's$ por c':

$$c' = a'rc' + b'sc'.$$

$x = rc'$, $y = sc'$ es, por tanto, una solución de la ecuación $a'x + b'y = c'$.

Por otro lado, restando las dos expresiones anteriores obtenemos

$$a'(x - rc') + b'(y - sc') = 0.$$

Por el Teorema de Euclides (3.1)

$$a' \mid y - sc' \qquad \text{y} \qquad b' \mid x - rc'.$$

Es decir, existen t y u tales que

$$\begin{aligned} y &= sc' + ta' \\ x &= rc' + ub'. \end{aligned}$$

Sustituyendo en la ecuación inicial,

$$c' = a'x + b'y = a'rc' + a'ub' + b'sc' + b'ta' = c' + a'b'(u + t),$$

ya que rc', sc' es una solución. Por tanto, $u + t = 0$. La solución general de la ecuación dada es, pues,

$$\begin{aligned} x &= rc' - tb' \\ y &= sc' + ta'. \end{aligned}$$

I.7 Nota histórica

La aritmética, que se inició con los babilonios hacia el año 2000 a. C. y se desarrolló entre los años 600 y 300 a. C. en las escuelas griegas de Pitágoras, Euclides y Diofanto, es todavía hoy una rama de intensa y atractiva actividad investigadora. Las propiedades de los números enteros y las relaciones entre ellos, los conceptos y propiedades de múltiplo, divisor, número primo, la descomposición de un entero (positivo) en producto de primos, el teorema de Euclides, etc., formaron ya parte del cuerpo de doctrina de los libros VII, VIII y IX de los *Elementos* de Euclides. Pierre de Fermat (1601?–1665), un

hombre de letras que leía matemáticas por afición (la *Aritmética* de Diofanto de Alejandría) es una de las figuras clave de la aritmética moderna; él fue quien se planteó el resolver la mayoría de los problemas aritméticos dando algunos criterios, demostrando teoremas, estableciendo conjeturas y asegurando haber demostrado un resultado (conocido ahora como el *último teorema de Fermat*) que, pese a los esfuerzos de los más ilustres matemáticos, sigue siendo una cuestión abierta: la ecuación $x^n + y^n = z^n$ con x, y, z enteros y $n > 2$, no tiene ninguna solución no trivial. Es el gran reto (o la gran espina) que tienen los investigadores en teoría de números. Sin pasar por alto la contribución de Leonhard Euler (1707–1783) que, entre otros, demostró en 1736 el *pequeño teorema de Fermat*: $a^p \equiv a\,(p)$, p primo, y la de Carl Friedrich Gaus (1777–1855), que en sus *Disquisitiones Arithmeticae* sistematizó las congruencias y desarrolló su teoría tal com la usamos hoy en día, conviene mencionar también a Ernst Eduard Kummer (1810–1893), Julius Wilhelm Richard Dedekind (1831–1916) y Leopold Kronecker (1823–1891), los cuales, en sus trabajos sobre números algebraicos, utilizan ya los conceptos de anillo, ideal y cuerpo, aunque las teorías abstractas no se han desarrollado hasta el siglo 20.

I.8 Ejercicios

1. Calcular m.c.d.$(28n + 5, 35n + 2)$ para todo $n \geq 1$.
2. Probar que en la sucesión de Fibonacci 0, 1, 1, 2, 3, 5, 8, 13, ... $(a_n = a_{n-1} + a_{n-2})$ dos términos consecutivos son siempre primos entre sí.
3. Demostrar que, si p es primo, $(p-1)! \equiv -1\,(p)$ (*congruencia de Wilson*).
4. Demostrar que, si p es primo, $a^p \equiv a\,(p)$ para todo a (*pequeño teorema de Fermat*).
5. Calcular 2001^{2001} módulo 17.
6. Demostrar los criterios de divisibilidad por 3, 4, 5, 9, 11, 13 y 19.
7. Resolver las ecuaciones diofánticas $111x + 36y = 15$, $10x + 26y = 1224$, $6x + 10y = 20$, $6x + 10y = 3$.
8. A una isla desierta —sólo habitada por un mono y muchos cocoteros— llegan cinco náufragos; recogen tantos cocos como pueden y se echan a descansar. A medianoche, un marinero desconfiado, temiendo que los otros se despierten y coman algún coco, se levanta, hace cinco partes iguales del total de cocos, separa su parte y deja el resto; pero le ha

sobrado un coco, que da al mono. Al cabo de una hora, un segundo marinero tiene la misma idea: hace cinco partes iguales del total de cocos (¡de los que quedan, por supuesto!), se guarda una parte, deja el resto y da al mono un coco que ha sobrado. Al cabo de otra hora, ... Cada uno de los cinco marineros efectúa la misma operación.

Al día siguiente por la mañana, al levantarse, deciden repartir los cocos (los del montón final) entre los cinco, cada uno de ellos disimulando la risa. Sobra un coco, que dan al mono. Pregunta: ¿cuántos cocos habían recogido como mínimo? (The Saturday Evening Post, $\simeq$ 1925).

9. Oliana Molls trabaja cuatro días consecutivos y descansa uno. Betty trabaja dos y descansa uno. Sólo se ven los días de luna llena (uno de cada veintiocho días). Betty tuvo fiesta ayer, Oliana la tendrá pasado mañana y hace diez días había luna llena. ¿Cuántos días faltan para que se vean? ¿Cuántos días libres comunes habrán perdido mientras tanto por falta de luna llena?

10. a) Encontrar las soluciones de la ecuación lineal $6x \equiv 14\,(16)$, y de la ecuación de segundo grado $x^2 - 3x - 3 \equiv 0\,(7)$.

b) Estudiar en general la resolución de las ecuaciones $ax \equiv b\,(m)$, $ax^2 + bx + c \equiv 0\,(p)$ con p primo.

11. (*Teorema chino del resto*) Demostrar que si $(m, n) = 1$ las ecuaciones $x \equiv a\,(m)$ y $x \equiv b\,(n)$ tienen una única solución módulo mn.

12. Determinar los $a \in \mathbf{Z}/(8)$ tales que el sistema $7x+5y = 2$, $5x+ay = 16$ tiene solución en $\mathbf{Z}/(8)$.

13. a) Demostrar que, si $(a, n) = (b, n) = 1$, la ecuación $ax + by = c$ tiene exactamente n soluciones en $\mathbf{Z}/(n)$.

b) Encontrar las soluciones de $3x+4y = 1$ en $\mathbf{Z}/(7)$ y de $3x+7y = 2$ en $\mathbf{Z}/(8)$.

14. Demostrar que en cualquier solución entera x, y, z de la ecuación $x^2 + y^2 = z^2$ (*terna pitagórica*),

a) x, y o z es múltiplo de 5,

b) x o y es múltiplo de 3,

c) x o y es múltiplo de 4.

15. Demostrar que las únicas relaciones de equivalencia en $\mathbf{Z}$ compatibles con la suma y el producto son las congruencias.

I.9 Ejercicios para programar

16. Cálculo del máximo común divisor y del mínimo común múltiplo de dos números enteros. (Indicación: utilizar las proposiciones I.2.1 y I.3.2.)

17. Resolución de la ecuación diofántica $ax + by = c$. (Indicación: utilizar como subprograma el ejercicio I.16 y seguir el proceso del apartado I.6.)

18. Factorización de un número entero en producto de primos.

19. Construcción de la tabla de los números primos más pequeños que 100.000. (Indicación: ir guardando los primos más pequeños o iguales que 313 en una variable dimensionada. Así estarán disponibles para ir efectuando las sucesivas divisiones.)

20. Cálculo de $1/a$ en $\mathbf{Z}/(p)$, $a \neq 0$, p primo.

21. Cálculo de $\sqrt{a}$ en $\mathbf{Z}/(p)$, p primo, si existe.

Capítulo II

Divisibilidad en el anillo de polinomios

II.1 Definición del anillo de polinomios

Sea K un cuerpo conmutativo. Recordemos que esto significa que en el conjunto K hay definidas dos operaciones, que normalmente denominaremos suma $(+)$ y producto $(\cdot)$, con unas propiedades que hemos explicitado en (I.5). Todos los cuerpos que utilizaremos en este curso serán conmutativos; es decir, cumplirán $a \cdot b = b \cdot a$ para todo par de elementos $a, b \in K$. Por este motivo, diremos simplemente "cuerpo" para indicar un cuerpo conmutativo.

Una *sucesión* de elementos de K es una aplicación

$$\{0, 1, 2, \ldots\} \longrightarrow K.$$

Si indicamos por a_n la imagen de n, está claro que la sucesión queda determinada dando

$$(a_0, a_1, \ldots, a_n, \ldots),$$

que denotaremos abreviadamente por (a_n).

Un *polinomio con coeficientes en* K es una sucesión (a_n) con $a_i = 0$ para todo i salvo un número finito. Si $a_m \neq 0$, pero $a_i = 0$ para todo $i > m$, diremos que m es el *grado* del polinomio (a_n): $m = \mathrm{gr}(a_n)$. Los a_i se llaman los *coeficientes* del polinomio (a_n).

Observaciones:

1. El polinomio $(0, 0, \ldots)$ no tiene grado.
2. Aquellos que recuerden la noción de polinomio como una expresión del tipo $a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$, deben pensar que lo único realmente significativo *a priori* en esta expresión son los coeficientes. En la definición dada más arriba, nos hemos fijado

simplemente en ellos, y hemos considerado como un polinomio $(a_0, a_1, \ldots, a_n, 0, \ldots)$.

Designaremos por $K[x]$ el conjunto de polinomios con coeficientes en K. El porqué de esta notación quedará claro más adelante. Definimos dos operaciones en $K[x]$ de la siguiente manera:

$$(a_0, a_1, \ldots, a_n, \ldots) + (b_0, b_1, \ldots, b_n, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots, a_n + b_n, \ldots)$$

$$(a_0, a_1, \ldots, a_n, \ldots) \cdot (b_0, b_1, \ldots b_n, \ldots) = (a_0 b_0, a_0 b_1 + a_1 b_0, \ldots, c_n, \ldots),$$

donde $c_n = \sum_{i+j=n} a_i b_j$.

Con estas operaciones, $K[x]$ es un anillo conmutativo con unidad. El cero de este anillo es $(0) = (0, 0, \ldots)$ y la unidad $(1, 0, \ldots)$. Se cumple también que si $(a_n) \neq (0)$, $(b_n) \neq (0)$,

$$\begin{aligned} \mathrm{gr}[(a_n) + (b_n)] &\leq \max[\mathrm{gr}(a_n), \mathrm{gr}(b_n)] \\ \mathrm{gr}[(a_n) \cdot (b_n)] &= \mathrm{gr}(a_n) + \mathrm{gr}(b_n). \end{aligned}$$

La segunda igualdad tiene como consecuencias interesantes:

- $(a_n) \cdot (b_n) = (0) \Rightarrow (a_n) = (0)$ o $(b_n) = (0)$.
- $(a_n) \cdot (b_n) = (a_n) \cdot (c_n) \neq 0 \Rightarrow (b_n) = (c_n)$.
- Los únicos elementos invertibles de $K[x]$ son los de grado 0. (Demostrarlo.)

Definimos ahora una nueva operación; si $a \in K$ y $(a_n) \in K[x]$,

$$a \cdot (a_0, a_1, a_2, \ldots) = (aa_0, aa_1, \ldots, aa_n, \ldots).$$

Esto nos permite escribir

$$(a_n) = a_0 \cdot (1, 0, \ldots) + a_1 \cdot (0, 1, 0 \ldots) + a_2 \cdot (0, 0, 1, 0, \ldots) + \ldots + a_n \cdot (0, \ldots, 0, 1, 0 \ldots) + \ldots$$

Esta suma es siempre finita. Además,

$$\begin{aligned} (0,0,1,0,\ldots) &= (0,1,0,\ldots) \cdot (0,1,0,\ldots) \\ (0,0,0,1,0,\ldots) &= (0,0,1,0,\ldots) \cdot (0,1,0,\ldots) \\ &\ldots \qquad \ldots \\ (0,0,\overset{n}{\ldots},0,1,0,\ldots) &= (0,\overset{n-1}{\ldots},1,0,\ldots) \cdot (0,1,0,\ldots). \end{aligned}$$

Si denotamos $(0, 1, 0, \ldots)$ por x, obtenemos

$$(a_n) = a_0 \cdot (1, 0, \ldots) + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots.$$

Si a cada $a \in K$ le hacemos corresponder el polinomio $(a, 0, \ldots)$, obtenemos una aplicación inyectiva

$$K \longrightarrow K[x]$$

que conserva las operaciones; es decir,

$$\begin{aligned} a+b &\longmapsto (a+b,0,\ldots) = (a,0,\ldots)+(b,0,\ldots) \\ ab &\longmapsto (ab,0,\ldots) = (a,0,\ldots)\cdot(b,0,\ldots). \end{aligned}$$

El conjunto $\{(a,0,\ldots) \mid a \in K\}$ de los polinomios de grado 0 juntamente con el (0) está pues en correspondencia biyectiva con K y se comporta igual que K respecto a la suma y al producto. Esto justifica el que designemos $(a,0,\ldots)$ simplemente por a. En particular, $(1,0,\ldots) = 1$ y $(a_0,0,\ldots) = a_0$.

Con esta notación,

$$(a_n) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots$$

o abreviadamente $a(x)$, lo cual motiva la notación usual y la expresión $K[x]$.

II.2 División entera e ideales en $K[x]$

Este apartado debe irse comparando con (I.1), observando especialmente que el papel del valor absoluto en $\mathbf{Z}$ lo juega aquí, en $K[x]$, el grado de un polinomio.

Teorema 2.1 (de la división entera) *Dados dos polinomios $a(x)$ y $b(x)$ diferentes de cero de $K[x]$, existen dos únicos polinomios $q(x)$ y $r(x)$ tales que*

$$a(x) = b(x)\cdot q(x) + r(x)$$

con $r(x) = 0$ o bien $\operatorname{gr} r(x) < \operatorname{gr} b(x)$. Estos polinomios $q(x)$ y $r(x)$ se llaman el cociente *y el* resto *de la división entera de $a(x)$ por $b(x)$.*

DEMOSTRACIÓN: Veremos primero que existen $q(x)$ y $r(x)$.

Sean

$$\begin{aligned} a(x) &= a_0 + a_1x + \ldots + a_nx^n, \qquad a_n \neq 0 \\ b(x) &= b_0 + b_1x + \ldots + b_mx^m, \qquad b_m \neq 0. \end{aligned}$$

Si $n < m$, entonces $a(x) = b(x)\cdot 0 + a(x)$.

Si $n \geq m$, entonces

$$a(x) - b(x)\cdot\left(\frac{a_n}{b_m}x^{n-m}\right) = r_1(x)$$

es cero o tiene grado $n_1 < n$. Por convenio escribimos $x^0 = 1$.

Si $r_1(x) = 0$ o $n_1 < m$, tenemos

$$q(x) = \frac{a_n}{b_m} x^{n-m}$$

y $r(x) = r_1(x)$.

Si $n_1 \geq m$, sea $r_1(x) = c_0 + \ldots + c_{n_1} x^{n_1}$;

$$r_1(x) - b(x) \cdot \left(\frac{c_{n_1}}{b_m} x^{n_1 - m} \right) = r_2(x)$$

es cero o tiene grado $n_2 < n_1$.

Sustituyendo más arriba, obtenemos

$$a(x) = b(x) \cdot \left(\frac{a_n}{b_m} x^{n-m} + \frac{c_{n_1}}{b_m} x^{n_1 - m} \right) + r_2(x).$$

Si $r_2(x) = 0$ o $n_2 < m$, esta expresión nos da un cociente y un resto. Si $n_2 \geq m$, volvemos a repetir el proceso, restando a $r_2(x)$ un múltiplo conveniente de $b(x)$.

Después de un número finito de pasos, obtendremos

$$a(x) = b(x) \cdot \left(\frac{a_n}{b_m} x^{n-m} + \frac{c_{n_1}}{b_m} x^{n_1 - m} + \ldots \right) + r_k(x)$$

con $r_k(x) = 0$ o $\operatorname{gr} r_k(x) < m$.

Veamos ahora que $q(x)$ y $r(x)$ son únicos. Si

$$\begin{aligned} a(x) &= b(x) \cdot q(x) + r(x) \\ a(x) &= b(x) \cdot q_1(x) + r_1(x), \end{aligned}$$

entonces $b(x) \cdot [q(x) - q_1(x)] = r_1(x) - r(x)$. Aquí tenemos $r_1(x) - r(x) = 0$ o $\operatorname{gr}[r_1(x) - r(x)] < m$ y también $q(x) - q_1(x) = 0$ o $\operatorname{gr}[b(x) \cdot [q(x) - q_1(x)]] \geq m$.

Las segundas posibilidades son incompatibles. Por tanto, tendremos $r_1(x) = r(x)$ y $q_1(x) = q(x)$. □

Si el resto de la división entera de $a(x)$ por $b(x)$ es 0, se dice que $a(x)$ es un *múltiplo* de $b(x)$ (y se escribe $a(x) = \dot{b(x)}$), o que $b(x)$ es un *divisor* de $a(x)$ (y se escribe $b(x) \mid a(x)$). Indicaremos por

$$(b(x))$$

el conjunto de los múltiplos de $b(x)$.

Observemos que $(b(x))$ cumple las dos propiedades siguientes:

- es cerrado por la suma $(a(x), c(x) \in (b(x)) \Rightarrow a(x) + c(x) \in (b(x)))$,
- si $a(x) \in (b(x))$ y $c(x) \in K[x]$, entonces $a(x) \cdot c(x) \in (b(x))$.

Denominaremos *ideal* de $K[x]$ a todo subconjunto $I \subset K[x]$ que cumpla:

1. $a(x), b(x) \in I \Rightarrow a(x) + b(x) \in I$,
2. $a(x) \in I$ y $c(x) \in K[x] \Rightarrow a(x) \cdot c(x) \in I$.

Así pues, $(b(x))$ es un ideal. La proposición siguiente nos dice que todos los ideales de $K[x]$ son de este tipo.

Proposición 2.2 *Si I es un ideal de $K[x]$, existe siempre un polinomio $b(x)$ tal que $(b(x)) = I$.*

DEMOSTRACIÓN: Puede ser que $I = \{0\}$. En ese caso, $I = (0)$. Si I contiene algún polinomio diferente del cero, escojamos uno de grado mínimo: $b(x) \in I$. La condición 2 de ideal nos dice que $(b(x)) \subset I$. Por otro lado, dado $a(x) \in I$, podemos efectuar la división entera por $b(x)$:

$$a(x) = b(x) \cdot q(x) + r(x).$$

Entonces $r(x)$ es suma de dos polinomios de I: $r(x) = a(x) + b(x) \cdot [-q(x)]$ y, por tanto, $r(x) \in I$.

Si $r(x) \neq 0$, tendríamos un polinomio en I de grado menor que el grado de $b(x)$, lo cual es imposible. Por tanto, tenemos que $r(x) = 0$ y $a(x) = b(x) \cdot q(x) \in (b(x))$.

Esto nos dice que también $I \subset (b(x))$ y, por tanto, $I = (b(x))$. □

Observaciones:

- $(b(x)) = (k \cdot b(x))$ con $0 \neq k \in K$.
- Si $(a(x)) = (b(x))$, entonces $a(x) = k \cdot b(x)$ con $k \in K$. (Demostrarlo.)
- $(a(x)) \subset (b(x)) \Leftrightarrow b(x) \mid a(x)$.

II.3 Mínimo común múltiplo y máximo común divisor

La intersección

$$(a_1(x)) \cap (a_2(x)) \cap \ldots \cap (a_n(x))$$

es el conjunto de múltiplos comunes de los polinomios $a_1(x), \ldots, a_n(x)$. Este conjunto cumple las dos condiciones de ideal y, por (2.2), $(a_1(x)) \cap \ldots \cap (a_n(x)) = (m(x))$ para un cierto $m(x)$. El polinomio $m(x)$ está caracterizado por las dos propiedades siguientes:

- $m(x)$ es múltiplo común de $a_1(x), \ldots, a_n(x)$;
- cualquier otro polinomio múltiplo común de $a_1(x), \ldots, a_n(x)$ es múltiplo de $m(x)$.

Se le denomina *mínimo común múltiplo* de $a_1(x), \ldots, a_n(x)$:

$$m(x) = \text{m.c.m.}[a_1(x), \ldots, a_n(x)].$$

¡Atención!:

Observemos que $k \cdot m(x)$ (donde $0 \neq k \in K$) es también mínimo común múltiplo de $a_1(x), \ldots, a_n(x)$.

Consideremos ahora la unión: $(a_1(x)) \cup \ldots \cup (a_n(x))$. Este conjunto no es en general un ideal. Consideremos el menor de los ideales que contienen a $(a_1(x)) \cup \ldots \cup (a_n(x))$. Éste debe contener todas las sumas de múltiplos de $a_1(x), \ldots, a_n(x)$. Con ellas ya es suficiente, dado que

$$\{a_1(x) \cdot c_1(x) + \ldots + a_n(x) \cdot c_n(x) \mid c_1(x), \ldots, c_n(x) \in K[x]\}$$

es ya un ideal de $K[x]$, que designaremos por

$$(a_1(x), \ldots, a_n(x)).$$

Por (2.2), $(a_1(x), \ldots, a_n(x)) = (d(x))$ para un polinomio $d(x)$ conveniente. Este polinomio $d(x)$ está caracterizado por:

- $d(x)$ es un divisor común de $a_1(x), \ldots, a_n(x)$, ya que

$$a_i(x) = a(x) \cdot 0 + \ldots + a_i(x) \cdot 1 + \ldots + a_n(x) \cdot 0 \in (d(x)).$$

- Todo divisor común $D(x)$ de $a_1(x), \ldots, a_n(x)$ divide a $d(x)$. En efecto, $D(x) \mid a_i(x)$ para $i = 1, \ldots, n \Rightarrow (a_i(x)) \subset (D(x))$ para $i = 1, \ldots, n \Rightarrow (d(x)) = (a_1(x), \ldots, a_n(x)) \subset (D(x))$; es decir, $D(x) \mid d(x)$.

Diremos que $d(x)$ es el *máximo común divisor* de $a_1(x), \ldots, a_n(x)$:

$$d(x) = \text{m.c.d.}[a_1(x), \ldots, a_n(x)].$$

Observaciones:

1. Si $0 \neq k \in K$, $k \cdot d(x)$ es también m.c.d.$[a_1(x), \ldots, a_n(x)]$.
2. $d(x) = a_1(x) \cdot r_1(x) + \ldots + a_n(x) \cdot r_n(x)$.

La siguiente proposición nos proporciona un método práctico para calcular el máximo común divisor.

Proposición 3.1 *Si* $a(x) = b(x) \cdot q(x) + r(x)$ *es la división entera de* $a(x)$ *por* $b(x) \neq 0$, *entonces* $(a(x), b(x)) = (b(x), r(x))$.

DEMOSTRACIÓN: Comparar con (I.2.1) y adaptar aquella demostración al caso de polinomios. □

Apliquemos ahora esta proposición hasta que el resto obtenido sea 0:

$$\begin{aligned}
a(x) &= b(x) \cdot q(x) + r(x), & (a(x), b(x)) &= (b(x), r(x)),\\
b(x) &= r(x) \cdot q_1(x) + r_1(x), & (b(x), r(x)) &= (r(x), r_1(x)),\\
&\ldots & &\ldots\\
r_{i-1}(x) &= r_i(x) \cdot q_{i+1}(x) + r_{i+1}(x), & (r_{i-1}(x), r_i(x)) &= (r_i(x), r_{i+1}(x)),\\
r_i(x) &= r_{i+1}(x) \cdot q_{i+2}(x) + 0, & (r_i(x), r_{i+1}(x)) &= (r_{i+1}(x), 0) = (r_{i+1}(x)).
\end{aligned}$$

Observemos que $\operatorname{gr} b(x) > \operatorname{gr} r(x) > \operatorname{gr} r_1(x) > \ldots$ y, por tanto, siempre llega un momento en que el resto es 0. Tenemos pues

$$(a(x), b(x)) = (r_{i+1}(x));$$

es decir, $r_{i+1}(x) = \text{m.c.d.}(a(x), b(x))$.

Este método para hallar el m.c.d. se conoce como *algoritmo de Euclides.*

Ejercicios:

- m.c.d. $(a_1(x), \ldots, a_n(x)) = \text{m.c.d.}\,(\text{m.c.d.}(a_1(x), \ldots, a_{n-1}(x)),\ a_n(x))$.
- m.c.m. $(a_1(x), \ldots, a_n(x)) = \text{m.c.m.}\,(\text{m.c.m.}(a_1(x), \ldots, a_{n-1}(x)),\ a_n(x))$.
- Si $d(x) = \text{m.c.d.}(a(x), b(x))$, hallar polinomios $r(x)$, $s(x)$ tales que

$$d(x) = a(x) \cdot r(x) + b(x) \cdot s(x).$$

El algoritmo de Euclides, juntamente con el hecho de que el producto del m.c.m. y el m.c.d. de dos polinomios coincide con el producto de los polinomios salvo factores de K (ver apartado 4), nos proporciona también una manera de calcular el m.c.m..

II.4 Polinomios irreducibles y polinomios primos entre sí

Dos polinomios $a(x)$, $b(x)$ son *primos entre sí* cuando m.c.d.$(a(x), b(x)) = 1$.

Ejemplos:

1. m.c.d.$(x^2 - 1, x^2 + x - 6) = 1$. Observemos que
$$1 = (x^2 + x - 6)\left(-\frac{1}{24}x - \frac{5}{24}\right) + (x^2 - 1)\left(\frac{1}{4} + \frac{1}{24}x\right).$$
2. Si $d(x) = \text{m.c.d.}(a(x), b(x))$, $a(x) = d(x) \cdot r(x)$ y $b(x) = d(x) \cdot s(x)$ entonces m.c.d.$(r(x), s(x)) = 1$.

Proposición 4.1 (Teorema de Euclides) *Si* m.c.d.$(a(x), b(x)) = 1$ *y* $a(x) \mid b(x) \cdot c(x)$, *entonces* $a(x) \mid c(x)$.

DEMOSTRACIÓN: Si m.c.d.$(a(x), b(x)) = 1$,
$$1 = a(x) \cdot r(x) + b(x) \cdot s(x),$$
de donde
$$c(x) = a(x) \cdot c(x) \cdot r(x) + b(x) \cdot c(x) \cdot s(x).$$
Pero $a(x)$ divide a los dos sumandos y, por tanto, $a(x) \mid c(x)$. □

Proposición 4.2 *Si* $m(x) = \text{m.c.m.}(a(x), b(x))$ *y* $d(x) = \text{m.c.d.}(a(x), b(x))$, *entonces* $m(x) \cdot d(x) = k\ a(x) \cdot b(x)$ *con* $k \in K$.

DEMOSTRACIÓN: Si $a(x) = d(x) \cdot r(x)$ y $b(x) = d(x) \cdot s(x)$, basta ver que $d(x) \cdot r(x) \cdot s(x)$ es un m.c.m.$(a(x), b(x))$. Pero $d(x) \cdot r(x) \cdot s(x)$ es claramente múltiplo de $a(x)$ y de $b(x)$. Si $M(x)$ es también un múltiplo común,
$$M(x) = a(x) \cdot c(x) = b(x) \cdot h(x),$$
de donde $d(x) \cdot r(x) \cdot c(x) = d(x) \cdot s(x) \cdot h(x)$.

Tenemos pues $r(x) \cdot c(x) = s(x) \cdot h(x)$ y, por tanto, $r(x) \mid s(x) \cdot h(x)$. Puesto que $(r(x), s(x)) = 1$, (4.1) nos dice que $r(x) \mid h(x)$; pongamos $h(x) = r(x) \cdot t(x)$. Entonces,
$$M(x) = b(x) \cdot h(x) = d(x) \cdot s(x) \cdot r(x) \cdot t(x);$$
es decir, $M(x)$ es un múltiplo de $d(x) \cdot s(x) \cdot r(x)$. □

Un polinomio $p(x)$ de grado diferente de cero se llama *irreducible* o *primo* si sus únicos divisores son k, $k \cdot p(x)$ con $k \in K$.

Proposición 4.3 *Todo polinomio* $a(x) \neq 0$ *de grado* > 0 *es producto de polinomios irreducibles.*

DEMOSTRACIÓN: Si $a(x)$ es primo, el resultado es cierto. En caso contrario, sea $p_1(x)$ un divisor de grado mínimo entre los de $a(x)$. Entonces $p_1(x)$ es primo (ya que todos sus divisores lo son también de $a(x)$). Pongamos $a(x) = p_1(x) \cdot a_1(x)$. Si $a_1(x)$ no es primo, consideremos uno de sus divisores, $p_2(x)$, de grado mínimo. Entonces $p_2(x)$ es primo y

$$a(x) = p_1(x) \cdot p_2(x) \cdot a_2(x).$$

Observemos que $\operatorname{gr} a(x) > \operatorname{gr} a_1(x) > \operatorname{gr} a_2(x) > \dots$. Llegará un momento, pues, en que

$$a(x) = p_1(x) \cdots p_{r-1}(x) \cdot a_r(x)$$

y $a_r(x)$ será primo. □

Las descomposiciones de un polinomio en factores irreducibles son, hasta cierto punto, únicas. Concretamente:

Proposición 4.4 *Si*

$$p_1(x) \cdots p_n(x) = q_1(x) \cdots q_m(x)$$

y todos los factores son polinomios irreducibles, entonces $n = m$ *y los polinomios* $\{p_i(x)\}$ *son los mismos que los* $\{q_j(x)\}$*, salvo factores del cuerpo* K*.*

DEMOSTRACIÓN: Procedemos por inducción sobre n. Si $n = 1$, claramente $m = 1$ y $p_1(x) = q_1(x)$.

Supongamos ahora que el resultado es cierto siempre que $n \leq r - 1$. Dada la expresión

$$p_1(x) \cdots p_r(x) = q_1(x) \cdots q_m(x),$$

tenemos que $p_r(x)$ divide a

$$q_1(x) \cdot (q_2(x) \cdots q_m(x)).$$

Si $p_r(x)$ no coincide con $q_1(x)$ (salvo factores de K), entonces es primo con él y, por tanto, $p_r(x)$ divide a $q_2(x) \cdots q_m(x)$.

Repitiendo el razonamiento, llegaremos hasta un $q_j(x)$ que será igual a $p_r(x)$ salvo un factor de K; $p_r(x) = q_j(x) \cdot k$, $0 \neq k \in K$. Suprimiendo este factor común, tenemos

$$p_1(x) \cdots p_{r-1}(x) = q_1(x) \cdots q_{j-1}(x) \cdot (k^{-1} q_{j+1}(x)) \cdots q_m(x)$$

y podemos aplicar la hipótesis de inducción para deducir que también

$$p_1(x), \ldots, p_{r-1}(x)$$

coinciden con los $q_i(x)$ restantes (salvo factores de K) y, en particular, que $r-1 = m-1$, de donde $r = m$. □

II.5 Ceros de un polinomio

Si $a(x) = a_0 + a_1 x + \ldots + a_n x^n$ es un polinomio de $K[x]$ y $k \in K$, denominaremos *valor de* $a(x)$ *en* k a

$$a(k) = a_0 + a_1 k + \ldots + a_n k^n \in K.$$

Observemos que el valor de $a(x) + b(x)$ en k es $a(k) + b(k)$, y el valor de $a(x) \cdot b(x)$ en k es $a(k) \cdot b(k)$.

Si $a(k) = 0$, diremos que k es un *cero* o una *raíz* de $a(x)$.

Proposición 5.1 *k es un cero del polinomio $a(x) \neq 0$ si y sólo si $a(x)$ es divisible por $x - k$.*

DEMOSTRACIÓN: Si $\operatorname{gr} a(x) = 0$, tendremos $a(x) = a_0 \neq 0$ y k no será un cero: $a(k) = a_0 \neq 0$.

Si $\operatorname{gr} a(x) \geq 1$, efectuemos la división entera por $(x - k)$. El resto deberá ser 0 o tener grado 0:

$$a(x) = (x - k) \cdot q(x) + r, \qquad r \in K.$$

Entonces $0 = a(k) = r$ y, por tanto, $x - k$ divide a $a(x)$. □

Diremos que $k \in K$ es un *cero de multiplicidad* p del polinomio $a(x) \in K[x]$ si $a(x) = (x-k)^p \cdot b(x)$ y $b(k) \neq 0$; es decir, si $a(x)$ es divisible por $(x-k)^p$ pero no lo es por $(x-k)^{p+1}$.

Corolario 5.2 *Si $\operatorname{gr} a(x) = n$, la suma de las multiplicidades de los ceros de $a(x)$ es menor o igual que n.* □

¿Puede ocurrir que dos polinomios diferentes $a(x) \neq b(x)$ tomen el mismo valor sobre todos los $k \in K$? Tendríamos un polinomio $a(x) - b(x) \neq 0$ del cual todos los $k \in K$ serían ceros. Pero (5.2) nos dice que, si K tiene suficientes elementos ($> \operatorname{gr}[a(x) - b(x)]$), esto no es posible. En particular:

Proposición 5.3 *Si K es infinito y $a(k) = b(k)$ para todo $k \in K$, entonces $a(x) = b(x)$.* □

Ejemplo:

Consideremos los polinomios $a(x) = x - 2$ y $b(x) = x^3 - 2$ con coeficientes en $\mathbf{Z}/(3)$. Como polinomios, $a(x) \neq b(x)$; ahora bien, $a(0) = -2 = b(0)$, $a(1) = -1 = b(1)$, $a(2) = 0 = b(2)$.

Acabaremos este apartado dando un criterio muy sencillo para encontrar los ceros racionales de un polinomio de $\mathbf{Q}[x]$. Consideremos

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

con

$$a_0, \ldots, a_n \in \mathbf{Q}.$$

Podemos siempre encontrar un entero $m \neq 0$ tal que $ma_0, \ldots, ma_n \in \mathbf{Z}$. El polinomio $ma(x)$ tiene los mismos ceros que $a(x)$ y sus coeficientes son enteros. El problema queda reducido, por tanto, a encontrar los ceros racionales de un polinomio con coeficientes enteros.

Consideremos, pues,

$$b(x) = b_0 + b_1 x + \ldots + b_n x^n$$

con

$$b_0, \ldots, b_n \in \mathbf{Z}.$$

Sea p/q un cero de $b(x)$ con p, q primos entre sí. De

$$b_0 + b_1 \frac{p}{q} + \ldots + b_n \frac{p^n}{q^n} = 0$$

obtenemos

$$b_0 q^n + b_1 p q^{n-1} + \ldots + b_{n-1} p^{n-1} q + b_n p^n = 0.$$

Puesto que m.c.d.$(p, q) = 1$, aplicando el teorema de Euclides resulta $p \mid b_0$ y $q \mid b_n$. Por tanto,

Proposición 5.4 *Si p/q, con p, q primos entre sí, es un cero del polinomio*

$$b(x) = b_0 + b_1 x + \ldots + b_n x^n \in \mathbf{Z}[x],$$

entonces $p \mid b_0$ y $q \mid b_n$. □

Corolario 5.5 *Si $k \in \mathbf{Z}$ es un cero de*

$$b_0 + b_1 x + \ldots + b_n x^n \in \mathbf{Z}[x],$$

entonces $k \mid b_0$. □

II.6 Polinomios irreducibles de $\mathbf{R}[x]$

El estudio de los polinomios irreducibles en $\mathbf{R}[x]$ y en $\mathbf{C}[x]$ se basa en el siguiente teorema:

Teorema 6.1 (fundamental del Álgebra) *Todo polinomio de grado* ≥ 1 *con coeficientes complejos tiene un cero.*

No daremos la demostración de este teorema, que va más allá del objetivo del libro. No obstante, haremos algunos comentarios sobre él y sacaremos consecuencias. En primer lugar cabe decir que, pese a su nombre, no se trata de un teorema "algebraico" sino de un teorema "topológico"; en otras palabras, este teorema es consecuencia de las propiedades de completitud de $\mathbf{C}$ (y de $\mathbf{R}$) y no de las propiedades de sus operaciones. Observemos también que del teorema se deduce que todo polinomio de $\mathbf{C}[x]$ de grado ≥ 1 es producto de factores lineales (de grado 1) y, por tanto,

Corolario 6.2 *Los polinomios irreducibles de* $\mathbf{C}[x]$ *son los de grado* 1.

Este corolario proporciona, de hecho, otro enunciado del teorema, ya que si $\alpha + \beta x$ es un factor lineal del polinomio $a(x)$, entonces $a(x)$ tiene el cero $-\alpha/\beta$.

Estudiaremos ahora los polinomios irreducibles de $\mathbf{R}[x]$. Todo polinomio real $a(x) = a_0 + a_1 x + \ldots + a_n x^n$ puede considerarse también como un polinomio con coeficientes complejos. En general, si $a(x) = a_0 + a_1 x + \ldots + a_n x^n$ es de $\mathbf{C}[x]$, denotaremos por $\bar{a}(x)$ el polinomio

$$\bar{a}_0 + \bar{a}_1 x + \ldots + \bar{a}_n x^n.$$

Si $z = a + bi \in \mathbf{C}$, $\bar{z} = a - bi$ indica su conjugado. Entonces $a(x)$ tiene coeficientes reales si y sólo si

$$\bar{a}(x) = a(x).$$

Por otra parte, si $z \in \mathbf{C}$,

$$\bar{a}(\bar{z}) = \bar{a}_0 + \bar{a}_1 \bar{z} + \ldots + \bar{a}_n \bar{z}^n = \overline{(a_0 + a_1 z + \ldots + a_n z^n)} = \overline{a(z)}$$

y, en particular, si z es un cero de $a(x)$, $(a(z) = 0)$, entonces $\bar{z}$ es un cero de $\bar{a}(x)$, $(\bar{a}(\bar{z}) = 0)$.

Cuando $a(x)$ tiene coeficientes reales, resulta que siempre que z sea un cero, $\bar{z}$ también lo es. Entonces, o bien $z = \bar{z}$ (es decir, z es un cero real de $a(x)$), o bien $z \neq \bar{z}$ y $a(x)$ es divisible por

$$(x - z)(x - \bar{z}) = x^2 - (z + \bar{z})x + z\bar{z},$$

que es un polinomio con coeficientes reales. Además, $x^2 - (z + \bar{z})x + z\bar{z}$ es irreducible en $\mathbf{R}[x]$, ya que en caso contrario tendría un divisor de primer grado y por tanto un cero real.

Así pues, los polinomios irreducibles de $\mathbf{R}[x]$ son de grado ≤ 2.

Nota:

En el anillo $\mathbf{Q}[x]$ se pueden encontrar polinomios irreducibles de grado tan grande como se desee.

II.7 Los anillos $K[x]/(m(x))$

Sea $m(x)$ un polinomio de $K[x]$. Diremos que dos polinomios $a(x)$ y $b(x)$ son *congruentes módulo* $m(x)$ si $a(x) - b(x) \in (m(x))$. Esto equivale a decir que los restos de las divisiones enteras de $a(x)$ y $b(x)$ por $m(x)$ son iguales (comparar con (I.4)). Escribiremos entonces

$$a(x) \equiv b(x) \quad (m(x)).$$

Esta relación es claramente de equivalencia. Designemos por $[a(x)]$ la clase de equivalencia de $a(x)$, es decir, el conjunto de polinomios congruentes con $a(x)$ módulo $m(x)$.

El conjunto de estas clases de equivalencia será denotado por

$$K[x]/(m(x))$$

y llamado *cociente* de $K[x]$ por $(m(x))$. Observemos que hay tantas clases de equivalencia como restos posibles en las divisiones enteras por $m(x)$. Estos restos son precisamente los polinomios de grado menor que el grado de $m(x)$. En otras palabras, en cada clase de equivalencia hay un polinomio de grado menor que el de $m(x)$, y solamente uno.

En el conjunto $K[x]/(m(x))$ podemos definir dos operaciones; suma:

$$[a(x)] + [b(x)] = [a(x) + b(x)]$$

y producto:

$$[a(x)] \cdot [b(x)] = [a(x) \cdot b(x)].$$

Debe comprobarse, sin embargo, que la clase suma y la clase producto no dependen de los representantes $a(x)$, $b(x)$ escogidos. Es decir, que si $a'(x)$, $b'(x)$ son otros representantes de $[a(x)]$, $[b(x)]$ respectivamente, entonces

$$[a'(x) + b'(x)] = [a(x) + b(x)], \qquad [a'(x) \cdot b'(x)] = [a(x) \cdot b(x)].$$

La comprobación se hace exactamente igual que en el caso de las clases de restos en $\mathbf{Z}$ (I.5).

$K[x]/(m(x))$ tiene, con estas operaciones, estructura de anillo conmutativo con unidad; ahora bien, este anillo posee algunas propiedades que no tenía $K[x]$. Por ejemplo,

Proposición 7.1 *Si* $(a(x), m(x)) = (1)$, *entonces* $[a(x)]$ *tiene un inverso en* $K[x]/(m(x))$. *Si* $(a(x), m(x)) = (d(x))$ *con* $\operatorname{gr} d(x) \geq 1$, *entonces* $[a(x)]$ *es un divisor de* 0 *en* $K[x]/(m(x))$.

La demostración es análoga a la de (I.5.1). □

En particular:

Corolario 7.2 *Si* $p(x) \in K[x]$ *es irreducible,* $K[x]/(p(x))$ *es un cuerpo.* □

Si $\operatorname{gr} p(x) \geq 1$,

$$\begin{array}{ccc} K & \longrightarrow & K[x]/(p(x)) \\ k & \mapsto & [k] \end{array}$$

es inyectiva y conserva la suma y el producto. Este hecho justifica que denotemos los elementos $[k]$ simplemente por k y el subconjunto de $K[x]/(p(x))$ imagen de la aplicación, por la letra K. Con esta notación, escribiremos

$$K \subset K[x]/(p(x)).$$

Cuando $p(x)$ es irreducible obtenemos, pues, un cuerpo $K[x]/(p(x))$ que "contiene" a K.

Todo polinomio $a(x)$ con coeficientes en K puede considerarse también un polinomio con coeficientes en $K[x]/(p(x))$. En particular, el polinomio

$$p(x) = p_0 + p_1 x + \ldots + p_n x^n$$

puede considerarse con coeficientes en $K[x]/(p(x))$.

Resulta, entonces, que si ponemos $\alpha = [x] \in K[x]/(p(x))$

$$p(\alpha) = p([x]) = p_0 + p_1[x] + \ldots + p_n[x^n] = [p_0 + p_1 x + \ldots + p_n x^n] = [0];$$

es decir, el polinomio $p(x)$, que era irreducible en $K[x]$, tiene un cero (y, por tanto, tiene un divisor lineal) en $K[x]/(p(x))$.

El cuerpo $K[x]/(p(x))$ se denota por $K(\alpha)$ y se llama una *extensión algebraica* de K.

Ejemplo:

Consideremos $x^2 + 1$, que es un polinomio irreducible en $\mathbf{R}[x]$. Los elementos del cuerpo $\mathbf{R}[x]/(x^2 + 1)$ tienen, cada uno, un único representante de primer grado

$$[a + bx].$$

La suma de dos clases es

$$[a + bx] + [c + dx] = [(a + c) + (b + d)x]$$

y el producto

$$[a + bx] \cdot [c + dx] = [ac + (ad + bc)x + bdx^2] = [(ac - bd) + (ad + bc)x].$$

Entonces, si "identificamos" cada $a \in \mathbf{R}$ con $[a] \in \mathbf{R}[x]/(x^2 + 1)$ y denotamos $[x]$ por i, obtenemos que los elementos del cociente son de la forma

$$[a + bx] = a + b[x] = a + bi.$$

Con esta notación,

$$i^2 = [x^2] = [-1] = -1$$

y las dos operaciones se expresan así:

$$\begin{aligned} (a + bi) + (c + di) &= (a + c) + (b + d)i \\ (a + bi) \cdot (c + di) &= (ac - bd) + (ad + bc)i. \end{aligned}$$

Existe, por tanto, una correspondencia biyectiva entre el cuerpo $\mathbf{R}[x]/(x^2 + 1)$ y el cuerpo $\mathbf{C}$ de los números complejos, que conserva las operaciones. Podemos decir, pues, que el cuerpo $\mathbf{R}[x]/(x^2 + 1)$ no es otra cosa que el cuerpo $\mathbf{C}$ de los números complejos. Con más precisión, se dice que $\mathbf{R}[x]/(x^2 + 1)$ y $\mathbf{C}$ son dos cuerpos *isomorfos*.

Ejemplo:

Consideremos

$$K = \mathbf{Q}[x]/(x^2 - 2);$$

$x^2 - 2$ es irreducible en $\mathbf{Q}[x]$ y, por tanto, K es un cuerpo que contiene a $\mathbf{Q}$. Todo elemento de $\mathbf{Q}$ tiene un representante (y sólo uno) de primer grado $ax + b$.

Las dos operaciones son

$$[ax + b] + [cx + d] = [(a + c)x + (b + d)]$$

$$[ax + b] \cdot [cx + d] = [acx^2 + (ad + bc)x + bd] = [(ad + bc)x + 2ac + bd].$$

El elemento $\alpha = [x] \in K$ cumple $\alpha^2 = 2$ y es un cero del polinomio $X^2 - 2 \in K[X]$. La extensión algebraica $K = \mathbf{Q}(\alpha)$ es isomorfa a su imagen por la aplicación

$$\begin{array}{ccc} \mathbf{Q}(\alpha) & \longrightarrow & \mathbf{R} \\ a\alpha + b & \mapsto & a\sqrt{2} + b. \end{array}$$

II.8 Nota histórica

El simbolismo usado en los polinomios y ecuaciones se ha ido elaborando a lo largo de la historia y no tomó su forma actual hasta principios del siglo 18. Parece ser que los signos "+" y "−" fueron usados por primera vez por J. Widman en el siglo 16 desplazando las letras "p" y "m", abreviaciones de "plus" y "minus". François Viète (1540–1603), un parlamentario que dedicaba su tiempo libre a las matemáticas, dio un gran impulso al álgebra simbólica, utilizando letras (las primeras del abecedario) para las variables. Escribía nuestra ecuación "$5BA^2 - 2CA + A^3 = D$" como "B5 in A quadratum − C plano 2 in A + A cubum aequator D solido" (y eso fue un gran avance respecto a sus predecesores). La obra de René Descartes (1596–1650) contiene ya la notación actual con dos variantes menores: "xx" por "x^2" y "$\propto$" por "=". En la resolución de ecuaciones, y especialmente en lo que se refiere a los apartados 5 y 6, hay que mencionar a Paolo Ruffini (1765–1822) y Carl Friedrich Gauss (1777–1855), el "príncipe de las matemáticas" según la inscripción que el rey George V de Hannover ordenó grabar, quien demostró el teorema fundamental del álgebra, y proporcionó cuatro demostraciones de él en su búsqueda de una que fuera puramente algebraica.

Augustin-Louis Cauchy (1789–1857) fue el primero en observar, en 1847, que los números complejos se pueden considerar como clases de equivalencia de $\mathbf{R}[x]$ módulo $x^2 + 1$. Resulta curioso que, pese a que desde Gauss ya se trabaja con relaciones de equivalencia en $\mathbf{Z}$ y en $K[x]$, pasa casi todo el siglo 19 antes de que se introduzca sistemáticamente el conjunto cociente correspondiente.

Dos matemáticos destacan por sus aportaciones iniciales a la teoría de cuerpos, considerando extensiones de un cuerpo por una raíz de un polinomio: Niels Henrik Abel (1802–1829) y Évariste Galois (1811–1831), ambos estudiando la resolubilidad de las ecuaciones de grado ≥ 5. Tanto Abel como Galois murieron muy jóvenes y trágicamente; uno de ellos tuberculoso y en la miseria, el otro en un duelo.

II.9 Ejercicios

1. Calcular el máximo común divisor $d(x)$ de los polinomios $p(x) = x^5 - 5x^3 + 4x$ y $q(x) = x^3 - 2x^2 - 5x + 6$. Encontrar dos polinomios $a(x)$ y $b(x)$ de manera que $p(x) \cdot a(x) + q(x) \cdot b(x) = d(x)$.

2. Si $p, q \in \mathbf{R}[x]$ son polinomios tales que $(p, q) = (1)$, demostrar que $(p + q, p \cdot q) = (1)$.

3. Factorizar como producto de polinomios irreducibles:

   a) $x^3 - 2$, $x^{12} - 4$, $x^p - 1$ en $\mathbf{Q}[x]$, $\mathbf{R}[x]$ y $\mathbf{C}[x]$.

   b) $x^p - x$ en $\mathbf{Z}/(p)\,[x]$, $x^3 + 2x^2 + 5x + 1$ en $\mathbf{Z}/(7)\,[x]$.

4. Determinar un polinomio $p(x)$ de grado mínimo tal que $x^2 + 1 \mid p(x)$ y $x^3 + 1 \mid p(x) - 1$.

5. Si $p(x) \in \mathbf{Z}[x]$ y $p(r/s) = 0$ con $(r, s) = 1$, demostrar que $r - s \mid p(1)$ y $r + s \mid p(-1)$.

6. Calcular todos los ceros racionales de $20x^3 - 56x^2 - 33x + 9$ y de $12x^5 - 17x^4 + 7x^3 - 5x^2 - 22x - 5$.

7. Determinar un polinomio $p(x)$ de grado 5 tal que $p(0) = p(1) = p(2) = p(3) = p(4) = 1$.

8. Descomponer $x^4 + a^2 \in \mathbf{R}[x]$ en factores irreducibles.

9. Descomponer $(x + 1)^n + (x - 1)^n \in \mathbf{C}[x]$ en factores lineales.

10. Demostrar que $2 + \sqrt[3]{3}$, $\sqrt{2} + \sqrt{3}$, $\sqrt[3]{2} + \sqrt{3}$ y $\sqrt{2} + \sqrt{3} + \sqrt{5}$ son cada uno de ellos cero de un polinomio de $\mathbf{Z}[x]$. Determinar esos polinomios.

11. Racionalizar las expresiones

$$\frac{1}{\sqrt[4]{8} - \sqrt[4]{4} + \sqrt[4]{2} - 3} \quad \text{y} \quad \frac{1}{\sqrt[3]{7} + \sqrt[4]{7}}.$$

12. Dado el polinomio $p(x) = 3x^3 + 5x^2 + 5x + 2$,

   a) encontrar todos los ceros complejos de $p(x)$;

   b) determinar los divisores de cero de los anillos $\mathbf{C}[x]/(p(x))$ y $\mathbf{R}[x]/(p(x))$.

13. Resolver en $\mathbf{R}[x]/(x^2 + 2x + 1)$ la ecuación $z^2 + z + \dfrac{1}{4} = 0$.

14. ¿Para qué valores $a \in \mathbf{C}$ la ecuación $z^2 + z + a = 0$ en el anillo $A = \mathbf{C}[x]/(x^3)$ tiene infinitas soluciones?

15. Para cada elemento $a \in A = \mathbf{R}[x]/(x^2 + x)$, determinar cuántas soluciones tiene la ecuación $z^2 = a$. Representar el anillo A sobre el plano y dividirlo en regiones según el número de soluciones de la ecuación anterior.

II.10 Ejercicios para programar

16. Resolución en $\mathbf{Z}/(p)$ de la ecuación de segundo grado

$$ax^2 + bx + c = 0.$$

(Indicación: utilizar como subprogramas accesorios los ejercicios I.20 y I.21.)

17. División entera de dos polinomios de $\mathbf{Z}/(p)\,[x]$. (Indicación: utilizar como subprograma accesorio el ejercicio I.20.)

18. Factorización de un polinomio de $\mathbf{Z}/(p)\,[x]$ en polinomios irreducibles. (Indicación: utilizar como subprogramas accesorios los ejercicios I.20, I.21 dentro de (II.16) y (II.17).)

19. Cálculo de los ceros racionales de un polinomio de $\mathbf{Q}[x]$. (Indicación: utilizar la proposición II.5.4 y el ejercicio II.5.)

Capítulo III

Grupos

III.1 Definición y ejemplos

Un *grupo* es un conjunto G junto con una operación $\cdot$ que cumple las propiedades:

- Asociativa: $g \cdot (g' \cdot g'') = (g \cdot g') \cdot g'' \qquad \forall g, g', g'' \in G.$
- Existe un elemento e, al que llamaremos *elemento neutro*, tal que

$$g \cdot e = g = e \cdot g \qquad \forall g \in G.$$

- Para cada $g \in G$ existe un elemento, al que denominaremos el *inverso* de g y denotaremos por g^{-1}, tal que

$$g \cdot g^{-1} = e = g^{-1} \cdot g.$$

Si se cumple también la propiedad conmutativa:

$$g \cdot g' = g' \cdot g \qquad \forall g, g' \in G,$$

diremos que el grupo es *conmutativo* o *abeliano*. En este caso, la operación se denota a menudo por $+$, el elemento neutro por 0 (y se denomina *cero*) y el elemento inverso por $-g$ (y se denomina el *opuesto* de g).

Cuando indicamos la operación por $\cdot$ (notación multiplicativa), el elemento neutro se acostumbra a llamar *unidad* y a escribir 1. Con esta notación multiplicativa, es costumbre suprimir el punto que indica la operación y escribir simplemente gg' para indicar $g \cdot g'$.

Ejemplos:

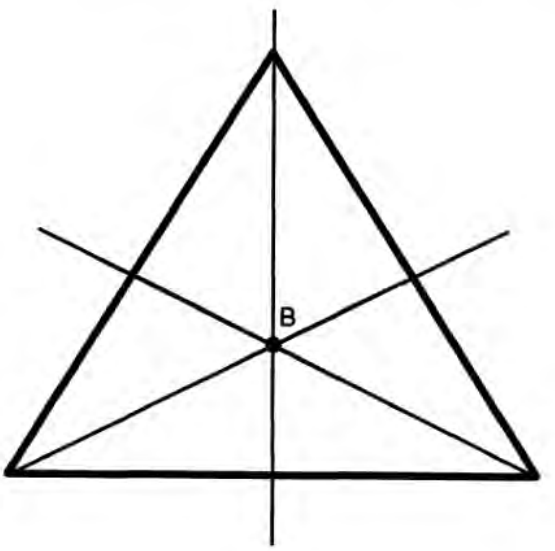
1. Los números enteros $\mathbf{Z}$ con la suma forman un grupo conmutativo. Lo mismo vale para los racionales $\mathbf{Q}$ y los reales $\mathbf{R}$. Los números naturales $\mathbf{N} = \{1, 2, \ldots\}$ no son un grupo con la suma.

   Los números racionales no nulos, $\mathbf{Q} - \{0\}$, con el producto forman un grupo conmutativo. Lo mismo vale para $\mathbf{R} - \{0\}$. Ni $\mathbf{Z} - \{0\}$ ni $\mathbf{N}$ son grupos con el producto.

2. Los números complejos $\mathbf{C}$ con la suma son un grupo conmutativo. $\mathbf{C} - \{0\}$ con el producto es un grupo conmutativo.

   $S^1 = \{z \in \mathbf{C} \mid |z| = 1\}$ con el producto es un grupo conmutativo.

3. Todos los ejemplos anteriores son grupos conmutativos. Los ejemplos más sencillos de grupos *no* conmutativos surgen en la geometría al estudiar determinados conjuntos de movimientos. Así, por ejemplo, el conjunto de movimientos del plano que dejan fijo un triángulo equilátero está formado por tres simetrías respecto a ejes que pasan por un vértice y el punto medio del lado opuesto, los giros de 120° y 240° alrededor del baricentro, y la identidad o giro de 0°. En estos ejemplos geométricos, la operación es *la composición*: la composición de los movimientos g' y g es el movimiento $g \circ g'$ que resulta de efectuar sucesivamente los movimientos g' y g. (¡Atención al orden!) Esta operación no es conmutativa.



Esos grupos de movimientos aparecerán de manera natural al estudiar la geometría. A continuación, en el apartado 2, vamos a ocuparnos de otros grupos no conmutativos sencillos: los grupos de permutaciones.

III.2 Permutaciones

Sea $A = \{a_1, \ldots, a_n\}$ un conjunto con n elementos. Una *permutación de A* es una aplicación biyectiva

$$\sigma : A \longrightarrow A.$$

La composición de permutaciones es, claramente, una permutación. Además, se cumplen las propiedades siguientes:

- Asociativa: $\sigma \circ (\rho \circ \tau) = (\sigma \circ \rho) \circ \tau \qquad \forall \sigma, \rho, \tau.$
- Existe una permutación I tal que: $\sigma \circ I = \sigma = I \circ \sigma \qquad \forall \sigma.$
  I es la aplicación identidad: $I(a) = a \qquad \forall a \in A.$
- Para toda permutación σ existe una permutación σ^{-1} tal que $\sigma \circ \sigma^{-1} = I = \sigma^{-1} \circ \sigma$. Esta permutación σ^{-1} es la aplicación inversa de σ.

El conjunto $\mathcal{S}_A$ de las permutaciones de A con la composición es, pues, un grupo. Llamaremos también *producto* a la composición y escribiremos, a veces, $\sigma\tau$ por $\sigma \circ \tau$.

Para simplificar la notación, supondremos desde ahora, si no indicamos lo contrario, que $A = \{1, \ldots, n\}$. El conjunto de permutaciones de $\{1, \ldots, n\}$ será designado por $\mathcal{S}_n$. Si A es cualquier conjunto con n elementos, las propiedades de $\mathcal{S}_A$ son exactamente las mismas que las de $\mathcal{S}_n$ (véase el apartado 4).

Para dar una permutación concreta σ, hemos de especificar cuáles son las imágenes de cada uno de los elementos $1, 2, \ldots, n$. Una manera cómoda de hacerlo es escribir estos elementos en fila y debajo de cada uno de ellos su imagen:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Proposición 2.1 *Si $n \geq 3$, $\mathcal{S}_n$ no es conmutativo.*

DEMOSTRACIÓN: En efecto, se tiene

$$\begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 3 & 2 & 1 & 4 \ldots n \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 1 & 3 & 2 & 4 \ldots n \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 3 & 1 & 2 & 4 \ldots n \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 1 & 3 & 2 & 4 \ldots n \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 3 & 2 & 1 & 4 \ldots n \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \ldots n \\ 2 & 3 & 1 & 4 \ldots n \end{pmatrix}. \quad \square$$

Para $n = 2$, $\mathcal{S}_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$ es un grupo conmutativo.

Un elemento j se denomina *fijo* por una permutación σ si $\sigma(j) = j$. Si j no es fijo, formamos la sucesión

$$j, \sigma(j), \sigma^2(j), \ldots, \sigma^r(j), \ldots$$

donde $\sigma^2 = \sigma \circ \sigma$ y, en general, $\sigma^r = \sigma \circ \sigma^{r-1}$. Dado que $\{1, 2, \ldots, n\}$ es finito, en algún momento un elemento $\sigma^k(j)$ coincidirá con uno de los anteriores. El primer elemento que vuelve a aparecer es precisamente j: en efecto, si $\sigma^k(j) = \sigma^h(j)$ con $h < k$, por ser σ biyectiva,

$$\sigma^k(j) = \sigma^h(j) \Rightarrow \sigma^{k-1}(j) = \sigma^{h-1}(j) \Rightarrow \ldots \Rightarrow \sigma^{k-h}(j) = j\ ;$$

es decir, j ya habría salido.

Sean, pues,

$$j, \sigma(j), \ldots, \sigma^{r-1}(j)$$

diferentes y $\sigma^r(j) = j$. Diremos que la permutación σ es un *ciclo de orden r* si deja fijos todos los elementos que no aparecen en la sucesión anterior. Escribiremos entonces

$$\sigma = (j, \sigma(j), \sigma^2(j), \ldots, \sigma^{r-1}(j)).$$

Ejemplo:

En $\mathcal{S}_5$,

$$(2,3,1,5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 4 & 2 \end{pmatrix}.$$

En caso de que la permutación σ no sea un ciclo, consideremos un j_1 no fijo por σ y diferente de $j, \sigma(j), \ldots, \sigma^{r-1}(j)$. Sean

$$j_1, \sigma(j_1), \ldots, \sigma^{r_1-1}(j_1)$$

diferentes y $\sigma^{r_1}(j_1) = j_1$. Un momento de reflexión nos convencerá de que ninguno de esos elementos había aparecido en la sucesión $j, \sigma(j), \ldots$. Repitamos este proceso tantas veces como sea necesario hasta agotar todos los elementos no fijos por σ. En cada paso tomemos un elemento j_m no fijo y que no haya salido antes y formemos la sucesión

$$j_m, \sigma(j_m), \ldots, \sigma^{r_m-1}(j_m); \quad \sigma^{r_m}(j_m) = j_m.$$

Sus elementos son todos diferentes de los que hemos obtenido con anterioridad. Supongamos que, después de efectuar este paso, no queda ya ningún otro elemento no fijo. Entonces σ es producto de ciclos:

$$(j_m, \sigma(j_m), \ldots, \sigma^{r_m-1}(j_m)) \cdots (j_1, \sigma(j_1), \ldots, \sigma^{r_1-1}(j_1))(j, \sigma(j), \ldots, \sigma^{r-1}(j))$$

Observemos que estos ciclos conmutan entre sí, ya que afectan a elementos distintos. De esta forma queda demostrada la proposición siguiente:

Proposición 2.2 *Toda permutación es producto de ciclos.* □

Los ciclos de orden 2 se llaman *trasposiciones.*

Proposición 2.3 *Todo ciclo es producto de trasposiciones.*

DEMOSTRACIÓN:

$$(a_1, a_2, \ldots, a_m) = (a_1, a_2)(a_2, a_3)\cdots(a_{m-1}, a_m). \square$$

Ejemplos:

1. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 7 & 5 & 2 & 6 & 3 \end{pmatrix} = (1,4,5,2)(3,7) = (1,4)(4,5)(5,2)(3,7).$

2. $I = (1,2)(1,2) = (3,4)(2,3)(1,2)(2,4)(3,2)(1,4).$

La identidad I, y por tanto cualquier permutación, se puede expresar de muchas maneras como producto de trasposiciones. Vamos a ver, no obstante, que el número de trasposiciones en tales productos tiene siempre la misma paridad.

Proposición 2.4 *La permutación identidad no se puede expresar como producto de un número impar de trasposiciones.*

DEMOSTRACIÓN: La demostración que vamos a dar se basa en un hecho aparentemente anecdótico: si en la expresión del producto

$$P = \prod_{i,j}(j - i)$$

donde $1 \leq i < j \leq n$, $\quad i, j \in \{1, 2, \ldots, n\}$, permutamos las i, j según una trasposición, obtenemos la misma expresión con signo contrario. La explicación es la siguiente. Si σ es una permutación de $\{1, 2, \ldots, n\}$, escribiremos

$$\sigma P = \prod_{i,j}(\sigma(j) - \sigma(i)).$$

En caso de que $\sigma = (h, k)$, $h < k$, ¿cuáles son los factores de σP? Observemos que

- si i, j son diferentes de h, k, $\quad \sigma(j) - \sigma(i) = j - i$;

- si $i < h < k$, el factor $h - i$ de P pasa a ser $k - i$ en σP, el factor $k - i$ de P pasa a ser $h - i$ en σP; es decir, el único cambio en este caso es un cambio en la posición de los factores;
- si $h < k < j$, el factor $j - h$ de P pasa a ser $j - k$ en σP, el factor $j - k$ de P pasa a ser $j - h$ en σP; como antes, solamente ha habido un cambio de posición;
- si $h < i < k$, el factor $i - h$ de P pasa a ser $i - k$ en σP, el factor $k - i$ de P pasa a ser $h - i$ en σP; ahora, el cambio es de posición y de signo; el signo cambia, sin embargo, dos veces y, por tanto, no afecta al producto. Finalmente,
- si $i = h < k = j$, el factor $k - h$ de P pasa a ser $h - k$ en σP. Este es el único cambio de signo que afecta al producto.

Obtenemos, pues,

$$\sigma P = -P.$$

Supongamos ahora que

$$I = \tau_n \circ \ldots \circ \tau_2 \circ \tau_1,$$

donde las τ_i son trasposiciones. Apliquemos a P sucesivamente las trasposiciones $\tau_1, \tau_2, \ldots, \tau_n$. Obtendremos $(-1)^n P$. Por otro lado, aplicar $\tau_1, \ldots, \tau_n$ equivale a aplicar la identidad y, por tanto, el resultado ha de ser P. Es decir, $(-1)^n P = P$, de donde resulta que n es par. □

Corolario 2.5 *Si $\sigma = \tau_p \circ \ldots \circ \tau_1 = \rho_q \circ \ldots \circ \rho_1$ son dos descomposiciones de la permutación σ como producto de trasposiciones, entonces p y q tienen la misma paridad.*

DEMOSTRACIÓN: Multiplicando los dos productos por ρ_1 a la derecha y teniendo en cuenta que $\rho_1 \circ \rho_1 = I$, obtenemos

$$\tau_p \circ \ldots \circ \tau_1 \circ \rho_1 = \rho_q \circ \ldots \circ \rho_2.$$

Multipliquemos a la derecha por $\rho_2, \ldots, \rho_q$ sucesivamente; obtenemos

$$\tau_p \circ \ldots \circ \tau_1 \circ \rho_1 \circ \ldots \circ \rho_q = I.$$

Entonces (2.4) nos dice que $p + q$ es par y, por tanto, p y q son ambos pares o ambos impares. □

Una permutación se llama *par* si se descompone en un número par de trasposiciones; una permutación se llama *impar* si se descompone en un número impar.

El producto de trasposiciones pares e impares sigue la regla de los signos: el producto de dos permutaciones pares o de dos impares es par; el producto de una permutación par y una impar es impar. Este hecho motiva la asignación a las permutaciones pares del signo "+" y a las permutaciones impares del signo "−". Denominaremos *aplicación signo* a la aplicación

$$\varepsilon : \mathcal{S}_n \longrightarrow \{+1, -1\}$$

tal que

$$\varepsilon(\sigma) = +1 \quad \text{si } \sigma \text{ es par,}$$
$$\varepsilon(\sigma) = -1 \quad \text{si } \sigma \text{ es impar.}$$

Se cumple

$$\varepsilon(I) = 1, \quad \varepsilon(\sigma \circ \tau) = \varepsilon(\sigma) \cdot \varepsilon(\tau).$$

En particular, por (2.3), si $(a_1, \ldots, a_m)$ es un ciclo de orden m,

$$\varepsilon(a_1, \ldots, a_m) = (-1)^{m-1}.$$

III.3 Subgrupos

Sea S un subconjunto no vacío de un grupo G. Si se cumple que

(1) para todo par $g, g' \in S$, $gg' \in S$,

entonces la operación de G da lugar a una operación en S, que denominaremos la "operación inducida" por la de G. Nos interesan los subconjuntos S de G que cumplen (1) y que con la operación inducida son a su vez un grupo. A esos subconjuntos los llamaremos "subgrupos".

Supongamos, pues, que S cumple (1) y tiene, por tanto, una operación inducida. Esa operación será automáticamente asociativa (por serlo la de G); si tiene un elemento neutro $e' \in S$, entonces

$$e'g = g \qquad \forall g \in S;$$

multiplicando a la derecha por el inverso de g (en G), obtenemos $e' = e$. Es decir, si S tiene elemento neutro, éste debe ser el mismo elemento neutro e de G. De manera parecida se ve que, si un elemento g de S tiene inverso por la operación inducida en S, éste debe coincidir con el inverso g^{-1} que g tiene en G. Por tanto, las condiciones que debe cumplir S para ser un grupo son:

- $e \in S$.
- $g \in S \Rightarrow g^{-1} \in S$.

La primera de estas dos condiciones es consecuencia de la segunda y de (1). Hemos justificado así la definición siguiente de subgrupo: diremos que un subconjunto S, no vacío, de un grupo G es un *subgrupo de* G, si cumple

1. $g, g' \in S \Rightarrow gg' \in S$,
2. $g \in S \Rightarrow g^{-1} \in S$.

De hecho, estas dos condiciones se pueden sintetizar en una:

Proposición 3.1 *Un subconjunto $S \neq \emptyset$ de un grupo G es un subgrupo de G si y sólo si cumple*

$$g', g \in S \Rightarrow g'g^{-1} \in S.$$

DEMOSTRACIÓN: Si S es subgrupo,

$$g', g \in S \Rightarrow g', g^{-1} \in S \Rightarrow g'g^{-1} \in S.$$

Si S cumple la condición del enunciado y $g \in S$ es arbitrario, $e = gg^{-1} \in S$. Entonces, para todo $g \in S$, $g^{-1} = eg^{-1} \in S$. Esto demuestra 2, que, a su vez, nos permite demostrar 1: $g', g \in S \Rightarrow g', g^{-1} \in S \Rightarrow g'g \in S$, ya que $(g^{-1})^{-1} = g$. □

Ejemplos:

1. $S^1 = \{z \in \mathbf{C} \mid |z| = 1\}$ es un subgrupo de $\mathbf{C} - \{0\}$ con el producto.
2. $\mathbf{Z}$ con $+$ es un grupo. Si S es un subgrupo de $\mathbf{Z}$, la primera condición de la definición de subgrupo nos dice que si $a, b \in S$, entonces $a+b \in S$. Si $a \in S$ y $n \in \mathbf{Z}$, entonces na es 0, o suma de varias a, o suma de varias $-a$. En cualquier caso, $na \in S$. Así pues, S es un ideal de $\mathbf{Z}$ y, por (I.1.2), es de la forma $S = (m)$.
3. Estudiemos los subgrupos del grupo de permutaciones $\mathcal{S}_3$. Los elementos de $\mathcal{S}_3$ son

   $$I,\ A = (1,2,3),\ B = (1,3,2),\ \tau_1 = (2,3),\ \tau_2 = (1,3),\ \tau_3 = (1,2).$$

   Todo subgrupo debe contener I. Por tanto, el único subgrupo formado por un solo elemento es $\{I\}$. Los conjuntos

   $$\{I, \tau_1\},\ \{I, \tau_2\} \text{ y } \{I, \tau_3\}$$

   son subgrupos. En cambio, ni $\{I, A\}$ ni $\{I, B\}$ lo son; de hecho, si un subgrupo contiene a A, debe también contener a $A^2 = B$ y si un subgrupo contiene a B, debe contener a $B^2 = A$. El conjunto

   $$\{I, A, B\}$$

es un subgrupo. Ningún otro subconjunto propio de $\mathcal{S}_3$ es subgrupo. En otras palabras, si un subgrupo S de $\mathcal{S}_3$ contiene, aparte de I, dos permutaciones que no sean A y B, entonces $S = \mathcal{S}_3$. Consideremos por ejemplo el caso en que $\tau_1 \in S$ y $A \in S$; entonces

$$B = A^2 \in S, \quad \tau_2 = \tau_1 A \in S, \quad \tau_3 = A\tau_1 \in S$$

y S contiene todos los elementos. De manera parecida se comprueban los otros casos.

Dado un subconjunto S de un grupo G, denominaremos *subgrupo generado por* S al "menor" subgrupo de G que contiene a S. Lo designaremos por (S). Aquí, "menor" significa que (S) está contenido en cualquier otro subgrupo que contenga a S. Debemos preguntarnos, no obstante, si existe siempre (S) y, en tal caso, cómo se forma.

Observemos, en primer lugar, que (S) debe contener los elementos de S, los inversos de los elementos de S y los productos de unos y otros. No es necesario añadir más elementos; el conjunto de productos

$$\{s_1 \cdots s_n \mid s_i \in S \text{ o } s_i^{-1} \in S\}$$

es ya un subgrupo. (¡Demostrarlo!) Este conjunto es, pues, (S).

Ejemplos:

1. El conjunto $(a_1, \ldots, a_n)$ construido en (I.2) es precisamente el subgrupo generado por $\{a_1, \ldots, a_n\}$.
2. El subgrupo de $\mathcal{S}_3$ generado por $\{I, \tau_1, A\}$ es todo $\mathcal{S}_3$.

III.4 Homomorfismos

Sean G y G' dos grupos. Una aplicación

$$f : G \longrightarrow G'$$

se llama un *homomorfismo* (o *morfismo*) *de grupos* si para todo par de elementos $g_1, g_2 \in G$ se cumple $f(g_1 \cdot g_2) = f(g_1) \cdot f(g_2)$.

Ejemplos:

1. Consideremos el grupo de los números reales con la suma, $(\mathbf{R}, +)$, y el grupo de los números reales positivos con el producto, $(\mathbf{R}^+, \cdot)$. La aplicación

$$\begin{array}{ccc} \mathbf{R} & \longrightarrow & \mathbf{R}^+ \\ x & \longmapsto & e^x \end{array}$$

es un homomorfismo de grupos, ya que $e^{x+y} = e^x \cdot e^y$.

2. La aplicación signo definida en el apartado 2

$$\varepsilon : \mathcal{S}_n \longrightarrow \{+1, -1\}$$

es un homomorfismo.

Proposición 4.1 *Sea $f : G \longrightarrow G'$ un homomorfismo de grupos. Sean e y e' los elementos neutros de G y G' respectivamente. Entonces,*

a) $f(e) = e'$,

b) $f(g^{-1}) = (f(g))^{-1}, \quad \forall g \in G.$

DEMOSTRACIÓN: (a) Si $g \in G, tenemos f(g) = f(ge) = f(g)f(e)$, de donde $f(e) = e'$.

(b) $f(g)f(g^{-1}) = f(gg^{-1}) = f(e) = e'$, de donde $f(g^{-1}) = (f(g))^{-1}$. □

Proposición 4.2 *Si $f : G \longrightarrow G'$ y $h : G' \longrightarrow G''$ son dos homomorfismos de grupos, entonces $h \circ f : G \longrightarrow G''$ es también un homomorfismo.*

Demostrarlo. □

Un homomorfismo inyectivo se llama un *monomorfismo*; un morfismo exhaustivo se llama un *epimorfismo*; un morfismo biyectivo se llama un *isomorfismo*; si $f : G \longrightarrow G'$ es un isomorfismo, diremos que G y G' son *isomorfos* y escribiremos $G \cong G'$.

Dos grupos isomorfos tienen las mismas propiedades, "los mismos subgrupos", etc.

Ejercicio:

Si A y B son dos conjuntos de n elementos, demostrar que $\mathcal{S}_A$ y $\mathcal{S}_B$ son isomorfos.

Denominaremos *núcleo* de un homomorfismo $f : G \longrightarrow G'$ al conjunto

$$\operatorname{Nuc} f = \{g \in G \mid f(g) = e'\}.$$

Denominaremos *imagen* de un homomorfismo f al conjunto

$$\operatorname{Im} f = \{g' \in G' \mid \text{ existe } g \in G \text{ tal que } f(g) = g'\}.$$

Es fácil comprobar que $\operatorname{Nuc} f$ es un subgrupo de G y que $\operatorname{Im} f$ es un subgrupo de G'.

Proposición 4.3 *Sea $f : G \longrightarrow G'$ un homomorfismo de grupos.*

a) f es inyectiva si y sólo si $\operatorname{Nuc} f = \{e\}$.

b) f es exhaustiva si y sólo si $\operatorname{Im} f = G'$.

DEMOSTRACIÓN: La segunda afirmación no es otra cosa que la definición de exhaustividad. Demostremos, pues, la primera.

Si $g \in \operatorname{Nuc} f$, entonces $f(g) = e' = f(e)$ (por (4.1)); por ser f inyectiva, esto implica que $g = e$.

Recíprocamente, supongamos que $f(g_1) = f(g_2)$; entonces

$$f(g_1 g_2^{-1}) = f(g_1) f(g_2)^{-1} = e',$$

de donde $g_1 g_2^{-1} \in \operatorname{Nuc} f = \{e\}$ y, por tanto, $g_1 g_2^{-1} = e$; es decir, $g_1 = g_2$. □

III.5 Grupo cociente. Subgrupos normales

Recordemos que los cocientes $\mathbf{Z}/(m)$ (I.4) estaban definidos a partir de la relación de equivalencia: $a \equiv b \Leftrightarrow a - b \in (m)$. En general, si G es un grupo conmutativo con una operación $+$ y H es un subgrupo de G, la relación "$g_1 \sim g_2 \Leftrightarrow g_1 - g_2 \in H$" es de equivalencia. Cuando G no es conmutativo hay dos posibles generalizaciones:

I. $g_1 \sim g_2 \Leftrightarrow g_1 \cdot g_2^{-1} \in H$

II. $g_1 \approx g_2 \Leftrightarrow g_2^{-1} \cdot g_1 \in H$.

Tanto una como la otra son relaciones de equivalencia. Demostrémoslo para la primera:

- es reflexiva: para todo $g \in G$, $gg^{-1} = e \in H$, de donde $g \sim g$,
- es simétrica: $g_1 \sim g_2 \Rightarrow g_1 g_2^{-1} \in H \Rightarrow g_2 g_1^{-1} = (g_1 g_2^{-1})^{-1} \in H \Rightarrow g_2 \sim g_1$,
- es transitiva: $g_1 \sim g_2$, $g_2 \sim g_3 \Rightarrow g_1 g_2^{-1} \in H$, $g_2 g_3^{-1} \in H \Rightarrow g_1 g_3^{-1} = (g_1 g_2^{-1})(g_2 g_3^{-1}) \in H \Rightarrow g_1 \sim g_3$.

Las clases de equivalencia para esta relación son

$$[g] = \{g_1 \in G \mid g_1 \sim g\} = \{g_1 \in G \mid g_1 = hg,\ h \in H\}.$$

Pondremos $[g] = Hg$ y el conjunto cociente será denotado por $H\backslash G$.

Las clases de equivalencia para la relación II son

$$\{g\} = \{g_1 \in G \mid g_1 \approx g\} = \{g_1 \in G \mid g_1 = gh,\ h \in H\}.$$

Pondremos $\{g\} = gH$ y el conjunto cociente será denotado por G/H.

Una manera lógica de definir las operaciones en $H\backslash G$ y en G/H sería

$$[g_1][g_2] = [g_1 g_2] \text{ y } \{g_1\}\{g_2\} = \{g_1 g_2\},$$

respectivamente. Esta no es siempre, sin embargo, una buena definición; veámoslo en un ejemplo.

Ejemplo:

Consideremos el subgrupo $H = \{I, \tau_1\}$ de $G = S_3$. Con las notaciones del apartado 3, las clases de $H\backslash G$ son

$$H = \{I, \tau_1\},\ HA = \{A, \tau_2\},\ HB = \{B, \tau_3\}.$$

Las clases de G/H son

$$H = \{I, \tau_1\},\ AH = \{A, \tau_3\},\ BH = \{B, \tau_2\}.$$

El producto de las clases HA y HB se debería obtener efectuando el producto de un representante de HA por uno de HB. Ahora bien,

$AB = I$, de donde el producto sería $H = [I]$
$\tau_2\tau_3 = A$, de donde el producto sería $HA = [A]$.

La operación no queda, pues, bien determinada. Algo parecido pasa con las clases de G/H.

¿Bajo qué condiciones el producto $[g_1][g_2] = [g_1 g_2]$ está bien determinado? Siempre que, para todo par $h_1, h_2 \in H$,

$$(h_1 g_1)(h_2 g_2) = h g_1 g_2$$

para un cierto $h \in H$. Ahora bien,

$$h_1 g_1 h_2 g_2 = h g_1 g_2 \Leftrightarrow h_1 g_1 h_2 g_1^{-1} = h \Leftrightarrow g_1 h_2 g_1^{-1} = h_1^{-1} h \in H.$$

Por tanto, existe un tal $h \in H$ si y sólo si $g_1 h_2 g_1^{-1} \in H$. La operación está, pues, bien definida si para todo $g_1 \in G$ y $h_2 \in H$, $g_1 h_2 g_1^{-1} \in H$; es decir, si

$$g_1 H g_1^{-1} \subset H \quad \forall g_1 \in G,$$

donde $g_1 H g_1^{-1} = \{g_1 h_2 g_1^{-1} \mid h_2 \in H\}$. En particular, para todo $g \in G$, aplicando la inclusión anterior a g y a g^{-1}, tenemos

$$gHg^{-1} \subset H$$
$$g^{-1}Hg \subset H; \text{ es decir, } H \subset gHg^{-1},$$

de donde resulta

$$gHg^{-1} = H.$$

Diremos que H es un *subgrupo normal* si es un subgrupo que cumple la igualdad anterior para todo $g \in G$. Si H es un subgrupo normal de G, en el conjunto $H\backslash G$ hay una operación bien definida: $[g_1][g_2] = [g_1 g_2]$. Con esta operación, $H\backslash G$ es un grupo.

Observemos que $gHg^{-1} = H$ equivale a

$$gH = Hg \quad \forall g \in G;$$

es decir, si H es normal, las clases para las dos relaciones I y II coinciden y, por tanto, los conjuntos cocientes también: $H\backslash G = G/H$. Naturalmente, en este caso, la operación de G/H también está bien definida y coincide con la de $H\backslash G$.

Nota:

El mismo resultado se obtiene, naturalmente, si empezamos estudiando en qué condiciones la operación de G/H está bien definida.

Ejercicio:

Demostrar que, si H es normal, G/H es un grupo. Su elemento neutro es $e = H$; el inverso de $[g]$ es $[g^{-1}]$.

Sea H un subgrupo normal de G. La aplicación

$$\begin{array}{ccc} G & \longrightarrow & G/H \\ g & \longmapsto & [g] \end{array}$$

es un epimorfismo de núcleo H.

Proposición 5.1 *Un subgrupo H es normal si y sólo si es núcleo de un homomorfismo.*

DEMOSTRACIÓN: Todo subgrupo normal H de un grupo G es el núcleo del epimorfismo $G \longrightarrow G/H$ que acabamos de definir.

Supongamos, ahora, que $H = \operatorname{Nuc} f$ para un cierto homomorfismo de grupos $f : G \longrightarrow G'$. Para todo $g \in G$ y todo $h \in H$,

$$f(ghg^{-1}) = f(g)f(h)f(g^{-1}) = f(g)f(g)^{-1} = e,$$

de donde resulta que $ghg^{-1} \in \operatorname{Nuc} f = H$ y, por tanto, H es un subgrupo normal. □

Ejemplo:

El conjunto $\mathcal{A}_n$ de las permutaciones pares de $\mathcal{S}_n$ es un subgrupo normal, ya que es el núcleo de la aplicación signo

$$\varepsilon : \mathcal{S}_n \longrightarrow \{+1, -1\}.$$

$\mathcal{A}_n = \operatorname{Nuc} \varepsilon$ se llama el *grupo alternado de orden* n. En particular, $\mathcal{A}_3 = \{I, A, B\}$ es un subgrupo normal de $\mathcal{S}_3$.

Teorema 5.2 (de isomorfismo) *Sea* $f : G \longrightarrow G'$ *un homomorfismo de grupos; entonces*

$$G/\operatorname{Nuc} f \cong \operatorname{Im} f.$$

DEMOSTRACIÓN: Por (5.1), $\operatorname{Nuc} f$ es normal y $G/\operatorname{Nuc} f$ es un grupo. Todos los elementos de una misma clase de $G/\operatorname{Nuc} f$ tienen la misma imageu por f; en efecto, si $gh \in [g]$ con $h \in \operatorname{Nuc} f$, entonces

$$f(gh) = f(g)f(h) = f(g).$$

Podemos definir, pues, una aplicación

$$\begin{array}{ccc} G/\operatorname{Nuc} f & \longrightarrow & \operatorname{Im} f \\ g & \longmapsto & f(g). \end{array}$$

Esta aplicación es, claramente, un homomorfismo exhaustivo. Para ver que es inyectivo, basta comprobar que la única clase que se aplica en el elemento neutro es $[e] = \operatorname{Nuc} f$ (4.3); en efecto, si $f(g) = e'$, $g \in \operatorname{Nuc} f$ y, por tanto, $[g] = \operatorname{Nuc} f$. □

III.6 Producto directo de grupos

Se dice que un grupo G es *producto directo* de sus subgrupos H_1 y H_2 si

(a) H_1 y H_2 son subgrupos normales de G;

(b) $H_1 \cap H_2 = \{e\}$ (donde e es el elemento neutro de G);

(c) $G = \{h_1 h_2 \mid h_1 \in H_1,\ h_2 \in H_2\}$.

Proposición 6.1 *G es producto directo de sus subgrupos H_1 y H_2 si y sólo si*

1. *todo $g \in G$ se expresa de manera única como producto $g = h_1 h_2$ con $h_1 \in H_1$ y $h_2 \in H_2$;*
2. $h_1 h_2 = h_2 h_1 \quad \forall h_1 \in H_1,\ h_2 \in H_2$.

DEMOSTRACIÓN: Supongamos que G es producto directo de H_1 y H_2. Entonces se cumple 1, ya que si $g = h_1 h_2 = h_1' h_2'$ con $h_1, h_1' \in H_1$ y $h_2, h_2' \in H_2$, $h_1^{-1} h_1' = h_2 (h_2')^{-1} \in H_1 \cap H_2 = \{e\}$, de donde resulta que $h_1^{-1} h_1' = e$ y $h_2 (h_2')^{-1} = e$ y, por tanto, $h_1 = h_1'$ y $h_2 = h_2'$.

También se cumple 2, ya que

$$(h_1 h_2)(h_2 h_1)^{-1} = (h_1 h_2)(h_1^{-1} h_2^{-1});$$

pero, por ser H_1 normal, $h_2 h_1^{-1} h_2^{-1} \in H_1$, de donde $h_1 (h_2 h_1^{-1} h_2^{-1}) \in H_1$ y, por ser H_2 normal, $h_1 h_2 h_1^{-1} \in H_2$, de donde $(h_1 h_2 h_1^{-1}) h_2^{-1} \in H_2$. Ahora bien, dado que $H_1 \cap H_2 = \{e\}$, debe ser $(h_1 h_2)(h_2 h_1)^{-1} = h_1 h_2 h_1^{-1} h_2^{-1} = e$; es decir, $h_1 h_2 = h_2 h_1$.

Recíprocamente, supongamos ahora que G cumple 1 y 2. La condición (c) es parte de 1 y, por tanto, ya se cumple. Para demostrar que H_1 es normal, consideremos elementos cualesquiera $g = h_1 h_2 \in G$ ($h_1 \in H_1$, $h_2 \in H_2$) y $h_1' \in H_1$; utilizando 2, obtenemos

$$g h_1' g^{-1} = h_1 h_2 h_1' h_2^{-1} h_1^{-1} = h_1 h_2 h_2^{-1} h_1' h_1^{-1} = h_1 h_1' h_1^{-1} \in H_1.$$

Análogamente se comprueba que H_2 es normal.

Finalmente, para probar (b), supongamos que $h \in H_1 \cap H_2$; por 1, las dos expresiones $he = eh$ han de ser la misma; así pues, $h = e$. □

Sean ahora G_1 y G_2 dos grupos (pueden ser el mismo). Llamaremos *producto directo de G_1 y G_2* al conjunto $G_1 \times G_2$ junto con la operación

$$(g_1, g_2)(g_1', g_2') = (g_1 g_1', g_2 g_2') \quad \forall g_1, g_1' \in G_1,\ g_2, g_2' \in G_2.$$

Como casi siempre que hablamos en abstracto, utilizamos la notación multiplicativa para G_1 y G_2. Se entiende, sin embargo, que en cada caso particular los elementos g_1 y g_1' se operan con la operación concreta de G_1, y los elementos g_2 y g_2' con la operación concreta de G_2.

El hecho de utilizar el nombre de "producto directo" también para el grupo $G_1 \times G_2$ que acabamos de definir no es casual. La siguiente proposición lo justifica.

Proposición 6.2 *Sea $G_1 \times G_2$ el producto directo de los grupos G_1 y G_2. Existen dos subgrupos de $G_1 \times G_2$, G_1' y G_2', isomorfos a G_1 y G_2 respectivamente, tales que $G_1 \times G_2$ es el producto directo de G_1' y G_2'.*

DEMOSTRACIÓN: Tomemos

$$G_1' = \{(g_1, e) \in G_1 \times G_2\}, \qquad G_2' = \{(e, g_2) \in G_1 \times G_2\},$$

donde indicamos por e tanto el elemento neutro de G_1 como el de G_2. Es fácil ver que G_1' y G_2' son subgrupos de $G_1 \times G_2$ y que las aplicaciones

$$\begin{array}{ccc} G_1 & \longrightarrow & G_1' \\ g_1 & \longmapsto & (g_1, e) \end{array} \qquad\qquad \begin{array}{ccc} G_2 & \longrightarrow & G_2' \\ g_2 & \longmapsto & (e, g_2) \end{array}$$

son isomorfismos. Comprobaremos que $G_1 \times G_2$ es producto directo de G_1' y G_2' viendo que cumple las condiciones 1 y 2 de (6.1). Todo elemento $(g_1, g_2) \in G_1 \times G_2$ se puede escribir como

$$(g_1, g_2) = (g_1, e)(e, g_2)$$

de manera única; esto demuestra 1. Además, se tiene siempre

$$(g_1, e)(e, g_2) = (e, g_2)(g_1, e),$$

lo que demuestra 2. □

Ejemplo:

Estudiemos el producto directo $\mathbf{Z}/(a) \times \mathbf{Z}/(b)$ cuando a y b son primos entre sí (I.5). (En $\mathbf{Z}/(a)$ y en $\mathbf{Z}/(b)$ consideramos la operación suma.) Sumando el elemento $([1], [1])$ consigo mismo suficientes veces, podemos obtener todos los elementos de $\mathbf{Z}/(a) \times \mathbf{Z}/(b)$; en efecto, esto equivale a decir que todo elemento $([m], [q])$ es de la forma

$$([m], [q]) = ([1], [1]) + \overset{n}{\cdots} + ([1], [1]) = ([n], [n]);$$

n ha de ser, pues, tal que

$$n = m + at = q + br.$$

Dado que a y b son primos entre sí, existen t y r tales que $m - q = -at + br$ y, por tanto, existe el número n que buscábamos.

Este hecho nos lleva de manera natural a definir la aplicación exhaustiva

$$\begin{array}{ccc} \mathbf{Z} & \longrightarrow & \mathbf{Z}/(a) \times \mathbf{Z}/(b) \\ n & \longmapsto & ([n], [n]). \end{array}$$

Esta aplicación es, claramente, un morfismo de núcleo $(a) \cap (b)$; ahora bien, puesto que m.c.d.$(a, b) = 1$, su mínimo común múltiplo es ab: $(a) \cap (b) = (ab)$. El teorema 5.2 nos dice entonces que

$$\mathbf{Z}/(ab) \cong \mathbf{Z}/(a) \times \mathbf{Z}/(b).$$

III.7 Grupos cíclicos

Un grupo G se llama *cíclico* si está generado por un elemento g (que se llama un *generador* de G). Escribiremos

$$G = (g).$$

Tal como hemos visto en el apartado 3, el subgrupo generado por g está formado por g, g^{-1} y productos de esos elementos: g^n, $(g^{-1})^n$. Si usamos la notación $g^0 = e$, $g^{-n} = (g^{-1})^n$, tenemos, pues,

$$(g) = \{g^n \mid n \in \mathbf{Z}\}.$$

Ejemplos:

1. $\mathbf{Z}$ con la suma es un grupo cíclico generado por 1.

2. $\mathbf{Z}/(m)$ con la suma es un grupo cíclico generado por [1].

La siguiente proposición nos dice que estos son los únicos ejemplos de grupos cíclicos, salvo isomorfismos.

Proposición 7.1 *Todo grupo cíclico es isomorfo a* $\mathbf{Z}$ *o a un* $\mathbf{Z}/(m)$.

DEMOSTRACIÓN: Sea $G = (g) = \{g^n \mid n \in \mathbf{Z}\}$ un grupo cíclico. La operación en G es $g^n g^m = g^{n+m}$, lo que nos dice que la aplicación

$$\begin{array}{ccc} \mathbf{Z} & \longrightarrow & G \\ n & \longmapsto & g^n \end{array}$$

es un epimorfismo. Su núcleo es un subgrupo de $\mathbf{Z}$ que, como hemos visto en el ejemplo 2 del apartado 3, será de la forma (m). Entonces (5.2) nos dice que $\mathbf{Z}/(m) \cong G$. El caso $m = 0$ corresponde a $\mathbf{Z} \cong G$. □

Un grupo cíclico $G = (g)$ se llama de *orden* $m \neq 0$ si es isomorfo a $\mathbf{Z}/(m)$; en este caso, $g^n = e$ siempre que $n = \dot{m}$. Un grupo cíclico $G = (g)$ se llama de *orden infinito* si es isomorfo a $\mathbf{Z}$; entonces $g^n = e$ sólo cuando $n = 0$.

Proposición 7.2 *Todo subgrupo de un grupo cíclico es cíclico.*

DEMOSTRACIÓN: Sea S un subgrupo de $G = (g)$.

Si $S = \{e\} = (e)$, S es cíclico.

Si $S \neq \{e\}$, sea g^k un elemento de S. Entonces también $g^{-k} \in S$ y, por tanto, S contiene potencias de g con exponente positivo. Sea $g^m \in S$ con exponente positivo mínimo. S contiene el subgrupo generado por g^m: $(g^m) \subset S$. Vamos a ver que $(g^m) = S$; en efecto, sea $g^k \in S$ y efectuemos la división entera de k por m:

$$k = mq + r \text{ con } 0 \leq r < m.$$

Entonces

$$g^r = g^{k-mq} = g^k (g^m)^{-q} \in S,$$

ya que $g^k \in S$ y $(g^m)^{-q} \in (g^m) \subset S$. Como el exponente m era mínimo, debe ser $r = 0$ y, por tanto, $k = \dot{m}$, de donde resulta que $g^k \in (g^m)$. □

III.8 Grupos finitos

Llamaremos *orden de un grupo finito* G al número de sus elementos, y lo denotaremos por $|G|$. Observemos que, si G es cíclico, este orden coincide con el definido en el apartado anterior.

Llamaremos *orden de un elemento* $g \in G$ al orden del subgrupo cíclico generado por g.

Proposición 8.1 *Si S es un subgrupo del grupo finito G, $|S|$ divide a $|G|$.*

DEMOSTRACIÓN: Formemos el conjunto cociente G/S (apartado 5). Todas las clases gS tienen el mismo número de elementos que S, ya que la aplicación $S \to gS$ tal que $x \longmapsto gx$ es biyectiva; por tanto, si en G/S hay i clases, $|G| = i \cdot |S|$. □

De la demostración de (8.1) se deduce que el número de clases de G/S y de $S\backslash G$ es el mismo; lo llamaremos *índice de S en G* y lo denotaremos por $[G : S]$. Por (8.1),

$$[G : S] = |G|/|S|.$$

Corolario 8.2 *El orden de un elemento divide al orden del grupo.* □

Corolario 8.3 *Si $|G| = p$ es primo, G es un grupo cíclico de orden p.*

DEMOSTRACIÓN: Sea $g \in G$, $g \neq e$. Por (8.2), g es de orden 1 o p. Si g fuese de orden 1, g sería igual a e; así pues, g es de orden p. Por tanto, $|(g)| = p = |G|$, de donde resulta $(g) = G$. □

Ejemplo:

Recordemos que al estudiar los subgrupos de $\mathcal{S}_3$ en el apartado 3, hemos encontrado subgrupos de orden $1, 2, 3$ y 6, que son los divisores de $|\mathcal{S}_3| = 6$. Respecto al orden de los elementos, I es de orden 1, τ_1, τ_2 y τ_3 son de orden 2 y A y B son de orden 3. No hay ningún elemento de orden 6.

Uno de los objetivos de la teoría de grupos finitos es determinar qué grupos hay de cada orden (salvo isomorfismos, naturalmente). El problema no está ni de lejos resuelto, pese a que existen listas de todos los grupos finitos hasta órdenes muy elevados (por ejemplo, en el libro *Group Tables* de A. D. Thomas y G. V. Wood). El corolario 8.3 proporciona una información importante: de cada orden primo p hay un único grupo, y posee una estructura muy simple: $\mathbf{Z}/(p)$.

Estudiemos, ahora, cuántos grupos hay de orden 4.

De entrada podemos considerar dos casos:

I. Hay un elemento de orden 4. Entonces se trata del grupo cíclico $\mathbf{Z}/(4)$.

II. Todos los elementos son de orden 2 (excepto el elemento neutro, que siempre es de orden 1). Sea, pues, $G = \{e, a, b, c\}$ con las relaciones $a^2 = b^2 = c^2 = e$. Formemos una tabla o cuadro donde escribiremos todos los productos de dos elementos de G: en cada casilla colocaremos el producto del elemento situado a la misma "altura" en la primera columna y el situado

a la misma "altura" en la primera fila (en este orden). Este cuadro se llama la "tabla de la operación" de G. En este caso II tenemos, de momento,

	e	a	b	c
e	e	a	b	c
a	a	e		
b	b		e	
c	c			e.

Los elementos que aparecen en una misma fila (o en una misma columna) han de ser todos diferentes (¿por qué?). Por tanto, el producto ab debe ser forzosamente c. Por el mismo motivo resulta $ac = b$, $ba = c, \ldots$. Obtenemos así como única tabla posible

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e.

Observemos que este grupo también es conmutativo. Además, se ve fácilmente que es el producto directo de sus subgrupos $\{e, a\}$ y $\{e, b\}$. Estos subgrupos son isomorfos a $\mathbf{Z}/(2)$. La aplicación

$$G \longrightarrow \mathbf{Z}/(2) \times \mathbf{Z}/(2)$$

que aplica e, a, b, c en $([0],[0]), ([1],[0]), ([0],[1]), ([1],[1])$, respectivamente, es un isomorfismo de grupos.

Hemos demostrado así la

Proposición 8.4 *Sólo hay dos grupos de orden* 4 : $\mathbf{Z}/(4)$ *y* $\mathbf{Z}/(2) \times \mathbf{Z}/(2)$. *Ambos son conmutativos.* □

Pasemos a estudiar, ahora, los grupos de orden 6. Nos podemos encontrar con los tres casos siguientes (que no se excluyen):

I. Hay un elemento de orden 6. Entonces se trata del grupo cíclico $\mathbf{Z}/(6)$.

II. Hay un elemento g de orden 3. Por (8.1), $[G : (g)] = 2$. En general, tenemos

Proposición 8.5 *Todo subgrupo S de un grupo G de índice* 2 *es normal.*

DEMOSTRACIÓN: Si S tiene índice 2, G/S consta de dos elementos S y $gS = G - S$. Análogamente, $S\backslash G$ consta de dos elementos, S y $Sg = G - S$. Por tanto, las clases por la derecha y por la izquierda son las mismas y S es un subgrupo normal. □

Volvamos al caso II. (g) es normal y $G/(g)$ es un grupo con 2 elementos:

$$[e] = (g) = \{e, g, g^2\} \qquad \text{y} \qquad [g_1] = g_1(g) = \{g_1, g_1g, g_1g^2\}.$$

El elemento $[g_1]$ tiene que ser de orden 2; es decir, $g_1^2 \in [e] = \{e, g, g^2\}$.

- Si $g_1^2 = g$, entonces $g_1^3 = gg_1 \neq e$, ya que en caso contrario $g_1 = g^{-1} = g^2$, pero $g_1 \notin (g)$; así pues, el orden de g_1 es mayor que 3 y, por tanto, es 6 (por (8.2)). G es, pues, cíclico e isomorfo a $\mathbf{Z}/(6)$.

- Si $g_1^2 = g^2$, entonces $g_1^3 = g^2g_1 \neq e$, ya que en caso contrario $g_1 = g^{-2} = g$, pero $g_1 \notin (g)$. Así pues, el orden de g_1 es mayor que 3 y, como antes, $G \cong \mathbf{Z}/(6)$.

- Si $g_1^2 = e$, entonces, dado que $\{g_1, g_1g, g_1g^2\} = \{g_1, gg_1, g^2g_1\}$ (ya que $g_1(g) = (g)g_1$), tenemos

o bien $g_1g = gg_1$ (y $g_1g^2 = g^2g_1$), de donde $(g_1g)^2 = g_1^2g^2 = g^2 \neq e$ y $(g_1g)^3 = g_1^3g^3 = g_1 \neq e$; es decir, el orden de g_1g es mayor que 3 y, por tanto, como antes, $G \cong \mathbf{Z}/(6)$,

o bien $g_1g = g^2g_1$ (y $g_1g^2 = gg_1$). En este caso, la tabla de la operación del grupo es

	e	g	g^2	g_1	g_1g	g_1g^2
e	e	g	g^2	g_1	g_1g	g_1g^2
g	g	g^2	e	g_1g^2	g_1	g_1g
g^2	g^2	e	g	g_1g	g_1g^2	g_1
g_1	g_1	g_1g	g_1g^2	e	g	g^2
g_1g	g_1g	g_1g^2	g_1	g^2	e	g
g_1g^2	g_1g^2	g_1	g_1g	g	g^2	e.

Observemos que esta tabla es "la misma" que la del grupo de permutaciones $\mathcal{S}_3$, cambiando sólo g, g^2 por A, B y g_1, g_1g, g_1g^2 por las tres trasposiciones τ_1, τ_2, τ_3. En otras palabras, la aplicación

$$G \longrightarrow \mathcal{S}_3$$

que transforma e, g, g^2, g_1, g_1g, g_1g^2 en I, A, B, τ_1, τ_2, τ_3, respectivamente, es un isomorfismo: $G \cong \mathcal{S}_3$.

III. El último caso a considerar es aquel en que todos los elementos de G son de orden 2.

Proposición 8.6 *Si todos los elementos de un grupo G son de orden 2, entonces G es un grupo conmutativo.*

DEMOSTRACIÓN: Un elemento es de orden 2 si y sólo si es inverso de sí mismo. Por tanto, para todo par $g_1, g_2 \in G$, $(g_1g_2)(g_2g_1)^{-1} = g_1g_2g_2g_1 = g_1g_1 = e$, de donde $g_1g_2 = g_2g_1$. □

Consideremos, pues, $g \in G$; (g) es normal por ser G conmutativo. Formemos el grupo $G/(g)$, que tendrá 3 elementos. Por (8.3), $G/(g)$ es cíclico. Sea $[g_1]$ un generador de $G/(g)$; $[g_1]$ debería ser de orden 3, pero $[g_1]^2 = [g_1^2] = [e]$. Esta contradicción nos asegura que este tercer caso no puede darse nunca.

Hemos demostrado

Proposición 8.7 *Sólo hay dos grupos de orden 6: uno conmutativo, $\mathbf{Z}/(6)$, y uno no conmutativo, $\mathcal{S}_3$.* □

Acabamos dando sin demostración un resultado muy importante. Hemos dicho antes que el problema de determinar todos los grupos de un cierto orden no estaba resuelto. No obstante, sí lo está el problema de determinar todos los grupos conmutativos de un cierto orden.

Aún más, se conocen todos los grupos conmutativos con un número finito de generadores. Concretamente, tenemos

Teorema 8.8 (de estructura de los grupos conmutativos) *Todo grupo conmutativo G con un número finito de generadores es producto directo de un número finito de grupos $\mathbf{Z}$ y $\mathbf{Z}/(m)$:*

$$G \cong \mathbf{Z} \times \ldots \times \mathbf{Z} \times \mathbf{Z}/(m_1) \times \ldots \times \mathbf{Z}/(m_r).$$

Esta descomposición es única si $m_1 \mid m_2 \mid \ldots \mid m_r$.

Existen demostraciones elementales de este teorema. (Ver, por ejemplo, el libro *Algebra*, volumen I, segunda edición, de P. M. Cohn (John Wiley & Sons, 1982).)

III.9 Nota histórica

Los inicios de la teoría de grupos se pueden situar en el estudio que Joseph-Louis Lagrange (1736–1813) hizo de la resolución de las ecuaciones de grado n. La idea de Lagrauge fue escoger una función racional de las raíces de

la ecuación que fuese invariante por todas las permutaciones de las raíces. Aunque el método de Lagrange no da el resultado esperado, sí que motiva la aparición de los primeros resultados de la teoría de grupos de permutaciones (y, haciendo abstracción, de grupos finitos).

Carl Friedrich Gauss (1777–1855), en sus *Disquisitiones Arithmeticae*, dedica una sección al estudio de las formas cuadráticas $ax^2 + 2bxy + cy^2$; define una composición de formas, una relación de equivalencia entre ellas y comprueba que las clases de equivalencia tienen la estructura de un grupo conmutativo (evidentemente, no utiliza este lenguaje). Es razonable pensar que Gauss tenía ya la idea del teorema de estructura de los grupos abelianos finitos, aunque no fue demostrado hasta el año 1870 por Leopold Kronecker (1823–1891).

Es, no obstante, el trabajo de Évariste Galois (1811–1832) el que da un impulso extraordinario a la teoría de grupos (de grupos de permutaciones en aquella época), introduciendo en ella nuevos conceptos y resultados, motivados por su investigación sobre la resolución por radicales de las ecuaciones de grado ≥ 5, recuperando así la herencia de Lagrange. Galois introduce las clases módulo un subgrupo, el producto de grupos, el concepto de isomorfismo, etc.. La obra de Galois, que comienza cuando con 16 años lee los trabajos de Lagrange y acaba 4 años más tarde después de una juventud turbulenta y de una muerte trágica, es uno de los capítulos más apasionantes de las matemáticas. (Léase *Obra d'Évariste Galois*, Institut d'Estudis Catalans, Monografies de la Secció de Ciències n. 1, 1984.)

A partir de la segunda mitad del siglo 19, la teoría de grupos se va configurando poco a poco, al principio de la mano de Arthur Cayley (1821–1895), quien dio la definición abstracta de grupo finito, y después con Camille Jordan (1838–1922), con su libro *Mémoire sur les groupes de mouvements*, y con Walter von Dyck (1856–1934), que considera ya los grupos abstractos a finales del siglo pasado e introduce, entre otros, el concepto de sistema de generadores.

III.10 Ejercicios

1. Demostrar que un ciclo de orden n no se puede expresar nunca como producto de menos de $n-1$ trasposiciones.

2. Demostrar que $\mathcal{S}_n$ admite los siguientes sistemas de generadores:

a) $(1,2),(1,3),\ldots,(1,n)$.

b) $(1,2),(2,3),\ldots,(n-1,n)$.

c) $(1,2,\ldots,n),(1,2)$.

3. Dada la permutación

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 7 & 8 & 9 & 4 & 5 & 2 & 1 & 6 \end{pmatrix},$$

calcular σ^{100}.

4. Sea G un subgrupo de $\mathcal{S}_n$ no contenido en $\mathcal{A}_n$. Demostrar que exactamente la mitad de las permutaciones de G son pares.

5. Una permutación se llama *regular* si, al descomponerla en producto de ciclos disjuntos, todos los ciclos tienen el mismo orden. Demostrar que una permutación es regular si y sólo si es una potencia de un ciclo de orden máximo.

6. Demostrar que las raíces de la ecuación $z^6+1=0$ con el producto de C forman un grupo cíclico. Encontrar sus generadores y sus subgrupos.

7. Demostrar que la aplicación

$$\begin{array}{ccc} \mathbf{R} & \longrightarrow & \mathbf{C}-\{0\} \\ t & \longmapsto & e^{2\pi it} = \cos 2\pi t + i \operatorname{sen} 2\pi t \end{array}$$

es un morfismo de grupos y explicitar el teorema de isomorfismo.

8. Demostrar que, para un grupo G, las siguientes afirmaciones son equivalentes:

a) G es abeliano.
b) $x \mapsto x^{-1}$ es un morfismo.
c) $x \mapsto x^2$ es un morfismo.

9. Si G es un grupo cíclico de orden n y r es un divisor de n, demostrar que G tiene como máximo $r-1$ elementos de orden r. Si r es primo, entonces hay exactamente $r-1$ elementos de orden r.

10. Demostrar que en todo grupo G de orden p^r, p primo, existe al menos un subgrupo de orden p.

11. Si G es un grupo de orden n tal que todo elemento (diferente del neutro) tiene orden 2, demostrar que n es una potencia de 2.

12. Sea G el grupo generado por las matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

con el producto de matrices. Demostrar que G es un grupo no abeliano de 8 elementos.

13. Sea G un grupo y $G' = \{x^2 \mid x \in G\}$. Demostrar:

a) Si G es abeliano, G' es un subgrupo normal de G. (Dar un contraejemplo en el caso G no abeliano.)

b) Si G es abeliano, G/G' no tiene cuadrados aparte del 0.

c) Si H es un subgrupo de G que contiene G', H es normal y G/H es un grupo abeliano.

d) Calcular G' y G/G' en los casos siguientes: $\mathbf{Z}$, $\mathbf{Z}/(n)$, $\mathbf{Q}$, $\mathcal{S}_3$.

14. Se define el *centro* $\mathrm{Z}G$ de un grupo G como el conjunto de los elementos que conmutan con todos los elementos de G. Demostrar:

a) $\mathrm{Z}G$ es un subgrupo normal de G.

b) Si G contiene un único elemento de orden 2, éste pertenece a $\mathrm{Z}G$.

c) Si H es un subgrupo normal de G contenido en $\mathrm{Z}G$ y G/H es cíclico, entonces G es abeliano.

d) $\mathrm{Z}(\mathcal{S}_n) = \{e\}$.

15. Se define el *conmutador* G' de un grupo G como el subgrupo generado por los elementos de la forma $xyx^{-1}y^{-1}$ con $x, y \in G$:

$$G' = (xyx^{-1}y^{-1} \mid x, y \in G).$$

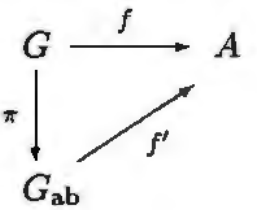
Demostrar:

a) G' es un subgrupo normal de G.

b) $G_{\mathrm{ab}} = G/G'$ es un grupo abeliano (se llama el *abelianizado* de G).

c) Para todo subgrupo normal H de G que contenga a G', G/H es abeliano.

d) Si $f : G \longrightarrow A$ es un morfismo de G en un grupo abeliano A, existe un único morfismo $f' : G_{\mathrm{ab}} \longrightarrow A$ tal que el diagrama siguiente es conmutativo:



donde π es el epimorfismo canónico que aplica cada elemento de G en su clase módulo G'.

e) La propiedad expresada en (d) caracteriza al grupo G_{ab}.

f) Determinar G' y G_{ab} para $G = \mathcal{S}_3$.

III.11 Ejercicios para programar

16. Producto de dos permutaciones. (Indicación: para guardar una permutación $\sigma \in \mathcal{S}_n$ guardar $\sigma(1), \ldots, \sigma(n)$.)

17. Orden de una permutación. (Indicación: generar $\sigma^2, \sigma^3, \ldots$ y combinar el ejercicio anterior con un pequeño dispositivo que reconozca cuándo una permutación es la identidad.)

18. Descomposición de una permutación en ciclos disjuntos. Signo. (Indicación: aplicar el método de la proposición 2.2.)

19. Subgrupo generado por un conjunto de permutaciones. (Indicación: preparar primero un subprograma que, dada una permutación, la compare con todas las de una lista $\tau_1, \ldots, \tau_k$ previamente obtenida.)

Sugerencia: utilizar variables alfanuméricas para obtener, en la lista $\tau_1, \ldots, \tau_k$, además de las diferentes permutaciones, sus expresiones en función de los generadores. Esto permite ahorrar operaciones y conocer la expresión de cualquier permutación del subgrupo en función de los generadores.

Capítulo IV

Espacios vectoriales

IV.1 Definición y ejemplos

De ahora en adelante, si no especificamos lo contrario, K indicará un cuerpo conmutativo. Un *espacio vectorial sobre* K es un conjunto E no vacío junto con

1. una operación $+$, a la que llamaremos *suma,* que cumple las siguientes propiedades:
   - es asociativa: $u+(v+w)=(u+v)+w \quad \forall u,v,w \in E$,
   - es conmutativa: $u+v=v+u \quad \forall u,v \in E$,
   - existe un elemento $\vec{0}$ tal que $u+\vec{0}=u \quad \forall u \in E$,
   - para todo $u \in E$ existe otro elemento, que se denota por $-u$, tal que $u+(-u)=\vec{0}$;
2. una aplicación
$$\begin{array}{ccc} K \times E & \longrightarrow & E \\ (a,u) & \longmapsto & au \end{array}$$
que denominaremos *producto por elementos de* K, que cumple
   - $a(u+v)=au+av \quad \forall a \in K, u,v \in E$,
   - $(a+b)u=au+bu \quad \forall a,b \in K, u \in E$,
   - $(ab)u=a(bu) \quad \forall a,b \in K, u \in E$,
   - $1u=u \quad \forall u \in E$, donde 1 es la unidad del cuerpo K.

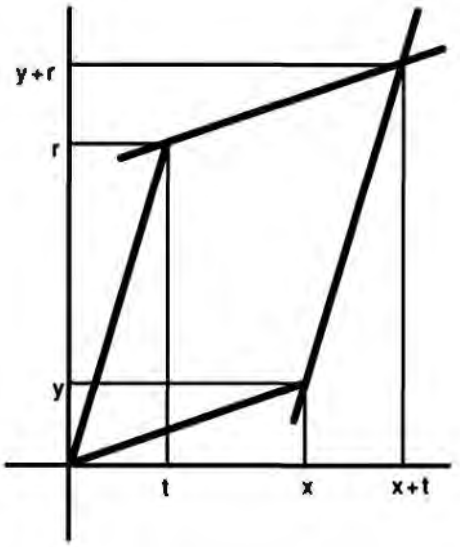
Observemos que la condición 1 asegura que $(E,+)$ es un grupo conmutativo (III.1). A los elementos de E los llamaremos *vectores;* a los de K, *escalares.* Usaremos la notación $u-v$ para indicar $u+(-v)$.

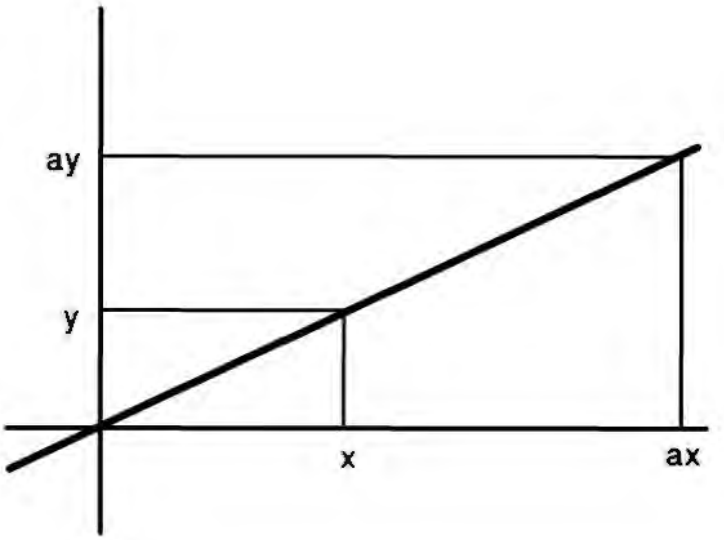
De la definición se deduce fácilmente:

- $0v=\vec{0}$. En efecto, $0v=(0+0)v=0v+0v \Rightarrow 0v=\vec{0}$.
- $a\vec{0}=\vec{0}$. En efecto, $a\vec{0}=a(\vec{0}+\vec{0})=a\vec{0}+a\vec{0} \Rightarrow a\vec{0}=\vec{0}$.
- $av=\vec{0} \Rightarrow a=0$ o $v=\vec{0}$. En efecto, si $a\neq 0$, a tiene un inverso a^{-1}. Entonces $v=1v=(a^{-1}a)v=a^{-1}(av)=a^{-1}\vec{0}=\vec{0}$.
- $(-1)v=-v$. En efecto, $v+(-1)v=1v+(-1)v=(1+(-1))v=0v=\vec{0}$.

Ejemplos:

1. Sea
$$\begin{cases} a_1^1x^1+\ldots+a_n^1x^n=0 \\ \cdots\cdots\cdots\cdots \\ a_1^mx^1+\ldots+a_n^mx^n=0 \end{cases}$$
un sistema homogéneo de m ecuaciones lineales con n incógnitas, y con coeficientes en un cuerpo K. Una solución es una n-pla $(s^1,\ldots,s^n)\in K^n$ tal que, al sustituir $x^1,\ldots,x^n$ por estos elementos en el sistema, todas las igualdades son ciertas. Si $(r^1,\ldots,r^n)$ es otra solución, $(s^1+r^1,\ldots,s^n+r^n)$ también lo es; y si $a\in K$, $(as^1,\ldots,as^n)$ es también una solución. El conjunto de soluciones del sistema dado, con las dos operaciones suma y producto por elementos de K que acabamos de definir, es un espacio vectorial sobre K.
2. Consideremos en $\mathbf{R}^2$ las dos operaciones
$$(x,y)+(t,r)=(x+t,y+r)$$
$$a(x,y)=(ax,ay).$$
Con estas operaciones, $\mathbf{R}^2$ es un espacio vectorial sobre $\mathbf{R}$. Los pares (x,y) de $\mathbf{R}^2$ se representan a menudo como puntos de un plano. La suma que hemos definido coincide con la conocida "ley del paralelogramo", según la cual se suman las fuerzas en Física. El producto por un $a\in\mathbf{R}$ coincide, igualmente, con el producto de una fuerza por un $a\in\mathbf{R}$.





3. K^n con las operaciones

$$(x^1, \ldots, x^n) + (y^1, \ldots, y^n) = (x^1 + y^1, \ldots, x^n + y^n)$$

$$a(x^1, \ldots, x^n) = (ax^1, \ldots, ax^n)$$

es un espacio vectorial sobre K. El ejemplo 2 es un caso particular de éste.

4. K es un espacio vectorial sobre sí mismo. El producto es el producto ordinario de K. $\mathbf{C}$ es un espacio vectorial sobre $\mathbf{C}$. $\mathbf{C}$ es también un espacio vectorial sobre $\mathbf{R}$, ya que existe un producto de elementos de $\mathbf{R}$ por elementos de $\mathbf{C}$ con las propiedades necesarias. Y también es un espacio vectorial sobre los racionales $\mathbf{Q}$, por el mismo motivo.
5. El conjunto de polinomios $K[x]$ es un espacio vectorial sobre K con las operaciones usuales.
6. Llamaremos *matriz* $m \times n$ a un cuadro de elementos de K:

$$\begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix}, \quad a_i^j \in K.$$

Designaremos por $M_{m\times n}(K)$ el conjunto de las matrices $m \times n$ sobre K. En este conjunto definimos una suma y un producto por elementos de K de la manera siguiente:

$$\begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} + \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & & \vdots \\ b_1^m & \dots & b_n^m \end{pmatrix} = \begin{pmatrix} a_1^1 + b_1^1 & \dots & a_n^1 + b_n^1 \\ \vdots & & \vdots \\ a_1^m + b_1^m & \dots & a_n^m + b_n^m \end{pmatrix}$$

$$k \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} = \begin{pmatrix} ka_1^1 & \dots & ka_n^1 \\ \vdots & & \vdots \\ ka_1^m & \dots & ka_n^m \end{pmatrix}.$$

Con estas operaciones $M_{m\times n}(K)$ es un espacio vectorial sobre K.

IV.2 Subespacios vectoriales

Sea E un espacio vectorial sobre K. Un subconjunto no vacío $F \subset E$ se llama *subespacio vectorial de* E si

1. $u, v \in F \Rightarrow u + v \in F$,
2. $u \in F,\ k \in K \Rightarrow ku \in F$.

Estas dos condiciones nos dicen que las operaciones de E permiten definir unas operaciones en F. Observemos que con estas operaciones F es automáticamente un espacio vectorial sobre K.

Ejemplo:

El conjunto de soluciones de un sistema homogéneo con n incógnitas es un subespacio vectorial de K^n (ejemplos 1 y 3 del §1).

Un vector u es *combinación lineal* de los vectores $v_1, \dots, v_n$ si existen $a^1, \dots, a^n \in K$ tales que

$$u = a^1 v_1 + \dots + a^n v_n.$$

De las condiciones 1 y 2 de la definición de subespacio vectorial resulta que toda combinación lineal de vectores $v_1, \dots, v_n$ de F es un vector de F. Supongamos que S es un subconjunto cualquiera de E. Designemos por $\langle S \rangle$ el conjunto de las combinaciones lineales de elementos de S. Todo subespacio vectorial F que contenga a S deberá contener a $\langle S \rangle$. Por otra parte $\langle S \rangle$ es, él mismo, un subespacio vectorial. Tenemos pues la siguiente

Proposición 2.1 *Si S es un subconjunto de un espacio vectorial E, el conjunto $\langle S \rangle$ es el menor subespacio vectorial de E que contiene a S.* □

Si $\langle S \rangle = F$, se dice que S *genera* F, que F *está generado por* S o que S es un *sistema de generadores* de F.

Ejemplos:

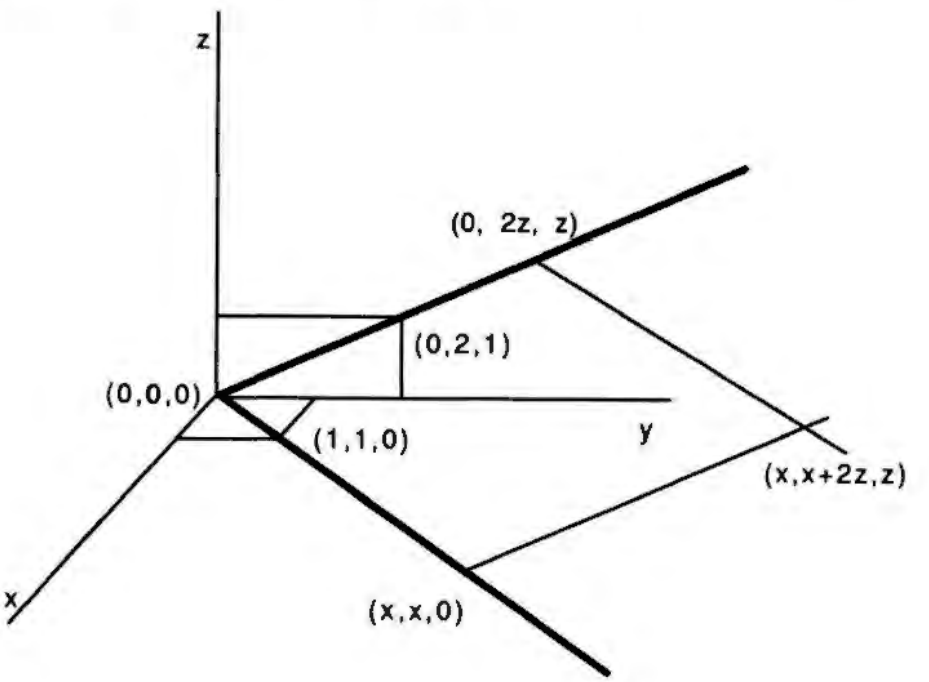
1. $\mathbf{R}^2 = \langle (1,0),(0,1) \rangle$, ya que todo par $(x,y) \in \mathbf{R}^2$ es de la forma

$$(x,y) = x(1,0) + y(0,1).$$

2. $K^n = \langle (1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1) \rangle$.
3. $K[x] = \langle 1, x, x^2, \ldots, x^n, \ldots \rangle$.
4. $F = \{(x, 2z+x, z) \mid x, z \in \mathbf{R}\}$ es un subespacio vectorial de $\mathbf{R}^3$. Sus elementos son de la forma

$$(x, 2z+x, z) = x(1,1,0) + z(0,2,1).$$

Resulta pues que $F = \langle (1,1,0),(0,2,1) \rangle$. Representemos los elementos de R^3 como puntos del espacio de la manera usual. Los elementos de F son, en esta representación, los puntos del plano que pasa por $(0,0,0)$, $(1,1,0)$ y $(0,2,1)$.



IV.3 Bases de un espacio vectorial

Un conjunto de vectores S se llama *linealmente independiente* si toda combinación lineal de vectores de S nula tiene todos los coeficientes nulos: $a^1v_1 + \ldots + a^mv_m = \vec{0}$, $v_i \in S$, $i = 1,\ldots,m \Rightarrow a^1 = \ldots = a^m = 0$. En caso contrario, diremos que S es *linealmente dependiente.*

Proposición 3.1 *$v_1,\ldots,v_m$ son linealmente dependientes si y sólo si uno de ellos es combinación lineal de los restantes.*

DEMOSTRACIÓN: Si $v_1,\ldots,v_m$ son linealmente dependientes hay una combinación lineal

$$a^1v_1 + \ldots + a^mv_m = \vec{0}$$

con algún coeficiente no nulo. Podemos suponer que $a^1 \neq 0$ (reordenando $v_1,\ldots,v_m$ si es necesario). Entonces existe $(a^1)^{-1}$ y

$$v_1 = -(a^1)^{-1}a^2v_2 - (a^1)^{-1}a^3v_3 - \ldots - (a^1)^{-1}a^mv_m.$$

Recíprocamente,

$$v_1 = a^2v_2 + \ldots + a^mv_m \Rightarrow 1v_1 - a^2v_2 - \ldots - a^mv_m = \vec{0}.$$

El primer coeficiente no es nulo y, por tanto, $v_1,\ldots,v_m$ son linealmente dependientes. □

Ejemplos:

1. Un único vector v es linealmente independiente si y sólo si $v \neq \vec{0}$.
2. En $\mathbf{R}^2$, $(1,0)$ y $(0,1)$ son linealmente independientes. En general, en K^n
$$(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)$$
son linealmente independientes.
3. En $\mathbf{R}^3$, $(-1,1,0)$, $(5,2,3)$, $(0,7,3)$ son linealmente dependientes, ya que
$$5(-1,1,0) + (5,2,3) - (0,7,3) = (0,0,0).$$
4. En $K[x]$, $S = \{1, x, x^2, \ldots, x^n, \ldots\}$ es linealmente independiente.

Una *base* de un espacio vectorial E es un sistema de generadores linealmente independientes.

Proposición 3.2 *$B \subset E$ es una base de E si y sólo si todo $u \in E$ se expresa de manera única como combinación lineal de elementos de B.*

DEMOSTRACIÓN: Demostraremos sucesivamente los dos sentidos de la implicación.

$\Rightarrow$) Dado $u \in E$, u es combinación de elementos de B, ya que B genera E. Sean $u = a^1v_1 + \ldots + a^nv_n$ y $u = b^1u_1 + \ldots + b^mu_m$ dos expresiones de u como combinación lineal de vectores de B. Si un vector v_i no aparece en la segunda expresión, añadimos a ésta el sumando $0v_i$; análogamente, si un u_i no aparece en la primera expresión, añadimos a ésta el sumando $0u_i$. Obtenemos así dos combinaciones lineales de los mismos vectores. Sean

$$u = a^1v_1 + \ldots + a^rv_r = b^1v_1 + \ldots + b^rv_r.$$

Restando, obtenemos

$$(a^1 - b^1)v_1 + \ldots + (a^r - b^r)v_r = \vec{0}.$$

Los vectores $v_1, \ldots, v_r$ son de B y, por tanto, linealmente independientes. Los coeficientes de esta combinación lineal han de ser, pues, nulos. Por tanto,

$$a^1 = b^1, \ldots, a^r = b^r.$$

$\Leftarrow$) Todo vector de E se expresa como combinación lineal de vectores de B. Es decir, $\langle B \rangle = E$. Si $a^1v_1 + \ldots + a^mv_m = \vec{0}$ con $v_i \in B$, $i = 1, \ldots, n$, dado que también $0v_1 + \ldots + 0v_m = \vec{0}$ y la expresión debe ser única,

$$a^1 = 0, \ldots, a^n = 0.$$

B es, pues, linealmente independiente. □

Observación:

Si S es linealmente independiente, S es base de $\langle S \rangle$.

Proposición 3.3 *Si S es linealmente independiente y $u \notin \langle S \rangle$, entonces $S \cup \{u\}$ es linealmente independiente.*

DEMOSTRACIÓN: Consideremos

$$au + a^1v_1 + \ldots + a^mv_m = \vec{0}$$

con $v_i \in S$, $i = 1, \ldots, m$. Si $a \neq 0$, existe a^{-1} y

$$u = -a^{-1}a^1v_1 - \ldots - a^{-1}a^mv_m \in \langle S \rangle,$$

en contra de la hipótesis hecha. Por tanto, $a = 0$. Pero entonces tenemos $a^1v_1 + \ldots + a^mv_m = \vec{0}$ y todos los vectores de esta expresión son de S. Por ser S linealmente independiente, $a^1 = \ldots = a^m = 0$. □

Teorema 3.4 *Todo espacio vectorial $E \neq \{\vec{0}\}$ generado por un número finito de vectores tiene una base finita.*

DEMOSTRACIÓN: Sea $E = \langle v_1, \ldots, v_m \rangle$. Si los generadores son linealmente independientes, forman base. En caso contrario, hay uno, pongamos v_i, que es combinación lineal de los otros. Entonces

$$E = \langle v_1, \ldots, v_m \rangle = \langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m \rangle.$$

Si este nuevo conjunto de generadores es linealmente independiente, forman base. En caso contrario podemos suprimir uno de ellos, obteniendo un nuevo conjunto de generadores. Repitamos el proceso tantas veces como sea necesario, eliminando siempre aquellos generadores que sean combinación lineal de los restantes. Llegaremos de esta manera a un conjunto linealmente independiente (es decir, a una base), o los eliminaremos todos. En este segundo caso será $E = \{\vec{0}\}$. □

En realidad, esta demostración prueba más de lo que establece el enunciado: prueba que todo sistema de generadores contiene una base.

Ejemplos:

1. Los vectores $(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)$ forman una base de K^n.
2. El conjunto $\{1, x, x^2, \ldots, x^n, \ldots\}$ es una base de $K[x]$.
3. Sea E_i^j la matriz de $M_{m\times n}(K)$ formada por 0 en todas las posiciones, excepto en la columna i, fila j, en que aparece 1. El conjunto E_i^j $(i = 1,\ldots,n,\ j = 1,\ldots,m)$ es una base de $M_{m\times n}(K)$.
4. Los vectores $(1,1,0)$, $(0,2,1)$ forman una base del subespacio vectorial $F = \{(x, 2z + x, z) \mid x, z \in \mathbf{R}\}$. (Ver ejemplo 4 del apartado 2.)

Ejercicio:

Demostrar la veracidad de todos los ejemplos que acabamos de dar.

Por sus consecuencias, el siguiente teorema tiene una especial importancia.

Teorema 3.5 (de Steinitz) *Sea $u_1, \ldots, u_n$ una base del espacio vectorial E y sean $v_1, \ldots, v_m$ vectores linealmente independientes. Entonces se pueden sustituir m vectores de la base $u_1, \ldots, u_n$ por $v_1, \ldots, v_m$ obteniendo una nueva base. En particular, $m \leq n$.*

DEMOSTRACIÓN: Se trata de introducir uno por uno los $v_1, \ldots, v_m$ en sustitución de vectores de la base dada.

1. Introducción de v_1. Por ser $u_1, \ldots, u_n$ una base, tendremos

$$v_1 = \sum_{i=1}^{n} a^i u_i;$$

v_1 no es nulo y, por tanto, uno de los coeficientes a^i no es nulo. Podemos suponer, reordenando en caso necesario la base $u_1, \ldots, u_n$, que $a^1 \neq 0$. Entonces

$$u_1 = (a^1)^{-1} v_1 - \sum_{i=2}^{n} (a^1)^{-1} a^i u_i.$$

Esta expresión nos dice que

$$\langle u_1, u_2, \ldots, u_n \rangle = \langle v_1, u_2, \ldots, u_n \rangle.$$

Además, $v_1, u_2, \ldots, u_n$ son linealmente independientes. En efecto,

$$b^1 v_1 + b^2 u_2 + \ldots + b^n u_n = \vec{0} \Rightarrow b^1 (\sum_{i=1}^{n} a^i u_i) + b^2 u_2 + \ldots + b^n u_n = \vec{0}$$

$$\Rightarrow \quad b^1 a^1 u_1 + \sum_{i=2}^{n} (b^1 a^i + b^i) u_i = \vec{0} \Rightarrow b^1 a^1 = 0, \; b^1 a^i + b^i = 0, \; i \geq 2,$$

ya que $u_1, \ldots, u_n$ es una base. Pero $a^1 \neq 0$. Por tanto, $b^1 = 0$ y $b^i = 0$ $(i = 2, \ldots, n)$. Así pues, $v_1, u_2, \ldots, u_n$ forman una base de E.

2. Supongamos que ya hemos sustituido h vectores de la base por los vectores $v_1, \ldots, v_h$. Reordenando si es necesario la base $u_1, \ldots, u_n$ podemos suponer que hemos sustituido los h primeros, y tenemos que

$$v_1, \ldots, v_h, u_{h+1}, \ldots, u_n$$

es una base de E. Procedamos igual que en 1 y expresemos v_{h+1} como combinación lineal de los vectores de esta base:

$$v_{h+1} = \sum_{i=1}^{h} a^i v_i + \sum_{h+1}^{n} a^i u_i.$$

Entonces 1 nos asegura que podemos sustituir por v_{h+1} cualquier vector que, en esta expresión, tenga coeficiente no nulo. Todo queda

reducido, pues, a comprobar que uno de los coeficientes del segundo sumatorio no es nulo. Pero, en efecto, si $a^{h+1} = \ldots = a^n = 0$, entonces v_{h+1} sería combinación lineal de $v_1, \ldots, v_h$ y esto no es cierto, ya que los vectores $v_1, \ldots, v_m$ son linealmente independientes. □

Corolario 3.6 *Si el espacio vectorial E tiene una base finita, todas las bases de E tienen el mismo número de vectores.*

DEMOSTRACIÓN: Sean $u_1, \ldots, u_n$ y $\{v_j \mid j \in J\}$ dos bases de E. De (3.5) se deduce que toda familia finita $v_{j_1}, \ldots, v_{j_k}$ satisface $k \leq n$. Por tanto, J ha de ser finito. Entonces, si J tiene m elementos, tenemos $m \leq n$ y también $n \leq m$ (3.5). Por tanto, $m = n$. □

La *dimensión* de un espacio vectorial E sobre un cuerpo K es el número de elementos de sus bases, si son finitas. Si no lo son, diremos que E es de *dimensión infinita.*

Corolario 3.7 *La dimensión de un espacio coincide con el número máximo de elementos linealmente independientes, y también con el número mínimo de generadores.*

DEMOSTRACIÓN: La primera afirmación es consecuencia inmediata del teorema de Steinitz. La segunda resulta de la demostración de (3.4). □

Corolario 3.8 *Todo conjunto de vectores linealmente independientes puede completarse hasta obtener una base.* □

Ejemplos:

1. K^n es de dimensión n sobre K.
2. $K[x]$ es de dimensión infinita sobre K.
3. $M_{m\times n}(K)$ es de dimensión nm sobre K.
4. Los complejos $\mathbf{C}$ son un espacio vectorial sobre $\mathbf{C}$ de dimensión 1, y un espacio vectorial sobre $\mathbf{R}$ de dimensión 2. En el segundo caso, $\{1, i\}$ es una base.

Proposición 3.9 *Sea F un subespacio del espacio vectorial E. Si la dimensión de E es finita, la de F también lo es y*

$$\dim F \leq \dim E.$$

Además,

$$\dim F = \dim E \Leftrightarrow F = E.$$

DEMOSTRACIÓN: Si $F = \{\vec{0}\}$, no hay nada que decir. En caso contrario, sea $\vec{0} \neq v_1 \in F$; si $F = \langle v_1 \rangle$, v_1 es base. En caso contrario, sea $v_2 \in F$, $v_2 \notin \langle v_1 \rangle$; si $F = \langle v_1, v_2 \rangle$, $\{v_1, v_2\}$ forman base por (3.3). En caso contrario, sea $v_3 \in F$, $v_3 \notin \langle v_1, v_2 \rangle, \ldots$. Por (3.7), este proceso tiene que acabar. Habremos hallado, entonces, una base de F que tendrá como máximo n elementos (donde n es la dimensión de E). Si $\dim F = n$ y $v_1, \ldots, v_n$ es una base de F, por (3.5) también es una base de E. Entonces

$$F = \langle v_1, \ldots, v_n \rangle = E. \square$$

IV.4 Fórmula de Grassmann. Suma directa de subespacios

Sea E un espacio vectorial y F, G dos subespacios de E.

Proposición 4.1 *$F \cap G$ es un subespacio vectorial de E.*

Ejercicio:

Demostrar (4.1).

En general, $F \cup G$ no es un subespacio vectorial de E. El motivo es que la suma de un vector de F y un vector de G puede no pertenecer ni a F ni a G.

Ejemplo:

Consideremos los subespacios $F = \{(x, 0) \mid x \in \mathbf{R}\}$ y $G = \{(0, y) \mid y \in \mathbf{R}\}$ de $\mathbf{R}^2$. La suma $(1, 0) + (0, 1)$ no pertenece ni a F ni a G.

Para evitar trabajar con conjuntos que no son subespacios vectoriales, normalmente consideramos, en lugar de la unión $F \cup G$, el subespacio vectorial generado por esta unión. Este subespacio es precisamente

$$\{u + v \mid u \in F, v \in G\}.$$

En efecto, es fácil ver que este es el menor de los subespacios que contienen a F y a G. Lo llamaremos *suma* de F y G y lo designaremos por $F + G$.

Teorema 4.2 (Fórmula de Grassmann) *Sean F y G dos subespacios vectoriales de E y supongamos que la dimensión de E es finita. Entonces F, G, $F \cap G$ y $F + G$ son todos ellos de dimensión finita y*

$$\dim F + \dim G = \dim(F + G) + \dim(F \cap G).$$

DEMOSTRACIÓN: Por (3.9) todos ellos son de dimensión finita. Sea $u_1, \ldots, u_m$ una base de $F \cap G$. Podemos completar esta base hasta obtener una base de F y una base de G (por (3.8)): $u_1, \ldots, u_m, u_{m+1}, \ldots, u_r$ base de F, $u_1, \ldots, u_m, v_{m+1}, \ldots, v_s$ base de G. Todo vector de la forma $u+v$ con $u \in F$ y $v \in G$ será, pues, combinación lineal de $u_1, \ldots, u_m, u_{m+1}, \ldots, u_r$, $v_{m+1}, \ldots, v_s$. Si demostramos que todos estos vectores son linealmente independientes, habremos obtenido una base de $F+G$ con un número de vectores que demuestra la igualdad del enunciado. Sea pues

$$\sum_{i=1}^{r} a^i u_i + \sum_{i=m+1}^{s} b^i v_i = \vec{0}.$$

Entonces,

$$\sum_{i=1}^{r} a^i u_i = -\sum_{i=m+1}^{s} b^i v_i \in F \cap G,$$

de donde $\sum_{i=m+1}^{s} b^i v_i = \sum_{j=1}^{m} c^j u_j$; es decir, $\sum_{j=1}^{m} c^j u_j + \sum_{i=m+1}^{s} b^i v_i = \vec{0}$ y, dado que $u_1, \ldots, u_m, v_{m+1}, \ldots, v_s$ es una base de G, $c^j = 0$ $(j = 1, \ldots, m)$ y $b^i = 0$ $(i = m+1, \ldots, s)$. Por tanto, $\sum_{i=1}^{r} a^i u_i = \vec{0}$ y, dado que $u_1, \ldots, u_r$ es una base de F,

$$a^i = 0, \qquad i = 1, \ldots, r.$$

Es decir, en la combinación lineal inicial todos los coeficientes son 0. □

Si $F \cap G = \{\vec{0}\}$, diremos que la suma $F+G$ es una *suma directa* y escribiremos

$$F \oplus G.$$

El teorema 4.2 nos dice que la dimensión de $F \oplus G$ es la suma de las dimensiones de los dos subespacios F y G. La proposición que sigue da una caracterización de las sumas directas.

Proposición 4.3 *La suma $F+G$ es directa si y sólo si la expresión de un vector de $F+G$ como suma de un vector de F y un vector de G es única.*

DEMOSTRACIÓN: Demostremos la implicación ($\Rightarrow$): si tenemos dos expresiones $u+v = u_1+v_1$ con $u, u_1 \in F$ y $v, v_1 \in G$, entonces $u-u_1 = v_1 - v \in F \cap G = \vec{0}$, de donde $u-u_1 = v_1 - v = \vec{0}$ y, por tanto, $u = u_1$, $v = v_1$.

Demostremos la implicación ($\Leftarrow$): si $w \in F \cap G$, resulta que $w+\vec{0} = \vec{0}+w$ son dos expresiones del mismo vector de $F+G$. Las dos expresiones deben coincidir. Por tanto, $w = \vec{0}$. □

Proposición 4.4 *Si la dimensión de E es finita, para todo subespacio F hay otro subespacio G tal que $E = F \oplus G$.*

DEMOSTRACIÓN: Sea $u_1, \ldots, u_m$ una base de F. Completémosla a una base de E, $u_1, \ldots, u_m, u_{m+1}, \ldots, u_n$ (3.8). El subespacio $G = \langle u_{m+1}, \ldots, u_n \rangle$ cumple el enunciado de la proposición. □

El subespacio mencionado en (4.4) se llama un *complementario de F*. Un subespacio F tiene, en general, muchos complementarios.

Todo lo que hemos hecho en este apartado puede ser generalizado a un número finito de subespacios $F_1, \ldots, F_k$. En el caso de la intersección, la generalización de (4.1) es clara: $F_1 \cap \ldots \cap F_k$ es siempre un subespacio vectorial. La unión $F_1 \cup \ldots \cup F_k$ no es siempre un subespacio vectorial; definimos la *suma*

$$F_1 + \ldots + F_k$$

como el subespacio generado por $F_1 \cup \ldots \cup F_k$. Resulta que

$$F_1 + \ldots + F_k = \{v_1 + \ldots + v_k \mid v_i \in F_i,\ i = 1, \ldots, k\}.$$

La generalización de la suma directa presenta más dificultades. La forma correcta de hacerlo es por la vía de (4.3). Así pues, diremos que la suma $F_1 + \ldots + F_k$ es *suma directa* y escribiremos

$$F_1 \oplus \ldots \oplus F_k$$

si la expresión de todo vector de $F_1 + \ldots + F_k$ como suma de vectores de $F_1, \ldots, F_k$ es única.

Ejercicio:

Demostrar que $F_1 + \ldots + F_k$ es una suma directa si y sólo si, para todo $i = 1, \ldots, k$, $F_i \cap (F_1 + \ldots + F_{i-1} + F_{i+1} + \ldots + F_k) = \{\vec{0}\}$.

IV.5 Suma directa de espacios vectoriales

Sean E y F dos espacios vectoriales sobre el mismo cuerpo K. Llamaremos *suma directa* de E y F al conjunto $E \times F$ con las operaciones

$$(u, v) + (u_1, v_1) = (u + u_1, v + v_1)$$

$$k(u, v) = (ku, kv),$$

donde $u, u_1 \in E$, $v, v_1 \in F$ y $k \in K$. Con estas operaciones $E \times F$ es un espacio vectorial, que designaremos por

$$E \oplus F.$$

Ejemplo:

$$K^n = K \overbrace{\oplus \ldots \oplus} K.$$

Proposición 5.1 *Si E y F son de dimensión finita, $E \oplus F$ también lo es, y* $\dim E \oplus F = \dim E + \dim F$.

DEMOSTRACIÓN: Sea $u_1, \ldots, u_n$ una base de E y $v_1, \ldots, v_m$ una base de F. Entonces $(u_1, \vec{0}), \ldots, (u_n, \vec{0}), (\vec{0}, v_1), \ldots, (\vec{0}, v_m)$ es una base de $E \oplus F$. En efecto: estos vectores generan $E \oplus F$, ya que si $(u, v) \in E$ tenemos

$$(u, v) = (u, \vec{0}) + (\vec{0}, v) = (\sum_{i=1}^{n} a^i u_i, \vec{0}) + (\vec{0}, \sum_{j=1}^{m} b^j v_j) = \sum_{i=1}^{n} a^i (u_i, \vec{0}) + \sum_{j=1}^{m} b^j (\vec{0}, v_j),$$

y son linealmente independientes, ya que si

$$\sum_{i=1}^{n} a^i (u_i, \vec{0}) + \sum_{j=1}^{m} b^j (\vec{0}, v_j) = (\vec{0}, \vec{0})$$

entonces

$$(\sum_{i=1}^{n} a^i u_i, \sum_{j=1}^{m} b^j v_j) = (\vec{0}, \vec{0}),$$

lo que implica $\sum_{i=1}^{n} a^i u_i = \vec{0}$, de donde $a^i = 0 \; \forall i$, y $\sum_{j=1}^{m} b^j v_j = \vec{0}$, de donde $b^j = 0 \; \forall j$. □

Tanto el nombre como la proposición 5.1 sugieren una relación entre la suma directa de espacios y la suma directa de subespacios vectoriales de un espacio E. Hablaremos de esta relación en el ejemplo 6 de (V.1).

IV.6 Espacio vectorial cociente

Sea E un espacio vectorial y F un subespacio vectorial de E. Diremos que $u, v \in E$ están *relacionados módulo* F si $u - v \in F$. Esta relación es de equivalencia (I.4) y se puede formar el correspondiente conjunto cociente, que designaremos por E/F.

La clase $[u]$ de un vector $u \in E$ es el conjunto $\{u + v \mid v \in F\}$ y la denotaremos también por $u + F$.

Si u y v son equivalentes módulo F a u_1 y v_1, respectivamente, entonces $(u + v)$ y $(u_1 + v_1)$ son equivalentes módulo F. Podemos, pues, definir una operación suma

$$[u] + [v] = [u + v],$$

que no depende de los representantes. Análogamente, la operación

$$k[u] = [ku], \qquad k \in K,$$

no depende de los representantes. Estas dos operaciones hacen de E/F un espacio vectorial sobre K.

Proposición 6.1 *Si E es de dimensión finita, E/F también lo es y*

$$\dim E/F = \dim E - \dim F.$$

DEMOSTRACIÓN: Completemos una base de F, $u_1, \ldots, u_m$, hasta obtener una base de E : $u_1, \ldots, u_m, u_{m+1}, \ldots, u_n$. Para $i = 1, \ldots, m$, $[u_i] = [\vec{0}]$; para $i > m$, las clases

$$[u_{m+1}], \ldots, [u_n]$$

forman una base de E/F; en efecto, toda clase $[u]$ de un vector $u = \sum_{i=1}^{n} a^i u_i$ se puede escribir como $[u] = \sum_{i=m+1}^{n} a^i [u_i]$ y, por tanto, estas clases generan E/F. Para ver que son linealmente independientes, consideremos

$$\sum_{i=m+1}^{n} a^i [u_i] = [\vec{0}].$$

Entonces $[\sum_{i=m+1}^{n} a^i u_i] = [\vec{0}]$, de donde $\sum_{i=m+1}^{n} a^i u_i \in F$. Este vector se puede expresar en la base de F: $\sum_{i=m+1}^{n} a^i u_i = \sum_{j=1}^{m} b^j u_j$; es decir,

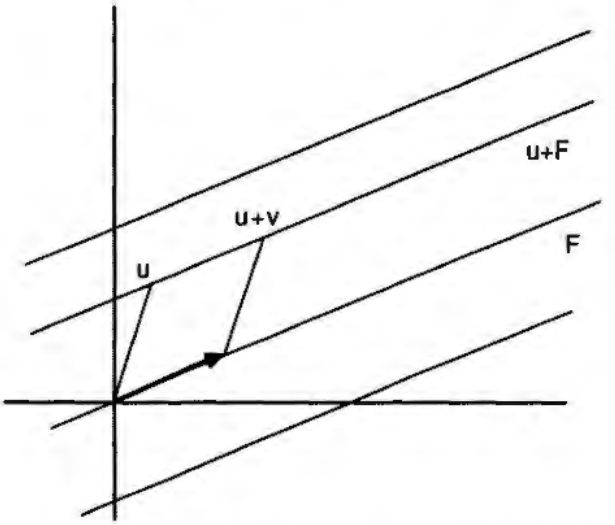
$$\sum_{j=1}^{m} b^j u_j - \sum_{i=m+1}^{n} a^i u_i = \vec{0}.$$

Los vectores que aparecen aquí forman una base de E y, por tanto, los coeficientes han de anularse. En particular, $a^i = 0$ para $i = m+1, \ldots, n$.

Hemos obtenido una base de E/F, demostrando así la igualdad del enunciado. □

Ejemplos:

1. Consideremos un subespacio $F = \langle v \rangle$ de $\mathbf{R}^2$. Si representamos los elementos de $\mathbf{R}^2$ como puntos del plano, los vectores de F corresponden a los puntos de una recta que pasa por el origen. Cada una de las clases $u + F$ corresponde, entonces, a puntos de una recta paralela a F que pasa por u. El conjunto $\mathbf{R}^2/F$ es, en este caso, el conjunto de rectas paralelas a F. Análogamente, si $F = \langle v \rangle \subset \mathbf{R}^3$, $\mathbf{R}^3/F$ es el conjunto de rectas de $\mathbf{R}^3$ paralelas a una recta F que pasa por el origen. Si $F = \langle u, v \rangle \subset \mathbf{R}^3$ es de dimensión 2, F corresponde a un plano que pasa por el origen y $\mathbf{R}^3/F$ es el conjunto de planos paralelos a F.



2. Designemos por $\mathbf{R}_n[x]$ los polinomios de grado $\leq n$ de $\mathbf{R}[x]$. Dos polinomios $p(x) = p_0 + p_1 x + \ldots + p_r x^r$ y $q(x) = q_0 + q_1 x + \ldots + q_s x^s$ determinan la misma clase módulo $\mathbf{R}_n[x]$ si y sólo si $p_i = q_i$ para $i > n$.

IV.7 Coordenadas

Fijada una base $u_1, \ldots, u_n$ en un espacio vectorial E, la expresión de cada vector v de E

$$v = a^1 u_1 + \ldots + a^n u_n$$

es única (3.2). Diremos entonces que $(a^1, \ldots, a^n)$ son las *coordenadas de v en la base $u_1, \ldots, u_n$*. Naturalmente, las coordenadas de v en otra base $e_1, \ldots, e_n$ serán distintas:

$$v = b^1 e_1 + \ldots + b^n e_n.$$

¿Qué relación hay entre $(a^1, \ldots, a^n)$ y $(b^1, \ldots, b^n)$? Supongamos que

$$e_i = \sum_{j=1}^{n} p_i^j u_j, \quad i = 1, \ldots, n.$$

Tenemos

$$v = \sum_{i=1}^{n} b^i e_i = \sum_{i=1}^{n} b^i \left(\sum_{j=1}^{n} p_i^j u_j\right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} b^i p_i^j\right) u_j,$$

de donde

$$a^j = \sum_{i=1}^{n} b^i p_i^j, \quad j = 1, \ldots, n.$$

Podemos escribir estas igualdades de una manera mucho más cómoda utilizando el producto de matrices. Concretamente, si $A = (a_i^j)$ es una matriz de n columnas y m filas, y $B = (b_i^j)$ es una matriz de r columnas y n filas, el *producto de A por B* es una matriz $C = (c_i^j)$,

$$C = AB,$$

de r columnas y m filas, en la cual

$$c_i^j = \sum_{k=1}^{n} a_k^j b_i^k.$$

Consideremos en nuestro caso las matrices

$$P = \begin{pmatrix} p_1^1 & \dots & p_n^1 \\ \vdots & & \vdots \\ p_1^n & \dots & p_n^n \end{pmatrix}, \quad A = \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}, \quad B = \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix}.$$

Las igualdades que relacionan las coordenadas de V en una y otra base se pueden resumir así:

$$A = PB.$$

La matriz P se llama la *matriz del cambio de base.*

Efectuemos ahora un nuevo cambio de base. Sea $v_1, \dots, v_n$ una nueva base:

$$v_j = \sum_{i=1}^{n} q_j^i e_i, \qquad j = 1, \dots, n.$$

$Q = (q_j^i)$ es la matriz del nuevo cambio de base. ¿Qué relación hay entre las matrices P y Q de estos dos cambios de base sucesivos y la matriz del cambio directo de la base $u_1, \dots, u_n$ a la base $v_1, \dots, v_n$? Solamente hace falta hallar las coordenadas de los vectores v_j en la base $u_1, \dots, u_n$:

$$v_j = \sum_{i=1}^{n} q_j^i e_i = \sum_{i=1}^{n} q_j^i \left(\sum_{k=1}^{n} p_i^k u_k \right) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} p_i^k q_j^i \right) u_k.$$

Estos coeficientes son, precisamente, los de la matriz PQ; esta es, pues, la matriz del cambio directo.

¿Qué pasa si, en estos dos cambios sucesivos, la tercera base vuelve a ser la primera, $v_1 = u_1, \dots, v_n = u_n$? El cambio directo tiene, evidentemente, la matriz

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

A esta matriz la designaremos por I y la llamaremos *matriz identidad*, ya que cumple

$$MI = IM = M$$

para toda matriz $n \times n$ M. En este caso resulta, pues, que

$$PQ = I.$$

Análogamente, si efectuamos primero el cambio de $e_1, \ldots, e_n$ a $u_1, \ldots, u_n$ de matriz Q, y después el cambio de $u_1, \ldots, u_n$ a $e_1, \ldots, e_n$ de matriz P, obtenemos

$$QP = I.$$

Dos matrices cuyo producto, en cualquier orden, es siempre I, se llaman *inversas* una de la otra. P y Q son una inversa de la otra. Escribiremos

$$P = Q^{-1} \quad \text{y} \quad Q = P^{-1}$$

Hemos demostrado, pues, la

Proposición 7.1 *Las matrices de los cambios de base son siempre invertibles.* □

IV.8 Nota histórica

Las primeras ideas sobre coordenadas son de René Descartes (1596–1650), quien introduce la notación $x^1, x^2, x^3, \ldots$ y la regla de los signos para polinomios en el *Discours de la méthode*, "...pour bien conduire la raison et chercher la vérité dans les sciences". La consideración de conceptos en dimensión n, así como la suma de vectores (al identificar $\mathbf{R}^2$ con $\mathbf{C}$ en la demostración del teorema fundamental del álgebra) son debidas a Carl Friedrich Gauss (1777–1855). Arthur Cayley (1821–1895) y Hermann G. Grassmann (1809–1877) —este último, un maestro de escuela sin formación científica— ya utilizan los espacios vectoriales (que no serían definidos axiomáticamente hasta 1888 por Giuseppe Peano (1858–1932)). Grassmann introduce los conceptos de subespacio, generadores, dimensión, suma, intersección, así como las fórmulas para el cambio de coordenadas. El concepto y el nombre de matriz fueron introducidos por James Joseph Sylvester (1814–1897) en el año 1850, con posterioridad al estudio de los determinantes (ver Cap. VI). (Léase "Els orígens físics de l'anàlisi vectorial" en *El desenvolupament de les matemàtiques al segle 19*, Institut d'Estudis Catalans, Arxius de la Secció de Ciències, LXXV, 1984.)

IV.9 Ejercicios

1. Demostrar que el conjunto E de las aplicaciones $f : \mathbf{R} \longrightarrow \mathbf{R}$ puede dotarse de manera natural de estructura de espacio vectorial real.

a) ¿Cuáles de las siguientes familias de elementos de E son linealmente independientes?

(a) $\operatorname{sen} x$, $\cos x$, 1.
(b) e^x, e^{x+2}.
(c) 2, $x+2$, x^2.
(d) 0, 1, $x+1$.

b) Expresar (si es posible) los siguientes elementos de E como combinación lineal de las familias correspondientes:

(a) $\operatorname{sen} x$; $1, x, x^2, \ldots$
(b) x^2+x-1; $1, x-1, (x-1)^2$.
(c) 1; $x+1, x^2-1, x^3+1$.
(d) 0; $(x-1)^2, x, x^2+2, \sqrt{3}$.

c) ¿Cuáles de los siguientes subconjuntos de E son subespacios vectoriales?

$$E_1 = \{f \in E \mid f(-x) = f(x) \quad \forall x \in \mathbf{R}\}$$
$$E_2 = \{f \in E \mid f(-x) = -f(x) \quad \forall x \in \mathbf{R}\}$$
$$E_3 = \{f \in E \mid f \text{ es continua}\}$$
$$E_4 = \{f \in E \mid f(0) = f(1)\}$$
$$E_5 = \{f \in E \mid f \text{ es dos veces derivable y } f'' - f' + f = 0\}$$

2. Sea $E = M_{n\times m}(\mathbf{R})$. ¿Cuáles de los siguientes subconjuntos de E son subespacios vectoriales?

$$E_1 = \{A \in E \mid a_1^1 = 0\}$$
$$E_2 = \{A \in E \mid a_1^1 + a_1^2 = 0\}$$
$$E_3 = \{A \in E \mid \textstyle\sum_{i=1}^{n} a_i^i = 0,\ n = m\}$$
$$E_4 = \{A \in E \mid a_j^i = a_i^j \ \ \forall i, j,\ n = m\}$$
$$E_5 = \{A \in E \mid a_i^j = a_r^s \ \ \forall i, j, r, s\}$$
$$E_6 = \{A \in E \mid \textstyle\sum_{j=1}^{m} a_i^j = 2\pi,\ i = 1, \ldots, n\}.$$

3. Sea $p(x)$ un polinomio de $K[x]$ de grado n. Demostrar que $K[x]/(p(x))$ tiene una estructura natural de espacio vectorial en la que el conjunto $[1], [x], \ldots, [x^{n-1}]$ forma base.

4. Demostrar que el conjunto de las sucesiones de elementos de un cuerpo K con las operaciones

$$(x_1, x_2, \ldots, x_n, \ldots) + (y_1, y_2, \ldots, y_n, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n, \ldots)$$

$$k(x_1, x_2, \ldots, x_n, \ldots) = (kx_1, kx_2, \ldots, kx_n, \ldots)$$

es un espacio vectorial sobre K. Consideremos el conjunto S de las sucesiones con todos los elementos 0 salvo uno que valga 1. ¿Es S linealmente independiente? ¿Es S una base?

5. Sea F el subespacio de las sucesiones (a_n) tales que $a_k = a_{k-1} + 2a_{k-2}$. Hallar las progresiones geométricas contenidas en F y comprobar que existen bases de F formadas por progresiones geométricas. Expresar cualquier otra sucesión de F como combinación lineal de una de esas bases y encontrar así el término general de las sucesiones de F.

6. Sea $K_n[x]$ el conjunto de los polinomios de grado $\leq n$ junto con el 0. Dado $0 \neq p(x) \in K_n[x]$, designemos por $p'(x), p''(x), \ldots, p^{(n)}(x)$ sus derivadas. Demostrar que $p(x), p'(x), p''(x), \ldots, p^{(n)}(x)$ forman una base de $K_n[x]$. Escribir $1, x, x^2, \ldots, x^n$ como combinación lineal de esta base, en el caso $p(x) = 1 + x + x^2 + \ldots + x^n$.

7. Dada una matriz $A = (a_i^j) \in M_{n\times n}(K)$, llamaremos *traspuesta* de A a la matriz $A^t = (b_i^j)$ tal que $b_i^j = a_j^i$ para todo i, j. Consideremos los conjuntos

$$S = \{A \in M_{n\times n}(K) \mid A = A^t\} \quad \text{(matrices } \textit{simétricas}\text{)}$$
$$H = \{A \in M_{n\times n}(K) \mid A = -A^t\} \quad \text{(matrices } \textit{hemisimétricas}\text{)}.$$

Demostrar que S y H son subespacios vectoriales de $M_{n\times n}(K)$ y que, si en K se cumple $2 = 1 + 1 \neq 0$, entonces $M_{n\times n}(K) = S \oplus M$. ¿Qué pasa si $K = \mathbf{Z}/(2)$?

8. Hallar un sistema homogéneo de ecuaciones lineales que tenga como soluciones los elementos de $F = ((1,1,1,1),(2,1,0,3))$.

9. Hallar una base del espacio vectorial de las soluciones del sistema de tres ecuaciones: $x - y = 0$, $2x + y + z = 0$, $x + y - z = 0$. Suponer que los coeficientes son elementos de: a) $\mathbf{R}$; b) $\mathbf{Z}/(2)$; c) $\mathbf{Z}/(5)$.

10. Considerar las inclusiones $\mathbf{Q} \subset \mathbf{Q}(\sqrt[3]{5}) \subset \mathbf{R} \subset \mathbf{C}$ (ver el ejemplo de (II.7)). Demostrar que cada uno de estos cuerpos es un espacio vectorial sobre el anterior y determinar sus dimensiones.

11. Sean

$$M = \{\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid a,b,c \in \mathbf{R}\},\ M' = \{\begin{pmatrix} a-b & 2a \\ b & a+2b \end{pmatrix} \mid a,b \in \mathbf{R}\}.$$

Hallar una base de $M_{2\times 2}(\mathbf{R})/M$ y de $M_{2\times 2}(\mathbf{R})/M'$.

12. Sean F, G, H subespacios del espacio vectorial E. Demostrar o dar contraejemplos de las afirmaciones siguientes:

a) $F \cap (G + H) = (F \cap G) + (F \cap H)$.

b) $F + (G \cap H) = (F + G) \cap (F + H)$.

c) $\dim(F \cap (G+H)) = \dim(F \cap G) + \dim(F \cap H) + \dim(F \cap H \cap G)$.

13. ¿Qué condiciones deben cumplir a, b, c para que los vectores de $\mathbf{R}^3$ (a, a^2, a^3), (b, b^2, b^3), (c, c^2, c^3) sean linealmente independientes?

14. Consideremos en $\mathbf{R}^4$ los subespacios $F = \langle a, b, c\rangle$ y $G = \langle d, e\rangle$ donde $a = (1,2,3,4)$, $b = (2,2,2,6)$, $c = (0,2,4,4)$, $d = (1,0,-1,2)$ y $e = (2,3,0,1)$.

a) Determinar las dimensiones de $F, G, F \cap G$ y $F + G$ y dar una base de cada subespacio.

b) Dar sendas bases de $\mathbf{R}^4/F$ y de $\mathbf{R}^4/G$.

IV.10 Ejercicios para programar

15. Preparar un programa para calcular, sobre $\mathbf{R}$,

a) la suma de dos matrices;

b) el producto de una matriz por un escalar;

c) el producto de dos matrices.

16. Modificar el programa anterior de manera que en lugar de trabajar con $K = \mathbf{R}$ lo haga con

a) $K = \mathbf{C}$;

b) $K = \mathbf{Z}/(p)$.

Nota:

Conviene preparar estos programas en forma de subprogramas que puedan ser utilizados cuando sea necesario dentro de programas mayores.

17. Descomposición *LU*

Sea $A = (a_i^j) \in M_{n\times n}(\mathbf{R})$ una matriz dada. Preparar un programa que descomponga $A = LU$, donde $L = (l_i^j)$ es una matriz triangular inferior ($l_i^j = 0\ \forall i > j$) con unos en la diagonal principal ($l_i^i = 1$, $i = 1,\ldots,n$) y $U = (u_i^j)$ es una matriz triangular superior ($u_i^j = 0$ $\forall i < j$). (Indicación: efectuando el producto de las matrices L y U obtenemos unas ciertas fórmulas, que son las que hay que programar.) (Ver Cap. VI, Ejercicios de programación, para una ampliación de este ejercicio.)

El ejercicio siguiente utiliza el ejercicio VII.12.

18. Si $e_1,\ldots,e_m$ es una familia cualquiera de vectores de $\mathbf{R}^n$, preparar un programa que permita extraer una familia linealmente independiente y completarla a una base de $\mathbf{R}^n$ con vectores de la base canónica. (Indicación: seguir los pasos de la demostración del teorema de Steinitz.)

Capítulo V

Aplicaciones lineales

V.1 Definición y ejemplos

En el capítulo anterior nos hemos ocupado del estudio de unas ciertas estructuras, los espacios vectoriales. En éste estudiaremos aplicaciones entre esas estructuras, las aplicaciones lineales. Resulta lógico pensar que las aplicaciones interesantes entre espacios vectoriales son aquellas que respetan la estructura de espacio vectorial. Un espacio vectorial es un conjunto con dos operaciones; una aplicación entre los conjuntos que "conserva" las dos operaciones es una "aplicación que respeta la estructura vectorial".

Sean E y F dos espacios vectoriales sobre el mismo cuerpo K. Una aplicación

$$f : E \longrightarrow F$$

se llama una *aplicación lineal* si para todo $u, v \in E$ y todo $a \in K$,

$$\begin{aligned} f(u+v) &= f(u) + f(v) \\ f(au) &= af(u). \end{aligned}$$

Si f es lineal, se cumple

- $f(au + bv) = af(u) + bf(v)$ y, en general,
- $f\left(\sum_{i=1}^{m} a^i u_i\right) = \sum_{i=1}^{m} a^i f(u_i)$.
- $f(\vec{0}) = \vec{0}$, ya que de $\vec{0} = 0v$ se deduce $f(\vec{0}) = 0f(v) = \vec{0}$.
- $f(-v) = -f(v)$, ya que $f(v) + f(-v) = f(v + (-v)) = f(\vec{0}) = \vec{0}$.
- Si $f : E \longrightarrow F$ y $g : F \longrightarrow G$ son lineales, $g \circ f : E \longrightarrow G$ es lineal.

Ejemplos:

1. Sea E un espacio vectorial sobre K y $u_1, \ldots, u_n$ una base de E. La aplicación
$$\begin{array}{ccc} E & \to & K^n \\ v & \mapsto & (a^1, \ldots, a^n) \end{array}$$
que hace corresponder a cada vector sus coordenadas en la base dada es lineal.
2. Sea $E_1 \oplus E_2$ la suma directa de los espacios vectoriales E_1 y E_2. Las aplicaciones
$$\begin{array}{ccc} E_1 & \to & E_1 \oplus E_2 \\ u & \mapsto & (u, \vec{0}) \end{array} \qquad \begin{array}{ccc} E_2 & \to & E_1 \oplus E_2 \\ v & \mapsto & (\vec{0}, v) \end{array}$$
$$\begin{array}{ccc} E_1 \oplus E_2 & \to & E_1 \\ (u, v) & \mapsto & u \end{array} \qquad \begin{array}{ccc} E_1 \oplus E_2 & \to & E_2 \\ (u, v) & \mapsto & v \end{array}$$
son lineales.
3. Sea F un subespacio del espacio vectorial E. La aplicación
$$\begin{array}{ccc} E & \to & E/F \\ u & \mapsto & [u] \end{array}$$
que hace corresponder a cada vector su clase es lineal.
4. Sean $a_1, \ldots, a_n$ números reales fijos. La aplicación
$$\begin{array}{ccc} \mathbf{R}^n & \longrightarrow & \mathbf{R} \\ (x^1, \ldots, x^n) & \longmapsto & a_1 x^1 + \ldots + a_n x^n \end{array}$$
es lineal. Más en general, si a_i^j, $i = 1, \ldots, n$, $j = 1, \ldots, m$, son números reales fijos, la aplicación
$$\begin{array}{ccc} \mathbf{R}^n & \longrightarrow & \mathbf{R}^m \\ (x^1, \ldots, x^n) & \longmapsto & (\sum_{i=1}^{n} a_i^1 x^i, \ldots, \sum_{i=1}^{n} a_i^m x^i) \end{array}$$
es lineal.

Sea $f : E \to F$ una aplicación lineal.

- Denominaremos *núcleo de* f al subespacio vectorial de E
$$\operatorname{Nuc} f = \{u \in E \mid f(u) = \vec{0}\}.$$
- Denominaremos *imagen de* f al subespacio vectorial de F
$$\operatorname{Im} f = \{v \in F \mid \text{ existe } u \in E \text{ tal que } v = f(u)\}.$$

Ejercicio:

Demostrar que, efectivamente, Nuc f e Im f son subespacios vectoriales.

Proposición 1.1 *Si la aplicación $f : E \to F$ es lineal y E es de dimensión finita, entonces* Nuc f *e* Im f *son de dimensión finita y* $\dim E = \dim \operatorname{Nuc} f + \dim \operatorname{Im} f$.

DEMOSTRACIÓN: Nuc f es un subespacio vectorial de E y, por (IV.3.9), es de dimensión finita. Tomemos una base de Nuc f, $v_1, \ldots, v_k$, y completémosla hasta obtener una base de E (IV.3.8): $v_1, \ldots, v_k, \ldots, v_n$. Las imágenes por f de los k primeros vectores son $\vec{0}$. Las imágenes

$$f(v_{k+1}), \ldots, f(v_n)$$

forman una base de Im f. En efecto, generan Im f, ya que si $v \in \operatorname{Im} f$, existe un $u = \sum_{i=1}^{n} a^i v_i \in E$ tal que

$$v = f(u) = f\left(\sum_{i=1}^{n} a^i v_i\right) = \sum_{i=1}^{n} a^i f(v_i) = \sum_{i=k+1}^{n} a^i f(v_i).$$

Además, son linealmente independientes, ya que

$$\vec{0} = \sum_{i=k+1}^{n} a^i f(v_i) = f\left(\sum_{i=k+1}^{n} a^i v_i\right) \Rightarrow \sum_{i=k+1}^{n} a^i v_i \in \operatorname{Nuc} f \Rightarrow$$

$$\Rightarrow \sum_{i=k+1}^{n} a^i v_i = \sum_{i=1}^{k} b^i v_i \Rightarrow \sum_{i=1}^{k} b^i v_i - \sum_{i=k+1}^{n} a^i v_i = \vec{0}.$$

Los vectores que aparecen en esta combinación lineal son linealmente independientes; por tanto, los coeficientes son 0. En particular, $a^i = 0$ para $i = k+1, \ldots, n$. □

Se llama *rango de una aplicación lineal* f a la dimensión de su imagen.

Una aplicación lineal inyectiva se llama un *monomorfismo*. Una aplicación lineal exhaustiva se llama un *epimorfismo*. Una aplicación lineal biyectiva se llama un *isomorfismo*. Una aplicación lineal de un espacio E en sí mismo, $E \to E$, se llama un *endomorfismo*. Un endomorfismo biyectivo se llama un *automorfismo*.

Proposición 1.2 *Una aplicación lineal f es inyectiva si y sólo si* $\operatorname{Nuc} f = \vec{0}$. *Una aplicación lineal f es exhaustiva si y sólo si* $\operatorname{Im} f = F$.

DEMOSTRACIÓN: La segunda afirmación no es más que la definición de exhaustividad. Demostremos la primera:

$\Longrightarrow$) $u \in \operatorname{Nuc} f \Rightarrow f(u) = \vec{0} = f(\vec{0}) \Rightarrow$ (por ser f inyectiva) $u = \vec{0}$.

$\Longleftarrow$) $f(u) = f(v) \Rightarrow f(u - v) = \vec{0}$ (por ser f lineal) $\Rightarrow$
$\Rightarrow u - v \in \operatorname{Nuc} f \Rightarrow u - v = \vec{0} \Rightarrow u = v$. □

Proposición 1.3 *Si f es lineal y biyectiva, entonces f^{-1} también es lineal.*

DEMOSTRACIÓN: Para probar que $f^{-1}(u + v) = f^{-1}(u) + f^{-1}(v)$, basta ver que, por la aplicación biyectiva f, los dos miembros de esta expresión tienen la misma imagen,

$$f(f^{-1}(u + v)) = u + v = ff^{-1}(u) + ff^{-1}(v) = f(f^{-1}(u) + f^{-1}(v)).$$

De manera parecida se ve que $f^{-1}(au) = af^{-1}(u)$. □

Observación:

Esta proposición nos dice que si $f : E \to F$ es un isomorfismo, hay una aplicación lineal $g : F \to E$ ($g = f^{-1}$) tal que $g \circ f = I_E$ y $f \circ g = I_F$. (Aquí, $I_E : E \to E$ es la aplicación identidad, que envía todo elemento a sí mismo; análogamente para I_F.)

Recíprocamente, si existe una aplicación lineal $g : F \to E$ tal que $g \circ f = I_E$ y $f \circ g = I_F$, resulta que f es inyectiva y exhaustiva y, por tanto, es biyectiva e inversa de g. Entonces (1.3) nos dice que f es un isomorfismo.

Dos espacios vectoriales E y F se llaman *isomorfos* si existe un isomorfismo $E \to F$. Escribiremos entonces

$$E \cong F.$$

Ejemplos:

5. Si E es un espacio vectorial de dimensión n sobre K, entonces $E \cong K^n$ (ejemplo 1).
6. Sea $E_1 \oplus E_2$ la suma directa de dos espacios vectoriales E_1 y E_2. Consideremos los subespacios

$$E_1' = \{(u, \vec{0}) \mid u \in E_1\}, \qquad E_2' = \{(\vec{0}, v) \mid v \in E_2\}.$$

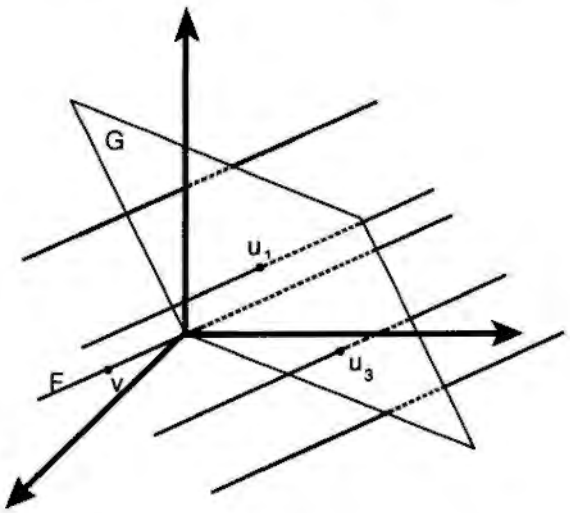
Las aplicaciones

$$\begin{array}{ccc} E_1 & \to & E_1' \\ u & \mapsto & (u, \vec{0}) \end{array} \qquad \begin{array}{ccc} E_2 & \to & E_2' \\ v & \mapsto & (\vec{0}, v) \end{array}$$

son isomorfismos. Los subespacios E_1' y E_2' tienen intersección $\{(\vec{0}, \vec{0})\}$ y su suma es $E_1 \oplus E_2$; es decir,

$$E_1 \oplus E_2 = E_1' \oplus E_2' \quad \text{con} \quad E_i' \cong E_i,\ i = 1, 2.$$

Toda suma directa de espacios es también, pues, suma directa de subespacios isomorfos a los espacios de partida.



7. Sean F y G dos subespacios complementarios en E; es decir, $F \oplus G = E$. En la figura hemos representado el caso particular en que $F = \langle v \rangle$ es una recta y $G = \langle u_1, u_2 \rangle$ un plano. Así pues, $\mathbf{R}^3/F$ es el conjunto de rectas paralelas a F (IV.6). Este conjunto está claramente en correspondencia biyectiva con los puntos del plano G: a cada punto de G le asociamos la recta paralela a F que pasa por ese punto. En general,

$$\begin{array}{ccc} G & \to & E/F \\ u & \mapsto & [u] \end{array}$$

es una aplicación lineal de núcleo $G \cap F = \{\vec{0}\}$ y exhaustiva, ya que la clase de un $w \in E, w = u + v, u \in G, v \in F$, es $[w] = [u]$ y, por tanto, es imagen del vector u de G. Los espacios G y E/F son, pues, isomorfos.

Proposición 1.4 *Dos espacios vectoriales de dimensión finita E y F son isomorfos si y sólo si* $\dim E = \dim F$.

DEMOSTRACIÓN: Si $E \cong F$, tienen la misma dimensión como consecuencia inmediata de (1.2) y (1.1).

Supongamos que $\dim E = \dim F$. Escojamos bases $u_1, \ldots, u_n$ de E y $v_1, \ldots, v_n$ de F. La aplicación

$$f : E \to F$$

definida por

$$f\left(\sum_{i=1}^{n} a^i u_i\right) = \sum_{i=1}^{n} a^i v_i$$

es claramente lineal y exhaustiva. Para probar la inyectividad, utilizamos (1.2). □

Ejemplo:

$\mathbf{R}^2$ y $\mathbf{C}$ son espacios vectoriales sobre $\mathbf{R}$ isomorfos. Un isomorfismo es

$$(a, b) \mapsto a + bi.$$

V.2 Matriz asociada a una aplicación lineal

Una aplicación lineal queda totalmente determinada por las imágenes de los vectores de una base. Esas imágenes pueden ser, no obstante, arbitrarias. Demostremos estas afirmaciones.

Proposición 2.1 *Sea $B = \{u_i \mid i \in I\}$ una base de un espacio vectorial E sobre el cuerpo K. Sean $\{w_i \mid i \in I\}$ vectores cualesquiera de un espacio vectorial F sobre K. Existe una aplicación lineal, y sólo una,*

$$f : E \to F,$$

tal que $f(u_i) = w_i$ para cada $i \in I$.

DEMOSTRACIÓN: Puesto que una aplicación lineal cumple

$$f\left(\sum_{i=1}^{n} a^i u_i\right) = \sum_{i=1}^{n} a^i f(u_i),$$

la aplicación que buscamos debe definirse por

$$f\left(\sum_{i=1}^{n} a^i u_i\right) = \sum_{i=1}^{n} a^i w_i.$$

Sólo falta demostrar que, así definida, f es lineal. Se deja como ejercicio. □

Los siguientes hechos son fáciles de demostrar (notación como en (2.1)):

f es inyectiva $\Leftrightarrow$ $\{w_i \mid i \in I\}$ es linealmente independiente;

f es exhaustiva $\Leftrightarrow$ $\{w_i \mid i \in I\}$ genera F;

f es biyectiva $\Leftrightarrow$ $\{w_i \mid i \in I\}$ es una base de F.

Sea $f : E \to F$ lineal con E y F de dimensión finita. Sean $u_1, \ldots, u_n$ y $v_1, \ldots, v_n$ bases de E y F respectivamente. Entonces (2.1) nos dice que f queda determinada si conocemos las coordenadas de $f(u_1), \ldots, f(u_n)$:

$$f(u_i) = \sum_{j=1}^{m} a_i^j v_j, \qquad i = 1, \ldots, n.$$

Se llama *matriz asociada a f respecto a las bases $\{u_i\}, \{v_j\}$* a

$$A = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix}.$$

Observemos que (2.1) nos dice, también, que toda matriz $m \times n$ es la matriz asociada a una aplicación lineal respecto a las bases $\{u_i\}$ y $\{v_j\}$.

La matriz asociada es la herramienta con que, muy a menudo, estudiamos una aplicación lineal. Veamos, de momento, cómo esta matriz nos permite calcular las coordenadas de la imagen de un vector. Supongamos que $w = \sum_{i=1}^{n} w^i u_i \in E$; entonces,

$$f(w) = \sum_{i=1}^{n} w^i f(u_i) = \sum_{i=1}^{n} w^i \left(\sum_{j=1}^{m} a_i^j v_j\right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_i^j w^i\right) v_j.$$

Las coordenadas $(\bar{w}^1, \ldots, \bar{w}^m)$ de $f(w)$ en la base $\{v_j\}$ son, pues,

$$\bar{w}^j = \sum_{i=1}^{n} a_i^j w^i, \qquad j = 1, \ldots, m.$$

Escribamos las coordenadas $(w^1, \ldots, w^n)$ y $(\bar{w}^1, \ldots, \bar{w}^m)$ como matrices de una columna W y $\bar{W}$, respectivamente. Entonces las igualdades anteriores se pueden resumir en

$$\bar{W} = AW.$$

Ejemplos:

1. La aplicación

$$\begin{array}{rcl} f : \mathbf{R}^n & \longrightarrow & \mathbf{R}^m \\ (x^1, \ldots, x^n) & \longmapsto & (\sum_{i=1}^n a_i^1 x^i, \ldots, \sum_{i=1}^n a_i^m x^i) \end{array}$$

en las bases $\{e_i = (0, \ldots, 1, \ldots, 0)\}$ tiene por matriz

$$\begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix}.$$

2. Sea F un subespacio de E y

$$f : E \to E/F$$

la aplicación que hace corresponder a cada vector $u \in E$ su clase $[u]$ módulo F. Sea $v_1, \ldots, v_k$ una base del subespacio F y $v_1, \ldots, v_k, \ldots, v_n$ una base de E; sabemos que, en esta situaciòn, $[v_{k+1}], \ldots, [v_n]$ es una base de E/F. Respecto a estas bases, la matriz de f es

$$\begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

k columnas $\quad n - k$ columnas.

3. Consideremos la matriz de la identidad $I_E : E \to E$. Tomemos en el primer espacio E una base $u_1, \ldots, u_n$ y en el segundo una base

$v_1, \ldots, v_n$. La matriz de I_E respecto a estas bases está formada por los coeficientes de

$$I_E(u_i) = u_i = \sum_{j=1}^{n} a_i^j v_j.$$

Si w tiene coordenadas $(w^1, \ldots, w^n)$ en la base $\{u_i\}$, las coordenadas de $w = I_E(w)$ en la base $\{v_i\}$, $(\bar{w}^1, \ldots, \bar{w}^n)$, cumplen

$$\bar{W} = AW.$$

De esta manera volvemos a encontrar la expresión del cambio de coordenadas de (IV.7). La matriz A es la matriz del cambio.

4. La matriz de $I_E : E \to E$, considerando la misma base $\{u_i\}$ en los dos espacios, es la matriz identidad

$$I = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

5. Sea $f : E \to F$ un isomorfismo. Consideremos una base $\{u_i\}$ de E y la base $\{f(u_i)\}$ de F. La matriz de f respecto a estas bases es la matriz identidad.

Proposición 2.2 *Sean $f : E \to F$ y $g : F \to H$ aplicaciones lineales. Sean $\{u_1, \ldots, u_n\}$, $\{v_1, \ldots, v_m\}$, $\{e_1, \ldots, e_s\}$ bases de E, F, H, respectivamente. Sean A, B, C las matrices de f, g, $g \circ f$ respecto a estas bases. Entonces*

$$C = BA.$$

DEMOSTRACIÓN: Tenemos

$$f(u_i) = \sum_{j=1}^{m} a_i^j v_j, \qquad g(v_j) = \sum_{k=1}^{s} b_j^k e_k,$$

$$\begin{aligned} (g \circ f)(u_i) &= g\left(\sum_j a_i^j v_j\right) = \sum_j a_i^j g(v_j) = \\ &= \sum_j a_i^j \left(\sum_k b_j^k e_k\right) = \sum_k \left(\sum_j a_i^j b_j^k\right) e_k. \end{aligned}$$

Por tanto,

$$c_i^j = \sum_j a_i^j b_j^k,$$

es decir,

$$C = BA. \square$$

Corolario 2.3 *El producto de matrices es asociativo.*

DEMOSTRACIÓN: Debemos decir, en primer lugar, que hemos cometido un abuso de lenguaje al usar la expresión "producto de matrices" como si se tratara de una operación. Recordemos (IV.7) que este producto sólo está definido cuando el número de columnas de la primera matriz coincide con el de filas de la segunda. Sean, pues,

$$A \in M_{m\times n}(K), \qquad B \in M_{s\times m}(K), \qquad C \in M_{t\times s}(K);$$

consideremos espacios vectoriales E, F, H, G sobre K de dimensiones n, m, s, t respectivamente. Fijemos bases en cada uno de estos espacios. Existen, entonces, aplicaciones lineales f, g, h con matrices A, B, C respecto a estas bases. El hecho de que $h \circ (g \circ f) = (h \circ g) \circ f$ implica, por (2.2), que

$$C(BA) = (CB)A. \square$$

Queremos ahora relacionar las matrices de una misma aplicación lineal respecto a diferentes bases. Sea, pues,

$$f : E \to F$$

con matriz A respecto a las bases $u_1, \ldots, u_n$ de E y $v_1, \ldots, v_m$ de F, y con matriz B respecto a las bases $\bar{u}_1, \ldots, \bar{u}_n$ de E y $\bar{v}_1, \ldots, \bar{v}_m$ de F. Escribamos f como la composición

$$E \overset{I_E}{\to} E \overset{f}{\to} F \overset{I_F}{\to} F.$$

Consideremos en cada uno de estos cuatro espacios las bases $\{\bar{u}_i\}$, $\{u_i\}$, $\{v_i\}$, $\{\bar{v}_i\}$ respectivamente. La proposición 2.2 nos dice que

$$B = QAP,$$

donde P es la matriz de I_E

$$\bar{u}_i = \sum_{k=1}^{n} p_i^k u_k, \qquad i = 1, \ldots, n$$

y Q es la matriz de I_F

$$v_j = \sum_{k=1}^{m} q_j^k \bar{v}_k, \qquad j = 1, \ldots, m.$$

En muchas ocasiones, lo que conocemos son las coordenadas de la nueva base $\bar{v}_k$ en la base v_j,

$$\bar{v}_k = \sum_{j=1}^{m} r_k^j v_j, \qquad k = 1, \ldots, m;$$

recordemos, sin embargo, que, como vimos en (IV.7), las matrices $R = (r_k^j)$ y Q son inversas una de la otra. Así pues,

$$B = R^{-1} A P.$$

V.3 Teorema de isomorfismo

Teorema 3.1 (de isomorfismo) *Si $f : E \to F$ es lineal, entonces*

$$\operatorname{Im} f \cong E/\operatorname{Nuc} f.$$

DEMOSTRACIÓN: La aplicación f envía todos los elementos de una clase $u + \operatorname{Nuc} f$ al mismo vector de F; en efecto, si $w \in \operatorname{Nuc} f$, $f(w) = \vec{0}$ y

$$f(u + w) = f(u) + f(w) = f(u).$$

Esto nos permite definir una aplicación

$$\begin{array}{ccc} E/\operatorname{Nuc} f & \longrightarrow & \operatorname{Im} f \\ [u] & \longmapsto & f(u) \end{array}$$

que es, claramente, lineal y exhaustiva. Aplicaremos (1.2) para ver que es inyectiva; el que $[u]$ vaya a $\vec{0}$ significa que $f(u) = \vec{0}$, es decir, $u \in \operatorname{Nuc} f$ y $[u] = [\vec{0}]$. □

Corolario 3.2 *Si F y G son subespacios de E, se cumple*

$$(F + G)/F \cong G/F \cap G.$$

DEMOSTRACIÓN: La aplicación

$$\begin{array}{rcl} f : G & \longrightarrow & (F+G)/F \\ v & \longmapsto & [v] \end{array}$$

es lineal; su núcleo está formado por aquellos $v \in G$ tales que $[v] = F$, es decir, tales que $v \in F$; por tanto, $\operatorname{Nuc} f = F \cap G$. Además, f es exhaustiva, ya que, dada $[u]$, podemos escribir $u = w + v$ con $w \in F$, $v \in G$; entonces, $[u] = [v] = f(v)$. Aplicando (3.1), obtenemos

$$G/F \cap G \cong (F+G)/F. \square$$

Corolario 3.3 *Si $F \subset G$ son subespacios de E, entonces*

$$(E/F)/(G/F) \cong E/G.$$

DEMOSTRACIÓN: Observemos primero que $F \subset G$ implica que si una clase $[u]$ de E/F tiene un representante en G, todos sus elementos son de G. Por tanto, G/F es un subconjunto de E/F.

Definimos, ahora,

$$\begin{array}{rcl} f : E/F & \longrightarrow & E/G \\ [u] & \longmapsto & \{u\}, \end{array}$$

donde $\{u\}$ indica la clase de u módulo G. Esta aplicación está bien definida, ya que elementos equivalentes respecto a F también son equivalentes respecto a G. f es lineal y exhaustiva y su núcleo está formado por las clases $[u]$ tales que $\{u\} = \{\vec{0}\}$, es decir, tales que $u \in G$. Así pues, $\operatorname{Nuc} f = G/F$ y, aplicando (1.3),

$$(E/F)/(G/F) \cong E/G. \square$$

Observación:

Si la dimensión del espacio E es finita, $\operatorname{Im} f$ y $E/\operatorname{Nuc} F$ tienen la misma dimensión ((1.1) y (IV.6.1)). Son, por tanto, espacios isomorfos (por (1.4)). ¿Qué nos dice de nuevo, pues, la proposición 3.1? En primer lugar, nos dice que el resultado es válido para cualquier espacio vectorial, de dimensión finita o no. Aún más importante es, sin embargo, el hecho de que se pueda definir un isomorfismo de una manera muy natural y sin intervención de bases. Una aplicación lineal definida sin utilizar ninguna base, o que permita una definición de ese tipo, se llama una aplicación *canónica*. Todos los isomorfismos establecidos en las tres proposiciones de este apartado son isomorfismos canónicos.

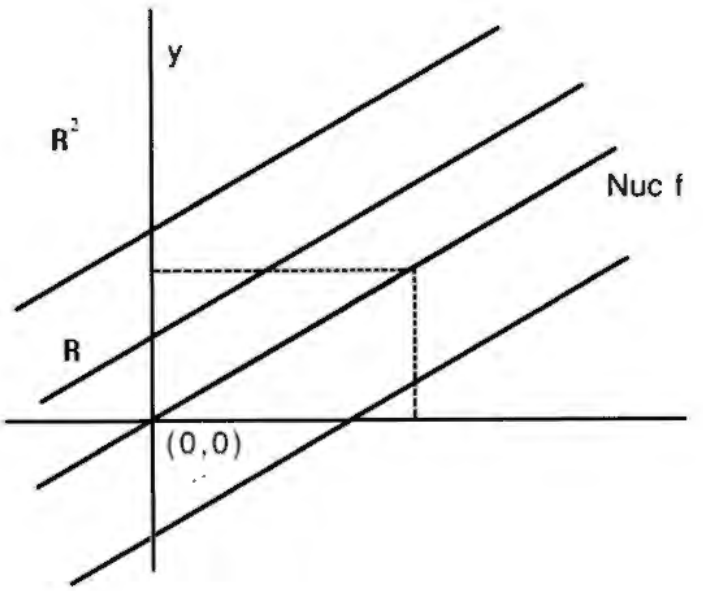
Queremos dar, antes de concluir este apartado, unos ejemplos que ilustren el contenido de esas tres proposiciones.

Ejemplos:

1. Consideremos la aplicación lineal

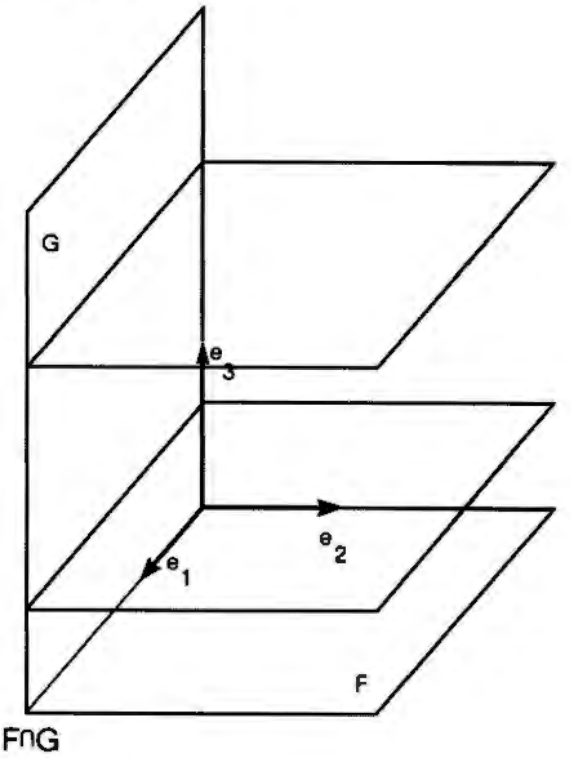
$$\begin{array}{rccc} f: & \mathbf{R}^2 & \longrightarrow & \mathbf{R} \\ & (x,y) & \longmapsto & x-y. \end{array}$$

El núcleo de f corresponde en el plano a una recta que pasa por $(0,0)$; $\mathbf{R}^2/\operatorname{Nuc} f$ es, entonces, el conjunto de rectas paralelas a $\operatorname{Nuc} f$. Cada una de esas rectas corta al eje de las x en un punto $(a,0)$, y su imagen por f es, precisamente, a. El isomorfismo de (3.1) hace corresponder a cada recta de $\mathbf{R}^2/\operatorname{Nuc} f$ su intersección con el eje de las x.



2. Consideremos $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ en $\mathbf{R}^3$. Sean F y G los subespacios vectoriales de $\mathbf{R}^3$ definidos por $F = \langle e_1, e_2\rangle$, $G = \langle e_1, e_3\rangle$. Entonces $F \cap G = \langle e_1\rangle$ y $F + G = \mathbf{R}^3$. Por tanto, $(F+G)/F$ es el conjunto de planos de $\mathbf{R}^3$ paralelos a F y $G/F \cap G$ es el conjunto de rectas del plano G paralelas a $F \cap G$. El isomorfismo de (3.2) hace corresponder a cada recta de G paralela a $F \cap G$ el plano de $\mathbf{R}^3$ paralelo a F que la contiene.
3. Sean e_1, e_2, e_3 como en el ejemplo anterior. Consideremos los subespacios de $\mathbf{R}^3$, $F = \langle e_1\rangle$ y $G = \langle e_1, e_2\rangle$. Entonces $\mathbf{R}^3/F$ es el conjunto de rectas de $\mathbf{R}^3$ paralelas a F y G/F es el conjunto de rectas de G paralelas a F. Cada clase de $(\mathbf{R}^3/F)/(G/F)$ está formada por todas las rectas (paralelas a F) situadas en un plano paralelo a G. El isomorfismo de (3.3) hace corresponder a cada

una de esas clases el plano paralelo a G que la contiene (que es un elemento de $\mathbf{R}^3/G$).



V.4 El espacio de las aplicaciones lineales

Consideremos el conjunto $L(E,F)$ de todas las aplicaciones lineales de E en F. Hay una manera natural de definir una suma y un producto por elementos del cuerpo K en $L(E,F)$. Concretamente, si $f,g \in L(E,F)$ y $a \in K$, definimos *la suma* $f+g$ por

$$(f+g)(u) = f(u) + g(u) \qquad \forall u \in E$$

y *el producto* af por

$$(af)(u) = af(u) \qquad \forall u \in E.$$

Las aplicaciones $f+g$ y af son, claramente, lineales. $L(E,F)$ con estas dos operaciones cumple todas las condiciones de espacio vectorial (IV.1); lo llamaremos *espacio vectorial de las aplicaciones de* E *en* F.

Proposición 4.1 *Si* E *y* F *son de dimensión finita,* $L(E,F)$ *también lo es y* $\dim L(E,F) = \dim E \cdot \dim F$.

DEMOSTRACIÓN: Sea $u_1,\ldots,u_n$ una base de E y $v_1,\ldots,v_m$ una base de F. Definimos

$$f_{ij} : E \to F, \qquad i = 1,\ldots,n,\ j = 1,\ldots,m,$$

por $f_{ij}(u_k) = \vec{0}$ si $k \neq i$, $f_{ij}(u_i) = v_j$.

La proposición quedará demostrada si probamos que estas nm aplicaciones forman una base de $L(E,F)$. Para ello, tenemos que ver que

- $\{f_{ij}\}$ genera $L(E,F)$. En efecto, sea $f : E \to F$ lineal; supongamos que $f(u_k) = \sum_{j=1}^{m} a_k^j v_j$. Entonces $f = \sum_{i,j} a_i^j f_{ij}$, ya que para todo u_k

$$\left(\sum_{i,j} a_i^j f_{ij}\right)(u_k) = \sum_{i,j} a_i^j f_{ij}(u_k) = \sum_j a_k^j v_j = f(u_k).$$

- $\{f_{ij}\}$ son linealmente independientes. En efecto,

$$\sum_{i,j} a_i^j f_{ij} = 0 \Rightarrow \sum_{i,j} a_i^j f_{ij}(u_k) = \vec{0} \qquad \forall k = 1, \ldots, n.$$

Mediante un cálculo como el efectuado más arriba, obtenemos

$$\sum_j a_k^j v_j = \vec{0}, \qquad k = 1, \ldots, n$$

y, por tanto, $a_k^j = 0$ para todo $k = 1, \ldots, n$ y todo $j = 1, \ldots, m$. □

La matriz de la aplicación f_{ij} respecto a las bases $u_1, \ldots, u_n$ y $v_1, \ldots, v_m$ es la matriz E_i^j formada por 0 en todas las posiciones, excepto en la columna i, fila j, donde aparece un 1. En (IV.3) vimos que esas matrices forman una base de $M_{m\times n}(K)$. La aplicación

$$\begin{array}{ccc} L(E,F) & \longrightarrow & M_{m\times n}(K) \\ f_{ij} & \longmapsto & E_i^j \qquad i = 1, \ldots, n, \quad j = 1, \ldots, m \end{array}$$

es un isomorfismo entre estos dos espacios vectoriales. De la demostración de (4.1) se deduce que este isomorfismo asocia a cada aplicación lineal f su matriz en las bases consideradas.

V.5 El álgebra de endomorfismos

Un caso particular del espacio estudiado en el apartado anterior es $L(E,E)$, el espacio de los endomorfismos de E, que denotaremos por $\mathrm{End}(E)$.

Dos elementos $f, g \in \mathrm{End}(E)$ se pueden componer siempre y la composición $g \circ f$ es también un elemento de $\mathrm{End}(E)$, que denominaremos *producto*, o *producto interno* si hay peligro de confusión con el producto por elementos del cuerpo. Este producto cumple las propiedades siguientes:

- Asociativa: $h \circ (g \circ f) = (h \circ g) \circ f \qquad \forall f, g, h \in \mathrm{End}(E)$.
- Existe un elemento neutro, que es la aplicación identidad I_E:

$$I_E \circ f = f \circ I_E = f \qquad \forall f \in \mathrm{End}(E).$$

En general, no obstante, el producto no es conmutativo y los únicos elementos que tienen inverso son los endomorfismos biyectivos (automorfismos).

El producto interno de $\mathrm{End}(E)$ está relacionado con las dos operaciones de la estructura vectorial por las propiedades siguientes:

- Distributivas:

$$\begin{aligned} (h+g) \circ f &= h \circ f + g \circ f \\ h \circ (g+f) &= h \circ g + h \circ f \qquad \forall f, g, h \in \mathrm{End}(E). \end{aligned}$$

- $(ag) \circ f = a(g \circ f) = g \circ (af) \quad \forall a \in K \quad \forall f, g \in \mathrm{End}(E)$.

En particular, $\mathrm{End}(E)$ con las operaciones $+$ y $\circ$ es un anillo con unidad.

Un conjunto A con tres operaciones —una suma $+$, un producto $\cdot$, y un producto por elementos de un cuerpo K— se llama una *álgebra sobre* K si A con $+$ y $\cdot$ es un anillo, A con $+$ y el producto por elementos de K es un espacio vectorial y $k(a \cdot b) = (ka) \cdot b = a \cdot (kb)$ para todo $k \in K$, $a, b \in A$.

Ejemplos:

1. El conjunto de polinomios $K[x]$ es una álgebra conmutativa y con unidad sobre K.
2. $\mathrm{End}(E)$ es una álgebra con unidad. También es una álgebra con unidad el conjunto de las matrices cuadradas $M_{n\times n}(K)$. Si E es de dimensión n, en el apartado anterior hemos establecido una aplicación

$$\begin{array}{rccc} \mathrm{End}(E) = & L(E,E) & \longrightarrow & M_{n\times n}(K) \\ & f & \longmapsto & A \end{array}$$

donde A es la matriz asociada a f en una base prefijada del espacio E. Esta aplicación es un isomorfismo de espacios vectoriales y, además, "conserva" los productos internos, por (2.2). Se dice entonces que es un isomorfismo de álgebras y que las álgebras $\mathrm{End}(E)$ y $M_{n\times n}(K)$ son isomorfas.

Denominaremos *homotecia vectorial de razón* a al endomorfismo aI_E de E.

Proposición 5.1 *Si E tiene dimensión 1, sus únicos endomorfismos son las homotecias vectoriales.*

DEMOSTRACIÓN: Sea $u \neq 0$ una base de E y $f \in \mathrm{End}(E)$. La imagen de u se expresa como combinación lineal de la base:

$$f(u) = au.$$

Entonces, la imagen de cualquier otro vector $v = cu \in E$ es

$$f(v) = f(cu) = cf(u) = c(au) = a(cu) = av,$$

de donde resulta que $f = aI_E$. □

Observemos que, en el caso de la proposición 5.1, la matriz de $f = aI_E$ es precisamente (a) y, por tanto, independiente de la base. El isomorfismo de álgebras del ejemplo 2 es, en este caso, un isomorfismo canónico

$$\begin{array}{ccc} \mathrm{End}(E) & \xrightarrow{\simeq} & K \\ f = aI_E & \longmapsto & a. \end{array}$$

V.6 El espacio dual

En este apartado vamos a estudiar otro caso particular del espacio de aplicaciones lineales: el caso en que el segundo espacio vectorial es K. Las aplicaciones lineales en K se llaman, también, *formas*; al espacio

$$E' = L(E, K)$$

lo llamaremos el *espacio dual de* E. Todas las consideraciones del apartado 4 se aplican, en particular, a este caso. Así pues, E' es un espacio vectorial de la misma dimensión que E (si $\dim E$ es finita). Dada una base $u_1, \ldots, u_n$ de E, las aplicaciones

$$\begin{array}{rccccll} u_i' : & E & \longrightarrow & K & & & \\ & u_j & \longmapsto & 0 & \text{si} & i \neq j & \\ & u_j & \longmapsto & 1 & \text{si} & i = j, & \qquad i = 1, \ldots, n, \end{array}$$

forman una base de E', que denominaremos *base dual* de $u_1, \ldots, u_n$.

¡Atención!:

Supongamos que $u_1, \ldots, u_n$ y $v_1, \ldots, v_n$ son dos bases diferentes de E, pero con algunos vectores comunes, por ejemplo $u_1 = v_1$; en las bases duales $u_1', \ldots, u_n'$, $v_1', \ldots, v_n'$ los elementos u_1' y v_1' no tienen por qué ser iguales.

Proposición 6.1 *Sea $u_1, \ldots, u_n$ una base del espacio E y $u'_1, \ldots, u'_n$ su base dual. Las coordenadas de una forma $\omega \in E'$ en la base $u'_1, \ldots, u'_n$ son $\omega(u_1), \ldots, \omega(u_n)$*

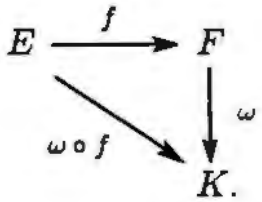
$$\omega = \omega(u_1)u'_1 + \ldots + \omega(u_n)u'_n.$$

DEMOSTRACIÓN: Para todo vector u_k de la base de E, tenemos

$$\left(\sum_{i=1}^{n} \omega(u_i)u'_i\right)(u_k) = \sum_{i=1}^{n} \omega(u_i)u'_i(u_k) = \omega(u_k)$$

y, por tanto, $\sum_{i=1}^{n} \omega(u_i)u'_i = \omega$. □

Fijada una aplicación lineal $f : E \to F$, cada elemento $\omega \in F'$ nos da, al componer con f, un elemento $\omega \circ f$ de E'



Tenemos, por tanto, una aplicación de F' en E', que designaremos por

$$f' : F' \to E'$$

y denominaremos *aplicación dual de* f. Es fácil ver que f' es lineal y que

$$(g \circ f)' = f' \circ g'.$$

Comprobamos esto último: para toda forma ω,

$$(g \circ f)'(\omega) = \omega \circ (g \circ f) = (\omega \circ g) \circ f = f'(\omega \circ g) = f'(g'(\omega)) = (f' \circ g')(\omega).$$

Supongamos, ahora, que $f : E \to F$ tiene por matriz asociada en unas determinadas bases $u_1, \ldots, u_n$, $v_1, \ldots, v_m$ la matriz $A = (a_i^j)$. Sea $B = (b_i^j)$ la matriz asociada a $f' : F' \to E'$ en las bases duales de las anteriores: $v'_1, \ldots, v'_m$, $u'_1, \ldots, u'_n$. ¿Cuál es la relación entre las matrices A y B? Por (6.1),

$$\begin{aligned} b_i^j &= (f'(v'_i))(u_j) = (v'_i \circ f)(u_j) = v'_i(f(u_j)) = \\ &= v'_i\left(\sum_{k=1}^{m} a_j^k v_k\right) = \sum_{k=1}^{m} a_j^k v'_i(v_k) = a_j^i. \end{aligned}$$

Así pues,

Proposición 6.2 *Si A es la matriz de la aplicación lineal f en unas determinadas bases, la matriz de su dual f' en las bases duales es la matriz* traspuesta *de A, que denotaremos por A^t.* □

Podemos considerar el espacio dual de cualquier espacio vectorial; en particular, podemos considerar el espacio dual de E', que denominaremos *bidual de* E y denotaremos por E''. Nos proponemos demostrar que, si la dimensión de E es finita, el bidual E'' es canónicamente isomorfo al espacio inicial E.

Consideremos la aplicación

$$\begin{array}{rccc} \langle\ ,\ \rangle: & E' \times E & \longrightarrow & K \\ & (\omega, u) & \longmapsto & \langle \omega, u \rangle = \omega(u). \end{array}$$

Si fijamos $u \in E$, obtenemos una aplicación

$$\begin{array}{rccc} \langle\ , u\rangle: & E' & \longrightarrow & K \\ & \omega & \longmapsto & \langle \omega, u \rangle \end{array}$$

qne es lineal y, por tanto, un elemento de E''.

Proposición 6.3 *Si la dimensión de E es finita, la aplicación*

$$\begin{array}{rccc} \varphi: & E & \longrightarrow & E'' \\ & u & \longmapsto & \langle\ , u\rangle \end{array}$$

es un isomorfismo.

DEMOSTRACIÓN: φ es lineal, ya que para todo $u, v \in E$ y todo $\omega \in E'$,

$$\begin{aligned} (\varphi(u+v))(\omega) &= \langle \omega, u+v \rangle = \omega(u+v) = \\ &= \omega(u) + \omega(v) = \langle \omega, u \rangle + \langle \omega, v \rangle = \\ &= \varphi(u)(\omega) + \varphi(v)(\omega) = \\ &= (\varphi(u) + \varphi(v))(\omega), \end{aligned}$$

de donde se obtiene que $\varphi(u+v) = \varphi(u) + \varphi(v)$.

Análogamente se demuestra que $\varphi(au) = a\varphi(u)$.

Para ver que φ es inyectiva, probaremos que el único vector del núcleo es $\vec{0}$ (1.2). Si $\varphi(u) = \vec{0}$, entonces, para todo $\omega \in E'$,

$$0 = \varphi(u)(\omega) = \langle \omega, u \rangle = \omega(u).$$

Si $u \neq \vec{0}$, hay una base de E de la forma $u, u_2, \ldots, u_n$; consideremos un $\omega \in E'$ tal que $\omega(u) = 1$, $\omega(u_i) = 0$, $i = 2, \ldots, n$. Entonces

$$\varphi(u)(\omega) = \langle \omega, u \rangle = 1 \neq 0,$$

en contra de lo que hemos obtenido antes. Así pues, $u = \vec{0}$ y la aplicación φ es inyectiva.

La exhaustividad de φ resulta de (1.1) y de que E y E'' tienen la misma dimensión finita. □

Observación:

En la demostración anterior solamente hemos utilizado que la dimensión de E es finita para probar la exhaustividad; para espacios de dimensión infinita, φ es un monomorfismo.

Proposición 6.4 *Sea $f : E \to F$ una aplicación lineal entre espacios de dimensión finita, y sea $f'' : E'' \to F''$ su bidual. Si $\varphi : E \cong E''$ y $\bar{\varphi} : F \cong F''$ son los isomorfismos de (6.3), entonces*

$$\bar{\varphi}^{-1} \circ f'' \circ \varphi : E \cong E'' \longrightarrow F'' \cong F$$

coincide con f.

DEMOSTRACIÓN: Si $u \in E$,

$$(f'' \circ \varphi)(u) = f''(\langle\ ,u\rangle) = \langle\ ,u\rangle \circ f' \in F'',$$

y si $\omega \in F'$,

$$\begin{aligned}(\langle\ ,u\rangle \circ f')(\omega) &= \langle\ ,u\rangle(f'(\omega)) = \langle f'(\omega), u\rangle = (f'(\omega))(u) = \\ &= \omega(f(u)) = \langle \omega, f(u)\rangle = \langle\ ,f(u)\rangle(\omega),\end{aligned}$$

de donde resulta que

$$\langle\ ,u\rangle \circ f' = \langle\ ,f(u)\rangle.$$

El vector de F que corresponde por $\bar{\varphi}$ a este elemento de F'' es $f(u)$ y, por tanto,

$$\left(\bar{\varphi}^{-1} \circ f'' \circ \varphi\right)(u) = \bar{\varphi}^{-1}(\langle\ ,f(u)\rangle) = f(u).$$

Así pues, $\bar{\varphi}^{-1} \circ f'' \circ \varphi = f$. □

V.7 Subespacios ortogonales

En este apartado supondremos que trabajamos únicamente con espacios vectoriales de dimensión finita. Al igual que en el apartado anterior, E' indicará el dual del espacio E, $E' = L(E, K)$.

Sea A un subconjunto de E. Definimos el *ortogonal de* A como el conjunto

$$A^{\perp} = \{\omega \in E' \mid \omega(u) = 0 \ \forall u \in A\}.$$

Se tienen, entonces, las propiedades siguientes:

1. $A^{\perp}$ es un subespacio vectorial de E'.
2. $A \subset B \Rightarrow B^{\perp} \subset A^{\perp}$.
3. Si F es un subespacio de E, $\dim F^{\perp} = \dim E - \dim F$.
4. $E^{\perp} = \{0\}$, $\{\vec{0}\}^{\perp} = E'$.

1, 2 y 4 son inmediatas; demostremos 3: tomemos una base $u_1, \ldots, u_k$ de F y completémosla hasta obtener una base $u_1, \ldots, u_k, u_{k+1}, \ldots, u_n$ de E. Sea $u'_1, \ldots, u'_k, u'_{k+1}, \ldots, u'_n$ la base dual correspondiente. Para $j = k+1, \ldots, n$, u'_j se anula sobre la base $u_1, \ldots, u_k$ de F y, por tanto, sobre todo F; es decir, $u'_j \in F^{\perp}$ para $j = k+1, \ldots, n$. Ahora bien, estos elementos forman una base de $F^{\perp}$, ya que son linealmente independientes (por formar parte de una base) y generan $F^{\perp}$, ya que si $\omega = a^1 u'_1 + \ldots + a^n u'_n \in F^{\perp} \subset E'$, como $u_h \in F$, $h = 1, \ldots, k$, se tiene $0 = \omega(u_h) = a^h$, de donde $\omega = a^{k+1} u'_{k+1} + \ldots + a^n u'_n$.

Queremos definir, ahora, el ortogonal de un subconjunto A de E'. Hay dos maneras de hacerlo.

1. $A^{\perp} = \{u \in E \mid \omega(u) = 0 \quad \forall \omega \in A\}$.

   Obtenemos, así, un subconjunto del espacio inicial E de manera muy parecida a como hemos obtenido el ortogonal de un $A \subset E$.

2. Aplicamos la definición que hemos dado al principio del apartado. Obtenemos así un subconjunto del bidual

$$A^{\perp} = \{\alpha \in E'' \mid \alpha(\omega) = 0 \quad \forall \omega \in A\}.$$

Estas dos maneras son, esencialmente, la misma. Con más precisión, estos dos ortogonales se corresponden por el isomorfismo de (6.3). En efecto, recordemos que en aquel isomorfismo un elemento $\alpha \in E''$ correspondía

a un vector $u \in E$ de forma que $\alpha = \langle\ ,u\rangle$. Por tanto, en 2,

$$\{\alpha \in E' \mid \alpha(\omega) = 0 \quad \forall \omega \in A\} = \{\langle\ ,u\rangle \in E'' \mid \langle \omega, u\rangle = 0 \quad \forall \omega \in A\} = \{\langle\ ,u\rangle \in E'' \mid \omega(u) = 0 \quad \forall \omega \in A\},$$

que corresponde a $\{u \in E \mid \omega(u) = 0 \quad \forall \omega \in A\}$ de 1.

La definición 2 es la dada al principio del apartado. Por tanto, las propiedades 1, 2, 3 y 4 son válidas en este caso. Las consideraciones anteriores nos dicen que estas propiedades son también válidas para ortogonales definidos según 1. De aquí en adelante trabajaremos siempre con la definición 1.

Otras propiedades de los ortogonales son:

5. Si F es un subespacio vectorial, $F^{\perp\perp} = F$.

6. Si F y G son subespacios vectoriales,

$$(F \cap G)^{\perp} = F^{\perp} + G^{\perp} \text{ y } (F + G)^{\perp} = F^{\perp} \cap G^{\perp}.$$

7. Si $E = F \oplus G$, $E' = F^{\perp} \oplus G^{\perp}$.

DEMOSTRACIÓN DE 5:

$$u \in F \Rightarrow \langle \omega, u\rangle = \omega(u) = 0\ \forall \omega \in F^{\perp} \Rightarrow u \in F^{\perp\perp},$$

de donde $F \subset F^{\perp\perp}$; pero la propiedad 3 nos dice que estos dos espacios tienen la misma dimensión y, por tanto, $F = F^{\perp\perp}$. □

DEMOSTRACIÓN DE 6:

$$\begin{aligned} F \cap G \subset F,\ F \cap G \subset G &\Rightarrow \text{(por 2) } F^{\perp} \subset (F \cap G)^{\perp}, G^{\perp} \subset (F \cap G)^{\perp} \Rightarrow \\ &\Rightarrow F^{\perp} + G^{\perp} \subset (F \cap G)^{\perp}; \qquad (*) \\ F + G \supset F, F + G \supset G &\Rightarrow \text{(por 2) } F^{\perp} \supset (F + G)^{\perp}, G^{\perp} \supset (F + G)^{\perp} \Rightarrow \\ &\Rightarrow (F + G)^{\perp} \subset F^{\perp} \cap G^{\perp}; \qquad (**) \end{aligned}$$

Entonces, por 5, (*) y (**),

$$F \cap G = (F \cap G)^{\perp\perp} \subset (F^{\perp} + G^{\perp})^{\perp} \subset F^{\perp\perp} \cap G^{\perp\perp} = F \cap G,$$

y todas las inclusiones son igualdades; en particular, $F \cap G = (F^{\perp} + G^{\perp})^{\perp}$, de donde, por 5, $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.

Para demostrar la otra igualdad se procede de manera parecida. □

DEMOSTRACIÓN DE 7:

$$\begin{aligned} E = F \oplus G &\Leftrightarrow E = F + G \text{ y } F \cap G = \{\vec{0}\} \Leftrightarrow \\ &\Leftrightarrow \{\vec{0}\} = E^{\perp} = (F + G)^{\perp} = F^{\perp} \cap G^{\perp} \text{ y} \\ &\quad E' = \{\vec{0}\}^{\perp} = (F \cap G)^{\perp} = F^{\perp} + G^{\perp} \Leftrightarrow \\ &\Leftrightarrow E' = F^{\perp} \oplus G^{\perp}. \ \square \end{aligned}$$

Proposición 7.1 *Sea $f : E \to F$ una aplicación lineal entre espacios vectoriales de dimensión finita y $f' : F' \to E'$ su dual. Entonces,*

$$(\operatorname{Im} f)^{\perp} = \operatorname{Nuc} f' \quad y \quad (\operatorname{Nuc} f)^{\perp} = \operatorname{Im} f'.$$

DEMOSTRACIÓN:

$$\begin{aligned}
(\operatorname{Im} f)^{\perp} &= \{\omega \in F' \mid \omega(v) = 0 \ \forall v \in Imf\} = \{\omega \in F' \mid \omega(fu) = 0 \ \forall u \in E\} = \\
&= \{\omega \in F' \mid (f'\omega)(u) = 0 \ \forall u \in E\} = \{\omega \in F' \mid f'\omega = 0\} = \operatorname{Nuc} f'. \\
(\operatorname{Im} f')^{\perp} &= \{u \in E \mid \rho(u) = 0 \ \forall \rho \in Imf'\} = \{u \in E \mid (f'\omega)(u) = 0 \ \forall \omega \in F'\} = \\
&= \{u \in E \mid \omega(fu) = 0 \ \forall \omega \in F'\} = \\
&\quad \text{(por un razonamiento hecho en (6.3))} \\
&= \{u \in E \mid (fu) = 0\} = \operatorname{Nuc} f.
\end{aligned}$$

La propiedad 5 nos da, ahora, la segunda igualdad. □

V.8 Nota histórica

Para Leonhard Euler (1707–1783) una función era una fórmula o ecuación que contenía variables y constantes. Euler y Joseph-Louis Lagrange (1736–1813) ya sabían que las soluciones de un sistema homogéneo forman un espacio vectorial, pero este hecho no fue explotado hasta Augustin-Louis Cauchy (1789–1857). La teoría de aplicaciones lineales se desarrolla a mediados del siglo 19 (aunque una definición precisa como la actual no fue dada hasta finales de siglo por Giuseppe Peano (1858–1932)) y la conexión entre matrices y aplicaciones lineales fue establecida y desarrollada por Arthur Cayley (1821–1895) en 1855. Georg Ferdinand Frobenius (1849–1917) considera en 1879 el rango de una matriz y lo utiliza en el estudio de los sistemas de ecuaciones lineales (ver Cap. VII).

Se puede observar en las notas históricas de estos capítulos que, contrariamente a lo que hemos hecho nosotros, en el desarrollo histórico de las matemáticas las definiciones precisas llegan después de la utilización de las herramientas y de la obtención de buena parte de los resultados.

V.9 Ejercicios

1. Demostrar que, dada cualquier aplicación lineal $f : E \longrightarrow F$, existen bases de E y F tales que la matriz de f en esas bases es

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

¿Qué significado tiene r?

2. Sea

$$E = \{\begin{pmatrix} 0 & a \\ b & c \end{pmatrix} \in M_{2\times 2}(\mathbf{R}) \mid c = a + b\}.$$

Consideremos el endomorfismo $f : E \longrightarrow E$ dado por

$$f\begin{pmatrix} 0 & a \\ b & c \end{pmatrix} = \begin{pmatrix} 0 & 3c + 3a \\ -2a - b & a - b + 3c \end{pmatrix}.$$

Hallar una base de E, la matriz de f en esa base y sendas bases de $\operatorname{Nuc} f$ e $\operatorname{Im} f$.

3. ¿Cuál es la matriz del cambio de base entre

$$\{1, x, x^2, \ldots, x^n\} \quad \text{y} \quad \{1, x - a, (x - a)^2, \ldots, (x - a)^n\}$$

como bases del espacio vectorial de los polinomios reales de grado menor o igual que n? Utilizarla para probar *la fórmula de Taylor*:

$$p(x) = p(a) + p'(a)(x - a) + \frac{1}{2!}p''(a)(x - a)^2 + \ldots + \frac{1}{n!}p^{(n)}(a)(x - a)^n.$$

4. Sea $f : E \longrightarrow F$ una aplicación lineal. Demostrar

a) f es inyectiva $\Longleftrightarrow$ existe una $g : F \longrightarrow E$ lineal tal que $g \circ f = I$.

b) f es exhaustiva $\Longleftrightarrow$ existe una $g : F \longrightarrow E$ lineal tal que $f \circ g = I$.

5. Dadas dos aplicaciones lineales $f : E \longrightarrow F$ y $g : E \longrightarrow G$, hallar sendas condiciones necesarias y suficientes para que

a) Exista $h : F \longrightarrow G$ lineal tal que $h \circ f = g$.

b) Exista una única $h : F \longrightarrow G$ lineal tal que $h \circ f = g$.

c) Exista un monomorfismo $h : F \longrightarrow G$ tal que $h \circ f = g$.

6. Dado un endomorfismo f de un espacio vectorial de dimensión finita n, demostrar que los conjuntos

$$\mathcal{F} = \{g \in \operatorname{End}(E) \mid f \circ g = 0\} \quad \text{y} \quad \mathcal{G} = \{g \in \operatorname{End}(E) \mid g \circ f = 0\}$$

son subespacios vectoriales de $\operatorname{End}(E)$ y determinar sus dimensiones.

7. Demostrar que todo endomorfismo f de un espacio vectorial E de dimensión finita puede expresarse como diferencia de dos automorfismos.

8. Un endomorfismo $f : E \longrightarrow E$ se llama un *proyector* si $f^2 = f$. Demostrar:

a) f es un proyector si y sólo si $I - f$ lo es.

b) Si f es un proyector, $E = \operatorname{Nuc} f \oplus \operatorname{Im} f$.

c) Si f y g son proyectores, determinar condiciones necesarias y suficientes para que $f + g$ también lo sea.

d) Si f es un proyector, encontrar las relaciones existentes entre $\operatorname{Nuc} f$, $\operatorname{Im} f$, $\operatorname{Nuc}(I - f)$, $\operatorname{Im}(I - f)$.

9. Sea $e_1, \ldots, e_n$ una base del espacio vectorial E y sean $(a_i^1, \ldots, a_i^n)$ las coordenadas de los vectores $v_i \in E$, $i = 1, \ldots, k$. ¿Qué condiciones deben cumplir las coordenadas de una forma $\omega \in E'$ en la base dual, $e'_1, \ldots, e'_n$, de $e_1, \ldots, e_n$, para que $\omega \in \langle v_1, \ldots, v_k \rangle^\perp$?

Sean ahora $(b_j^1, \ldots, b_j^n)$ coordenadas de formas $\omega_j \in E'$, $j = 1, \ldots, k$, en la base $e'_1, \ldots, e'_n$. ¿Qué condiciones deben cumplir las coordenadas de un vector $v \in E$ para que $v \in \langle \omega_1, \ldots, \omega_k \rangle^\perp$?

10. Demostrar que $\omega_1, \ldots, \omega_m \in E'$ son linealmente independientes si y sólo si para cada m-pla $(a_1, \ldots, a_m) \in K^m$ existe un vector $u \in E$ tal que $\omega_i(u) = a_i$, $i = 1, \ldots, m$.

11. Sea $f \in \operatorname{End}(E)$ tal que $f^2 = I$. Sean $E_1 = \{x \in E \mid f(x) = x\}$, $E_2 = \{x \in E \mid f(x) = -x\}$. Demostrar que $E = E_1 \oplus E_2$. ¿Significa esto que para todo $x \in E$ se cumple $f(x) = x$ o $f(x) = -x$?

12. Sea $f \in \operatorname{End}(E)$. Demostrar que $\operatorname{Nuc} f = \operatorname{Im} f$ si y sólo si $\dim E$ es par, $f^2 = 0$ y $\operatorname{rang} f = n/2$.

13. Demostrar que $f \in \operatorname{End}(E)$ conmuta con todos los endomorfismos de E si y sólo si $f = aI$, $a \in K$.

14. Sea $f \in \operatorname{End}(E)$ tal que $f^2 + f + I = 0$. Demostrar que f es un isomorfismo y determinar su inverso.

15. Sea E un espacio vectorial de dimensión finita y f un subespacio vectorial de E. La inclusión $F^\perp \subset E'$ permite definir una aplicación

$$f : E'' \longrightarrow (F^\perp)'$$

por restricción. Demostrar que f es un epimorfismo.

Sea ahora $\varphi : E \cong E''$ el isomorfismo de la proposición 6.3. Comprobar que $\operatorname{Nuc}(f \circ \varphi) = F$ y explicitar un isomorfismo $E/F \cong (F^\perp)'$.

V.10 Ejercicios para programar

Los ejercicios que proponemos a continuación utilizan todos, en algún momento, el ejercicio VII.12.

16. Elaborar un programa que permita

a) Cambiar de base las coordenadas de un vector dado.

b) Cambiar de base la matriz de una aplicación lineal dada.

(Indicación: si $u_1, \ldots, u_n$ es la base original y $e_1, \ldots, e_n$ la nueva base, construir la matriz del cambio y su inversa —utilizando el ejercicio VII.12. Utilizar el subprograma del ejercicio IV.15 c.)

17. Elaborar un programa que, dados $e_1, \ldots, e_n$ vectores de $\mathbf{R}^n$, compruebe que forman una base y calcule su base dual. (Indicación: si P es la matriz que expresa la base $e_1, \ldots, e_n$ en función de la base canónica, entonces $(P^t)^{-1}$ expresa las formas $e'_1, \ldots, e'_n$ en función de la base dual de la canónica.)

Capítulo VI

Determinantes

VI.1 Determinante de n vectores

Consideremos el conjunto de las matrices $n \times n$ sobre K. Queremos asociar a cada matriz un elemento del cuerpo K, su "determinante", de forma que se cumplan las siguientes propiedades: si multiplicamos por $a \in K$ los elementos de una columna, el determinante queda multiplicado por a; si una columna es suma de dos, el determinante es suma de los determinantes calculados con cada una de las columnas-sumandos; si dos columnas son iguales, el determinante es cero. Estas tres condiciones son, de hecho, suficientemente restrictivas como para no permitir mucho margen al escoger (¡definir!) qué será el determinante de una matriz. Vamos a verlo.

Observemos, en primer lugar, que las condiciones impuestas se refieren a las columnas; por ello, conviene considerar cada matriz $n \times n$ como un elemento de $\overbrace{K^n \times \ldots \times K^n}^{n}$, interpretando cada columna como una n-pla de K^n. Así pues, un determinante ha de ser una aplicación

$$\begin{array}{rccc} \det: & \overbrace{K^n \times \ldots \times K^n}^{n} & \longrightarrow & K \\ & (a_1, \ldots, a_n) & \longmapsto & \det(a_1, \ldots, a_n) \end{array}$$

que cumpla:

- $\det(a_1, \ldots, aa_i, \ldots, a_n) = a \det(a_1, \ldots, a_n) \qquad \forall a \in K,\ i = 1, \ldots, n.$
- $\det(a_1, \ldots, a_i + a_i', \ldots, a_n) = \det(a_1, \ldots, a_i, \ldots, a_n) + \det(a_1, \ldots, a_i', \ldots, a_n), \quad i = 1, \ldots, n.$
- $\det(a_1, \ldots, a_n) = 0 \quad$ si $\ a_i = a_j \ $ con $\ i \neq j.$

Esto nos conduce a la primera definición que vamos a dar para una situación algo más general.

Sea E un espacio vectorial de dimensión n sobre un cuerpo K. Una *n-forma lineal alternada* es una aplicación

$$D : \overbrace{E \times \ldots \times E}^{n} \longrightarrow K$$

que cumple

(a) $D(v_1, \ldots, av_i, \ldots, v_n) = a\, D(v_1, \ldots, v_i, \ldots, v_n)\ \forall a \in K,\ i = 1, \ldots, n.$

(b) $D(v_1, \ldots, v_i + v_i', \ldots, v_n) = D(v_1, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_i', \ldots, v_n),\quad i = 1, \ldots, n.$

(c) $D(v_1, \ldots, v_n) = 0$ si $v_i = v_j$ con $i \neq j$.

Las condiciones (a) y (b) se resumen diciendo que D es *multilineal* o bien lineal en cada factor; el nombre de "alternada" se refiere a la tercera condición o, más exactamente, a la primera de las propiedades que enunciaremos. El nombre de "forma" se reserva para aplicaciones lineales o multilineales en K.

Pasemos a enunciar las propiedades de las n-formas lineales alternadas.

1. $D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$.

   En efecto, por (c),

$$\begin{aligned} 0 &= D(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n) = \quad \text{(por (b))} \\ &= D(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + \\ &\quad + D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n) = \\ &\quad \text{(por (c))} \\ &= D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n). \ \square \end{aligned}$$

   Esta propiedad 1 equivale en muchos casos a la condición (c). Concretamente, si una aplicación $D : E \times \ldots \times E \longrightarrow K$ cumple (a), (b) y 1, entonces, para $v_i = v_j$,

$$D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

   si y sólo si

$$2\, D(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) = 0.$$

   Siempre que en K sea $2 \neq 0$, esto equivale a

$$D(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) = 0.$$

2. Para toda permutación $\sigma \in \mathcal{S}_n$ de signo $\varepsilon(\sigma)$ (III.2),

$$D(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \varepsilon(\sigma)\, D(v_1, \ldots, v_n).$$

Esta propiedad se reduce a 1 si σ es una trasposición. En general, podemos descomponer σ en producto de trasposiciones y aplicar 1 reiteradamente. □

3. Si un vector $v_i = \vec{0}$, entonces $D(v_1, \ldots, v_i, \ldots, v_n) = 0$.
En efecto, como $v_i = 0v_i$, tenemos

$$D(v_1, \ldots, v_i, \ldots, v_n) = 0 \cdot D(v_1, \ldots, v_i, \ldots, v_n) = 0. \square$$

4. Si $v_j = \sum_{k \neq j} a^k v_k$, entonces $D(v_1, \ldots, v_j, \ldots, v_n) = 0$.
En efecto,

$$D(v_1, \ldots, v_j, \ldots, v_n) = \sum_{k \neq j} a^k D(v_1, \ldots, v_k, \ldots, v_n).$$

En cada sumando, v_k ocupa las posiciones j y k, y, por tanto, el sumando se anula. □

5. $D(v_1, \ldots, v_i + \sum_{k \neq i} a^k v_k, \ldots, v_n) = D(v_1, \ldots, v_i, \ldots, v_n)$.
Es consecuencia inmediata de 4. □

6. D está determinada por los valores que toma sobre una base $e_1, \ldots, e_n$ de E.
En efecto, calculemos $D(v_1, \ldots, v_n)$. Si $v_i = \sum_{h=1}^{n} a_i^h e_h$,

$$\begin{aligned} D(v_1, \ldots, v_n) &= D\left(\sum_h a_1^h e_h, \ldots, \sum_h a_n^h e_h\right) = \\ &= \sum_{h_1, \ldots, h_n = 1}^{n} D(a_1^{h_1} e_{h_1}, \ldots, a_n^{h_n} e_{h_n}) = \\ &= \sum_{h_1, \ldots, h_n = 1}^{n} a_1^{h_1} \ldots a_n^{h_n} D(e_{h_1}, \ldots, e_{h_n}). \end{aligned}$$

En $D(e_{h_1}, \ldots, e_{h_n})$, los subíndices $h_1, \ldots, h_n$ pueden tomar valores arbitrarios en $\{1, \ldots, n\}$, pero el sumando se anulará siempre que dos de

los subíndices sean iguales. Quedarán solamente, pues, los sumandos en que $h_1, \ldots, h_n$ sean precisamente $1, \ldots, n$ permutados. Designemos por h la permutación

$$h = \begin{pmatrix} 1 & \ldots & n \\ h_1 & \ldots & h_n \end{pmatrix};$$

tenemos entonces

$$\begin{aligned} D(v_1, \ldots, v_n) &= \sum_{h \in S_n} a_1^{h_1} \ldots a_n^{h_n} D(e_{h_1}, \ldots, e_{h_n}) = \quad \text{(por 2)} \\ &= \sum_{h \in S_n} \varepsilon(h) a_1^{h_1} \ldots a_n^{h_n} D(e_1, \ldots, e_n). \end{aligned}$$

Denominaremos *determinante de los vectores* $v_1, \ldots, v_n$ *en la base* $e_1, \ldots, e_n$ al elemento de K

$$\det_{(e_i)}(v_1, \ldots, v_n) = \sum_{h \in S_n} \varepsilon(h) a_1^{h_1} \ldots a_n^{h_n}.$$

Este elemento solamente depende de las coordenadas de los vectores $v_1, \ldots, v_n$ en la base $e_1, \ldots, e_n$, y cumple la condición de que, para toda n-forma lineal alternada D,

$$D(v_1, \ldots, v_n) = \det_{(e_i)}(v_1, \ldots, v_n)\, D(e_1, \ldots, e_n).$$

Esto demuestra 6. □

7. Sea $e_1, \ldots, e_n$ una base de E. Dado $k \in K$, existe una n-forma lineal alternada D, y sólo una, tal que $D(e_1, \ldots, e_n) = k$.

DEMOSTRACIÓN: La propiedad 6 nos dice que, si existe, D ha de ser tal que

$$D(v_1, \ldots, v_n) = \det_{(e_i)}(v_1, \ldots, v_n) k.$$

Sólo debemos comprobar, por tanto, que esto es siempre una n-forma lineal alternada. Con las notaciones de 6, tenemos

$$\begin{aligned} D(v_1, \ldots, av_i, \ldots, v_n) &= \det_{(e_i)}(v_1, \ldots, av_i, \ldots, v_n)\, k = \\ &= \left(\sum_{h \in S_n} \varepsilon(h) a_1^{h_1} \ldots (a a_i^{h_i}) \ldots a_n^{h_n} \right) k = \\ &= a \left(\sum_{h \in S_n} \varepsilon(h) a_1^{h_1} \ldots a_i^{h_i} \ldots a_n^{h_n} \right) k = \\ &= a\, D(v_1, \ldots, v_i, \ldots, v_n), \end{aligned}$$

lo que demuestra la condición (a). Comprobemos (b):

$$D(v_1,\dots,v_i+w_i,\dots,v_n)=\det_{(e_i)}(v_1,\dots,v_i+w_i,\dots,v_n)k=$$
$$=\left(\sum_{h\in S_n}\varepsilon(h)a_1^{h_1}\dots(a_i^{h_i}+b_i^{h_i})\dots a_n^{h_n}\right)k=$$
$$=\left(\sum_{h\in S_n}\varepsilon(h)a_1^{h_1}\dots a_i^{h_i}\dots a_n^{h_n}+\sum_{h\in S_n}\varepsilon(h)a_1^{h_1}\dots b_i^{h_i}\dots a_n^{h_n}\right)k=$$
$$=D(v_1,\dots,v_i,\dots,v_n)+D(v_1,\dots,w_i,\dots,v_n).$$

Falta, por último, demostrar (c). Supongamos que $v_i=v_j$; luego sus coordenadas son iguales, $a_i^k=a_j^k$ para todo k. Entonces,

$$\det_{(e_i)}(v_1,\dots,v_i,\dots,v_j,\dots,v_n)=\sum_{h\in S_n}\varepsilon(h)a_1^{h_1}\dots a_i^{h_i}\dots a_j^{h_j}\dots a_n^{h_n}.$$

Consideremos un sumando cualquiera

$$\varepsilon(\sigma)a_1^{\sigma_1}\dots a_i^{\sigma_i}\dots a_j^{\sigma_j}\dots a_n^{\sigma_n},$$

y comparémoslo con el sumando correspondiente a $\tau=\sigma\circ(i,j)$,

$$\varepsilon(\tau)a_1^{\tau_1}\dots a_i^{\tau_i}\dots a_j^{\tau_j}\dots a_n^{\tau_n}=-\varepsilon(\sigma)a_1^{\sigma_1}\dots a_i^{\sigma_j}\dots a_j^{\sigma_i}\dots a_n^{\sigma_n}=$$
$$=-\varepsilon(\sigma)a_1^{\sigma_1}\dots a_j^{\sigma_j}\dots a_i^{\sigma_i}\dots a_n^{\sigma_n}.$$

Estos dos sumandos suman 0 y podemos eliminarlos del sumatorio. Este proceso puede repetirse tantas veces como sea necesario; al final quedará

$$\det_{(e_i)}(v_1,\dots,v_i,\dots,v_j,\dots,v_n)=0,$$

de donde

$$D(v_1,\dots,v_i,\dots,v_j,\dots,v_n)=0$$

si $v_i=v_j$. □

Así pues, existen tantas n-formas lineales alternadas como elementos del cuerpo K. Hay una manera más atractiva y precisa de expresar este hecho. Es la siguiente: consideremos el conjunto $\mathcal{A}(E)$ de las n-formas lineales alternadas y, en él, las operaciones dadas por

$$(D_1+D_2)(v_1,\dots,v_n)=D_1(v_1,\dots,v_n)+D_2(v_1,\dots,v_n)$$
$$(aD)(v_1,\dots,v_n)=a\,D(v_1,\dots,v_n)\qquad\forall a\in K.$$

Es muy fácil probar que si D_1, D_2 y D son de $\mathcal{A}(E)$, entonces $D_1 + D_2$ y aD también lo son. Se ve también sin dificultad que $\mathcal{A}(E)$, con estas operaciones, es un espacio vectorial sobre el cuerpo K.

Sea ahora $e_1, \ldots, e_n$ una base de E. La aplicación

$$\begin{array}{ccc} \mathcal{A}(E) & \longrightarrow & K \\ D & \longmapsto & D(e_1, \ldots, e_n) \end{array}$$

es lineal y, por 7, biyectiva. Tenemos, pues, un isomorfismo

$$\mathcal{A}(E) \cong K$$

y, en particular, $\mathcal{A}(E)$ tiene dimensión 1. La n-forma correspondiente al 1 de K es precisamente

$$\begin{array}{rccc} \det_{(e_i)}: & \overbrace{E \times \ldots \times E}^{n} & \longrightarrow & K \\ & (v_1, \ldots, v_n) & \longmapsto & \det_{(e_i)}(v_1, \ldots, v_n). \end{array}$$

El valor de una n-forma D no idénticamente cero sobre n vectores linealmente independientes es, por 6, diferente de cero. El valor de D sobre vectores linealmente dependientes es siempre cero (por 4). Tomando, en particular, $D = \det_{(e_i)}$, tenemos

Proposición 1.1 *$v_1, \ldots, v_n$ son linealmente independientes si y sólo si*

$$\det_{(e_i)}(v_1, \ldots, v_n) \neq 0. \ \square$$

Proposición 1.2 *Sean $e_1, \ldots, e_n$ y $u_1, \ldots, u_n$ dos bases de un espacio vectorial E. Entonces,*

$$\det_{(u_i)}(v_1, \ldots, v_n) = \det_{(e_i)}(v_1, \ldots, v_n) \det_{(u_i)}(e_1, \ldots, e_n).$$

DEMOSTRACIÓN: Para cualquier n-forma lineal alternada D tenemos

$$\begin{aligned} D(v_1, \ldots, v_n) &= \det_{(e_i)}(v_1, \ldots, v_n) D(e_1, \ldots, e_n) = \\ &= \det_{(e_i)}(v_1, \ldots, v_n) \det_{(u_i)}(e_1, \ldots, e_n) D(u_1, \ldots, u_n), \end{aligned}$$

y también $D(v_1, \ldots, v_n) = \det_{(u_i)}(v_1, \ldots, v_n) D(u_1, \ldots, u_n)$.

Podemos tomar D tal que $D(u_1, \ldots, u_n) \neq 0$ y obtenemos la igualdad deseada. $\square$

VI.2 Determinante de una matriz

Dada una matriz $n \times n$ sobre K

$$A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \dots & a_n^n \end{pmatrix},$$

llamaremos *determinante de* A al elemento de K

$$\det A = \sum_{h \in S_n} \varepsilon(h) a_1^{h_1} \cdots a_n^{h_n}.$$

Usaremos también la notación

$$\det A = \begin{vmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^n & \dots & a_n^n \end{vmatrix}.$$

Fijemos un espacio vectorial E sobre K y una base $e_1, \dots, e_n$ de E. Podemos interpretar las columnas de A como las coordenadas de los vectores de E

$$a_i = \sum_{j=1}^{n} a_i^j e_j.$$

Tenemos así una correspondencia biyectiva entre matrices $n \times n$ y n-plas de vectores de E:

$$\begin{array}{ccc} M_{n\times n}(K) & \longrightarrow & \overbrace{E \times \dots \times E}^{n} \\ A = (a_i^j) & \longmapsto & (a_1, \dots, a_n) \end{array}$$

de forma que

$$\det A = \det_{(e_i)}(a_1, \dots, a_n).$$

Esto nos permite traducir las propiedades de los determinantes de n vectores en propiedades de los determinantes de las matrices. Por ejemplo, la condición (a) de n-forma lineal alternada nos dice que, si multiplicamos los términos de una columna por un elemento a de K, el valor del determinante queda multiplicado por a; la propiedad 4 dice que, si una columna es "combinación lineal" de las otras, el determinante de la matriz es 0. Y así todas.

La proposición siguiente nos da una propiedad que no se obtiene como "traducción" de ninguna propiedad de los determinantes de n vectores.

Proposición 2.1 *Sea A una matriz $n \times n$ y A^t su traspuesta; entonces*

$$\cdot \det A = \det A^t.$$

DEMOSTRACIÓN: Sean $A=(a_i^j)$ y $A^t=(b_i^j)$, de forma que $a_i^j=b_j^i$. Tenemos, entonces,

$$\begin{aligned}
\det A &= \sum_{h\in S_n} \varepsilon(h) a_1^{h(1)} \dots a_n^{h(n)} = \quad \text{(ordenando los superíndices)} \\
&= \sum_{h\in S_n} \varepsilon(h) a^1_{h^{-1}(1)} \dots a^n_{h^{-1}(n)} = \quad \text{(poniendo } \sigma = h^{-1}\text{)} \\
&= \sum_{\sigma\in S_n} \varepsilon(\sigma) a^1_{\sigma(1)} \dots a^n_{\sigma(n)} = \sum_{\sigma\in S_n} \varepsilon(\sigma) b_1^{\sigma(1)} \dots b_n^{\sigma(n)} = \\
&= \det A^t. \ \square
\end{aligned}$$

Esta proposición tiene como consecuencia que todas las propiedades de los determinantes de matrices $n \times n$ referentes a sus columnas dan lugar a propiedades referentes a sus filas. Por ejemplo, si multiplicamos los elementos de una fila por un elemento a de K, el valor del determinante queda multiplicado por a; si una fila es "combinación lineal" de las otras, el determinante es cero; etc.

VI.3 Determinante de un endomorfismo

Sea $f : E \longrightarrow E$ un endomorfismo. Para toda n-forma lineal alternada D, la aplicación

$$\begin{array}{rccc}
\hat{f}(D): & \overbrace{E \times \dots \times E}^{n} & \longrightarrow & K \\
& (v_1, \dots, v_n) & \longmapsto & D(f(v_1), \dots, f(v_n))
\end{array}$$

es una n-forma lineal alternada. Tenemos, pues, una aplicación

$$\begin{array}{rccc}
\hat{f}: & \mathcal{A}(E) & \longrightarrow & \mathcal{A}(E) \\
& D & \longmapsto & \hat{f}(D)
\end{array}$$

que resulta ser lineal. Ahora bien, $\mathcal{A}(E)$ es un espacio vectorial de dimensión 1 y toda aplicación lineal de $\mathcal{A}(E)$ en sí mismo es una homotecia (V.5.1). En particular, $\hat{f} = a\, I$, donde I indica la aplicación identidad y $a \in K$ es la razón de la homotecia. Llamaremos *determinante del endomorfismo* f a la razón de la homotecia $\hat{f}$

$$\hat{f} = (\det f)\, I.$$

Para calcular explícitamente $\det f$ consideramos una base $e_1, \ldots, e_n$ de E y una n-forma lineal alternada $D \neq 0$. $\hat{f} = (\det f) I$ implica $\hat{f}(D) = (\det f)\, D$ y, por tanto,

$$\hat{f}(D)\,(e_1, \ldots, e_n) = (\det f)\, D(e_1, \ldots, e_n).$$

Por la definición de $\hat{f}$,

$$\begin{aligned}\hat{f}(D)(e_1, \ldots, e_n) &= D(f(e_1), \ldots, f(e_n)) = \\ &= \det_{(e_i)}(f(e_1), \ldots, f(e_n)) D(e_1, \ldots, e_n).\end{aligned}$$

Igualando y teniendo en cuenta que $D(e_1, \ldots, e_n) \neq 0$, resulta que

$$\det f = \det_{(e_i)}(f(e_1), \ldots, f(e_n)).$$

La matriz A que tiene por columnas las coordenadas de las imágenes $f(e_1), \ldots, f(e_n)$ en la base $e_1, \ldots, e_n$ es la matriz asociada a f en la base $e_1, \ldots, e_n$; por tanto,

$$\det f = \det A.$$

Llegados a este punto, debemos preguntarnos por qué hemos escogido un camino tan "complicado" para definir el determinante de un endomorfismo y no nos hemos limitado a decir que es el determinante de su matriz asociada. Hay diversas razones para ello. La primera es que la definición dada es muy curiosa y elegante; y esto es importante porque, a menudo, como aquí, un razonamiento de este tipo permite ver la conexión que hay entre cosas aparentemente muy diferentes. La segunda razón es que, tal como lo hemos hecho, ha quedado bien claro que todas las matrices asociadas a f en diferentes bases tienen el mismo determinante. Naturalmente, este hecho puede demostrarse directamente, pero los cálculos necesarios son mucho más "complicados" que la definición dada; la demostración se basa en el hecho de que el determinante de un producto de matrices es el producto de los determinantes de las matrices (¡inténtese dar una demostración directa!). La tercera razón es que esa definición que hemos dado nos va a permitir demostrar que $\det AB = \det A \cdot \det B$ sin ningún cálculo.

Proposición 3.1 *Si f, g son dos endomorfismos de E,*

$$\widehat{g \circ f} = \hat{f} \circ \hat{g} \qquad y \qquad \hat{I}_E = I_{\mathcal{A}(E)},$$

donde I_E y $I_{\mathcal{A}(E)}$ son la identidad en E y en $\mathcal{A}(E)$, respectivamente.

DEMOSTRACIÓN: Para toda $D \in \mathcal{A}(E)$ y $v_1, \ldots, v_n$ de E,

$$\begin{aligned}\widehat{g \circ f}(D)(v_1, \ldots, v_n) &= D(g(f(v_1)), \ldots, g(f(v_n))) = \\ &= \hat{g}(D)(f(v_1), \ldots, f(v_n)) = \\ &= \hat{f}(\hat{g}(D))(v_1, \ldots, v_n) = (\hat{f} \circ \hat{g})(D)(v_1, \ldots, v_n),\end{aligned}$$

de donde $\widehat{g \circ f} = \hat{f} \circ \hat{g}$.

$$\hat{I}_E(D)(v_1, \ldots, v_n) = D(I_E(v_1), \ldots, I_E(v_n)) = \\ = D(v_1, \ldots, v_n) = I_{\mathcal{A}(E)}(D)(v_1, \ldots, v_n),$$

de donde $\hat{I}_E = I_{\mathcal{A}(E)}$. □

Corolario 3.2 *Si f, g son dos endomorfismos de E, e I es la identidad, se cumple*

$$\det(g \circ f) = \det f \cdot \det g \qquad y \qquad \det I = 1.$$

Si f es biyectiva,

$$\det f^{-1} = (\det f)^{-1}. \ \square$$

Corolario 3.3 *Si A, B son de $M_{n\times n}(K)$ e I es la matriz identidad, entonces*

$$\det AB = \det A \cdot \det B \qquad y \qquad \det I = 1. \ \square$$

De (3.2) se deduce que los automorfismos tienen determinante no nulo; el recíproco también es cierto.

Proposición 3.4 *Un endomorfismo f de E es automorfismo si y sólo si* $\det f \neq 0$.

DEMOSTRACIÓN: Sea $e_1, \ldots, e_n$ una base de E. f es automorfismo si y sólo si los vectores $f(e_1), \ldots, f(e_n)$ son linealmente independientes, lo cual, por (1.1), equivale a

$$\det_{(e_i)}(f(e_1), \ldots, f(e_n)) \neq 0;$$

es decir, $\det f \neq 0$. □

Recordemos, finalmente, que si A es la matriz asociada a un endomorfismo f en una cierta base, la traspuesta A^t es la matriz asociada a la aplicación dual f' en la base dual de la anterior (V.6). Esto, juntamente con (2.1), nos dice que f y f' tienen el mismo determinante.

Proposición 3.5 *Sea f un endomorfismo y f' su dual. Entonces se cumple* $\det f = \det f'$. □

VI.4 Regla de Laplace

En este apartado vamos a dar otra expresión del determinante de una matriz que permite, muchas veces, calcular más cómodamente ese determinante. También deduciremos de ella una manera de calcular la matriz inversa de A. Necesitamos, sin embargo, una notación apropiada.

$C_p(1,2,\ldots,n)$, o simplemente C_p, indicará el conjunto de todos los subconjuntos de p elementos de $\{1,2,\ldots,n\}$. Si $H \in C_p$, H' indicará el complementario de H, es decir, el conjunto de elementos de $\{1,2,\ldots,n\}$ que no están en H. f_H designará la permutación que cumple

$$\underset{f_H(1)<\ldots<f_H(p)}{\{f_H(1),\ldots,f_H(p)\}} = H, \qquad \underset{f_H(p+1)<\ldots<f_H(n)}{\{f_H(p+1),\ldots,f_H(n)\}} = H'.$$

Un *menor de orden p* de una matriz $A = (a_i^j)$ es una matriz $p \times p$ formada por los elementos de A situados en p filas y p columnas prefijadas. Es decir, para cada elección de p filas $H = \{i_1,\ldots,i_p\}$ y p columnas $L = \{j_1,\ldots,j_p\}$, hay un menor de orden p, que es

$$\begin{pmatrix} a_{j_1}^{i_1} & \ldots & a_{j_p}^{i_1} \\ \vdots & & \vdots \\ a_{j_1}^{i_p} & \ldots & a_{j_p}^{i_p} \end{pmatrix}.$$

Denotaremos por A_L^H el determinante de este menor.

Proposición 4.1 (Regla de Laplace) *Fijemos $L \in C_p(1,2,\ldots,n)$. Entonces,*

$$\det A = \sum_{H \in C_p} \varepsilon(f_L)\,\varepsilon(f_H)\,A_L^H A_{L'}^{H'}.$$

DEMOSTRACIÓN: Tenemos

$$\begin{aligned} \det A &= \sum_{h \in S_n} \varepsilon(h) a_1^{h_1} \ldots a_n^{h_n} = \text{(reordenando factores en cada sumando)} \\ &= \sum_{h \in S_n} \varepsilon(h) a_{f_L(1)}^{hf_L(1)} \ldots a_{f_L(n)}^{hf_L(n)} = \quad \text{(escribiendo } g = hf_L\text{)} \\ &= \sum_{g \in S_n} \varepsilon(g)\,\varepsilon(f_L)\, a_{f_L(1)}^{g(1)} \ldots a_{f_L(n)}^{g(n)}. \end{aligned}$$

Para cada $g \in \mathcal{S}_n$, sea $H = \{g(1),\ldots,g(p)\}$. Entonces $g = \sigma' \circ \sigma \circ f_{H'}$ donde σ permuta los elementos de H y deja fijos los de H', y σ' permuta los elementos de H' y deja fijos los de H. La última expresión de $\det A$ hallada se puede escribir así:

$$\sum_{\substack{H \in C_p \\ \sigma,\sigma'}} \varepsilon(f_L)\,\varepsilon(\sigma')\,\varepsilon(\sigma)\,\varepsilon(f_H)\, a_{f_L(1)}^{\sigma'\sigma f_H(1)} \ldots a_{f_L(n)}^{\sigma'\sigma f_H(n)}.$$

Observemos que si $j \in \{1,\ldots,p\}$, entonces $\sigma'\sigma f_H(j) = \sigma f_H(j)$, y si $j \in \{p+1,\ldots,n\}$, entonces $\sigma'\sigma f_H(j) = \sigma' f_H(j)$. Podemos reordenar, por tanto, la expresión anterior, obteniendo

$$\sum_{H\in C_p} \varepsilon(f_L)\,\varepsilon(f_H) \left(\sum_{\sigma} \varepsilon(\sigma) a^{\sigma f_H(1)}_{f_L(1)} \ldots a^{\sigma f_H(p)}_{f_L(p)}\right)\left(\sum_{\sigma'} \varepsilon(\sigma') a^{\sigma' f_H(p+1)}_{f_L(p+1)} \ldots a^{\sigma' f_H(n)}_{f_L(n)}\right).$$

En el sumatorio del primer paréntesis aparecen solamente las columnas $\{f_L(1),\ldots,f_L(p)\} = L$ y las filas $\{f_H(1),\ldots,f_H(p)\} = H$ permutadas de todas las maneras posibles. Este sumatorio es, por tanto, la expresión del determinante del menor formado por las columnas L y las filas H: A^H_L. De la misma manera, el sumatorio del segundo paréntesis resulta ser $A^{H'}_{L'}$. Hemos obtenido, pues,

$$\det A = \sum_{H\in C_p} \varepsilon(f_L)\,\varepsilon(f_H)\, A^H_L\, A^{H'}_{L'}. \quad \square$$

Corolario 4.2 $\det A = \sum_{i=1}^{n} (-1)^{i+j} a^i_j A^{i'}_{j'}$, *donde* $i' = \{i\}'$, $j' = \{j\}'$.

DEMOSTRACIÓN: Apliquemos la regla de Laplace (4.1) para $L = \{j\}$. Para cada $H = \{i\}$, $A^H_L = a^i_j$ y $A^{H'}_{L'} = A^{i'}_{j'}$. Además, se ve fácilmente que $\varepsilon(f_L) = (-1)^{j-1}$ y $\varepsilon(f_H) = (-1)^{i-1}$. Sustituyendo ahora en (4.1) obtenemos la expresión del enunciado. $\square$

La expresión del corolario 4.2 se conoce como el *desarrollo del determinante por los elementos de la columna* j.

Recordemos que si A^t es la matriz traspuesta de A, $\det A^t = \det A$ (2.1). De ahí que la regla de Laplace sea también válida fijando p filas ($H \in C_p$),

$$\det A = \sum_{L\in C_p} \varepsilon(f_H)\,\varepsilon(f_L)\, A^H_L\, A^{H'}_{L'}.$$

Análogamente, tenemos una expresión del *desarrollo de un determinante por los elementos de una fila* i

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a^i_j A^{i'}_{j'}.$$

Llamaremos *adjunto de un elemento* a^i_j de una matriz A a

$$X^i_j = (-1)^{i+j} A^{i'}_{j'}.$$

Corolario 4.3 $\sum_{i=1}^{n} a_j^i X_j^i = \det A.$

Para $j \neq k$, $\quad \sum_{i=1}^{n} a_j^i X_k^i = 0.$

DEMOSTRACIÓN: La primera afirmación es (4.2). Para demostrar la segunda, consideremos una matriz $\bar{A}$ con las mismas columnas que A, salvo la columna k, donde vuelve a aparecer la columna j. Claramente, $\det \bar{A} = 0$. Ahora bien, si desarrollamos $\bar{A}$ por los términos de la columna k, obtenemos

$$0 = \det \bar{A} = \sum_{i=1}^{n} \bar{a}_k^i \bar{X}_k^i \ ;$$

pero, tal como hemos definido $\bar{A}$, $\bar{a}_k^i = a_j^i$ y $\bar{X}_k^i = X_k^i$, de donde se obtiene el resultado. □

Este resultado (4.3) puede interpretarse de otra manera. Consideremos la matriz $B = (b_i^j)$ con $b_i^j = X_j^i$; entonces

$$BA = (\det A)\, I \,,$$

donde I indica la matriz identidad. Si $\det A = 0$, $BA = 0$; si $\det A \neq 0$, entonces $(\det A)^{-1} BA = I$.

Corolario 4.4 *La matriz* $C = (c_i^j)$ *con* $c_i^j = X_j^i (\det A)^{-1}$ *es la inversa de la matriz* A. □

Acabaremos este apartado aplicando la regla de Laplace a un caso muy típico. Supongamos que, en una matriz A, los elementos de las p primeras columnas son 0, excepto, quizás, los elementos que están situados en las p primeras filas:

$$A = \begin{pmatrix} a_1^1 & \dots & a_p^1 & a_{p+1}^1 & \dots & a_n^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^p & \dots & a_p^p & a_{p+1}^p & \dots & a_n^p \\ 0 & \dots & 0 & a_{p+1}^{p+1} & \dots & a_n^{p+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{p+1}^n & \dots & a_n^n \end{pmatrix}.$$

Tomando $L = \{1, \dots, p\}$ en (4.1), se obtiene

$$\det A = \begin{vmatrix} a_1^1 & \dots & a_p^1 \\ \vdots & & \vdots \\ a_1^p & \dots & a_p^p \end{vmatrix} \cdot \begin{vmatrix} a_{p+1}^{p+1} & \dots & a_n^{p+1} \\ \vdots & & \vdots \\ a_{p+1}^n & \dots & a_n^n \end{vmatrix}.$$

Un caso particular, para $p = 1$, es el siguiente:

$$\begin{vmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ 0 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ 0 & a_2^n & \dots & a_n^n \end{vmatrix} = a_1^1 \begin{vmatrix} a_2^2 & \dots & a_n^2 \\ \vdots & & \vdots \\ a_2^n & \dots & a_n^n \end{vmatrix}.$$

Así, por ejemplo, si todos los elementos por debajo de la "diagonal principal" son 0 ($a_i^j = 0$ cuando $i < j$), tenemos

$$\begin{vmatrix} a_1^1 & a_2^1 & a_3^1 & \dots & a_n^1 \\ 0 & a_2^2 & a_3^2 & \dots & a_n^2 \\ 0 & 0 & a_3^3 & \dots & a_n^3 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_n^n \end{vmatrix} = a_1^1 a_2^2 \cdots a_n^n.$$

El método general más práctico para calcular un determinante es reducirlo a un determinante "triangular" de este tipo, aplicando 2 y 5 de (VI.1) reiteradamente. El procedimiento es análogo al que explicaremos con detalle en (VII.5).

VI.5 Cálculo del rango de una matriz

Sea E un espacio vectorial de dimensión n. En (1.1) hemos dado un criterio para saber si n vectores de E son, o no, linealmente independientes: lo son si y sólo si su determinante es diferente de 0. En este apartado queremos ampliar este criterio para poder reconocer si k vectores $v_1, \dots, v_k$ de E son, o no, linealmente independientes, cuáles de ellos son combinación lineal del resto y cuáles son las combinaciones lineales que los relacionan.

Sea $e_1, \dots, e_n$ una base de E. Una vez fijada una base, cada vector vendrá representado por una n-pla de elementos del cuerpo K, sus coordenadas, y cada conjunto de vectores $v_1, \dots, v_k$ por una matriz de k columnas formadas por las coordenadas de los k vectores. Toda matriz A, $n \times k$, corresponde a k vectores de E; llamaremos *rango de* A al número de vectores-columna de A linealmente independientes. El problema planteado equivale, pues, al cálculo del rango de una matriz.

Proposición 5.1 *El rango de una matriz A es el máximo de los órdenes de los menores de A con determinante no nulo.*

DEMOSTRACIÓN: Indicaremos por a_j el vector correspondiente a la columna j de $A = (a_j^i)$; es decir, las coordenadas de a_j en la base prefijada son $(a_j^1, \dots, a_j^n)$. Pongamos $r = \operatorname{rang} A$.

Demostraremos, en primer lugar, que todo menor de orden $p > r$ tiene determinante cero. Sean $j_1, \ldots, j_p$ las columnas con las cuales se ha formado el menor. Los vectores $a_{j_1}, \ldots, a_{j_p}$ son linealmente dependientes y, por tanto, uno de ellos es combinación lineal del resto. Con más motivo, en el menor considerado una de las columnas será combinación lineal del resto y el determinante será cero.

Veamos ahora que hay un menor de orden r con determinante no nulo. En la matriz A hay r columnas linealmente independientes; sean $a_{j_1}, \ldots, a_{j_r}$. Completemos estos vectores con vectores de la base $e_1, \ldots, e_n$ en la que trabajamos hasta obtener una base del espacio E: $a_{j_1}, \ldots, a_{j_r}, e_{h_{r+1}}, \ldots, e_{h_n}$. El determinante de estos n vectores es diferente de cero:

$$0 \neq \det_{(e_i)}(a_{j_1}, \ldots, a_{j_r}, e_{h_{r+1}}, \ldots, e_{h_n}) = \begin{vmatrix} a^1_{j_1} & \ldots & a^1_{j_r} & 0 & 0 & \ldots & 0 \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & 1 & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & 1 \\ \cdot & & \cdot & \cdot & 1 & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ a^n_{j_1} & \ldots & a^n_{j_r} & 0 & 0 & \ldots & 0 \end{vmatrix}.$$

En cada una de las $n - r$ últimas columnas, todos los elementos son 0, excepto uno que vale 1; ese elemento aparece en cada columna en una fila distinta.

Apliquemos, ahora, la regla de Laplace al cálculo de este determinante. Fijemos, para ello, las $n - r$ últimas columnas; el único menor de determinante no nulo que podemos formar con ellas es el correspondiente a las $n - r$ filas en que aparece un 1. Si $i_1, \ldots, i_r$ son las filas restantes, queda

$$0 \neq \det_{(e_i)}(a_{j_1}, \ldots, e_{h_n}) = \pm \begin{vmatrix} a^{i_1}_{j_1} & \ldots & a^{i_1}_{j_r} \\ \vdots & & \vdots \\ a^{i_r}_{j_1} & \ldots & a^{i_r}_{j_r} \end{vmatrix}.$$

Obtenemos así un menor de orden r con determinante no nulo. □

Corolario 5.2 *Una matriz A y su traspuesta A^t tienen el mismo rango:*

$$\operatorname{rang} A = \operatorname{rang} A^t. \ \square$$

Corolario 5.3 *Una aplicación lineal f y su dual f' tienen el mismo rango.*

DEMOSTRACIÓN: El rango de una aplicación lineal $f : E \longrightarrow F$ es la dimensión de $\operatorname{Im} f$ (V.1). Si $e_1, \ldots, e_n$ es una base de E, entonces

$\operatorname{Im} f = \langle f(e_1), \ldots, f(e_n) \rangle$. Las coordenadas de $f(e_1), \ldots, f(e_n)$ forman las columnas de la matriz A asociada a f. Por tanto,

$$\operatorname{rang} f = \operatorname{rang} A.$$

El resultado se deduce ahora de (5.2) y de (V.6.2). □

Nota:

El corolario 5.3 se obtiene también como consecuencia de (V.7.1).

La proposición 5.1 proporciona un método para calcular el rango de una matriz A mediante el cálculo de los determinantes de los menores de A, empezando por los de orden máximo hasta encontrar uno no nulo. La proposición siguiente reduce sensiblemente, en muchas ocasiones, el número de determinantes a calcular.

Proposición 5.4 *Sean* $a_i = (a_i^1, \ldots, a_i^n)$, $i = 1, \ldots, r$, *vectores linealmente independientes, y sea*

$$M = \begin{vmatrix} a_1^{i_1} & \ldots & a_r^{i_1} \\ \vdots & & \vdots \\ a_1^{i_r} & \ldots & a_r^{i_r} \end{vmatrix} \neq 0.$$

Un vector $v = (v^1, \ldots, v^n)$ *es combinación lineal de* $a_1, \ldots, a_r$ *si y sólo si, para todo* j,

$$\begin{vmatrix} a_1^{i_1} & \ldots & a_r^{i_1} & v^{i_1} \\ \vdots & & \vdots & \vdots \\ a_1^{i_r} & \ldots & a_r^{i_r} & v^{i_r} \\ a_1^{j} & \ldots & a_r^{j} & v^{j} \end{vmatrix} = 0.$$

DEMOSTRACIÓN: El becho de que, si $a_1, \ldots, a_r, v$ son linealmente dependientes, esos determinantes sean 0, es consecuencia de (5.1). Supongamos, recíprocamente, que esos determinantes son todos cero. Desarrollando por la última fila, obtenemos

$$a_1^j \begin{vmatrix} a_2^{i_1} & \ldots & a_r^{i_1} & v^{i_1} \\ \vdots & & \vdots & \vdots \\ a_2^{i_r} & \ldots & a_r^{i_r} & v^{i_r} \end{vmatrix} - a_2^j \begin{vmatrix} a_1^{i_1} & a_3^{i_1} & \ldots & a_r^{i_1} & v^{i_1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^{i_r} & a_3^{i_r} & \ldots & a_r^{i_r} & v^{i_r} \end{vmatrix} + \ldots$$

$$\ldots \pm v^j \begin{vmatrix} a_1^{i_1} & \ldots & a_r^{i_1} \\ \vdots & & \vdots \\ a_1^{i_r} & \ldots & a_r^{i_r} \end{vmatrix} = 0 \qquad \text{para todo } j.$$

Denotemos por $M_1, M_2, \ldots$ los determinantes de esta expresión. El último es M y ninguno de ellos depende del índice j. Despejando v^j, obtenemos

$$v^j = \pm M_1 M^{-1} a_1^j \pm M_2 M^{-1} a_2^j \pm \ldots \pm M_r M^{-1} a_r^j \qquad \text{para todo } j,$$

de donde

$$v = \pm (M_1 M^{-1}) a_1 \pm (M_2 M^{-1}) a_2 \pm \ldots \pm (M_r M^{-1}) a_r \,,$$

que nos da v como combinación lineal de $a_1, \ldots, a_r$. □

Nota:

Observemos que en (5.4) la condición es obvia para $j = i_1, \ldots, i_r$.

Ejemplo:

Consideremos la matriz

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -2 & 1 & 1 \\ 4 & 3 & 6 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

$a_1 = (1, -1, 4, 0)$ y $a_2 = (1, -2, 3, 0)$ son claramente linealmente independientes, ya que sus coordenadas no son proporcionales. Escojamos

$$M = \begin{vmatrix} 1 & 1 \\ -1 & -2 \end{vmatrix} = -1 \neq 0.$$

Para ver si $a_3 = (1, 1, 6, 0)$ depende o no linealmente de a_1, a_2, debemos calcular

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 4 & 3 & 6 \end{vmatrix} = 0, \qquad \begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Así pues, a_3 es combinación lineal de a_1, a_2. Para encontrar esa combinación, hacemos

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ a_1^j & a_2^j & a_3^j \end{vmatrix} = 3a_1^j - 2a_2^j - a_3^j = 0,$$

de donde

$$a_3^j = 3a_1^j - 2a_2^j \qquad \text{para todo } j,$$

y, por tanto,

$$a_3 = 3a_1 - 2a_2.$$

Sigamos calculando el rango de A estudiando si $a_4 = (1, 1, -1, 2)$ es o no combinación lineal de a_1, a_2. Para ello, calculemos los determinantes de los menores de orden 3 formados a partir de M. Tenemos

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 0 & 2 \end{vmatrix} = -2 \neq 0.$$

a_1, a_2, a_4 son, pues, linealmente independientes y la matriz A tiene rango igual a 3.

VI.6 Nota histórica

Los determinantes aparecen por primera vez en la resolución de sistemas de ecuaciones lineales (ver Cap. VII) en 1772 de la mano de Alexandre-Théophile Vandermonde (1735–1796) y se aplican ya en el siglo 19 a la teoría de la eliminación, transformación de coordenadas, cambio de variable, etc.

La palabra "determinante" fue introducida por Carl Friedrich Gauss (1777–1855) en el estudio de ciertas formas cuadráticas. No obstante, el tratamiento sistemático y prácticamente actual es debido a Augustin-Louis Cauchy (1789–1857) en el año 1815, quien demuestra entre otras propiedades la regla de Laplace (demostrada ya por Pierre-Simon de Laplace (1749–1827) en 1772), demostrando casi todas las propiedades mencionadas en el presente capítulo, y a James Joseph Sylvester (1817–1897), quien la aplica a problemas de la teoría de ecuaciones.

Leopold Kronecker (1823–1891) y Karl Wilhelm Weierstrass (1815–1897) (según algunos, el mejor profesor que nunca haya tenido una Universidad) introdujeron en sus cursos en Berlín los determinantes como formas multilineales alternadas.

VI.7 Ejercicios

1. Calcular el determinante de la matriz $A = (a_j^i)$, donde $a_j^i = |i - j|$.

2. Probar que $(x-1)^3$ divide al polinomio

$$\begin{vmatrix} 1 & x & x^2 & x^3 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \end{vmatrix}.$$

3. Dada la matriz

$$M = \left(\begin{array}{c|c} 0 & A \\ \hline \operatorname{adj} A & 0 \end{array}\right),$$

calcular $\det M$ en función de $\det A$. ($\operatorname{adj} A$ indica la *matriz adjunta* de A, que se obtiene a partir de A sustituyendo cada elemento por su adjunto.)

4. Una matriz $n \times n$ $A = (a_i^j)$ se llama *hemisimétrica* si $a_i^j = -a_j^i$ para todo i, j. Probar que, si A es hemisimétrica, $\det A = (-1)^n \det A$. Deducir de ello que las matrices hemisimétricas de orden impar tienen determinante cero.

5. Descomponer el polinomio de $\mathbf{C}[x]$

$$\begin{vmatrix} 1+x+x^2 & 1 & 1 & 1 & 1 \\ 1 & 1+x+x^2 & 1 & 1 & 1 \\ 1 & 1 & 1+x+x^2 & 1 & 1 \\ 1 & 1 & 1 & 1+x+x^2 & 1 \\ 1 & 1 & 1 & 1 & 1+x+x^2 \end{vmatrix}$$

en factores irreducibles.

6. Sea E el espacio vectorial de las funciones reales de variable real generado por sen y cos. Calcular el determinante del endomorfismo de E definido por la derivación.

7. Demostrar que la aplicación que va de las matrices invertibles de $M_{n\times n}(K)$ (que denotaremos $GL(n, K)$) al grupo multiplicativo del cuerpo K:

$$\det : GL(n, K) \longrightarrow K - \{0\}$$

es un homomorfismo de grupos. Estudiarlo.

8. Sea A una matriz $n \times n$ y $\operatorname{adj} A$ la matriz adjunta de A (ejercicio 3). Demostrar:

a) $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

b) Si $\operatorname{rang} A = n - 1$, entonces $\operatorname{rang}(\operatorname{adj} A) = 1$.

9. Repetir ahora los ejercicios 13 y 14 del capítulo IV.

10. Consideremos el determinante

$$\mathrm{D}(n,k) = \begin{vmatrix} 1^k & 2^k & \dots & n^k \\ 2^k & 3^k & \dots & (n+1)^k \\ \vdots & \vdots & & \vdots \\ n^k & (n+1)^k & \dots & (2n-1)^k \end{vmatrix}.$$

a) Calcular $\mathrm{D}(1,1)$, $\mathrm{D}(2,1)$, $\mathrm{D}(3,1)$, $\mathrm{D}(4,1)$.

b) Demostrar que $\mathrm{D}(n,2) = 0 \quad \forall n > 3$.

c) Demostrar que $\mathrm{D}(n,k) = 0 \quad \forall n > k+1$.

11. Demostrar que, si A es invertible, $(A^{-1})^{\mathrm{t}} = (A^{\mathrm{t}})^{-1}$.

VI.8 Ejercicios para programar

12. Dada una matriz $A \in M_{n\times n}(\mathbf{R})$, elaborar un programa que calcule $\det A$ por el método explicado al final del §4.

(Indicación: permutando filas, si es necesario, se consigue $a_1^1 \neq 0$. Guardar en una variable el posible cambio de signo. Anular todos los elementos de la primera columna bajo la diagonal. Si X_1^1 es el adjunto de a_1^1, entonces $\det A = a_1^1 X_1^1$. Repetir el proceso para X_1^1.)

13. Contar cuántas sumas y multiplicaciones son necesarias para obtener $\det A$ por el método del ejercicio anterior. Contar también cuántas habría que hacer si se utilizara el desarrollo de (4.2). Observar la gran diferencia que hay cuando n es grande y sacar consecuencias.

14. (Ver el ejercicio IV.17.) Demostrar que, si $\det A \neq 0$, la descomposición LU es siempre posible, salvo que tal vez haya que permutar las filas de A. Observar que la descomposición $A = LU$, si es posible, es única.

15. Aplicar el ejercicio anterior para

a) Calcular $\det A$.

b) Resolver un sistema de ecuaciones $Ax = b$.

(Contar cuántas sumas y multiplicaciones son necesarias para resolver $Ax = b$ por este método. Observar que es esencialmente equivalente al método de Gauss (Cap. VII), siempre que $\det A \neq 0$.)

Capítulo VII

Sistemas de ecuaciones lineales

VII.1 Planteo del problema

Queremos resolver el siguiente problema: supongamos que tenemos un sistema de ecuaciones lineales

$$\begin{cases} a_1^1 x^1 + \ldots + a_n^1 x^n &= b^1 \\ \cdots\cdots\cdots & \\ a_1^m x^1 + \ldots + a_n^m x^n &= b^m \end{cases}$$

donde los a_i^j y b^j son elementos conocidos de un cuerpo K. Se trata de encontrar las n-plas $(x^1, \ldots, x^n)$ de K^n que satisfacen todas estas ecuaciones. Pongamos

$$A = \begin{pmatrix} a_1^1 & \ldots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \ldots & a_n^m \end{pmatrix}, \quad b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix}, \quad x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

Podemos escribir entonces el sistema anterior así:

$$Ax = b.$$

Este mismo problema se puede plantear de otra manera. Sean E y F dos espacios vectoriales sobre K con bases $u_1, \ldots, u_n$ y $v_1, \ldots, v_m$ respectivamente. Sabemos que existe una aplicación lineal $f : E \longrightarrow F$ que tiene como matriz asociada en esas bases la matriz A (V.2). Designemos por b el vector $b = b^1 v_1 + \ldots + b^m v_m$ de F y por x el vector $x = x^1 u_1 + \ldots + x^n u_n$ de E. La condición $Ax = b$ equivale a

$$f(x) = b.$$

El problema consiste, por tanto, en encontrar las antiimágenes del vector b por la aplicación f.

Recíprocamente, si suponemos dada una aplicación lineal $f : E \longrightarrow F$ y un vector $b \in F$, el problema de encontrar las antiimágenes x de b equivale a resolver el sistema de ecuaciones lineales

$$Ax = b,$$

donde A es la matriz asociada a f en unas ciertas bases y b, x son matrices de una columna formadas por las coordenadas de los vectores correspondientes. Tanto en la primera interpretación como en la segunda, nuestro objetivo es:

1. Saber cuándo el problema tiene solución y cuándo no.
2. Saber cuántas soluciones tiene.
3. Dar un método para encontrar todas las soluciones.

Para resolver cada una de estas cuestiones usaremos la interpretación que nos resulte más cómoda. En general, los razonamientos de tipo teórico son más simples en el lenguaje de aplicaciones lineales y la resolución de los casos concretos se lleva a cabo mediante el lenguaje de matrices.

VII.2 Existencia de soluciones

Observemos que existe un $x \in E$ tal que $f(x) = b$ si, equivalentemente,

$$\begin{array}{rcl} b \in \operatorname{Im} f & \Leftrightarrow & \langle f(u_1), \ldots, f(u_n)\rangle = \langle f(u_1), \ldots, f(u_n), b\rangle \quad \Leftrightarrow \\ \Leftrightarrow \quad \operatorname{rang} f & = & \dim \operatorname{Im} f = \dim\langle f(u_1), \ldots, f(u_n), b\rangle. \end{array}$$

En la demostración de (VI.5.3) vimos que $\operatorname{rang} f = \operatorname{rang} A$. Análogamente,

$$\dim\langle f(u_1), \ldots, f(u_n), b\rangle = \operatorname{rang}(A, b),$$

donde (A, b) indica la matriz que se obtiene añadiendo a A una columna formada por las coordenadas de b. Así pues, en lenguaje de matrices, tenemos que el sistema $Ax = b$ tiene solución si y sólo si $\operatorname{rang} A = \operatorname{rang}(A, b)$. En (VI.5) dimos un método para calcular el rango de una matriz. Tenemos, por tanto, resuelto el problema de saber si el sistema tiene o no soluciones. ¿Cuántas soluciones hay? Es decir, ¿cuántas antiimágenes tiene b? En la demostración del teorema de isomorfismo (V.3.1) vimos que todos los vectores de E que se aplican en el mismo vector de F forman una clase módulo el núcleo de f : $x_0 + \operatorname{Nuc} f$. Así pues, b tiene tantas antiimágenes como vectores tiene $\operatorname{Nuc} f$. Además, todas las antiimágenes se obtienen

sumando a una solución particular x_0 los vectores de Nuc f. Los vectores de Nuc f tienen por coordenadas las soluciones del sistema homogéneo asociado al sistema dado

$$Ax = 0.$$

Por tanto, las soluciones del sistema $Ax = b$ se obtienen sumando a una solución particular las soluciones del sistema homogéneo asociado.

VII.3 Regla de Cramer

Vamos a dar en este apartado un método para encontrar las soluciones de un sistema de ecuaciones lineales en un caso muy particular. Supongamos que en la expresión

$$Ax = b$$

la matriz A es cuadrada, es decir, $n \times n$, y $\det A \neq 0$. O bien, equivalentemente, supongamos que en la expresión $f(x) = b$ f es un isomorfismo. Hay, entonces, una solución y sólo una, que es

$$x = A^{-1}b.$$

Hemos visto en (VI.4.4) que si $A_j^{-1} = (c_j^i)$, $c_j^i = X_i^j(\det A)^{-1}$, donde X_i^j es el adjunto del elemento a_i^j de A. Por tanto,

$$x^i = \sum_{j=1}^{n} X_i^j b^j (\det A)^{-1}.$$

Ahora bien, $\sum_{j=1}^{n} X_i^j b^j$ es el desarrollo por los términos de la columna i del determinante de una matriz con los mismos elementos que A, salvo la columna i, que ha sido sustituida por $(b^1, \ldots, b^n)$. Designemos ese determinante por $\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)$. Entonces

$$x^i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det A}, \qquad i = 1, \ldots, n.$$

Esta fórmula se conoce con el nombre de *regla de Cramer*.

VII.4 Resolución de un sistema de ecuaciones lineales

Dado el sistema $Ax = b$,

$$\begin{cases} a_1^1 x^1 + \ldots + a_n^1 x^n &= b^1 \\ \cdots\cdots\cdots & \\ a_1^m x^1 + \ldots + a_n^m x^n &= b^m, \end{cases}$$

supongamos que existen soluciones, es decir, que

$$\operatorname{rang} A = \operatorname{rang}(A, b) = r.$$

Reordenando las ecuaciones y las incógnitas convenientemente podemos suponer (VI.5.1) que

$$M = \begin{vmatrix} a_1^1 & \dots & a_r^1 \\ \vdots & & \vdots \\ a_1^r & \dots & a_r^r \end{vmatrix} \neq 0.$$

En la matriz (A, b), las $m - r$ últimas filas son combinaciones lineales de las anteriores; es decir, en el sistema dado, las $m - r$ últimas ecuaciones son combinaciones lineales de las r primeras. Por tanto, el sistema original tiene exactamente las mismas soluciones que el sistema

$$\begin{cases} a_1^1 x^1 + \dots + a_n^1 x^n &= b^1 \\ \cdots\cdots\cdots & \\ a_1^r x^1 + \dots + a_n^r x^r &= b^r \end{cases}$$

formado por las r primeras ecuaciones. Basta, pues, encontrar las soluciones de este sistema parcial.

Escribiremos el sistema anterior en la forma

$$\begin{cases} a_1^1 x^1 + \dots + a_r^1 x^r &= b^1 - a_{r+1}^1 x^{r+1} - \dots - a_n^1 x^n \\ \cdots\cdots\cdots & \\ a_1^r x^1 + \dots + a_r^r x^r &= b^r - a_{r+1}^r x^{r+1} - \dots - a_n^r x^n. \end{cases}$$

Para cada uno de los conjuntos de valores que demos a $x^{r+1}, \dots, x^n$ arbitrariamente, esto es un sistema de r ecuaciones con r incógnitas. El determinante de su matriz es $M \neq 0$. Podemos aplicar, pues, la regla de Cramer y obtenemos unos valores únicos para las incógnitas $x^1, \dots, x^r$:

$$x^i = M^{-1} \det(a_1, \dots, a_{i-1}, b - a_{r+1} x^{r+1} - \dots - a_n x^n, a_{i+1}, \dots, a_r).$$

(Aquí $b = (b^1, \dots, b^r)$ y $a_j = (a_j^1, \dots, a_j^r)$.)

Ahora bien,

$$\det(a_1, \dots, b - a_{r+1} x^{r+1} - \dots - a_n x^n, \dots, a_r) = \det(a_1, \dots, b, \dots, a_r) - \\ - \det(a_1, \dots, a_{r+1}, \dots, a_r) x^{r+1} - \dots - \det(a_1, \dots, a_n, \dots, a_r) x^n.$$

Observemos que estos determinantes se obtienen sustituyendo en M la columna i por b, $a_{r+1}, \dots, a_n$ sucesivamente. Pongamos

$$M_b^i = M^{-1} \det(a_1, \dots, b, \dots a_r), \quad M_j^i = M^{-1} \det(a_1, \dots, a_j, \dots, a_r);$$

entonces

$$x^i = M_b^i - M_{r+1}^i x^{r+1} - \ldots - M_n^i x^n, \qquad i = 1, \ldots, r.$$

Esta expresión nos da la *solución general* del sistema en función de las $n - r$ incógnitas arbitrarias $x^{r+1}, \ldots, x^n$.

Observaciones:

1. Tomando $x^{r+1} = \ldots = x^n = 0$ obtenemos una solución particular con $x^i = M_b^i$ para $i = 1, \ldots, r$:

$$(M_b^1, \ldots, M_b^r, 0, \ldots, 0).$$

2. Para resolver el sistema homogéneo $Ax = 0$ asociado al nuestro, hemos de sustituir b por $(0, \ldots, 0)$ en todo lo anterior. Resulta entonces

$$x^i = -M_{r+1}^i x^{r+1} - \ldots - M_n^i x^n, \quad i = 1, \ldots, r.$$

El conjunto de estas soluciones forma un espacio vectorial (que corresponde a $\operatorname{Nuc} f$, tal como vimos en el apartado 1). Podemos obtener una base de este espacio de las soluciones del sistema homogéneo dando a las incógnitas arbitrarias $x^{r+1}, \ldots, x^n$ el valor 0, excepto una con valor 1:

$$(-M_{r+1}^1, \ldots, -M_{r+1}^r, 1, 0, \ldots, 0)$$

$$(-M_{r+2}^1, \ldots, -M_{r+2}^r, 0, 1, \ldots, 0)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$(-M_n^1, \ldots, -M_n^r, 0, 0, \ldots, 1).$$

Estas soluciones son, en efecto, linealmente independientes y su número es

$$n - r = n - \operatorname{rang} A = n - \dim \operatorname{Im} f = \dim \operatorname{Nuc} f.$$

3. La solución general obtenida es suma de la solución particular

$$(M_b^1, \ldots, M_b^r, 0, \ldots, 0)$$

y la solución general del sistema homogéneo asociado, lo que ya sabíamos desde el apartado 2.

VII.5 Método de Gauss

Otro método para resolver sistemas de ecuaciones es el de reducción o de Gauss-Jordan. Su justificación teórica está basada en unos razonamientos muy simples. Sea $f : E \longrightarrow F$ una aplicación lineal y sean $u_1, \ldots, u_n$ y $v_1, \ldots, v_m$ bases de E y F respectivamente. Denotamos, como siempre, por A la matriz asociada a f en estas bases y por b el vector de F cuyas coordenadas son $(b^1, \ldots, b^m)$. Queremos encontrar las antiimágenes x de b: $f(x) = b$. Los cambios en la base de F dan lugar a cambios en la matriz asociada y en las coordenadas de b. Las coordenadas de x permanecen invariables. Por tanto, si A_1 y b_1 son la nueva matriz asociada y las nuevas coordenadas de b, las soluciones del sistema $A_1 x = b_1$ coinciden con las del sistema original $Ax = b$. Efectuando cambios en la base de F podemos conseguir un sistema con una matriz lo bastante simple como para que el hecho de encontrar la solución general no implique ningún cálculo. Además, los cambios que efectuaremos son únicamente de tres tipos muy simples:

1. Permutación del orden de los vectores de la base de F. Naturalmente, entonces, las coordenadas de

$$f(u_i) = a_i^1 v_1 + \ldots + a_i^m v_m$$

y de

$$b = b^1 v_1 + \ldots + b^m v_m$$

quedan permutadas, lo cual equivale a que en (A, b) las filas queden permutadas.

2. Sustitución de un vector v_j de la base por kv_j con $k \neq 0$. Entonces

$$f(u_i) = a_i^1 v_1 + \ldots + (a_i^j k^{-1}) k v_j + \ldots + a_i^m v_m$$

$$b = b^1 v_1 + \ldots + (b^j k^{-1}) k v_j + \ldots + b^m v_m.$$

Es decir, en la matriz (A, b) la fila j queda multiplicada por k^{-1}.

3. Sustitución de un vector v_j de la base de F por $v_j + kv_h$ $(h \neq j)$. Entonces

$$f(u_i) = a_i^1 v_1 + \ldots + a_i^j (v_j + k v_h) + \ldots + (a_i^h - a_i^j k) v_h + \ldots + a_i^m v_m$$

$$b = b^1 v_1 + \ldots + b^j (v_j + k v_h) + \ldots + (b^h - b^j k) v_h + \ldots + b^m v_m.$$

Es decir, en la matriz (A, b), a la fila h se le resta la fila j multiplicada por k.

Haciendo cambios del tipo 1, 2 y 3 y permutando, si es necesario, el orden de las incógnitas, lo que equivale a permutar el orden de las n primeras columnas de (A, b), obtenemos una matriz de la forma

$$\begin{pmatrix} 1 & 0 & \dots & 0 & c_{r+1}^1 & \dots & c_n^1 & d^1 \\ 0 & 1 & \dots & 0 & c_{r+1}^2 & \dots & c_n^2 & d^2 \\ \cdot & \cdot & \dots & \cdot & \dots & \dots & \cdot & \cdot \\ 0 & 0 & \dots & 1 & c_{r+1}^r & \dots & c_n^r & d^r \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & d^{r+1} \\ \cdot & \cdot & \dots & \cdot & \dots & \dots & \cdot & \cdot \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & d^m \end{pmatrix}.$$

El sistema

$$\begin{cases} x^1 + c_{r+1}^1 x^{r+1} + \dots + c_n^1 x^n = d^1 \\ \dots\dots\dots\dots \\ x^r + c_{r+1}^r x^{r+1} + \dots + c_n^r x^n = d^r \\ 0 = d^{r+1} \\ \dots\dots \\ 0 = d^m \end{cases}$$

tiene, pues, las mismas soluciones que el original, salvo tal vez el orden de las incógnitas. Por tanto, el sistema es compatible si y sólo si

$$d^{r+1} = \dots = d^m = 0$$

y, en este caso, la solución general es

$$x^i = d^i - c_{r+1}^i x^{r+1} - \dots - c_n^i x^n, \quad i = 1, \dots, r.$$

Nota:

Los cambios en la matriz (A, b) se efectúan de la manera siguiente: si la primera columna es toda 0, se pasa al lugar n. Si hay un elemento no nulo, se permutan las filas de forma que quede en primer lugar. Con un cambio del tipo 2 se puede conseguir que este elemento pase a ser un 1 y con cambios del tipo 3 se puede consegnir que el resto de la columna sea 0. La primera columna queda, así, en la forma deseada. Supongamos que tenemos h columnas en la forma deseada. Si en la columna $h+1$ los elementos de las filas $h+1, \dots, m$ son 0, la situamos en el lugar n. En caso contrario, colocamos un elemento no nulo en la fila $h+1$, permutando únicamente las filas $h+1, \dots, m$. Con cambios del tipo 2 y 3 podemos conseguir que este elemento sea 1 y el resto de la columna sea 0. Observemos que de esta forma las columnas anteriores no varían. El proceso puede continuar hasta obtener una matriz como la que hemos escrito más arriba.

Una de las ventajas del método de Gauss es que se puede aplicar simultáneamente a sistemas de ecuaciones con la misma matriz y diferentes términos independientes. Sean, por ejemplo, $Ax = b$ y $Ax = c$ dos sistemas con matriz A. Si efectuamos los cambios necesarios en la matriz (A, b, c), resolveremos al mismo tiempo los dos sistemas.

Ejemplo:

Consideremos el sistema

$$\begin{cases} x^1 - 2x^2 + 3x^3 + 5x^4 - 4x^5 = b^1 \\ 2x^1 - 4x^2 + 6x^3 + 5x^4 + 2x^5 = b^2 \\ 2x^1 - 5x^2 + 7x^3 + 7x^4 + 3x^5 = b^3 \\ -x^1 + x^2 - 2x^3 - 3x^4 + 5x^5 = b^4 \end{cases}$$

y supongamos que nos interesan las soluciones cuando los términos (b^1, b^2, b^3, b^4) son $(2, -6, -7, -3)$ y $(-3, -1, 1, 2)$. Escribamos:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & & \\ \hline \left(\begin{array}{ccccccc} 1 & -2 & 3 & 5 & -4 & 2 & -3 \\ 2 & -4 & 6 & 5 & 2 & -6 & -1 \\ 2 & -5 & 7 & 7 & 3 & -7 & 1 \\ -1 & 1 & -2 & -3 & 5 & -3 & 2 \end{array}\right) \end{array}.$$

Hemos escrito una primera fila que indica el orden de las columnas correspondientes a la matriz A. Empecemos, por ejemplo, restando de la segunda, tercera y cuarta filas la primera multiplicada por 2, 2 y -1 respectivamente:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & & \\ \hline \left(\begin{array}{ccccccc} 1 & -2 & 3 & 5 & -4 & 2 & -3 \\ 0 & 0 & 0 & -5 & 10 & -10 & 5 \\ 0 & -1 & 1 & -3 & 11 & -11 & 7 \\ 0 & -1 & 1 & 2 & 1 & -1 & -1 \end{array}\right) \end{array}.$$

Para continuar, hemos de cambiar el orden, por ejemplo, de las filas segunda y tercera. Entonces, con cambios del tipo 2 y 3 obtenemos:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & & \\ \hline \left(\begin{array}{ccccccc} 1 & 0 & 1 & 11 & -26 & 24 & -17 \\ 0 & 1 & -1 & 3 & -11 & 11 & -7 \\ 0 & 0 & 0 & -5 & 10 & -10 & 5 \\ 0 & 0 & 0 & 5 & -10 & 10 & -8 \end{array}\right) \end{array}.$$

Para poder continuar con el método general explicado en la nota, tenemos que cambiar el orden de las columnas. Fijémonos, sin embargo, en que si aquí sumamos las dos últimas filas, obtenemos como última fila

$$(0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad -3).$$

Esto nos dice que el sistema, con la segunda serie de términos independientes, es incompatible. Continuemos, por tanto, sólo con la primera columna de términos independientes. Suprimamos también la última fila de ceros, que corresponde a la "ecuación" $0 = 0$. Queda

$$\begin{array}{c} \begin{array}{ccccc} 1 & 2 & 4 & 3 & 5 \end{array} \\ \left(\begin{array}{ccccc|c} 1 & 0 & 11 & 1 & -26 & 24 \\ 0 & 1 & 3 & -1 & -11 & 11 \\ 0 & 0 & -5 & 0 & 10 & -10 \end{array}\right), \end{array}$$

de donde resulta

$$\begin{array}{c} \begin{array}{cccccc} 1 & 2 & 4 & 3 & 5 \end{array} \\ \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -4 & 2 \\ 0 & 1 & 0 & -1 & -5 & 5 \\ 0 & 0 & 1 & 0 & -2 & 2 \end{array}\right) \end{array}$$

y, por tanto, las soluciones del primer sistema dado son las soluciones del sistema

$$\left\{\begin{array}{rcl} x^1 + x^3 - 4x^5 & = & 2 \\ x^2 - x^3 - 5x^5 & = & 5 \\ x^4 - 2x^5 & = & 2; \end{array}\right.$$

es decir,

$$\left\{\begin{array}{rcl} x^1 & = & 2 - x^3 + 4x^5 \\ x^2 & = & 5 + x^3 + 5x^5 \\ x^4 & = & 2 + 2x^5. \end{array}\right.$$

VII.6 Cálculo de la matriz inversa

Dada $A \in M_{n\times n}(K)$, se trata de encontrar, si existe, una matriz $B = (b_i^j)$ que cumpla $AB = I$. Esto equivale a buscar las n columnas $b_i = (b_i^1, \ldots, b_i^n)$ de B de forma que

$$Ab_i = e_i$$

donde $e_i = (0, \ldots, 1, \ldots, 0)$ es la columna i de I. En otras palabras, debemos resolver los n sistemas

$$Ax = e_i, \qquad i = 1, \ldots, n,$$

todos con la misma matriz. Hagámoslo por el método de Gauss. Consideremos la matriz $(A, e_1, \ldots, e_n)$ y modifiquémosla como en el apartado anterior. Observemos que $(e_1, \ldots, e_n)$ es precisamente la matriz identidad. Partimos, pues, de

$$(A, I)$$

y llegaremos a una matriz del tipo (I, B):

$$\begin{pmatrix} 1 & \ldots & 0 & b_1^1 & \ldots & b_n^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \ldots & 1 & b_1^n & \ldots & b_n^n \end{pmatrix} \neq 0.$$

La solución del primer sistema, es decir, la primera columna de B, es $x^i = b_1^i$, $i = 1, \ldots, n$, etc. En resumen, resulta que la matriz (b_i^j) que hemos obtenido es precisamente la matriz inversa buscada. Naturalmente, puede suceder que la matriz A no se pueda transformar en la matriz identidad I. Entonces uno de los sistemas $Ax = e_i$ resulta incompatible y A no tiene matriz inversa.

Ejercicio:

Probar que, si A no tiene inversa, uno de los sistemas $Ax = e_i$ es incompatible.

VII.7 Nota histórica

El estudio de los sistemas de ecuaciones lineales fue iniciado por Gottfried Wilhelm Leibniz (1646–1716). Ya en 1693, Leibniz considera un sistema de tres ecuaciones lineales con dos incógnitas, elimina las incógnitas y obtiene un determinante (la *resultante* del sistema).

La solución de sistemas de ecuaciones lineales utilizando lo que hoy llamamos determinantes fue ideada por Colin MacLaurin (1698–1746) en 1729. Gabriel Cramer (1704–1752) primero, y después en 1764 Étienne Bézout (1730–1783), demostraron que un sistema homogéneo cuadrado tiene solución no trivial si y sólo si el determinante del sistema se anula. También durante ese siglo Jean-Baptiste le Rond d'Alembert (1717–1783) demuestra que la solución general de un sistema se obtiene sumando una solución particular a las soluciones del sistema homogéneo asociado.

La existencia y número de soluciones fueron temas discutidos por Henry J. S. Smith (1826–1883) en 1861 en términos de los rangos de la matriz del sistema y de la matriz ampliada. La mayor parte de resultados en este sentido son debidos a Leopold Kronecker (1823–1891) y a Arthur Cayley (1821–1895) y aparecen ya en 1867 en el libro de Charles L. Dodgson (Lewis Carroll (1832-1898), el autor de *Alicia en el país de las maravillas*) *An elementary theory of determinants.*

VII.8 Ejercicios

1. Dado el sistema

$$\left\{\begin{array}{rcl} x+by+az &=& 1 \\ ax+by+z &=& a \\ x+aby+z &=& b, \end{array}\right.$$

a) ¿Para qué valores de a y b el sistema tiene solución?

b) Resolverlo y determinar cuándo tiene una única solución.

2. Discutir el sistema homogéneo

$$\left\{\begin{array}{rcl} -6x-6y+(11-a)z &=& 0 \\ 3x+(12-a)y-6z &=& 0 \\ (2-a)x+3y-6z &=& 0 \end{array}\right.$$

y encontrar sus soluciones.

3. Resolver los sistemas de congruencias

a)
$$\left.\begin{array}{rcl} x+2y+z &\equiv& 1 \\ 2x+y+2z &\equiv& 1 \\ y+2z &\equiv& 1 \end{array}\right\} \bmod 5$$

b)
$$\left.\begin{array}{rcl} x+2y+z &\equiv& -1 \\ 2x+y+2z &\equiv& 1 \\ y+2z &\equiv& 1 \end{array}\right\} \bmod 3.$$

4. Encontrar un sistema de ecuaciones lineales homogéneo cuyas soluciones sean exactamente los vectores del subespacio vectorial de $(\mathbf{Z}/(7))^4$ generado por

$$(1,0,1,-1), \quad (2,1,3,0) \quad \text{y} \quad (1,3,4,5).$$

5. Discutir según los valores del parámetro a el sistema de congruencias módulo 5

$$\left.\begin{array}{rcl} x+y+3z &\equiv& 2 \\ 2x+3y+4z &\equiv& 0 \\ 3x+4y+az &\equiv& 3 \end{array}\right\}.$$

6. Resolver el sistema de ecuaciones lineales complejas

$$\begin{cases} x+y+iz+t &= 0 \\ 2x-y+2z-t &= 1 \\ x+iy-z+it &= 2 \\ x+y+z-t &= 0. \end{cases}$$

7. Determinar $a \in \mathbf{R}$ para que el endomorfismo f de $\mathbf{R}^3$ definido por

$$f(x,y,z) = (ax+y+z,\ x+ay+z,\ x+y+az)$$

tenga núcleo de la máxima dimensión posible, y dar una base de ese núcleo.

8. Sea $A \in M_{n\times n}(K)$ una matriz de rango $r < n$. Demostrar que existe una matriz $B \in M_{n\times n}(K)$, como mínimo, $B \neq 0$, tal que $AB = 0$. ¿Cuál es el máximo rango de una tal matriz B?

9. Sea $A \in M_{n\times n}(K)$ una matriz de rango $n-1$. Demostrar que, eligiendo dos filas cualesquiera de A, las n-plas formadas por los adjuntos de sus elementos son proporcionales.

10. Discutir y resolver el sistema

$$\begin{cases} ax^i + a_i x^{n+1} &= b_i, \qquad i = 1,2,\ldots,n \\ a_1x^1 + \ldots + a_nx^n + ax^{n+1} &= b_{n+1} \end{cases}$$

para los diferentes valores de a, a_i, b_j.

VII.9 Ejercicios para programar

11. Escribir un programa que dé la solución de sistemas 2×2 o bien 3×3 (reales) con determinante diferente de cero. (Indicación: empezar calculando el determinante. Si es 0, parar el programa; si no, aplicar la regla de Cramer de la sección 3.)

Nota:

Para n grande, este método es muy ineficiente.

12. (Método de Gauss.) Este programa ha de servir para

a) Dar la solución general de un sistema de ecuaciones.

b) Calcular el rango de una matriz (VI.5).

c) Invertir una matriz (VII.6).

d) Encontrar una base del espacio de soluciones de un sistema homogéneo.

Preparación:

- Entrar una matriz real A no necesariamente cuadrada.
- Entrar una familia de vectores $b_1, \ldots, b_k$ de $\mathbf{R}^n$.
- Construir la matriz ampliada $(A, b_1, \ldots, b_k)$. Efectuar en ella los cambios 1, 2, 3 del §5 para reducir A a una matriz de la forma correspondiente.
- Aplicar este procedimiento a cada una de las tres primeras cuestiones planteadas.
- Si el sistema es homogéneo ($i = 1$, $b_1 = \vec{0}$), entonces la dimensión del espacio de soluciones es $m - r$. Una base de este espacio se obtiene dando sucesivamente el valor 1 a cada una de las incógnitas libres y 0 a las restantes.

Notas:

a) Este programa debe prepararse de manera que pueda ser utilizado dentro de otros programas siempre que convenga.

b) Si vamos guardando los cambios de signo y los escalares eliminados, podemos utilizar este programa para calcular $\det A$.

c) Los elementos que van quedando en la diagonal de A se llaman "pivotes". Al objeto de minimizar la propagación de errores de redondeo, es conveniente efectuar en cada etapa permutaciones de filas de manera que quede como pivote el elemento de mayor valor absoluto entre todos los disponibles en la columna correspondiente.

Capítulo VIII

Estructura de los endomorfismos

Dada una aplicación lineal $f : E \longrightarrow F$, podemos siempre escoger bases de E y F en las que la matriz de f sea extraordinariamente simple. En efecto, sea $u_1, \ldots, u_k$ una base de $\operatorname{Nuc} f$ y $u_1, \ldots, u_k, u_{k+1}, \ldots, u_n$ una base de E. Entonces, por (V.1.1), $f(u_{k+1}), \ldots, f(u_n)$ son vectores de F linealmente independientes. Completémoslos a una base de F: $f(u_{k+1}), \ldots, f(u_n)$, $v_{n-k+1}, \ldots, v_m$. La matriz de f en estas bases es:

$$\begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 1 \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Al estudiar endomorfismos $f : E \longrightarrow E$ es natural, sin embargo, exigir que los vectores $u \in E$ y sus imágenes estén expresados en la misma base. Entonces no se puede conseguir, en general, una matriz tan simple como la que acabamos de encontrar. En todo este capítulo, E denotará un espacio vectorial de dimensión n sobre un cuerpo K.

VIII.1 Vectores propios y valores propios. Polinomio característico

Sea $f \in \operatorname{End}(E)$. Un vector $v \in E$, $v \neq \vec{0}$, es un *vector propio* de f si

$$f(v) = kv, \quad k \in K.$$

Diremos, entonces, que k es un *valor propio* de f.

Ejemplos:

1. Los vectores de Nuc f diferentes de $\vec{0}$ son vectores propios de valor propio 0.
2. Si $f = kI$ (homotecia de razón k), todo $v \neq \vec{0}$ es un vector propio de f, y k es el único valor propio de f.

Ejercicio:

Si todo $v \in E$, $v \neq \vec{0}$, es vector propio de f, f es una homotecia.

Un vector $v \neq \vec{0}$ es vector propio de f de valor propio k si y sólo si $f(v) - kv = \vec{0}$, es decir, si y sólo si $v \in \text{Nuc}(f - kI)$. Un elemento $k \in K$ es un valor propio de f si y sólo si $\text{Nuc}(f - kI) \neq \{\vec{0}\}$. Se llama *multiplicidad* del valor propio k a la dimensión de $\text{Nuc}(f - kI)$.

Proposición 1.1 *$k \in K$ es valor propio de f si y sólo si* $\det(f - kI) = 0$.

DEMOSTRACIÓN: k es valor propio $\Leftrightarrow \text{Nuc}(f-kI) \neq \{\vec{0}\} \Leftrightarrow \det(f-kI) = 0$, por (VI.3.4). □

Sea $A = (a_i^j)$ la matriz de f en una cierta base $e_1, \ldots, e_n$ de E. Entonces,

$$\det(f - kI) = \begin{vmatrix} a_1^1 - k & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 - k & \ldots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \ldots & a_n^n - k \end{vmatrix} = 0.$$

Esta expresión es una ecuación de grado n en la incógnita k, el miembro izquierdo de la cual es el valor en k de un polinomio $p_A(x)$, que denominaremos *polinomio característico de A*. Si B es la matriz asociada a f en otra base, veremos que $p_B(x) = p_A(x)$. Esto nos permitirá hablar del *polinomio característico de f*, $p_f(x)$. Tenemos

$$\det(B - kI) = \det(f - kI) = \det(A - kI) \qquad \forall k \in K.$$

Es decir, $p_B(k) = p_A(k)$ para todo $k \in K$. Si K tiene más de n elementos, resulta que $p_B(x) = p_A(x)$ (II.5.3). Si K tiene n elementos o menos, esta demostración no sirve, pero el resultado continúa siendo cierto. Una manera

de demostrarlo es considerar $p_A(x)$ como un determinante con elementos en el anillo $K[x]$:

$$p_A(x) = \det(A - xI) = \begin{vmatrix} a_1^1 - x & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 - x & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \dots & a_n^n - x \end{vmatrix}.$$

La definición y la mayoría de las propiedades de los determinantes de matrices con elementos en un cuerpo K se pueden generalizar a matrices con elementos en $K[x]$. En particular, se cumple que el determinante de un producto de matrices es el producto de sus determinantes. Entonces, si A y B son matrices del mismo endomorfismo, $B = P^{-1}AP$ (donde P es una matriz invertible) y

$$\begin{aligned} &\det(B - xI) = \det(P^{-1}AP - xI) = \\ &\quad = \det(P^{-1}(A - xI)P) = \det P^{-1} \cdot \det(A - xI) \cdot \det P = \det(A - xI). \end{aligned}$$

Proposición 1.2 *El polinomio característico de A es*

$$p_A(x) = (-1)^n x^n + (-1)^{n-1}(a_1^1 + \ldots + a_n^n)x^{n-1} + \ldots + (-1)^r A_r x^r + \ldots + \det A,$$

donde A_r es la suma de los determinantes de los menores de orden $n - r$ formados por los elementos de A de $(n - r)$ filas y $(n - r)$ columnas correspondientes a los mismos índices: a_i^j, $i = i_1, \ldots, i_{n-r}$, $j = i_1, \ldots, i_{n-r}$. Es decir, son los menores de orden $n - r$ que tienen la diagonal principal sobre la diagonal principal de A.

No efectuaremos el cálculo de los A_r, que es largo y pesado. Por otra parte, existen maneras más cómodas de hallarlos que las que podríamos dar con los conocimientos que tenemos ahora. Nos limitaremos a calcular el coeficiente de x^{n-1} y el término independiente.

De la definición de determinante resulta

$$\begin{vmatrix} a_1^1 - k & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 - k & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n & \dots & a_n^n - k \end{vmatrix} = (a_1^1 - k)(a_2^2 - k)\cdots(a_n^n - k) + S,$$

donde S es una suma de productos en cada uno de los cuales hay a lo sumo $n - 2$ elementos de la diagonal. Los términos de grado n y $n - 1$ en k son, por tanto,

$$(-1)^n k^n + (-1)^{n-1}(a_1^1 + \ldots + a_n^n)k^{n-1}.$$

El término independiente de $p_A(x)$ es $p_A(0) = \det(A - 0I) = \det A$.

Dos matrices A y B asociadas al mismo endomorfismo, es decir, tales que $B = P^{-1}AP$ con P invertible, se llaman *equivalentes*. Entonces, de (1.2) y de la invariancia del polinomio característico, se deduce que $A_r = B_r$ para todo r y, en particular,

$$a_1^1 + \ldots + a_n^n = b_1^1 + \ldots + b_n^n.$$

La suma $a_1^1 + \ldots + a_n^n$ se llama la *traza de* A, $\operatorname{tr} A$. Dado que dos matrices equivalentes tienen la misma traza, tiene sentido referirse a la *traza de* f, $\operatorname{tr} f$.

VIII.2 Diagonalización de matrices

Sea $f \in \operatorname{End}(E)$. Si conseguimos una base de E con vectores propios de f, la matriz de f tendrá una forma muy simple:

$$\begin{pmatrix} k_1 & 0 & \ldots & 0 \\ 0 & k_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & k_n \end{pmatrix}.$$

Todos los elementos no situados sobre la diagonal serán 0 y $k_1, k_2, \ldots, k_n$ serán los valores propios de los vectores propios de la base.

Una matriz de este tipo se llama una matriz *diagonal*. Diagonalizar un endomorfismo f quiere decir encontrar una base de vectores propios de f; diagonalizar una matriz A quiere decir encontrar una matriz diagonal equivalente a A. Diremos que un endomorfismo f es *diagonalizable* si se puede encontrar una base de vectores propios de f. Una matriz se puede diagonalizar si y sólo si el endomorfismo asociado es diagonalizable.

En el apartado anterior hemos dado ya una manera de encontrar los valores propios y los vectores propios: los valores propios son los ceros del polinomio característico (1.1) y $\operatorname{Nuc}(f - kI)$ es el conjunto de vectores propios de valor propio k (más el $\vec{0}$). Sólo queda, pues, ver si hay n vectores propios linealmente independientes.

Proposición 2.1 *Vectores propios de valores propios diferentes son linealmente independientes.*

DEMOSTRACIÓN: Sean $v_1, \ldots, v_m$ vectores propios de valores propios distintos $k_1, \ldots, k_m$. Procederemos por inducción sobre m. Si $m = 1$, $v_1 \neq \vec{0}$ es linealmente independiente. En general, sea $a^1 v_1 + \ldots + a^m v_m = \vec{0}$; entonces

$$\vec{0} = (f - k_1 I)(a^1 v_1 + \ldots + a^m v_m) = a^2(k_2 - k_1)v_2 + \ldots + a^m(k_m - k_1)v_m.$$

Por hipótesis de inducción, $v_2, \ldots, v_m$ son linealmente independientes, de donde $a^j(k_j - k_1) = 0$, $j = 2, \ldots, m$. Puesto que $k_j \neq k_1$ si $j \neq 1$, será

$$a^j = 0 \qquad \text{para } j = 2, \ldots, m.$$

Entonces $a^1 v_1 = \vec{0}$, de donde también $a^1 = 0$. □

Corolario 2.2 *El número de valores propios diferentes es* $\leq n$. *Si hay exactamente n valores propios diferentes, el endomorfismo es diagonalizable.* □

Ejemplos:

1. Consideremos el endomorfismo $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ cuya matriz es

$$\begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix}$$

en la base usual $(1,0)$, $(0,1)$. Su polinomio característico es

$$x^2 - 1 = (x-1)(x+1).$$

f tiene, por tanto, dos valores propios: $+1$ y -1. El subespacio de vectores propios de valor propio $+1$ es $\mathrm{Nuc}(f - I)$. La matriz de $(f - I)$ es

$$\begin{pmatrix} -2/5 & 4/5 \\ 4/5 & -8/5 \end{pmatrix},$$

de donde resulta que $\mathrm{Nuc}(f - I) = \{(2y, y)\} = \langle (2,1) \rangle$. Análogamente se ve que $\mathrm{Nuc}(f + I) = \langle (-1,2) \rangle$ es el subespacio de vectores propios de valor propio -1. Los vectores $(2,1)$, $(-1,2)$ forman una base en la que la matriz de f es

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

La imagen de un $v \in E$ se puede encontrar geométricamente de la manera siguiente: descompongamos v como suma de un vector v_1 de $\langle (2,1) \rangle$ y un vector v_2 de $\langle (-1,2) \rangle$: $v = v_1 + v_2$. Entonces $f(v) = v_1 - v_2$. Esto es una simetría de eje $\langle (2,1) \rangle$.

2. Consideremos ahora $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ con matriz

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

en la base usual $(1,0),(0,1)$. El polinomio característico es $(1-x)^2$ y, por tanto, el único valor propio es 1. Si f diagonalizase, su matriz diagonal sería

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

f sería la identidad y eso no es cierto. De hecho, el subespacio de vectores propios tiene dimensión 1: $\mathrm{Nuc}(f-I) = \langle (1,0) \rangle$.

Proposición 2.3 *Si r es la multiplicidad del valor propio k, es decir, si se tiene $r = \dim \mathrm{Nuc}(f-kI)$, y s es la multiplicidad del cero k del polinomio característico, entonces $r \leq s$.*

DEMOSTRACIÓN: Sea $v_1, \ldots, v_r$ una base de $\mathrm{Nuc}(f-kI)$. Completémosla hasta obtener una base de E: $v_1, \ldots, v_n$. En esta base la matriz de f es de la forma

$$A = \begin{pmatrix} k & 0 & \ldots & 0 & a_{r+1}^1 & \ldots & a_n^1 \\ 0 & k & \ldots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & k & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & a_{r+1}^n & \ldots & a_n^n \end{pmatrix}.$$

El polinomio característico es, pues, $p(x) = (k-x)^r \cdot q(x)$, lo que demuestra el enunciado. □

Supongamos ahora que f es diagonalizable y sea

$$\begin{pmatrix} k_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & k_n \end{pmatrix}$$

su matriz diagonal ($k_1, \ldots, k_n$ no necesariamente distintos). El polinomio característico es

$$p(x) = (k_1 - x) \cdots (k_n - x)$$

y se descompone, por tanto, en factores lineales. Si un valor propio k_i aparece s veces en la diagonal, la multiplicidad del cero k_i de $p(x)$ es s. Por otro lado, $(f - k_i I)$ tendrá una matriz diagonal con exactamente s ceros en la diagonal, de donde $\dim \mathrm{Nuc}(f - k_i I) = s$. Estos hechos caracterizan los endomorfismos diagonalizables, como lo demuestra el siguiente teorema.

Teorema 2.4 (de diagonalización) *Un endomorfismo f es diagonalizable si y sólo si su polinomio característico se descompone en factores lineales y la multiplicidad de cada uno de sus ceros coincide con su multiplicidad como valor propio de f.*

DEMOSTRACIÓN: Hemos visto ya que estas condiciones son necesarias. Demostremos ahora que son suficientes para que f sea diagonalizable. Sea

$$p(x) = (-1)^n(x-k_1)^{n_1}\cdots(x-k_r)^{n_r}, \quad n_1+\ldots+n_r = n$$

el polinomio característico de f. Pongamos $E_{k_i} = \mathrm{Nuc}(f - k_i I)$. Vamos a ver que

$$E = E_{k_1} \oplus \ldots \oplus E_{k_r}.$$

Una vez visto esto, podremos obtener una base de vectores propios tomando bases en $E_{k_1}, \ldots, E_{k_r}$. En esa base, la matriz de f será diagonal. Demostremos, pues, que la expresión de los vectores de E como suma de vectores de los subespacios E_{k_i} es única (IV.4). En efecto,

$$v_1 + \ldots + v_r = w_1 + \ldots + w_r$$

con $v_i, w_i \in E_{k_i}$, $i = 1, \ldots, r$, implica $(v_1 - w_1) + \ldots + (v_r - w_r) = \vec{0}$. Los vectores $v_i - w_i$ son vectores propios de valores propios diferentes, o $\vec{0}$. Por (2.1), han de ser todos ellos cero: $v_i - w_i = \vec{0}$; es decir, $v_i = w_i$, $i = 1, \ldots, r$. La suma de los subespacios E_{k_i} es, pues, directa y su dimensión es $n_1 + \ldots + n_r = n$. Por tanto,

$$E_{k_1} \oplus \ldots \oplus E_{k_r} = E. \ \square$$

Una matriz $A = (a_i^j)$ se llama *triangular superior* si $a_i^j = 0$ para $i < j$. Diremos que un endomorfismo es *triangulable* si, en una base conveniente, su matriz es triangular.

Teorema 2.5 (de triangulación) *Un endomorfismo es triangulable si y sólo si su polinomio característico se descompone en factores de primer grado.*

DEMOSTRACIÓN: Si el endomorfismo f tiene una matriz triangular

$$\begin{pmatrix} a_1^1 & a_2^1 & a_3^1 & \ldots & a_n^1 \\ 0 & a_2^2 & a_3^2 & \ldots & a_n^2 \\ 0 & 0 & a_3^3 & \ldots & a_n^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_n^n \end{pmatrix},$$

su polinomio característico $p(x) = (a_1^1 - x)(a_2^2 - x)\cdots(a_n^n - x)$ se descompone en factores lineales. Probaremos el recíproco por inducción sobre n. Para $n = 1$, toda matriz es triangular. Sea $n \geq 2$ cualquiera. El polinomio característico tiene como mínimo una raíz; hay pues, como mínimo, un valor propio. Sea v_1 un vector propio, y $v_1, v_2, \ldots, v_n$ una base de E. La matriz de f en esta base es del tipo

$$\begin{pmatrix} k & a_2^1 & \ldots & a_n^1 \\ 0 & a_2^2 & \ldots & a_n^2 \\ \vdots & \vdots & & \vdots \\ 0 & a_2^n & \ldots & a_n^n \end{pmatrix}.$$

Consideremos ahora la aplicación

$$g : \langle v_2, \ldots, v_n \rangle \longrightarrow \langle v_2, \ldots, v_n \rangle$$

de matriz

$$\begin{pmatrix} a_2^2 & \ldots & a_n^2 \\ \vdots & & \vdots \\ a_2^n & \ldots & a_n^n \end{pmatrix}.$$

Tenemos que $\det(f - xI) = (k - x) \cdot \det(g - xI)$ y, por tanto, el polinomio característico de g, $\det(g - xI)$, se descompone también en factores de primer grado. Por hipótesis de inducción existe entonces una base $u_2, \ldots, u_n$ de $\langle v_2, \ldots, v_n \rangle$ en la cual la matriz de g es triangular:

$$\begin{pmatrix} b_2^2 & b_3^2 & \ldots & b_n^2 \\ 0 & b_3^3 & \ldots & b_n^3 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & b_n^n \end{pmatrix}.$$

Ahora bien, $f(v_i) = a_i^1 v_1 + g(v_i)$, $i = 2, \ldots, n$. Por tanto, si $u_j = \sum_{i=2}^n c_j^i v_i$,

$$f(u_j) = \sum_{i=2}^n c_j^i f(v_i) = \sum_{i=2}^n c_j^i (a_i^1 v_1 + g(v_i)) = \left(\sum_{i=2}^n c_j^i a_i^1 \right) v_1 + g(u_j),$$

para $j = 2, \ldots, n$. La matriz de f en la base $v_1, u_2, \ldots, u_n$ se obtiene, pues, añadiendo a la matriz de g una primera columna $(k, 0, \ldots, 0)$ y una primera fila $(k, b_2^1, \ldots, b_n^1)$, donde $b_j^1 = \sum_{i=2}^n c_j^i a_i^1$, $j \geq 2$. Es, por tanto, una matriz triangular. □

Corolario 2.6 *Todo endomorfismo de un espacio vectorial sobre los complejos es triangulable.* □

VIII.3 Polinomio mínimo

El estudio de los vectores propios nos ha permitido simplificar la matriz de un endomorfismo en muchos casos. Queda, sin embargo, sin resolver el caso general. Nuestros pasos se encaminan ahora hacia la obtención de unos teoremas de descomposición de E en suma directa de subespacios que permitan obtener bases convenientes en las que se pueda expresar el endomorfismo. Observemos que una descomposición de ese tipo es la que nos ha permitido demostrar el teorema de diagonalización (2.4).

Consideremos las potencias de f: f^r, $r = 0, 1, 2, \ldots$,

$$f^0 = I,\ f^1 = f,\ \ldots,\ f^r = f \circ f^{r-1},\ \ldots$$

Si la dimensión de E es n, el espacio vectorial $\mathrm{End}(E)$ tiene dimensión n^2, y las potencias f^r no pueden ser todas linealmente independientes. Las combinaciones lineales

$$a_0 I + a_1 f + \ldots + a_s f^s = 0$$

nos llevan a considerar el núcleo de la aplicación

$$\begin{array}{rcl} \Phi_f : K[x] & \longrightarrow & \mathrm{End}(E) \\ p(x) = a_0 + a_1 x + \ldots + a_r x^r & \longmapsto & p(f) = a_0 I + a_1 f + \ldots + a_r f^r. \end{array}$$

Se cumplen las siguientes propiedades:

- $\Phi_f(p(x) + q(x)) = p(f) + q(f) = \Phi_f(p(x)) + \Phi_f(q(x))$.
- $\Phi_f(p(x) \cdot q(x)) = p(f) \circ q(f) = \Phi_f(p(x)) \circ \Phi_f(q(x))$.
- $\Phi_f(kp(x)) = kp(f) = k\Phi_f(p(x))$.

Φ_f es, pues, un morfismo de álgebras (V.5). De la conmutatividad del producto de $K[x]$ se deduce que dos endomorfismos de la imagen siempre conmutan:

$$p(f) \circ q(f) = q(f) \circ p(f).$$

El núcleo de Φ_f es un ideal de $K[x]$ y, por (II.2.2),

$$\mathrm{Nuc}\,\Phi_f = \{p(x) \in K[x] \mid p(f) = 0\} = (m_f(x)).$$

Los polinomios de $\mathrm{Nuc}\,\Phi_f$ se llaman *polinomios anuladores de* f; $m_f(x)$ se llama el *polinomio mínimo de* f y está determinado salvo factores de K (II.2). En general se toma $m_f(x)$ mónico, es decir, con el coeficiente de grado máximo igual a 1.

Ejemplos:

1. Si $E = \{\vec{0}\}$ y f es el único endomorfismo de $\{\vec{0}\}$, entonces $\operatorname{Nuc}\Phi_f = K[x] = (a_0)$, $a_0 \neq 0$. Recíprocamente, si el polinomio mínimo de f es $a_0 \in K$, $a_0 \neq 0$, entonces $\operatorname{Nuc}\Phi_f = (a_0) = K[x]$ y, en particular, $0 = \Phi_f(1) = I_E$. Esto implica que $E = \{\vec{0}\}$.
2. Si $E \neq \{\vec{0}\}$ y $f = 0$, $\operatorname{Nuc}\Phi_f = (x)$.
3. Si $E \neq \{\vec{0}\}$ y $f = I$, $x - 1 \in \operatorname{Nuc}\Phi_f$ y $m_f(x) \mid (x-1)$. Pero puesto que $m_f(x)$ no es constante, $m_f(x) = x - 1$.
4. Si $E \neq \{\vec{0}\}$, $f = kI$ si y sólo si $m_f(x) = x - k$.

Fijado un vector $u \in E$, consideremos ahora la aplicación

$$\begin{array}{rccc} \Phi_u : & K[x] & \longrightarrow & E \\ & p(x) & \longmapsto & p(f)(u). \end{array}$$

El núcleo de Φ_u es un ideal de $K[x]$:

$$\operatorname{Nuc}\Phi_u = \{p(x) \in K[x] \mid p(f)(u) = \vec{0}\} = (m_u(x)).$$

$m_u(x)$ se llama el *polinomio mínimo de f en u* o simplemente el *polinomio mínimo de u* (si no hay confusión respecto a qué f nos referimos); está determinado salvo factores de K y generalmente se toma mónico.

Proposición 3.1 *Sea*

$$m_u(x) = a_0 + a_1 x + \ldots + a_s x^s$$

el polinomio mínimo de u. Entonces $u, f(u), \ldots, f^{s-1}(u)$ son linealmente independientes y $u, f(u), \ldots, f^{s-1}(u), f^t(u)$ $(t \geq s)$ son linealmente dependientes.

DEMOSTRACIÓN: Si $u, f(u), \ldots, f^{s-1}(u)$ fuesen linealmente dependientes, habría un polinomio $p(x)$ de grado $< s$ tal que $p(f)(u) = 0$. Esto contradice la definición de $m_u(x)$.

Si $t = s$, $m_u(f)(u) = a_0 u + a_1 f(u) + \ldots + a_s f^s(u) = 0$ nos dice que estos vectores son linealmente dependientes. Para $t > s$, procederemos por inducción. Así pues,

$$f^{t-1}(u) \in \langle u, f(u), \ldots, f^{s-1}(u)\rangle,$$

de donde

$$f^t(u) \in \langle f(u), f^2(u), \ldots, f^s(u)\rangle = \langle u, f(u), \ldots, f^{s-1}(u)\rangle$$

y $u, f(u), \ldots, f^{s-1}(u), f^t(u)$ son linealmente dependientes. □

VIII.4 Subespacios invariantes

Sea $f \in \mathrm{End}(E)$. Un subespacio F de E se llama *invariante por* f si $f(F) \subset F$. En ese caso, f induce un endomorfismo de F

$$\begin{aligned} f' = f \mid F : F &\longrightarrow F \\ v &\longmapsto f(v), \end{aligned}$$

al que llamaremos *restricción de* f *a* F.

Proposición 4.1 *$m_{f'}(x)$ divide a $m_f(x)$.*

DEMOSTRACIÓN: $m_f(f) = 0 \Rightarrow m_f(f)(u) = \vec{0} \ \ \forall u \in E \Rightarrow m_f(f')(v) = m_f(f)(v) = \vec{0} \ \ \forall v \in F \Rightarrow m_f(x) \in (m_{f'}(x))$. □

Corolario 4.2 *Si dos subespacios F y G de E, invariantes por f, tienen polinomios mínimos primos entre sí, entonces $F \cap G = \{\vec{0}\}$.*

DEMOSTRACIÓN: Claramente, $F \cap G$ es también invariante y, por (4.1), su polinomio mínimo es 1. En los ejemplos del §3 vimos que, en este caso, el espacio debe ser $\{\vec{0}\}$. □

Proposición 4.3 *Para todo polinomio $p(x) \in K[x]$, $\mathrm{Nuc}\, p(f)$ e $\mathrm{Im}\, p(f)$ son subespacios invariantes por f.*

DEMOSTRACIÓN: Si $u \in \mathrm{Nuc}\, p(f)$, $p(f)(f(u)) = f(p(f)(u)) = f(\vec{0}) = \vec{0}$, de donde $f(u) \in \mathrm{Nuc}\, p(f)$. Si $u = p(f)(v) \in \mathrm{Im}\, p(f)$, $f(u) = f(p(f)(v)) = p(f)(f(v))$, de donde $f(u) \in \mathrm{Im}\, p(f)$. □

Supongamos ahora que el polinomio mínimo de f, $m_f(x)$, se descompone en producto de dos factores primos entre sí:

$$m_f(x) = p(x) \cdot q(x).$$

Consideremos los subespacios invariantes $\mathrm{Nuc}\, p(f)$ y $\mathrm{Nuc}\, q(f)$. Los polinomios $p(x)$ y $q(x)$ son anuladores de la restricción de f a estos subespacios. Por tanto (4.2),

$$\mathrm{Nuc}\, p(f) \cap \mathrm{Nuc}\, q(f) = \{\vec{0}\}.$$

Ahora bien, es fácil ver que $\mathrm{Nuc}\, p(f) \supset \mathrm{Im}\, q(f)$ y que $\mathrm{Nuc}\, q(f) \supset \mathrm{Im}\, p(f)$. Comprobemos la primera inclusión: si $u = q(f)(v) \in \mathrm{Im}\, q(f)$, entonces

$p(f)(u) = p(f)q(f)(v) = m_f(f)(v) = \vec{0}$ y, por tanto, $u \in \operatorname{Nuc} p(f)$. Estas inclusiones indican que

$$\begin{aligned} n &= \dim \operatorname{Nuc} p(f) + \dim \operatorname{Im} p(f) \leq \\ &\leq \dim \operatorname{Nuc} p(f) + \dim \operatorname{Nuc} q(f) = \dim(\operatorname{Nuc} p(f) \oplus \operatorname{Nuc} q(f)) \leq n. \end{aligned}$$

Las dos inclusiones son, pues, igualdades y

$$E = \operatorname{Nuc} p(f) \oplus \operatorname{Nuc} q(f).$$

¿Cuáles son los polinomios mínimos de la restricción de f a esos dos subespacios invariantes en que se descompone E? Ya hemos dicho antes que tienen que ser divisores de $p(x)$ y de $q(x)$ (estos polinomios son anuladores): sean $\bar{p}(x)$ y $\bar{q}(x)$. Pero, entonces, $\bar{p}(x) \cdot \bar{q}(x)$ es un anulador de f: si $u \in E$, $u = u_1 + u_2$ con $u_1 \in \operatorname{Nuc} p(f)$, $u_2 \in \operatorname{Nuc} q(f)$ y, entonces,

$$\bar{p}(f)\bar{q}(f)(u) = \bar{q}(f)\bar{p}(f)(u_1) + \bar{p}(f)\bar{q}(f)(u_2) = \vec{0} + \vec{0} = \vec{0}.$$

Por tanto, por un lado $\bar{p}(x) \cdot \bar{q}(x) \in (m_f(x))$ y por el otro divide a $m_f(x) = p(x) \cdot q(x)$. Debe cumplirse, pues, $\bar{p}(x) \cdot \bar{q}(x) = m_f(x)$; es decir, $\bar{p}(x) = p(x)$ y $\bar{q}(x) = q(x)$ son los polinomios mínimos buscados.

Naturalmente, si ahora $p(x)$ (o $q(x)$) se descompone en factores primos, podemos descomponer $\operatorname{Nuc} p(f)$ (o $\operatorname{Nuc} q(f)$) en suma de subespacios invariantes y proceder así tantas veces como podamos. Tenemos de esta forma el

Teorema 4.4 (primer teorema de descomposición) *Si el polinomio mínimo de $f \in \operatorname{End}(E)$ es*

$$m_f(x) = m_1(x)^{n_1} \cdots m_r(x)^{n_r},$$

donde $m_1(x), \ldots, m_r(x)$ son factores irreducibles, el espacio E es suma directa de subespacios invariantes

$$E = E^1 \oplus \ldots \oplus E^r,$$

de forma que el polinomio mínimo de la restricción de f a E^i es $m_i(x)^{n_i}$. Esta descomposición es única:

$$E^i = \operatorname{Nuc}(m_i(f)^{n_i}), \qquad i = 1, \ldots, r.$$

DEMOSTRACIÓN: Lo único que falta por demostrar es la unicidad de la descomposición. Supongamos, pues, que tenemos dada una descomposición en subespacios invariantes

$$E = E^1 \oplus \ldots \oplus E^r,$$

de la cual solamente sabemos que el polinomio mínimo de la restricción de f a E^i es $m_i(x)^{n_i}$, para $i = 1, \ldots, r$. Esta última condición implica que $E^i \subset \operatorname{Nuc}(m_i(f)^{n_i})$, de donde

$$\begin{aligned} n &= \dim E^1 + \ldots + \dim E^r \leq \\ &\leq \dim \operatorname{Nuc}(m_1(f)^{n_1}) + \ldots + \dim \operatorname{Nuc}(m_r(f)^{n_r}) = n. \end{aligned}$$

La desigualdad tiene que ser, pues, una igualdad y todas las inclusiones anteriores tienen que ser igualdades:

$$E^i = \operatorname{Nuc}(m_i(f)^{n_i}), \qquad i = 1, \ldots, r. \ \square$$

La descomposición de E en suma directa de subespacios invariantes reduce el estudio del comportamiento de f al estudio de sus restricciones a cada uno de los subespacios. Si escribimos la matriz de f en una base de E formada por bases de cada uno de los subespacios, obtenemos

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & 0 \\ & & \ddots & \\ & 0 & & A_r \end{pmatrix}.$$

La matriz A está formada por matrices $A_1, \ldots, A_r$ con la diagonal sobre la de A, y 0 en el resto de posiciones. El estudio de A se reduce, pues, al de las matrices A_i, que son precisamente las matrices de las restricciones de f a cada uno de los subespacios en que se descompone E.

Aplicaremos ahora (4.4) a varios ejemplos concretos.

Ejemplos:

1. Consideremos el endomorfismo $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ cuya matriz es

$$A = \begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix}$$

en la base $(1,0),(0,1)$. (Ver §2.) El estudio de las combinaciones lineales entre sus potencias A^n resulta aquí trivial. Tenemos $A^2 = I$ y, por tanto, $x^2 - 1$ es un polinomio anulador. Si $m_f(x) = x - 1$, $m_f(f) = f - I = 0$, de donde $f = I$. Si $m_f(x) = x + 1$, $m_f(f) = f + I = 0$, de donde $f = -I$. Ninguno de los dos casos es el nuestro; por tanto,

$$m_f(x) = (x-1)(x+1) \qquad \text{y} \qquad E = E^1 \oplus E^2,$$

donde $E^1 = \mathrm{Nuc}(f - I) = \langle (2,1) \rangle$, $E^2 = \mathrm{Nuc}(f + I) = \langle (1,-2) \rangle$. La restricción de f a E^1 es I_{E^1}; la restricción a E^2 es $-I_{E^2}$. Es decir, E^1 y E^2 son precisamente los subespacios de los vectores propios de valores propios 1 y -1 que obtuvimos en el §2. Igual que allá, en la base $(2,1),(1,-2)$ la matriz resulta ser

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

2. Como generalización del ejemplo anterior, consideremos un endomorfismo f tal que $f^2 = I$ (se llama una *involución*). $m_f(x)$ divide a $x^2 - 1 = (x-1)(x+1)$ y, como en el ejemplo 1,

   si $m_f(x) = x - 1$, $f = I$;
   si $m_f(x) = x + 1$, $f = -I$;
   si $m_f(x) = (x-1)(x+1)$, $E = E^1 \oplus E^2$.

   El polinomio mínimo de la restricción de f a E^1 es $x - 1$ y, por tanto, sobre E^1 f es I_{E^1}. Análogamente, sobre E^2 f es $-I_{E^2}$. Si $e_1, \ldots, e_r$ es una base de E^1 y $e_{r+1}, \ldots, e_n$ una base de E^2, la matriz de E en la base que resulta de la unión de esas dos es

$$\left.\begin{pmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & & 1 & & \\ & & & -1 & \\ & 0 & & & \ddots \\ & & & & & -1 \end{pmatrix}\right\} \begin{array}{l} r \text{ filas} \\ \\ \\ \\ n - r \text{ filas.} \end{array}$$

   Observemos que aquí también E^1, E^2 son los subespacios de vectores propios de valores propios 1, -1 y que la matriz obtenida es una matriz diagonal.

3. Modifiquemos el ejemplo anterior estudiando endomorfismos f tales que $f^2 = -I$. El polinomio mínimo es ahora divisor de $x^2 + 1$. Si el cuerpo sobre el que trabajamos es el complejo $\mathbf{C}$, $x^2 + 1 = (x - i)(x + i)$, y obtenemos una situación muy similar a la del ejemplo 2. El polinomio $m_f(x)$ puede ser

   - $m_f(x) = x - i$, de donde $f = iI$;
   - $m_f(x) = x + i$, de donde $f = -iI$;
   - $m_f(x) = (x-i)(x+i)$, de donde $E = E^1 \oplus E^2$. La restricción de f a E^1 es iI_{E^1}; la de f a E^2 es $-iI_{E^2}$.

Tomando bases de E^1 y E^2 obtenemos una base de E en la cual f tiene una matriz diagonal:

$$\begin{pmatrix} i & & & & \\ & \ddots & & 0 & \\ & & i & & \\ & & & -i & \\ & 0 & & & \ddots \\ & & & & & -i \end{pmatrix}.$$

Si el cuerpo sobre el que trabajamos es el real $\mathbf{R}$, x^2+1 es irreducible: $m_f(x) = x^2+1$ y (4.4) no da ninguna descomposición propia. Podría ser, no obstante, que consiguiéramos simplificar la matriz de f tomando vectores propios para formar una base, tal como hemos estudiado en el §1. Tampoco es posible. ¿Por qué? Si k fuese un valor propio, el subespacio de vectores propios $\mathrm{Nuc}(f-kI) \neq \{\vec{0}\}$ sería invariante y con polinomio mínimo $x-k$. Esto implicaría que $x-k$ dividiría a $m_f(x)$, lo cual no es cierto para ningún k. Este razonamiento que acabamos de hacer es general y nos da:

Proposición 4.5 *Si k es un valor propio de f, $(x-k) \mid m_f(x)$.* □

Ejemplos:

1. Sea $f \in \mathrm{End}(E)$ tal que $f^3 = I$. El polinomio x^3-1 es un anulador y $m_f(x) \mid x^3-1$. Si el cuerpo es $\mathbf{C}$,

$$x^3-1 = (x-1)\left(x+\frac{1+i\sqrt{3}}{2}\right)\left(x+\frac{1-i\sqrt{3}}{2}\right).$$

Un estudio parecido al de los ejemplos anteriores nos da, para todos los posibles $m_f(x)$, bases de E en las que la matriz de f es diagonal y los valores propios son

$$\left\{1, \frac{-1+i\sqrt{3}}{2}, \frac{-1-i\sqrt{3}}{2}\right\}$$

o un subconjunto de éste.

Si el cuerpo es $\mathbf{R}$, $x^3 - 1 = (x-1)(x^2+x+1)$ y, por tanto, $E = E^1 \oplus E^2$. El polinomio mínimo de f sobre E^1 es $x-1$ y, por tanto, f sobre E^1 es I_{E^1}. El polinomio mínimo de f sobre E^2 es x^2+x+1, irreducible. Sea $e_1, \ldots, e_r$ una base de E^1, y $e_{r+1}, \ldots, e_n$ una base de E^2. La matriz de f en la base $e_1, \ldots, e_n$ es de la forma

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ & 1 & \\ 0 & & A_2 \end{pmatrix}.$$

Al igual que en el ejemplo 3, ningún valor propio permite simplificar A_2.

2. Sea $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ con matriz

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

en la base $(1,0),(0,1)$. Es fácil ver que $m_f(x) = (x-1)^2$. El teorema (4.4) no permite descomponer E en suma de subespacios invariantes. Hay, sin embargo, un subespacio invariante: el subespacio de vectores propios de valor propio 1, $\langle (1,0) \rangle$. Si f tuviese una matriz diagonal, tendría que ser

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

ya que 1 es el único valor propio de f. Pero $f \neq I$ y no puede tener nunca la matriz identidad.

3. Sea $f \in \mathrm{End}(E)$ con matriz

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

en la base e_1, e_2, e_3. Al calcular las potencias A^n se ve en seguida que $f^3 = f^2$ y, por tanto, $x^3 - x^2$ es un polinomio anulador. Así pues, $m_f(x) \mid x^3 - x^2 = x^2(x-1)$.

Si $m_f(x) = x-1$, $f = I$, lo cual es falso.

Si $m_f(x) = x$, $m_f(f) = f = 0$, lo cual es falso.

Si $m_f(x) = x^2$, $f^2 = 0$, lo cual es falso.

Si $m_f(x) = (x-1)x$, $f^2 = f$, lo cual es falso.

Así pues, $m_f(x) = (x-1)x^2$ y $E = E^1 \oplus E^2$. Sobre E^1, f es I_{E^1}. Sobre E^2, f tiene polinomio mínimo x^2. El cálculo de E^1 y E^2 nos da

$$E^1 = \mathrm{Nuc}(f - I) = \langle e_1 \rangle, \qquad E^2 = \mathrm{Nuc}\, f^2 = \langle e_2, e_3 \rangle.$$

En la base e_1, e_2, e_3, la matriz de f es la matriz A dada. Y los valores propios, ¿cuáles son? Pues son 1 y 0; los subespacios de vectores propios respectivos son $\langle e_1 \rangle$ y $\langle e_2 + e_3 \rangle$. Si completamos estos dos vectores hasta obtener una base de E: $e_1, e_2 + e_3, e_3$, obtenemos la matriz de f

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

que es triangular.

Todos los ejemplos estudiados nos llevan de manera natural a la siguiente conclusión.

Teorema 4.6 (de diagonalización) *Un endomorfismo es diagonalizable si y sólo si su polinomio mínimo se descompone en factores lineales no repetidos.*

DEMOSTRACIÓN: Sea $m_f(x) = (x - a_1) \cdots (x - a_r)$, donde $a_i \neq a_j$ si $i \neq j$. Por (4.4), $E = E^1 \oplus \ldots \oplus E^r$, donde

$$E^i = \mathrm{Nuc}(f - a_i I)$$

es el subespacio de vectores propios de valor propio a_i. Existe, pues, una base de vectores propios de E constituida por bases de cada uno de los subespacios $E^1, \ldots, E^r$. Recíprocamente, supongamos ahora que E tiene una base formada por vectores propios:

$$e_1^1, \ldots, e_{n_1}^1, e_1^2, \ldots, e_{n_2}^2, \ldots, e_1^r, \ldots, e_{n_r}^r.$$

Supongamos también que a_i es el valor propio de $e_1^i, \ldots, e_{n_i}^i$. Entonces, si ponemos $E^i = \langle e_1^i, \ldots, e_{n_i}^i \rangle$, tenemos

$$E = E^1 \oplus \ldots \oplus E^r$$

y el polinomio mínimo de E^i es $(x - a_i)$. Esto implica que

$$(x - a_1) \cdots (x - a_r) \mid m_f(x).$$

Ahora bien, $(x-a_1)\cdots(x-a_r)$ es un anulador, ya que si $u = u_1 + \ldots + u_r$ con $u_i \in E^i$, $i = 1, \ldots, r$, se tiene

$$\begin{aligned}(f-a_1I)\cdots(f-a_rI)(u) &= \\ &= (f-a_2I)\cdots(f-a_rI)(f-a_1I)(u_1)+ \\ &\quad +(f-a_1I)(f-a_3I)\cdots(f-a_rI)(f-a_2I)(u_2)+ \\ &\quad +\ldots+(f-a_1I)\cdots(f-a_rI)(u_r) = \\ &= \vec{0}+\vec{0}+\ldots+\vec{0} = \vec{0},\end{aligned}$$

de donde $m_f(x) = (x-a_1)\cdots(x-a_r)$. □

VIII.5 Grado del polinomio mínimo

Por definición (§3) el polinomio mínimo tiene grado $\leq n^2 = \dim \mathrm{End}(E)$. Esta cota es, sin embargo, muy grande cuando se trata de encontrar el polinomio mínimo de un endomorfismo.

Proposición 5.1 $\mathrm{gr}\, m_f(x) \leq n$.

DEMOSTRACIÓN: Sea $m_f(x) = m_1(x)^{n_1}\cdots m_r(x)^{n_r}$ y $E = E^1 \oplus \ldots \oplus E^r$ la descomposición de (4.4). Es suficiente ver que $\mathrm{gr}\, m_i(x)^{n_i} \leq \dim E^i$ para $i = 1, \ldots, r$.

Dado que $m_i(x)^{n_i}$ es el polinomio mínimo de la restricción de f a E^i, hay un $v_i \in E^i$ tal que $m_i(f)^{n_i}(v_i) = 0$ pero $m_i(f)^{n_i-1}(v_i) \neq 0$. Entonces el polinomio mínimo de v_i es $m_i(x)^{n_i}$ y, por (3.1), tendremos que los vectores $v_i, f(v_i), \ldots, f^{k_i-1}(v_i)$ son linealmente independientes, donde $k_i = \mathrm{gr}\, m_i(x)^{n_i}$. Por tanto, $k_i \leq \dim E^i$. □

VIII.6 El teorema de Cayley-Hamilton

En (4.5) hemos visto que los ceros del polinomio característico, que designaremos por $p_f(x)$, son también ceros del polinomio mínimo, $m_f(x)$. Demostraremos ahora el recíproco.

Proposición 6.1 *a es un cero de $m_f(x)$ si y sólo si es un cero de $p_f(x)$.*

DEMOSTRACIÓN: Si a es un cero de $m_f(x)$, $m_f(x) = (x-a)\cdot m_1(x)$. Existe un $u \in E$ tal que $m_1(f)(u) \neq \vec{0}$ (en caso contrario, el polinomio mínimo sería $m_1(x)$). Entonces el vector $w = m_1(f)(u)$ tiene valor propio a:

$$(f-aI)w = (f-aI)m_1(f)(u) = m_f(f)(u) = \vec{0}.$$

Por tanto, a es un cero de $p_f(x)$. □

Teorema 6.2 *Si el polinomio mínimo $m_f(x)$ y el polinomio característico $p_f(x)$ se descomponen en factores lineales, entonces $p_f(x)$ es un anulador de f; esto es, $p_f(f) = 0$.*

DEMOSTRACIÓN: Sea $m_f(x) = (x - a_1)^{n_1} \cdots (x - a_r)^{n_r}$ y sea $E = E^1 \oplus \ldots \oplus E^r$ la descomposición de (4.4). La matriz de f en una base formada por bases de los subespacios E^i es de la forma

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_r \end{pmatrix}.$$

Ahora bien, si $p_i(x)$ es el polinomio característico de A_i (es decir, de la restricción de f a E^i),

$$p_f(x) = p_1(x) \cdot p_2(x) \cdots p_r(x).$$

Cada $p_i(x)$ se descompone en factores lineales y sus ceros son ceros del polinomio mínimo de E^i, que es $(x - a_i)^{n_i}$. Por tanto, $p_i(x) = (x - a_i)^{m_i}$ con $m_i = \dim E^i$. Por (5.1), $n_i \leq m_i$, de donde resulta que

$$m_f(x) \mid p_f(x),$$

y $p_f(x)$ es un anulador. □

El resultado de (6.2) es válido en condiciones mucho más generales. Para verlo, observemos que si A es una matriz $n \times n$ sobre un cuerpo K, podemos referirnos al polinomio característico de A, $p_A(x)$, y al polinomio mínimo de A, $m_A(x)$, tal como lo hemos hecho en el caso de endomorfismos. Así pues, $m_A(x)$ será un polinomio de grado mínimo del ideal $\{p(x) \in K[x] \mid p(A) = 0\}$, donde, si $p(x) = a_0 + a_1 x + \ldots + a_n x^n$, ponemos

$$p(A) = a_0 I + a_1 A + \ldots + a_n A^n.$$

Si $p_A(x)$ y $m_A(x)$ se descomponen en factores lineales, (6.2) asegura que

$$p_A(A) = 0.$$

Sea ahora A una matriz real. A es también una matriz compleja y su polinomio característico es el mismo: $p_A(x) = \det(A - xI)$. Entonces, en $\mathbf{C}$,

$$p_A(A) = 0$$

y, naturalmente, esta igualdad vale también en $\mathbf{R}$. Este razonamiento hecho para $\mathbf{R}$ y $\mathbf{C}$ sirve para dos cuerpos cualesquiera $K \subset K'$ tales que todo

polinomio de K' se descomponga en factores lineales. Vimos en (II.7) que si un polinomio no tenía ceros en K podíamos construir un cuerpo $K_1 \supset K$ donde ese polinomio tuviera un cero. Se puede demostrar que la reiteración de este proceso conduce a un cuerpo $K' \supset K$ en el cual todo polinomio se descompone en factores lineales. El cuerpo K' se llama la *clausura algebraica* de K. Así pues, todo cuerpo tiene una clausura algebraica y tenemos en general

Teorema 6.3 (de Cayley-Hamilton) *El polinomio mínimo divide siempre al polinomio característico.*

Observación:

Este teorema proporciona un método práctico para calcular el polinomio mínimo de un endomorfismo f dado. Sea A la matriz de f en una base cualquiera. Calcular el polinomio característico $p_A(x)$, descomponerlo en factores irreducibles y buscar el menor de sus divisores $q(x)$ tales que $q(f) = 0$ (tal como se ha hecho en los ejemplos del §4). Éste será $m_f(x)$.

VIII.7 Matriz canónica (general) de un endomorfismo

El teorema de descomposición en subespacios invariantes (4.4) permite reducir el estudio de un endomorfismo f al estudio de sus restricciones a ciertos subespacios invariantes E^i. En casos muy particulares (4.6), los espacios E^i son subespacios de vectores propios y podemos obtener una matriz de f diagonal. Vamos a estudiar ahora las restricciones de f a los subespacios E^i en el caso general. Concretamente, vamos a descomponer cada subespacio E^i en suma de subespacios invariantes sobre los cuales la actuación de f es muy clara: los subespacios f-cíclicos.

Un subespacio F de E es *f-cíclico* si existe un vector $u \in E$ tal que $F = \langle u, f(u), f^2(u), \ldots \rangle$. F es invariante por f y, por (3.1), su dimensión es el grado del polinomio mínimo de f en u. Si este polinomio es

$$a_0 + a_1 x + \ldots + a_{s-1} x^{s-1} + x^s,$$

entonces $\{u, f(u), \ldots, f^{s-1}(u)\}$ es una base de F y, en esta base, la matriz de la restricción de f es

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{s-1} \end{pmatrix}.$$

Teorema 7.1 (segundo teorema de descomposición) *Si $f \in \mathrm{End}(E)$, E es suma directa de subespacios f-cíclicos.*

DEMOSTRACIÓN: Por (4.4), solamente hace falta considerar el caso en que el polinomio mínimo de f es una potencia de un polinomio primo: $m_f(x) = q(x)^s$, $q(x)$ primo.

Usaremos inducción sobre la dimensión de E. Si $\dim E = 1$, E ya es f-cíclico. Supongamos cierto el teorema para todos los espacios de dimensión menor o igual que $n-1$.

Sea $\dim E = n$. En la demostración de (5.1) probamos que existe un $u_0 \in E$ tal que el polinomio mínimo de f en u_0 es el mismo $m_f(x)$. Sea

$$F_0 = \langle u_0, f(u_0), f^2(u_0), \ldots \rangle$$

y $\bar{E} = E/F_0$. f induce un endomorfismo $\bar{f}$ de $\bar{E}$,

$$\begin{array}{rccc} \bar{f} : & \bar{E} & \longrightarrow & \bar{E} \\ & [v] & \longmapsto & [f(v)], \end{array}$$

bien definido, ya que la imagen por f de todos los representantes de $[v]$ está en $[f(v)]$:

$$f([v]) = f(v + F_0) = f(v) + f(F_0) \subset f(v) + F_0 = [f(v)].$$

Ahora bien, $\dim \bar{E} < n = \dim E$ y, por hipótesis de inducción,

$$\bar{E} = \bar{F}_1 \oplus \ldots \oplus \bar{F}_r$$

con $\bar{F}_i$ $\bar{f}$-cíclico, $i = 1, \ldots, r$. Sea $\bar{F}_i = \langle [u_i'], \bar{f}[u_i'], \ldots \rangle$.

Lema 7.2 *Existe un representante u_i de $[u_i']$ tal que, si denotamos por F_i el subespacio $\langle u_i, f(u_i), \ldots \rangle$, la proyección*

$$\begin{array}{rccc} \pi_i : & F_i & \longrightarrow & \bar{F}_i \\ & v & \longmapsto & [v] \end{array}$$

es un isomorfismo.

Supongamos demostrado, de momento, este lema. Entonces

$$E = F_0 \oplus F_1 \oplus \ldots \oplus F_r.$$

Comprobémoslo: si $v \in E$, $[v] = [v_1] + \ldots + [v_r]$ con $[v_i] \in \bar{F}_i$. El lema nos asegura entonces que podemos suponer que $v_i \in F_i$. Así pues, $E = F_0 + F_1 + \ldots + F_r$. Para ver que la suma es directa, supongamos que $v \in E$ se expresa de dos maneras como suma de vectores de $F_0, \ldots, F_r$:

$v = v_0 + v_1 + \ldots + v_r = w_0 + w_1 + \ldots + w_r$, $v_i, w_i \in F_i$, $i = 0, 1, \ldots, r \Rightarrow$ $[v] = [v_1] + \ldots + [v_r] = [w_1] + \ldots + [w_r]$ en $\bar{E} = \bar{F}_1 \oplus \ldots \oplus \bar{F}_r \Rightarrow [v_i] = [w_i]$, $i = 1, \ldots, r$. Pero, por el lema, $v_i = w_i$, $i = 1, \ldots, r$; de donde también $v_0 = w_0$.

Sólo queda demostrar el lema. Hagamos antes dos observaciones generales:

- Si $m(x)$ y $\bar{m}(x)$ son los polinomios mínimos de f y $\bar{f}$ en v y en $[v]$ respectivamente, $\bar{m}(x) \mid m(x)$, ya que

$$m(\bar{f})([v]) = [m(f)(v)] = [\vec{0}].$$

- El polinomio mínimo de f en v divide al polinomio mínimo de f. Esto implica, en nuestro caso, que aquel polinomio es una potencia de $q(x)$ con exponente $\leq s$.

DEMOSTRACIÓN DEL LEMA: Sean $q(x)^{s'}$ y $q(x)^{\bar{s}}$ los polinomios mínimos de f y $\bar{f}$ en u_i' y $[u_i']$, $\bar{s} \leq s' \leq s$. Entonces $q(\bar{f})^{\bar{s}}([u_i']) = [\vec{0}] \Rightarrow q(f)^{\bar{s}}(u_i') \in F_0$ $\Rightarrow q(f)^{\bar{s}}(u_i') = a(f)(u_0) \Rightarrow q(f)^{s-\bar{s}}a(f)(u_0) = q(f)^s(u_i') = \vec{0} \Rightarrow q(x)^s \mid q(x)^{s-\bar{s}}a(x)$ (ya que el polinomio mínimo de u_0 es $q(x)^s$) $\Rightarrow a(x) = q(x)^{\bar{s}}b(x)$ $\Rightarrow q(f)^{\bar{s}}(u_i') = a(f)(u_0) = q(f)^{\bar{s}}b(f)(u_0)$. El vector $u_i = u_i' - b(f)(u_0) \in u_i' + F_0 = [u_i']$ tiene un polinomio mínimo que divide a $q(x)^{\bar{s}}$, ya que

$$q(f)^{\bar{s}}(u_i' - b(f)(u_0)) = \vec{0}.$$

Por otra parte, el polinomio mínimo de u_i ha de ser múltiplo del de $[u_i] = [u_i']$, que es $q(x)^{\bar{s}}$. Así pues, el polinomio mínimo de u_i es $q(x)^{\bar{s}}$, el mismo que el de $[u_i]$. Los vectores $u_i, f(u_i), \ldots, f^{t-1}(u_i)$, donde $t = \bar{s} \cdot \operatorname{gr} q(x)$, forman, por (3.1), una base de F_i. Las clases

$$[u_i], \bar{f}([u_i]), \ldots, \bar{f}^{t-1}([u_i])$$

forman, también por (3.1), una base de $\bar{F}_i$, y por tanto la proyección $\pi_i : F_i \longrightarrow \bar{F}_i$ es un isomorfismo. Esto acaba la demostración de 7.2 y de 7.1. □

¿Hasta qué punto es única la descomposición obtenida en (7.1)? Supongamos que

$$E = G_0 \oplus G_1 \oplus \ldots \oplus G_m,$$

donde los subespacios G_i son f-cíclicos y el polinomio mínimo de la restricción de f a G_i es $q_i(x)^{n_i}$ con $q_i(x)$ irreducible. Agrupemos los sumandos que correspondan a potencias del mismo $q_i(x)$. Por ejemplo, supongamos $q_0(x) = q_1(x) = \ldots = q_s(x)$ y consideremos $E^0 = G_0 \oplus \ldots \oplus G_s$. El polinomio

mínimo de la restricción de f a E^0 es $q_0(x)^t$, donde $t = \max(n_0, \ldots, n_s)$. Agrupando de esta forma los G_i, obtenemos una descomposición de E,

$$E = E^0 \oplus E^1 \oplus \ldots \oplus E^r,$$

que es precisamente la de (4.4). Así pues, los subespacios E^i están unívocamente determinados y las diferentes descomposiciones de E en subespacios f-cíclicos corresponderán a las diferentes descomposiciones de los subespacios E^i.

Consideremos, pues, el caso en que el polinomio mínimo de $f \in \mathrm{End}(E)$ es $q(x)^s$, $q(x)$ primo. La primera observación que debe hacerse es que no podemos aspirar a demostrar la unicidad de la descomposición de (7.1). Lo veremos con un ejemplo.

Ejemplo:

Si $f = kI_E$, cualquier base $u_1, \ldots, u_n$ da lugar a una descomposición en subespacios f-cíclicos:

$$E = \langle u_1 \rangle \oplus \ldots \oplus \langle u_n \rangle.$$

Vamos a ver, no obstante, que en todas las posibles descomposiciones de E en subespacios f-cíclicos el número n_t de subespacios a los que corresponde un cierto polinomio mínimo $q(x)^t$ es el mismo. Recordemos que estamos considerando el caso en que el polinomio mínimo de f es $q(x)^s$. Supongamos que

$$E = F_0 \oplus \ldots \oplus F_r$$

es una descomposición de E en suma de subespacios f-cíclicos F_i. Sea $q(x)^{s_i}$ el polinomio mínimo de la restricción de f a F_i $(s_i \leq s)$.

Tenemos:

a) $q(f)^t(F_i) \subset F_i$; de donde

$$q(f)^t(E) = q(f)^t(F_0) \oplus \ldots \oplus q(f)^t(F_r)$$

y, por tanto,

$$\dim(q(f)^t(E)) = \sum_{i=0}^{r} \dim(q(f)^t(F_i)).$$

b) Si $s_i > t$, el polinomio mínimo de la restricción de f a $q(f)^t(F_i)$ es $q(x)^{s_i - t}$. Entonces,

$$\dim(q(f)^t(F_i)) = (s_i - t) \cdot \mathrm{gr}\, q(x) = \dim F_i - t \cdot \mathrm{gr}\, q(x).$$

c) Si $s_i \leq t$, $q(f)^t(F_i) = \{\vec{0}\}$. Entonces,

$$\dim(q(f)^t(F_i)) = 0 = \dim F_i - s_i \cdot \operatorname{gr} q(x).$$

Sustituyendo las expresiones de (b) y (c) en el sumatorio de (a), obtenemos

$$\dim(q(f)^t(E)) = \dim E - \sum_{i=0}^{r} \min(s_i, t) \cdot \operatorname{gr} q(x).$$

Observemos ahora que

$$\min(s_i, t) - \min(s_i, t-1)) = \begin{cases} 0 & \text{si } s_i \leq t-1 \\ 1 & \text{si } s_i \geq t, \end{cases}$$

de donde $\sum_{i=0}^{r}(\min(s_i, t) - \min(s_i, t-1)) = n_t + n_{t+1} + \ldots + n_s$. Denotemos por q_t la dimensión de $q(f)^t(E)$. Tenemos entonces

$$q_{t-1} - q_t = (n_t + n_{t+1} + \ldots + n_s) \cdot \operatorname{gr} q(x),$$

de donde resulta que

$$n_t = \frac{1}{\operatorname{gr} q(x)}(q_{t-1} - 2q_t + q_{t+1}).$$

Esta expresión de n_t no depende de la descomposición de E considerada.

Observación:

Se cumple

$$\{\vec{0}\} \subset \operatorname{Nuc} q(f) \subset \operatorname{Nuc} q(f)^2 \subset \ldots \subset \operatorname{Nuc} q(f)^s = \operatorname{Nuc} q(f)^{s+1} = \ldots = E.$$

Sea $\bar{q}_t = \dim \operatorname{Nuc} q(f)^t = n - q_t$ y designemos por

$$p_t = \dim\left(\operatorname{Nuc} q(f)^t / \operatorname{Nuc} q(f)^{t-1}\right) = \bar{q}_t - \bar{q}_{t-1}.$$

Resulta, entonces, que

$$n_t = \frac{1}{\operatorname{gr} q(x)}(p_t - p_{t+1}).$$

VIII.8 Matriz canónica de Jordan

Supongamos ahora que el polinomio mínimo de f se descompone en factores lineales. Sea

$$E = F_0 \oplus \ldots \oplus F_r$$

la descomposición de (7.1) en subespacios f-cíclicos y $(x-a_i)^{s_i}$ el polinomio mínimo de la restricción de f a F_i. Escojamos para cada F_i un vector u_i con polinomio mínimo $(x-a_i)^{s_i}$. Entonces,

$$u_i, (f-a_iI)(u_i), \ldots, (f-a_iI)^{s_i-1}(u_i)$$

son linealmente independientes. Ahora bien, por (3.1), $\dim F_i = s_i$ y por tanto estos vectores forman una base. La matriz de la restricción de f a F_i en esta base es

$$J(a_i, s_i) = \left.\begin{pmatrix} a_i & 0 & \cdots & 0 & 0 \\ 1 & a_i & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a_i & 0 \\ 0 & 0 & \cdots & 1 & a_i \end{pmatrix}\right\} s_i.$$

Tenemos así, como consecuencia de (7.1):

Teorema 8.1 *Si el polinomio mínimo de $f \in \mathrm{End}(E)$ se descompone en factores lineales, existe una base de E en la cual la matriz de f es de la forma*

$$J = \begin{pmatrix} \boxed{J(a_0, s_0)} & & & 0 \\ & \boxed{J(a_1, s_1)} & & \\ & & \ddots & \\ 0 & & & \boxed{J(a_r, s_r)} \end{pmatrix}. \quad \square$$

La matriz J se llama la *matriz canónica de Jordan de f*.

Observemos que los elementos a_i que aparecen en la diagonal de la matriz canónica de Jordan son los valores propios de f (posiblemente repetidos). Para obtener una base de E en la cual la matriz de f sea la matriz canónica de Jordan procederemos de la siguiente forma.

Sean

$$p(x) = (x-\lambda_1)^{\alpha_1} \ldots (x-\lambda_r)^{\alpha_r}$$

y

$$m(x) = (x-\lambda_1)^{s_1} \cdots (x-\lambda_r)^{s_r}$$

los polinomios característico y mínimo de f. Recordemos que $\alpha_1 + \cdots + \alpha_r = n = \dim E$ y $s_i \leq \alpha_i \ \forall i$. Los enteros s_i están caracterizados por el hecho de satisfacer

$$\{\vec{0}\} \subset \mathrm{Nuc}(f - \lambda_i I) \subset \mathrm{Nuc}(f - \lambda_i I)^2 \subset \\ \subset \ldots \subset \mathrm{Nuc}(f - \lambda_i I)^{s_i - 1} \subsetneq \mathrm{Nuc}(f - \lambda_i I)^{s_i}$$

y

$$\mathrm{Nuc}(f - \lambda_i)^{s_i} = \mathrm{Nuc}(f - \lambda_i)^t = E^i \qquad \forall t \geq s_i.$$

Además,

$$E = E^1 \oplus \ldots \oplus E^r.$$

La base que buscamos es unión de bases convenientes de los subespacios invariantes E^i. Restringiéndonos a estos espacios, podemos suponer que el polinomio característico de $f \in \mathrm{End}(E)$ es $(x - \lambda)^n$ y el polinomio mínimo $(x - \lambda)^s$, $s \leq n$. Para cualquier $u \in E$, designaremos $(f - \lambda I)(u)$ por $q(u)$. Consideremos el recuadro de vectores de la página 175. $u_{11}, \ldots, u_{1k_1}$ determinan clases que forman una base de $\mathrm{Nuc}(f - \lambda I)^s / \mathrm{Nuc}(f - \lambda I)^{s-1}$. Es fácil ver que $u_{11}, \ldots, u_{1k_1}$ son linealmente independientes y que $q(u_{11}), \ldots, q(u_{1k_1}) \in \mathrm{Nuc}(f - \lambda I)^{s-1}$ determinan clases linealmente independientes en $\mathrm{Nuc}(f - \lambda I)^{s-1} / \mathrm{Nuc}(f - \lambda I)^{s-2}$. Los vectores $u_{21}, \ldots, u_{2k_2}$ son vectores representantes de clases que, juntamente con las anteriores, forman una base de $\mathrm{Nuc}(f - \lambda I)^{s-1} / \mathrm{Nuc}(f - \lambda I)^{s-2}$. Repitamos ahora este proceso hasta obtener una base de $\mathrm{Nuc}(f - \lambda I)$ formada por los vectores situados en la última fila del recuadro.

Tenemos entonces:

i) El conjunto de todos los vectores que aparecen en el recuadro forman una base de E.

ii) El número de columnas es la multiplicidad del valor propio λ.

iii) El número de filas es el exponente del polinomio mínimo.

iv) El número de vectores en cada fila es

$$\dim(\mathrm{Nuc}(f - \lambda I)^t / \mathrm{Nuc}(f - \lambda I)^{t-1}) = \bar{q}_t - \bar{q}_{t-1} = p_t.$$

v) El número de matrices $J(\lambda, t)$ que aparecen en la matriz canónica de Jordan de f es $n_t = p_t - p_{t+1}$, que es la diferencia de vectores en filas consecutivas.

vi) Cada columna es la base de un subespacio f-cíclico de la descomposición de E.

$E = \mathrm{Nuc}(f-\lambda I)^s$	u_{11}	$\cdots$	u_{1k_1}													
$\mathrm{Nuc}(f-\lambda I)^{s-1}$	$q(u_{11})$	$\cdots$	$q(u_{1k_1})$		u_{21}	$\cdots$	u_{2k_2}									
$\cdots$	$\cdots$	$\cdots$	$\cdots$		$\cdots$	$\cdots$	$\cdots$									
$\mathrm{Nuc}(f-\lambda I)^2$	$q^{s-2}(u_{11})$	$\cdots$	$q^{s-2}(u_{1k_1})$		$q^{s-3}(u_{21})$	$\cdots$	$q^{s-3}(u_{2k_2})$	$\cdots$	$u_{s-1,1}$	$\cdots$	$u_{s-1,k_{s-1}}$					
$\mathrm{Nuc}(f-\lambda I)$	$q^{s-1}(u_{11})$	$\cdots$	$q^{s-1}(u_{1k_1})$		$q^{s-2}(u_{21})$	$\cdots$	$q^{s-2}(u_{2k_2})$	$\cdots$	$q(u_{s-1,1})$	$\cdots$	$q(u_{s-1,k_{s-1}})$		u_{s1}	$\cdots$	u_{sk_s}	

Observación:

$n_s \geq 1$ siempre, ya que $p_{s+1} \geq 1$. Esto significa que siempre aparece al menos una matriz $J(\lambda, s)$ de la máxima dimensión.

Ejemplo:

Consideremos un endomorfismo $f \in \mathrm{End}(E)$ con matriz

$$\begin{pmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & 3 & 1 & 1 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}.$$

Su polinomio característico es $(2-x)^5$. El rango de $(f-2I)$ es 2 y, por tanto, $\dim \mathrm{Nuc}(f-2I) = 3$. Además,

$$(f-2I)^2 = 0$$

y el polinomio mínimo de f es $(x-2)^2$. El recuadro considerado anteriormente tiene, en este caso, dos filas y tres elementos en la fila inferior:

u_{11}	u_{12}	
$q(u_{11})$	$q(u_{12})$	u_2

Podemos tomar $u_{11} = (1,0,0,0,0)$, $u_{12} = (0,0,0,1,0)$. Entonces,

$$\begin{aligned} q(u_{11}) &= (f-2I)(u_{11}) = (-1,1,0,0,0) \\ q(u_{12}) &= (f-2I)(u_{12}) = (-1,1,0,-1,1) \end{aligned}$$

son dos vectores propios que, juntamente con $u_2 = (0,-1,1,0,0)$, forman una base de vectores propios. En la base

$$u_{11},\ q(u_{11}),\ u_{12},\ q(u_{12}),\ u_2,$$

la matriz de f es

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

El problema de encontrar la matriz canónica de Jordan de un endomorfismo f dado queda, pues, resuelto, si el polinomio mínimo de f se descompone en factores lineales:

$$m(x) = (x - \lambda_1)^{s_1} \cdots (x - \lambda_r)^{s_r}.$$

Solamente nos hace falta conocer para cada valor propio λ_i los números n_t correspondientes ($1 \leq t \leq s_i$), los cuales nos indican cuántos subespacios f-cíclicos de dimensión t aparecen (submatrices con el valor propio λ_i en la diagonal y unos debajo de ella). En particular, si $s_i = 1$, el subespacio invariante correspondiente es un subespacio de vectores propios. El endomorfismo es diagonalizable si y sólo si $s_1 = \ldots = s_r = 1$ (4.6).

VIII.9 Nota histórica

En 1858 Arthur Cayley (1821–1895) enunció en general el teorema que hoy se conoce con el nombre de Cayley-Hamilton, demostrándolo para matrices 3×3 e introduciendo el polinomio característico de una matriz y sus raíces (valores propios).

Henry Taber (1860–?) enunció la proposición 1.2 y, en particular, que el término independiente del polinomio característico es el determinante de la matriz, introduciendo también la traza. La demostración del teorema la llevó a cabo William Henry Metzler (1863–?) en 1891.

Georg Ferdinand Frobenius (1849–1917) introdujo en 1878 el polinomio mínimo y Kurt Hensel (1861–1941) demostró, en 1904, sus principales propiedades; en particular, que cualquier otro anulador es múltiplo del polinomio mínimo.

Utilizando el concepto de matrices equivalentes y el polinomio característico, Camille Jordan (1838–1922) demostró en 1870 que toda matriz es equivalente a una en forma canónica (la forma canónica de Jordan).

El desarrollo algebraico de los temas de este capítulo fue de gran trascendencia para la física. Tal como profetizó Peter G. Tait (1831–1901), "Cayley is forging the weapons for future generations of physicists".

VIII.10 Ejercicios

1. Sea A la matriz de $M_{n\times n}(K)$ formada íntegramente por unos. Calcular los polinomios característico y mínimo de A. Probar que A es diagonalizable y encontrar una matriz diagonal D y una invertible M tales que $A = MDM^{-1}$.

2. a) Determinar todas las matrices $A \in M_{2\times 2}(\mathbf{R})$ tales que

$$\text{i) } A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \qquad \text{ii) } A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix};$$
$$\text{iii) } A^2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \quad \text{iv) } A^2 = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}.$$

b) Determinar todas las matrices $A \in M_{4\times 4}(\mathbf{R})$ tales que

$$A^2 - 3A + 2I = 0.$$

c) Determinar todas las matrices $A \in M_{n\times n}(\mathbf{R})$ tales que $A^2 = A$.

3. Dada

$$A = \begin{pmatrix} 5/2 & -1 \\ 3 & -1 \end{pmatrix},$$

calcular A^{1438} y $\lim_{n\to\infty} A^n$.

4. Sea $e_1, \ldots, e_n$ una base del espacio vectorial E y $f \in \mathrm{End}(E)$ tal que

$$f(e_1) = \ldots = f(e_n) = \sum_{i=1}^{n} a^i e_i.$$

Demostrar que f es diagonalizable si y sólo si $\sum_{i=1}^{n} a^i \neq 0$.

5. Determinar la forma general de las matrices que conmutan con las matrices diagonales y de las que conmutan con las diagonalizables.

6. Demostrar que $f \in \mathrm{End}_{\mathbf{C}}(E)$ es diagonalizable si y sólo si todo subespacio invariante por f admite un complementario también invariante por f.

7. Construir un endomorfismo f de $\mathbf{C}^3$ tal que

$$p_f(x) = m_f(x) = x^2(x-a), \quad a \neq 0.$$

Demostrar que si u es un vector propio de valor propio a, y v pertenece a $\mathrm{Nuc}\, f^2$ pero no a $\mathrm{Nuc}\, f$, entonces $u, v, f(v)$ es una base de $\mathbf{C}^3$. Hallar la matriz de f en esta base y calcular f^n.

8. Sea $f \in \mathrm{End}_{\mathbf{R}}(E)$. Demostrar que si $\dim E$ es impar, entonces f tiene algún valor propio y que si $\dim E$ es par y $\det f < 0$, f tiene al menos dos valores propios. Dar un ejemplo de un endomorfismo sin valores propios.

9. Demostrar que si $f \in \mathrm{End}(E)$ es diagonalizable y F es un subespacio invariante por f, entonces $f \mid F$ también es diagonalizable.

10. Demostrar que si $f, g \in \mathrm{End}(E)$ conmutan, los subespacios de vectores propios de g son invariantes por f y recíprocamente.

11. Se dice que dos endomorfismos $f, g \in \mathrm{End}(E)$ son *simultáneamente diagonalizables* si podemos encontrar una base de E en la cual las matrices de f y g sean ambas diagonales.

a) Demostrar que f y g son simultáneamente diagonalizables si y sólo si son diagonalizables y conmutan.
(Indicación: utilizar los dos ejercicios anteriores.)

b) Diagonalizar simultáneamente los dos endomorfismos de $\mathbf{R}^3$:

$$f(x,y,z) = (x+y+z, 2x+5y+2z, -2x-5y-2z)$$
$$g(x,y,z) = (-2y-2z, 0, 2y+2z).$$

12. Dada $A \in M_{n\times n}(\mathbf{R})$, consideramos la ecuación diferencial

$$A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix};$$

es decir, buscamos n funciones diferenciables $x_i : \mathbf{R} \to \mathbf{R}$ que satisfacen una determinada relación entre ellas y sus derivadas.

a) Convencerse de que si A es diagonal (y también si A es diagonalizable) sabemos resolver cualquier ecuación diferencial de este tipo.

b) Resolver las siguientes ecuaciones diferenciales:

i) $\begin{cases} x'=y \\ y'=x; \end{cases}$ ii) $\begin{cases} x'=-x-2z \\ y'=6x+y+6z \\ z'=x+2z; \end{cases}$ iii) $y''-y'=y.$

(Indicación: en (iii), introducir la nueva variable $z = y'$.)

13. Sea $(a_n), (b_n), \ldots$ una familia de m sucesiones tales que

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \\ \vdots \end{pmatrix} = A\begin{pmatrix} a_n \\ b_n \\ \vdots \end{pmatrix}$$

con $A \in M_{m\times m}(K)$. Hallar una expresión (no recurrente) del término general de cada una de estas sucesiones cuando A es diagonal. ¿Cómo se puede encontrar ese término general cuando A es diagonalizable?

Aplicarlo a los siguientes casos:

a) Encontrar todas las sucesiones (a_n) tales que $a_{n+1} = 2a_n + a_{n-1}$. (Indicación: introducir una nueva sucesión $b_n = a_{n-1}$.)

b) Encontrar el término general de la sucesión de Fibonacci 0, 1, 1, 2, 3, 5, 8, 13, ...

c) Discutir la convergencia de las sucesiones complejas (a_n), (b_n) dadas por

$$a_{n+1} = \alpha(a_n + \sqrt{3}b_n) \quad \text{y} \quad b_{n+1} = \alpha(-\sqrt{3}a_n + b_n)$$

según los valores de α, a_0 y b_0.

14. ¿Cuáles de las matrices

$$A = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \; C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

pueden estar asociadas al mismo endomorfismo? Para éstas, buscar la matriz del cambio de base.

15. Sea $f \in \mathrm{End}(E)$, $\dim E = n$. Demostrar:

a) Existe un $k \in \mathbf{N}$ tal que $\mathrm{Nuc}\, f^k = \mathrm{Nuc}\, f^{k+i}$ para todo $i \geq 0$.

b) $E = \mathrm{Nuc}\, f^k \oplus \mathrm{Im}\, f^k$ y $f \mid \mathrm{Im}\, f^k$ es un automorfismo de $\mathrm{Im}\, f^k$.

16. Estudiar los endomorfismos f que cumplen $f^3 = a^2 f$, $a \in \mathbf{R}$.

17. Un endomorfismo $f \in \mathrm{End}(E)$ se llama *nilpotente* si existe un $n \in \mathbf{N}$ tal que $f^n = 0$. Si f es nilpotente, demostrar:

a) $\mathrm{tr}\, f = 0$.

b) $\det(f + I) = 1$.

c) Para todo $g \in \mathrm{End}(E)$ que conmute con f, $\det(f + g) = \det g$. (Indicación: considerar por separado los casos g invertible y no invertible.)

d) Si $g \in \mathrm{End}(E)$, entonces $E = F \oplus G$ con $g \mid G$ isomorfismo y $g \mid F$ nilpotente (*descomposición de Fitting*).

18. Sea $f : \mathbf{R}_2[x] \longrightarrow \mathbf{R}_2[x]$ el endomorfismo que hace corresponder $f(p) = p + p'$ a cada polinomio real p de grado menor que 3.

a) Encontrar la forma canónica de Jordan de f.

b) Demostrar que f^{-1} es una expresión polinómica en f.

c) Encontrar la matriz de f^{-1} en la base $1, x, x^2$. (Indicación: utilizar b).)

19. Sea $q(x)$ un polinomio cualquiera. Demostrar que si el polinomio característico de un endomorfismo f es $(x - a_1)^{n_1} \cdots (x - a_r)^{n_r}$, el polinomio característico de $q(f)$ es $(x - q(a_1))^{n_1} \cdots (x - q(a_r))^{n_r}$.

20. Encontrar la forma canónica de Jordan del endomorfismo de $M_{2\times 2}(\mathbf{C})$ que envía $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ a $\begin{pmatrix} 2b + 5c - 6d & -a + 3b + 4c - 5d \\ -4c + 9d & -4c + 4d \end{pmatrix}$.

VIII.11 Ejercicios para programar

21. Preparar un programa que calcule los coeficientes del polinomio característico de una matriz $A \in M_{n\times n}(\mathbf{R})$ $(n \leq 5)$ usando la expresión dada en la proposición 1.2.

22. Preparar uu programa que, dada una matriz $A \in M_{n\times n}(\mathbf{R})$, calcule su polinomio mínimo por el siguiente procedimiento:

Calcular

$$r_k = \dim\langle I, A, A^2, \ldots, A^k\rangle \leq \dim M_{n\times n}(\mathbf{R}), \quad k = 1, 2, 3, \ldots.$$

Para el primer $k \leq n$ tal que $r_k = k$, resolver el sistema de ecuaciones

$$A^k = x_0 I + x_1 A + \ldots + x_{k-1} A^{k-1}$$

escogiendo como matriz del sistema aquella que había dado $r_{k-1} = k$. Se obtiene una única solución y los valores de las variables x_i son los coeficientes del polinomio mínimo.

23. Dada $A \in M_{2\times 2}(\mathbf{R})$, calcular sus valores propios k_1, k_2 (que pueden ser reales o complejos).

- Si $k_1 \neq k_2$, encontrar una base de vectores propios.
- Si $k_1 = k_2$ y A no es una homotecia, no puede ser diagonalizable. Encontrar el único subespacio de vectores propios.

24. Dada $A \in M_{3\times 3}(\mathbf{R})$, calcular todos sus valores propios por el siguiente procedimiento:

El polinomio característico debe tener por lo menos una raíz real α. Encontrarla, dividir por $x - \alpha$ y calcular las dos raíces restantes.

Para realizar un cálculo aproximado de α se puede usar el siguiente método: si $p(x) = a_0 + a_1 x + a_2 x^2 - x^3$ es el polinomio característico de A, todas sus raíces están en el intervalo $(-b, b)$ donde $b = 1 + |a_0| + |a_1| + |a_2|$. Por tanto, $p(-b) \cdot p(b) < 0$. Vayamos dividiendo sucesivamente el intervalo por la mitad, quedándonos en la parte que contenga un cambio de signo. (Los cálculos son más sencillos si se escribe $p(x) = a_0 + x(a_1 + x(a_2 - x))$.)

25. Dada $A \in M_{n\times n}(\mathbf{R})$, de la cual conocemos un valor propio real k, encontrar una base del subespacio de vectores propios correspondientes. (Indicación: plantear el sistema homogéneo $(A - kI)x = \vec{0}$ y resolverlo mediante el programa del ejercicio VII.12.)

En la práctica, para n grande, los valores propios reales de una matriz $A \in M_{n\times n}(\mathbf{R})$ no se buscan a partir del polinomio característico, sino que se aproximan por métodos iterativos. Describiremos a continuación uno de ellos y el resto en el capítulo XII.

26. Método *LU*

Escribamos $A = LU$ como en el ejercicio IV.17. Sea $A_1 = UL$. El algoritmo consiste en repetir este cálculo: descomponer $A_k = L_k U_k$ y tomar $A_{k+1} = U_k L_k$, $k = 1, 2, \ldots$

Las matrices que se van obteniendo son todas equivalentes a A, ya que $A_{k+1} = (L_k)^{-1} A_k L_k$ para todo k.

Bajo ciertas hipótesis, la sucesión $\{A_k\}$ tiende a una matriz triangular superior (y, por tanto, en la diagonal quedarán los valores propios de A).

Esas hipótesis son técnicas y no hay manera de comprobarlas *a priori*. El algoritmo fallará, pues, en algunos casos. Hay por lo menos dos condiciones necesarias obvias:

a) La descomposición LU tiene que ser posible en cada paso.

b) Todos los valores propios de A han de ser reales (p.e. si la matriz A es simétrica; ver XI.7.3).

Capítulo IX

Espacios afines

A estas alturas, estamos ya acostumbrados a asociar la "recta", el "plano" y el "espacio" con $\mathbf{R}$, $\mathbf{R}^2$ y $\mathbf{R}^3$, respectivamente, y tenemos, incluso, una cierta visión geométrica de $\mathbf{R}^n$ (IV.1, ejemplo 2). Un subespacio F de dimensión 1 es, "geométricamente", una recta que pasa por $(0,0)$; las rectas paralelas a ésta son, precisamente, los subconjuntos $u+F$ de $\mathbf{R}^2$ (recordemos que la suma de vectores corresponde geométricamente a la "ley del paralelogramo").

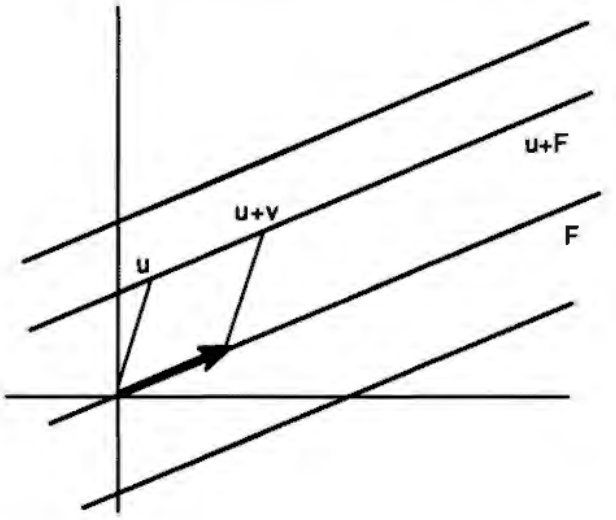
De la misma manera, las rectas (planos) de $\mathbf{R}^3$ son los subconjuntos $u+F$, donde F es un subespacio vectorial de dimensión 1 (2). Podemos considerar, por tanto, la "geometría afín" como el estudio de $\mathbf{R}^n$ y de sus subconjuntos $u+F$.

El hecho de identificar el estudio de la geometría con el estudio del espacio vectorial $\mathbf{R}^n$ tiene, no obstante, inconvenientes. En el plano, por ejemplo, no hay ningún criterio que permita diferenciar una recta o una familia de rectas entre ellas; todas las rectas tienen exactamente las mismas características y una elección sería, por tanto, totalmente arbitraria. En $\mathbf{R}^2$, en cambio, los subconjuntos $u+F$ se agrupan en dos tipos bien diferenciados: unos son subespacios vectoriales ($\vec{0}+F=F$) y los otros no. Esta diferenciación no corresponde a ningún hecho geométrico y la deberíamos evitar.

Si no hay ningún motivo para elegir una entre todas las rectas que pasan por un punto-origen, hagamos que cualquier punto pueda ser ese origen. Así pues, un *plano afín* será, por definición, un conjunto A y, para cada $p \in A$, una aplicación biyectiva

$$\varphi_p : A \longrightarrow \mathbf{R}^2$$

tal que $\varphi_p(p) = (0,0)$. Naturalmente, las aplicaciones φ_p no pueden ser totalmente arbitrarias; tendremos que exigir algunas condiciones que las



relacionen (en realidad, impondremos solamente una condición).

La geometría a la cual nos hemos referido hasta aquí es la "geometría afín real". En principio, no hay ningún inconveniente en considerar, en lugar de $\mathbf{R}$, cualquier cuerpo K y, en lugar de $\mathbf{R}^2$ (o $\mathbf{R}^n$), cualquier espacio vectorial sobre K.

IX.1 Definición de espacio afín

Un *espacio afín* sobre un cuerpo K es un conjunto $A \neq \emptyset$, un espacio vectorial E y una aplicación

$$\varphi : A \times A \longrightarrow E$$

que cumple:

1.
$$\varphi_p : \begin{array}{ccc} A & \longrightarrow & E \\ q & \longmapsto & \varphi(p,q) \end{array} \quad \text{es biyectiva } \forall p \in A.$$

2.
$$\varphi(p,q) + \varphi(q,r) = \varphi(p,r) \qquad \forall p,q,r \in A.$$

Escribiremos

$$\varphi(p,q) = \overrightarrow{pq}$$

y diremos que p y q son, respectivamente, *el origen* y *el extremo* del vector $\overrightarrow{pq}$. Con esta notación, la condición 2 establece que

$$\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}.$$

Los elementos de A se llaman *puntos*. E se llama el *espacio vectorial asociado a* A, y definimos la *dimensión* de A como la dimensión de E.

Ejemplo:

Sea K un cuerpo y $A = K^n$, $E = K^n$,

$$\begin{array}{rcl} \varphi : A \times A & \longrightarrow & E \\ ((a_1, \ldots, a_n), (b_1, \ldots, b_n)) & \longmapsto & (b_1 - a_1, \ldots, b_n - a_n). \end{array}$$

(A, E, φ) es un espacio afín que denominaremos *espacio afín estándar de dimensión n sobre K*.

Proposición 1.1 *a)* $\overrightarrow{pq} = \vec{0} \Leftrightarrow p = q$.

b) $\overrightarrow{pq} = -\overrightarrow{qp} \qquad \forall p, q \in A$.

c) $\overrightarrow{pq} = \overrightarrow{rs} \Leftrightarrow \overrightarrow{pr} = \overrightarrow{qs}$.

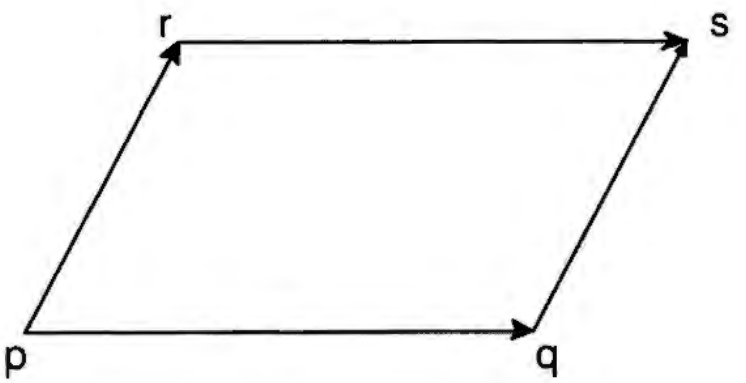
DEMOSTRACIÓN: La condición 2 en la definición de espacio afín nos dice que $\overrightarrow{pp} + \overrightarrow{pp} = \overrightarrow{pp}$. De ahí resulta $\overrightarrow{pp} = \vec{0}$. Recíprocamente, $\overrightarrow{pq} = \vec{0}$ implica $\varphi_p(q) = \vec{0}$; pero puesto que $\varphi_p(p) = \overrightarrow{pp} = \vec{0}$, de la biyectividad de φ_p resulta $p = q$, lo que demuestra (a).

Para probar (b), observemos que

$$\overrightarrow{pq} + \overrightarrow{qp} = \overrightarrow{pp} = \vec{0}; \qquad \text{es decir, } \overrightarrow{pq} = -\overrightarrow{qp}.$$

Finalmente, para demostrar (c), pongamos

$$\overrightarrow{pr} = \overrightarrow{pq} + \overrightarrow{qr}, \qquad \overrightarrow{qs} = \overrightarrow{qr} + \overrightarrow{rs}.$$



De ahí resulta que $\overrightarrow{pr} = \overrightarrow{qs}$ si y sólo si $\overrightarrow{pq} = \overrightarrow{rs}$. □

IX.2 Traslaciones. Otra definición de espacio afín

Sea (A, E, φ) un espacio afín. Dado $u \in E$, llamaremos *traslación de vector u* a la aplicación:

$$\begin{array}{rccc} T_u: & A & \longrightarrow & A \\ & p & \longmapsto & \varphi_p^{-1}(u); \end{array}$$

es decir, $T_u(p)$ es un punto q tal que $\overrightarrow{pq} = u$.

Ejemplo:

$T_{\vec{0}}$ es la aplicación identidad.

Proposición 2.1 *a)* T_u *es biyectiva para todo* $u \in E$.

b) Si existe un $p \in A$ *tal que* $T_u(p) = T_v(p)$, *entonces* $u = v$.

c) $T_u \circ T_v = T_{u+v}$

d) $T_{-u} = T_u^{-1}$.

e) Dados $p, q \in A$, *existe un* $u \in E$ *y sólo uno tal que* $T_u(p) = q$.

DEMOSTRACIÓN: T_u es inyectiva, ya que

$$\begin{aligned} T_u(p) = T_u(p') = q \quad &\Rightarrow \quad \overrightarrow{pq} = u = \overrightarrow{p'q} \Rightarrow \\ &\Rightarrow \quad \varphi_q(p) = \overrightarrow{qp} = \overrightarrow{qp'} = \varphi_q(p') \Rightarrow p = p'. \end{aligned}$$

T_u es exhaustiva, ya que, dado $q \in A$, el punto $p = \varphi_q^{-1}(-u)$ cumple $\overrightarrow{qp} = -u$; por tanto, $\overrightarrow{pq} = u$ y $q = T_u(p)$.

Para probar (b), observemos que de $T_u(p) = T_v(p) = q$ se deduce $\overrightarrow{pq} = u$, $\overrightarrow{pq} = v$ y, por tanto, $u = v$.

Sea $p \in A$ y pongamos $q = T_v(p)$, $r = T_u(q)$. Entonces $T_u \circ T_v(p) = r$. Ahora bien, $\overrightarrow{pr} = \overrightarrow{pq} + \overrightarrow{qr} = u + v$ y, por tanto, $T_u \circ T_v(p) = T_{u+v}(p)$. Este razonamiento se puede hacer para todo $p \in A$ y nos demuestra (c).

(d) resulta de (c) y de que $T_{\vec{0}} = I$.

En (e), $T_u(p) = q \Leftrightarrow u = \overrightarrow{pq}$ y, por tanto, este es el vector pedido. □

El último apartado de la proposición 2.1 nos dice que las traslaciones de un espacio afín A, $\{T_u; u \in E\}$, determinan la aplicación $\varphi : A \times A \longrightarrow E$; $\varphi(p, q)$ es precisamente el vector cuya existencia está asegurada por (e). Pero todavía hay más, como lo prueba la proposición siguiente.

Proposición 2.2 *Dado un conjunto A, un espacio vectorial E y una familia de aplicaciones $\mathcal{T} = \{T_u : A \longrightarrow A;\ \forall u \in E\}$ que cumplan*

1. $T_u \circ T_v = T_{u+v}$;
2. *dados $p, q \in A$, existe un $u \in E$ y sólo uno tal que $T_u(p) = q$;*

entonces existe un único espacio afín (A, E, φ) tal que $\mathcal{T}$ es precisamente su conjunto de traslaciones.

DEMOSTRACIÓN: Naturalmente, φ ha de ser la aplicación

$$\varphi : A \times A \longrightarrow E$$

tal que $\varphi(p, q) = u$ es el vector dado por 2. Sólo hace falta comprobar que (A, E, φ) es efectivamente un espacio afín.

φ_p es inyectiva, ya que

$$\varphi_p(q) = \varphi_p(q') = u \Rightarrow q = T_u(p) = q'.$$

φ_p es exhaustiva, ya que la antiimagen de $u \in E$ es $q = T_u(p)$.

Tenemos también: $\varphi(p, q) = u,\ \varphi(q, r) = v \Rightarrow T_u(p) = q, T_v(q) = r \Rightarrow r = T_v \circ T_u(p) = T_{v+u}(p) \Rightarrow \varphi(p, r) = u + v = \varphi(p, q) + \varphi(q, r)$.

Comprobemos finalmente que $\mathcal{T}$ es el conjunto de traslaciones de (A, E, φ); si $T'_u : A \longrightarrow A$ es la aplicación $T'_u(p) = \varphi_p^{-1}(u)$, entonces

$$T'_u(p) = q \Leftrightarrow \varphi(p, q) = u \Leftrightarrow T_u(p) = q.$$

Por tanto, $T'_u = T_u$ para todo $u \in E$. □

Esta proposición nos dice que un *espacio afín* también se puede definir como un conjunto A, un espacio vectorial E y una familia de aplicaciones $\mathcal{T} = \{T_u : A \longrightarrow A;\ \forall u \in E\}$ que cumplan las dos propiedades de (2.2).

IX.3 Variedades lineales

Sea (A, E, φ) un espacio afín. Sea $a \in A$ y F un subespacio vectorial de E. Se llama *variedad lineal que pasa por a y tiene la dirección F* al subconjunto de A

$$\{b \in A \mid \overrightarrow{ab} \in F\}.$$

Para indicar $u = \overrightarrow{ab}$, usaremos las notaciones

$$u = b - a, \qquad b = a + u.$$

Con esta notación, por ejemplo, $T_u(a) = a + u$.

Una variedad lineal es, pues, un conjunto del tipo $\{b \in A \mid b = a + u, u \in F\}$, que designaremos por

$$a + F.$$

Observemos que $A = a + E$ es también una variedad lineal.

La *dimensión* de una variedad lineal $a + F$ es la dimensión de su dirección F. Las variedades de dimensión 0 son los puntos de A. Las variedades de dimensiones 1, 2 y $(n-1)$ se llaman *rectas*, *planos* e *hiperplanos*, respectivamente ($n = \dim E$).

Proposición 3.1 $b \in a + F \Rightarrow a + F = b + F$.

DEMOSTRACIÓN: Por hipótesis, $\overrightarrow{ab} \in F$. Entonces

$$\begin{aligned} d \in a + F &\Leftrightarrow \overrightarrow{ad} \in F \Leftrightarrow \overrightarrow{ab} + \overrightarrow{bd} = \overrightarrow{ab} - \overrightarrow{db} \in F \Leftrightarrow \\ &\Leftrightarrow \overrightarrow{db} \in F \Leftrightarrow d \in b + F. \ \square \end{aligned}$$

Corolario 3.2 $p, q \in a + F \Rightarrow \overrightarrow{pq} \in F$.

DEMOSTRACIÓN: Si $p \in a + F$, por (3.1) $a + F = p + F$; puesto que $q \in p + F$, $\overrightarrow{pq} \in F$. $\square$

Proposición 3.3 *Dados $a_1, \ldots, a_k \in A$, existe una variedad lineal mínima que contiene a $a_1, \ldots, a_k$ y que denominaremos* variedad generada *por $a_1, \ldots, a_k$.*

Aquí "mínima" significa "contenida en cualquier otra variedad que contenga también a $a_1, \ldots, a_k$".

DEMOSTRACIÓN: Sea $a_1 + F$ una variedad que contenga a $a_1, \ldots, a_k$ (existe al menos una: todo A). Por (3.2), $\overrightarrow{a_1 a_i} \in F$, $i = 2, \ldots, k$, de donde $\langle \overrightarrow{a_1 a_2}, \ldots, \overrightarrow{a_1 a_k} \rangle \subset F$.

La variedad

$$a_1 + \langle \overrightarrow{a_1 a_2}, \ldots, \overrightarrow{a_1 a_k} \rangle$$

contiene a a_1, $a_i = a_1 + \overrightarrow{a_1 a_i}$ $(i = 2, \ldots, k)$ y está contenida en $a_1 + F$; es, pues, la variedad generada por $a_1, \ldots, a_k$. $\square$

Por cada punto $a \in A$ pasa una variedad lineal (¡y sólo una!) con una dirección dada F. Dos variedades lineales con la misma dirección diremos

que son paralelas. Más en general, diremos que dos variedades lineales $a+F$, $b+G$ son *paralelas* si

$$F \subset G \quad \text{o} \quad G \subset F.$$

Nota:

Cuando estudiábamos los espacios vectoriales o los grupos, nos ocupábamos también de los subconjuntos que eran ellos mismos espacios vectoriales o grupos con las "mismas" operaciones que el conjunto total: los subespacios vectoriales o los subgrupos, respectivamente. ¿Qué es ahora un "subespacio afín" de un espacio afín (A, E, φ)? Será un espacio afín (L, F, φ') donde $L \subset A$, F sea un subespacio de E y

$$\varphi' : L \times L \longrightarrow F$$

sea la restricción de φ. Es decir, se tendrá que cumplir

$$\overrightarrow{pq} = \varphi(p,q) \in F \qquad \forall p, q \in L$$

y, entonces, $\varphi'(p,q) = \varphi(p,q)$. En particular, si $p \in L$, cualquier otro $q \in L$ es de la forma $q = p + \overrightarrow{pq} \in p+F$, y $L \subset p+F$. Recíprocamente, todo $q = p+u \in p+F$ es tal que $q = \varphi_p^{-1}(u)$. Puesto que, claramente, $q = \varphi_p^{-1}(u) = \varphi'^{-1}_p(u)$, resulta que $q \in L$ y, por tanto, $p + F \subset L$.

Hemos probado de esta manera que las variedades lineales son precisamente los "subespacios afines" de un espacio afín.

IX.4 Intersección y suma de variedades lineales

Proposición 4.1 *Dos variedades lineales $a+F$, $b+G$ se cortan si y sólo si*

$$\overrightarrow{ab} \in F + G.$$

DEMOSTRACIÓN: Supongamos primero que se cortan y sea c un punto de la intersección $(a + F) \cap (b + G)$. Entonces $\overrightarrow{ac} \in F$ y $\overrightarrow{bc} \in G$, de donde $\overrightarrow{ab} = \overrightarrow{ac} - \overrightarrow{bc} \in F + G$.

Supongamos ahora que $\overrightarrow{ab} = u+v$ con $u \in F$ y $v \in G$. Sea c un punto tal que $\overrightarrow{ac} = u$ (en particular, $c \in a+F$). Entonces $v = \overrightarrow{ab} - u = \overrightarrow{ab} - \overrightarrow{ac} = \overrightarrow{cb}$ y, por tanto, $\overrightarrow{bc} = -v \in G$, de donde $c \in b+G$. El punto c es, pues, un punto común a las dos variedades. □

Corolario 4.2 *Dos variedades paralelas o no se cortan o una está contenida en la otra.*

DEMOSTRACIÓN: Sean $a+F$, $b+G$ con $F \subset G$.

$$(a+F)\cap(b+G)\neq\emptyset \Leftrightarrow \overrightarrow{ab}\in F+G=G \Leftrightarrow a\in b+G \Leftrightarrow$$

$$\Leftrightarrow a+F\subset a+G=b+G. \square$$

Proposición 4.3 *Si $a+F$ y $b+G$ tienen un punto c en común, entonces*

$$(a+F)\cap(b+G)=c+(F\cap G).$$

DEMOSTRACIÓN: Por (3.1), $a+F=c+F$ y $b+G=c+G$. Entonces

$$x\in(c+F)\cap(c+G) \Leftrightarrow \overrightarrow{cx}\in F \text{ y } \overrightarrow{cx}\in G \Leftrightarrow \overrightarrow{cx}\in F\cap G \Leftrightarrow$$

$$\Leftrightarrow x\in c+(F\cap G). \square$$

Se dice que dos variedades lineales *se cruzan* si no son paralelas ni se cortan.

Acabamos de ver que la intersección de dos variedades lineales, si no es vacía, es una variedad lineal (4.3). La unión de dos variedades lineales, en cambio, no es en general una variedad.

Ejemplo:

Sean u, v dos vectores linealmente independientes.

Los puntos $c=a+u$ y $d=a+v$ pertenecen a la unión $L=(a+\langle v\rangle)\cup(a+\langle u\rangle)$. Si L fuese una variedad lineal, el vector $\overrightarrow{cd}=v-u$ sería de su dirección (3.2) y el punto $p=a+(v-u)$ sería de L. Ahora bien, el punto p no es de $a+\langle v\rangle$ ni de $a+\langle u\rangle$ y, por tanto, no puede ser de L.

Dadas dos variedades lineales $a+F$ y $b+G$, ¿cuál es la variedad lineal "mínima" que las contiene? Supongamos que $a+H$ contiene a $a+F$ y $b+G$. Entonces

$$H\supset F+G+\langle\overrightarrow{ab}\rangle.$$

En efecto, $a, b \in a + H$, de donde $\overrightarrow{ab} \in H$. Además,

$$\begin{aligned} u \in F &\Rightarrow p = a + u \in a + F \subset a + H \Rightarrow u = \overrightarrow{ap} \in H; \\ v \in G &\Rightarrow q = b + v \in b + G \subset a + H \Rightarrow \overrightarrow{aq} \in H \Rightarrow \\ &\Rightarrow v = \overrightarrow{bq} = \overrightarrow{aq} - \overrightarrow{ab} \in H. \end{aligned}$$

Por otra parte, la variedad

$$a + F + G + \langle \overrightarrow{ab} \rangle$$

contiene claramente a $a + F$ y $b + G$. Esta es, por tanto, la variedad lineal "mínima" que contiene a $a + F$ y $b + G$. La llamaremos *variedad suma*: $(a + F) + (b + G)$.

Proposición 4.4 (Fórmulas de Grassmann) *Sean $L_1 = a + F$ y $L_2 = b + G$ dos variedades lineales y sea $L_1 + L_2$ su suma. Entonces, si $L_1 \cap L_2 \neq \emptyset$,*

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2),$$

y si $L_1 \cap L_2 = \emptyset$,

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(F \cap G) + 1.$$

DEMOSTRACIÓN: Claramente, $\dim(L_1 + L_2) = \dim(F + G + \langle \overrightarrow{ab} \rangle)$. Por (4.1), si $L_1 \cap L_2 \neq \emptyset$, $\overrightarrow{ab} \in F + G$ y

$$\dim(L_1 + L_2) = \dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

En este caso, $L_1 \cap L_2$ es una variedad lineal de dirección $F \cap G$ y $\dim(F \cap G) = \dim(L_1 \cap L_2)$. Entonces

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2).$$

Si $L_1 \cap L_2 = \emptyset$, por (4.1), $\overrightarrow{ab} \notin F + G$ y

$$\begin{aligned} \dim(L_1 + L_2) &= \dim(F + G) + 1 = \\ &= \dim F + \dim G - \dim(F \cap G) + 1 = \\ &= \dim L_1 + \dim L_2 - \dim(F \cap G) + 1. \ \square \end{aligned}$$

Ejemplo:

Sea A un espacio afín de dimensión 3. Sean

$$r = a + F, \qquad \pi = b + G,$$

una recta y un plano de A: $\dim F = 1$, $\dim G = 2$. Si $r \cap \pi = \emptyset$,

$$\dim(r + \pi) = 4 - \dim(F \cap G).$$

Pero esta dimensión ha de ser menor que 3; por tanto, $\dim(F \cap G) = 1$, de donde $F \subset G$ y r, π son paralelas.

Si $r \cap \pi \neq \emptyset$,

$$\dim(r + \pi) = 3 - \dim(r \cap \pi).$$

Por tanto,

o bien $\dim(r + \pi) = 3 \Rightarrow \dim(r \cap \pi) = 0 \Rightarrow r \cap \pi =$ un punto,
o bien $\dim(r + \pi) = 2 \Rightarrow \dim(r \cap \pi) = 1 \Rightarrow r \subset \pi$.

En este segundo caso ($r \subset \pi$), r y π son paralelas. Así pues, en un espacio afín de dimensión 3, una recta y un plano no paralelos siempre se cortan en un punto.

Ejercicio:

Demostrar que en un espacio afín de dimensión 2 dos rectas no paralelas siempre se cortan en un punto.

IX.5 Dependencia lineal de puntos

Proposición 5.1 *Sea (A, E) un espacio afín y sean $a_1, \ldots, a_k \in A$. Las siguientes condiciones son equivalentes:*

a) Los vectores $\overrightarrow{a_1a_2}, \ldots, \overrightarrow{a_1a_k}$ de E son linealmente independientes.

b) Fijado cualquier i, los vectores $\{\overrightarrow{a_ia_h},\ h \neq i\}$ son linealmente independientes.

c) Para todo $p \in A$,

$$\left.\begin{array}{c} \lambda^1\overrightarrow{pa_1} + \ldots + \lambda^k\overrightarrow{pa_k} = 0 \\ \lambda^1 + \ldots + \lambda^k = 0 \end{array}\right\} \Rightarrow \lambda^1 = \ldots = \lambda^k = 0.$$

DEMOSTRACIÓN: (a) $\Rightarrow$ (b). En efecto, de

$$\sum_{h\neq i} \lambda^h \overrightarrow{a_i a_h} = \vec{0}$$

resulta

$$\sum_{h\neq i} \lambda^h (\overrightarrow{a_1 a_h} - \overrightarrow{a_1 a_i}) = \sum_{h\neq i} \lambda^h \overrightarrow{a_1 a_h} - \left(\sum_{h\neq i} \lambda^h \right) \overrightarrow{a_1 a_i} = \vec{0}.$$

(a) nos dice que $\lambda^h = 0 \ \forall h \neq i, 1$. Pero $\sum_{h\neq i} \lambda^h = 0$ implica también $\lambda^1 = 0$.

(b) $\Rightarrow$ (a). En efecto, (a) coincide con (b) para $i = 1$.

(a) $\Rightarrow$ (c). En efecto, de

$$\begin{cases} \lambda^1 \overrightarrow{pa_1} + \ldots + \lambda^k \overrightarrow{pa_k} = \vec{0} \\ \lambda^1 + \ldots + \lambda^k = 0 \end{cases}$$

se deduce que

$$\begin{aligned} &(-\lambda^2 - \ldots - \lambda^k)\overrightarrow{pa_1} + \lambda^2 \overrightarrow{pa_2} + \ldots + \lambda^k \overrightarrow{pa_k} = \\ &= \lambda^2 (\overrightarrow{pa_2} - \overrightarrow{pa_1}) + \ldots + \lambda^k (\overrightarrow{pa_k} - \overrightarrow{pa_1}) = \\ &= \lambda^2 \overrightarrow{a_1 a_2} + \ldots + \lambda^k \overrightarrow{a_1 a_k} = \vec{0}. \end{aligned}$$

(a) nos dice entonces que $\lambda^2 = \ldots = \lambda^k = 0$ y, por tanto, también se tiene $\lambda^1 = -\lambda^2 - \ldots - \lambda^k = 0$.

(c) $\Rightarrow$ (a). En efecto, de

$$\sum_{h\geq 2} \lambda^h \overrightarrow{a_1 a_h} = \vec{0}$$

se obtiene

$$\sum_{h\geq 2} \lambda^h (\overrightarrow{pa_h} - \overrightarrow{pa_1}) = \left(-\sum_{h\geq 2} \lambda^h \right) \overrightarrow{pa_1} + \sum_{h\geq 2} \lambda^h \overrightarrow{pa_h} = \vec{0}.$$

En esta expresión, la suma de los coeficientes es 0 y, por tanto, (c) asegura que $\lambda^h = 0$ para todo $h \geq 2$. □

Diremos que los puntos $a_1, \ldots, a_k \in A$ son *linealmente independientes* si cumplen cualquiera de las condiciones de (5.1).

Observaciones:

1. Dos puntos son linealmente independientes si y sólo si son diferentes.
2. Si dim $A = n$, el número máximo de puntos linealmente independientes es $n + 1$.

Proposición 5.2 *Dados $a_1, \ldots, a_k$ puntos linealmente independientes de un espacio afín A de dimensión n, existen puntos $a_{k+1}, \ldots, a_{n+1}$ de A tales que $a_1, \ldots, a_k, \ldots, a_{n+1}$ son linealmente independientes.*

DEMOSTRACIÓN: Si $a_1, \ldots, a_k$ son linealmente independientes, entonces también lo son los vectores $\overrightarrow{a_1a_2}, \ldots, \overrightarrow{a_1a_k}$. Sean $u_k, \ldots, u_n$ vectores tales que

$$\overrightarrow{a_1a_2}, \ldots, \overrightarrow{a_1a_k}, u_k, \ldots, u_n$$

sea una base del espacio vectorial asociado a A. Consideremos puntos a_i tales que $\overrightarrow{a_1a_i} = u_{i-1}$, $i = k+1, \ldots, n+1$. Los puntos $a_1, \ldots, a_k, \ldots, a_{n+1}$ son claramente linealmente independientes. □

IX.6 Coordenadas baricéntricas

Sea (A, E) un espacio afín de dimensión n y sean $p_0, p_1, \ldots, p_n$ puntos linealmente independientes de A. Dado un $x \in A$, los puntos $p_0, p_1, \ldots, p_n, x$ no pueden ser linealmente independientes y, por tanto, existen $p \in A$ y elementos $k, k^0, \ldots, k^n \in K$, no todos cero, tales que

$$\left.\begin{array}{c} k\overrightarrow{px} + k^0\overrightarrow{pp_0} + \ldots + k^n\overrightarrow{pp_n} = \overrightarrow{0} \\ k + k^0 + \ldots + k^n = 0 \end{array}\right\}.$$

Si $k = 0$, estas igualdades implicarían que $p_0, \ldots, p_n$ fuesen linealmente dependientes, y no lo son por hipótesis. Por tanto, $k \neq 0$ y, poniendo $x^i = -k^i/k$, tenemos

$$\left.\begin{array}{c} \overrightarrow{px} = x^0\overrightarrow{pp_0} + \ldots + x^n\overrightarrow{pp_n} \\ x^0 + \ldots + x^n = 1 \end{array}\right\}.$$

Vamos a ver que estos $x^0, \ldots, x^n$ no dependen del punto p.

Proposición 6.1 *Sean $p_0, p_1, \ldots, p_k$ puntos de un espacio afín (A, E) de dimensión n. Supongamos que, dado $x \in A$, existen $p \in A$ y $x^0, \ldots, x^k \in K$ tales que*

$$\left.\begin{array}{c} \overrightarrow{px} = x^0\overrightarrow{pp_0} + \ldots + x^k\overrightarrow{pp_k} \\ x^0 + \ldots + x^k = 1 \end{array}\right\}.$$

Entonces

i) Estas expresiones también son ciertas si sustituimos p por cualquier otro $q \in A$.

ii) Los valores $x^0, \ldots, x^k$ están unívocamente determinados, siempre que $p_0, \ldots, p_k$ sean linealmente independientes.

DEMOSTRACIÓN:

$$\begin{aligned} \overrightarrow{qx} = \overrightarrow{qp} + \overrightarrow{px} &= (x^0 + \ldots + x^k)\overrightarrow{qp} + x^0\overrightarrow{pp_0} + \ldots + x^k\overrightarrow{pp_k} = \\ &= x^0(\overrightarrow{qp} + \overrightarrow{pp_0}) + \ldots + x^k(\overrightarrow{qp} + \overrightarrow{pp_k}) = \\ &= x^0\overrightarrow{qp_0} + \ldots + x^k\overrightarrow{qp_k}. \end{aligned}$$

Esto demuestra (i). Para probar (ii), aplicamos (i) al caso $q = p_0$; obtenemos

$$\overrightarrow{p_0x} = x^1\overrightarrow{p_0p_1} + \ldots + x^k\overrightarrow{p_0p_k}.$$

Pero los vectores $\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_k}$ son linealmente independientes y, por tanto, $x^1, \ldots, x^k$ están unívocamente determinados. De ahí resulta que $x^0 = 1 - x^1 - \ldots - x^k$ también está determinado. □

Llamaremos *sistema de referencia baricéntrico* o *sistema de coordenadas baricéntrico* de un espacio afín (A, E) de dimensión n a todo conjunto de $(n+1)$ puntos linealmente independientes $\{p_0, \ldots, p_n\}$. Llamaremos *coordenadas baricéntricas* de un punto $x \in A$ en el sistema $\{p_0, \ldots, p_n\}$ a la $(n+1)$-pla $(x^0, \ldots, x^n) \in K^{n+1}$ tal que

$$\left.\begin{array}{c} \overrightarrow{px} = x^0\overrightarrow{pp_0} + \ldots + x^n\overrightarrow{pp_n} \\ x^0 + \ldots + x^n = 1 \end{array}\right\}.$$

Por (6.1), esta $(n+1)$-pla es independiente del punto $p \in A$ y está bien determinada.

Notación:

Dados $p_0, \ldots, p_k$, si

$$\left.\begin{array}{c} \overrightarrow{px} = x^0\overrightarrow{pp_0} + \ldots + x^k\overrightarrow{pp_k} \\ x^0 + \ldots + x^k = 1 \end{array}\right\},$$

escribiremos

$$x = x^0 p_0 + \ldots + x^k p_k.$$

Tengamos en cuenta, sin embargo, que esta expresión no tiene ningún sentido si la suma de los coeficientes no es 1.

Se llama *baricentro* de m puntos $a_1, \ldots, a_m \in A$ a un punto $b = b^1 a_1 + \ldots + b^m a_m$, $\sum b^i = 1$, tal que $b^1 = \ldots = b^m$. Para que exista el baricentro de m puntos debe existir un elemento $d \in K$ tal que $\overbrace{d + \ldots + d}^{m} = 1$. Designaremos también por m el elemento de K que resulta de sumar m veces $1 \in K$:

$$m = \overbrace{1 + \ldots + 1}^{m} \in K.$$

Entonces

$$1 = \overbrace{d + \ldots + d}^{m} = d(\overbrace{1 + \ldots + 1}^{m}) = dm.$$

d existe sólo si $m \neq 0$ en K y entonces $d = m^{-1}$ (escribiremos también $d = 1/m$).

Nota:

El núcleo de la aplicación $f : \mathbf{Z} \longrightarrow K$ tal que $f(m) = m$ si $m > 0$, $f(m) = -f(-m)$ si $m < 0$ y $f(0) = 0$ es un ideal (p), $p \geq 0$. En esta situación, se dice que K es un *cuerpo de característica p*. Si $p \neq 0$, entonces p es el menor entero positivo que es 0 en K. Tiene que ser un número primo, ya que en caso contrario, si $p = rm$, tendríamos que en K $0 = rm$ con $r \neq 0$, $m \neq 0$. Los cuerpos $\mathbf{Q}$, $\mathbf{R}$ y $\mathbf{C}$ son de característica 0; para p primo, el cuerpo $\mathbf{Z}/(p)$ es de característica p.

Ejemplos:

1. En el sistema de referencia baricéntrico $\{p_0, \ldots, p_n\}$ las coordenadas baricéntricas del punto p_0 son $(1, 0, \ldots, 0)$; las de p_1 son $(0, 1, 0, \ldots, 0)$; ... ; las de p_n son $(0, \ldots, 0, 1)$.
2. Sea $(\mathbf{R}, \mathbf{R})$ la recta afín real. Sea $\{p_0, p_1\}$ una referencia baricéntrica de $\mathbf{R}$. Sea x un punto de coordenadas (x^0, x^1) (con $x^0 + x^1 = 1$) en este sistema.

   Entonces $\overrightarrow{px} = x^0 \overrightarrow{pp_0} + x^1 \overrightarrow{pp_1}$ para todo punto p, y, en particular, $\overrightarrow{p_0 x} = x^1 \overrightarrow{p_0 p_1}$.

   Esto nos permite conocer la posición de x respecto a los puntos p_0, p_1, según que sus coordenadas sean positivas o negativas (ver

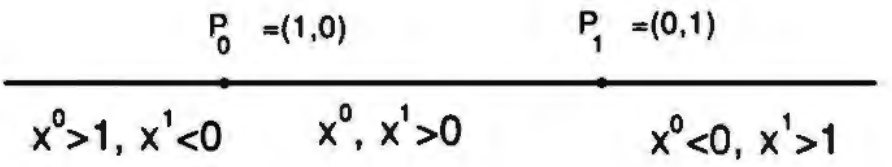


figura). En particular, el *segmento* determinado por p_0 y p_1 es el conjunto

$$\overline{p_0p_1} = \{x \in A \mid x = x^0p_0 + x^1p_1,\ x^0 + x^1 = 1,\ 0 \leq x^0, x^1 \leq 1\}.$$

Se llama *punto medio* del segmento $\overline{p_0p_1}$ al punto $\frac{1}{2}p_0 + \frac{1}{2}p_1$, es decir, al baricentro de $\{p_0, p_1\}$.

3. Sea $(\mathbf{R}^2, \mathbf{R}^2)$ el plano afín real y $\{p_0, p_1, p_2\}$ un sistema de referencia baricéntrico de $\mathbf{R}^2$. Sea x un punto, y (x^0, x^1, x^2) sus coordenadas en ese sistema:

$$\overrightarrow{px} = x^0\overrightarrow{pp_0} + x^1\overrightarrow{pp_1} + x^2\overrightarrow{pp_2}, \qquad x^0 + x^1 + x^2 = 1, \qquad \forall p.$$

Si x^0=1, $x^2 = -x^1$ y

$$\overrightarrow{px} = \overrightarrow{pp_0} + x^1(\overrightarrow{pp_1} - \overrightarrow{pp_2}) = \overrightarrow{pp_0} + x^1\overrightarrow{p_2p_1} \qquad \forall p.$$

Por tanto, $\overrightarrow{p_0x} = x^1\overrightarrow{p_2p_1}$; es decir, x está sobre la recta $p_0 + \langle\overrightarrow{p_2p_1}\rangle$ (ver figura).
Si $x^0 \neq 1$, resulta que, para todo p,

$$\overrightarrow{px} = x^0\overrightarrow{pp_0} + (1 - x^0)\left(\frac{x^1}{1 - x^0}\overrightarrow{pp_1} + \frac{x^2}{1 - x^0}\overrightarrow{pp_2}\right).$$

Sea q el punto tal que

$$\overrightarrow{pq} = \frac{x^1}{1 - x^0}\overrightarrow{pp_1} + \frac{x^2}{1 - x^0}\overrightarrow{pp_2}.$$

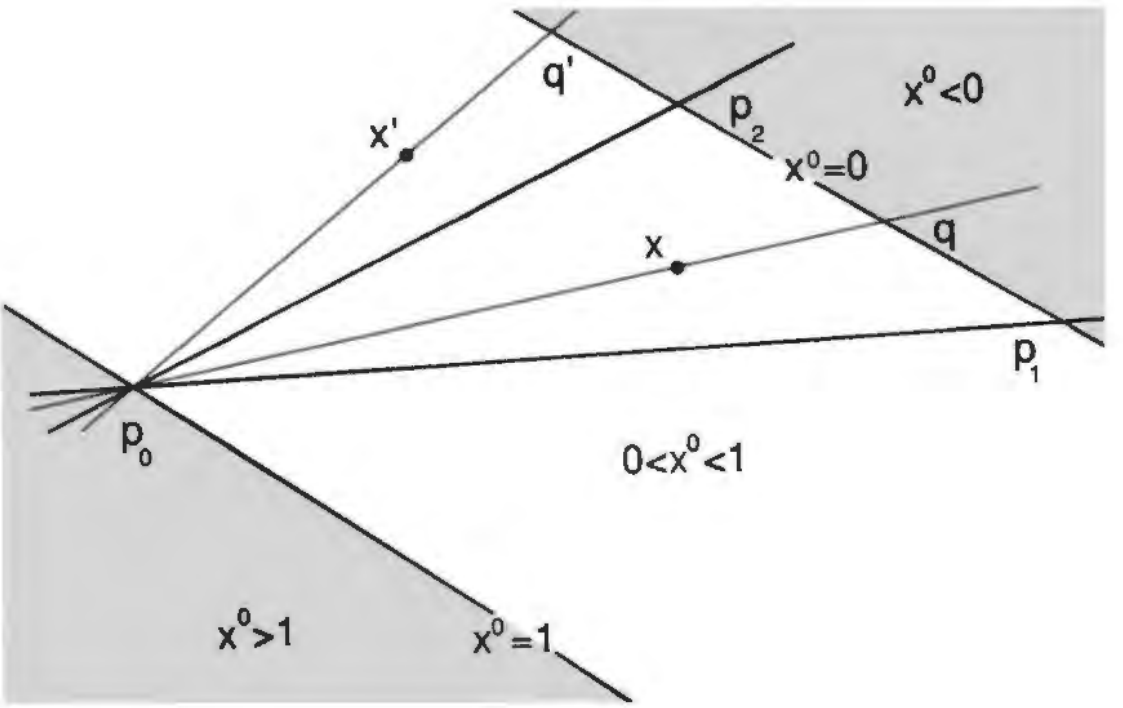
La suma de los coeficientes de esta expresión es 1 y, por tanto, de (6.1) resulta que

$$\overrightarrow{p_1 q} = \frac{x^2}{1 - x^0} \overrightarrow{p_1 p_2}$$

y q es un punto de la recta $p_1 + \langle \overrightarrow{p_1 p_2} \rangle$. Si aplicamos ahora los resultados del ejemplo anterior a

$$\overrightarrow{px} = x^0 \overrightarrow{pp_0} + (1 - x^0) \overrightarrow{pq},$$

podemos deducir la posición de x según que $x^0 < 0$, $0 \leq x^0 \leq 1$ o $x^0 > 1$, tal como está indicado en la figura. Naturalmente, podemos obtener resultados análogos para las otras dos coordenadas y, de esta manera, determinar la región del plano donde se encuentra el punto x, según los signos de sus coordenadas.



Ejercicio:

Generalizar el estudio llevado a cabo en el ejemplo 2 al espacio afín $\mathbf{R}^3$, estudiando la región donde se encuentra un punto x según los signos de sus coordenadas baricéntricas respecto a un sistema de referencia $\{p_0, p_1, p_2, p_3\}$.

Acabaremos este apartado estudiando cómo cambian las coordenadas baricéntricas de un punto al cambiar el sistema de referencia. Sean $\{p_0,\dots,p_n\}$ $\{q_0,\dots,q_n\}$ dos sistemas de referencia baricéntricos de un espacio afín A. Sean $(q_i^0,\dots,q_i^n)$ las coordenadas baricéntricas del punto q_i en el sistema $\{p_0,\dots,p_n\}$, $i=0,\dots,n$. Dado $x\in A$, indiquemos por $(x^0,\dots,x^n)$ y $(\bar{x}^0,\dots,\bar{x}^n)$ sus coordenadas baricéntricas en los sistemas $\{p_0,\dots,p_n\}$ y $\{q_0,\dots,q_n\}$ respectivamente. Entonces, para cada p,

$$\begin{aligned}\overrightarrow{px} &= \textstyle\sum_{i=0}^n \bar{x}^i\overrightarrow{pq_i} = \sum_{i=0}^n \bar{x}^i\left(\sum_{j=0}^n q_i^j\overrightarrow{pp_j}\right) = \\ &= \textstyle\sum_{j=0}^n\left(\sum_{i=0}^n \bar{x}^i q_i^j\right)\overrightarrow{pp_j},\end{aligned}$$

con suma de coeficientes

$$\sum_{j=0}^n\left(\sum_{i=0}^n \bar{x}^i q_i^j\right) = \sum_{i=0}^n \bar{x}^i\left(\sum_{j=0}^n q_i^j\right) = \sum_{i=0}^n \bar{x}^i = 1.$$

Por tanto,

$$\sum_{i=0}^n q_i^j\,\bar{x}^i = x^j, \qquad j=0,\dots,n.$$

Estas $(n+1)$ expresiones equivalen a la igualdad matricial

$$Q\bar{x} = x,$$

donde $Q=\left(q_i^j\right)$ es la matriz que tiene por columnas las coordenadas baricéntricas de $q_0,\dots,q_n$, y $\bar{x}$, x son las matrices de una columna formadas por las coordenadas de x en los sistemas $\{q_0,\dots,q_n\}$ y $\{p_0,\dots,p_n\}$, respectivamente.

Observación:

La matriz Q es invertible, ya que

$$\det Q = \begin{vmatrix} q_0^0 & \dots & q_n^0 \\ \vdots & & \vdots \\ q_0^n & \dots & q_n^n \end{vmatrix} =$$

(sumando todas las filas a la primera)

$$= \begin{vmatrix} 1 & \dots & 1 \\ q_0^1 & \dots & q_n^1 \\ \vdots & & \vdots \\ q_0^n & \dots & q_n^n \end{vmatrix} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ q_0^1 & q_1^1-q_0^1 & \dots & q_n^1-q_0^1 \\ \vdots & \vdots & & \vdots \\ q_0^n & q_1^n-q_0^n & \dots & q_n^n-q_0^n \end{vmatrix} =$$

$$= \det_{(\overrightarrow{p_0p_i})}(\overrightarrow{q_0q_1},\dots,\overrightarrow{q_0q_n}) \neq 0.$$

IX.7 Ecuaciones de una variedad en coordenadas baricéntricas

Sea (A, E) un espacio afín de dimensión n y $\{p_0, \ldots, p_n\}$ un sistema de referencia baricéntrico de A. Dados k puntos $a_1, \ldots, a_k$, sus coordenadas baricéntricas $(a_i^0, \ldots, a_i^n)$ cumplen

$$\operatorname{rang}\begin{pmatrix} a_1^0 & \ldots & a_k^0 \\ \vdots & & \vdots \\ a_1^n & \ldots & a_k^n \end{pmatrix} = \operatorname{rang}\begin{pmatrix} 1 & \ldots & 1 \\ a_1^1 & \ldots & a_k^1 \\ \vdots & & \vdots \\ a_1^n & \ldots & a_k^n \end{pmatrix} =$$

$$= \operatorname{rang}\begin{pmatrix} 1 & 0 & \ldots & 0 \\ a_1^1 & a_2^1 - a_1^1 & \ldots & a_k^1 - a_1^1 \\ \vdots & \vdots & & \vdots \\ a_1^n & a_2^n - a_1^n & \ldots & a_k^n - a_1^n \end{pmatrix} =$$

$$= \dim\langle \overrightarrow{a_1a_2}, \ldots, \overrightarrow{a_1a_k} \rangle + 1.$$

En particular, $a_1, \ldots, a_k$ son linealmente independientes si y sólo si este rango es k.

Designemos por L la variedad lineal generada por $a_1, \ldots, a_k$ (3.3). Tenemos

$$L = \{x \in A \mid \overrightarrow{a_1x} = \lambda^2\overrightarrow{a_1a_2} + \ldots + \lambda^k\overrightarrow{a_1a_k}\}.$$

Ahora bien,

$$\begin{aligned} \overrightarrow{a_1x} = & \ \lambda^2\overrightarrow{a_1a_2} + \ldots + \lambda^k\overrightarrow{a_1a_k} = \\ = & \ (1 - \lambda^2 - \ldots - \lambda^k)\overrightarrow{a_1a_1} + \lambda^2\overrightarrow{a_1a_2} + \ldots + \lambda^k\overrightarrow{a_1a_k} \end{aligned}$$

y, por (6.1),

$$\overrightarrow{px} = \lambda^1\overrightarrow{pa_1} + \ldots + \lambda^k\overrightarrow{pa_k} \qquad \forall p,$$

donde $\lambda^1 = 1 - \lambda^2 - \ldots - \lambda^k$. Es decir,

$$L = \{x \in A \mid x = \lambda^1 a_1 + \ldots + \lambda^k a_k,\ \lambda^1 + \ldots + \lambda^k = 1\}.$$

De ahí resulta inmediatamente que

$$x^i = \lambda^1 a_1^i + \ldots + \lambda^k a_k^i, \qquad i = 0, \ldots, n,$$

donde $(x^i), (a_1^i), \ldots, (a_k^i)$ son las coordenadas baricéntricas de $x, a_1, \ldots, a_k$, respectivamente, en un cierto sistema de referencia $\{p_0, \ldots, p_n\}$.

Supongamos que $a_1, \ldots, a_k$ son linealmente independientes y, por tanto, dim $L = k-1$ (3.3). Entonces $x \in L$ si y sólo si $x, a_1, \ldots, a_k$ son linealmente dependientes, es decir, si

$$k = \operatorname{rang} \begin{pmatrix} a_1^0 & \ldots & a_k^0 \\ \vdots & & \vdots \\ a_1^n & \ldots & a_k^n \end{pmatrix} = \operatorname{rang} \begin{pmatrix} a_1^0 & \ldots & a_k^0 & x^0 \\ \vdots & & \vdots & \vdots \\ a_1^n & \ldots & a_k^n & x^n \end{pmatrix}.$$

Fijemos un menor de orden k de la primera matriz, que tenga determinante diferente de cero. Por comodidad supondremos que está formado por las k primeras filas. Entonces una $(n+1)$-pla $(x^0, \ldots, x^n)$ es el conjunto de coordenadas baricéntricas de un punto $x \in L$ si y sólo si cumple:

$$\begin{vmatrix} a_1^0 & \ldots & a_k^0 & x^0 \\ \vdots & & \vdots & \vdots \\ a_1^{k-1} & \ldots & a_k^{k-1} & x^{k-1} \\ a_1^i & \ldots & a_k^i & x^i \end{vmatrix} = 0, \qquad i = k, \ldots, n$$

$$\textstyle\sum_{j=0}^{n} x^j = 1.$$

Estas son las *ecuaciones baricéntricas* de la variedad lineal L.

Ejemplo:

Las ecuaciones del hiperplano determinado por los puntos del sistema de referencia $p_0, \ldots, p_n$ salvo uno de ellos, p_i, son

$$\left.\begin{aligned} x^i &= 0 \\ x^0 + \ldots + x^n &= 1 \end{aligned}\right\}.$$

IX.8 Coordenadas cartesianas

Fijemos un punto p del espacio afín (A, E). Sabemos que la aplicación

$$\begin{aligned} \varphi_p : \ A &\longrightarrow E \\ q &\longmapsto \overrightarrow{pq} \end{aligned}$$

es biyectiva y, por tanto, fijada una base $e_1, \ldots, e_n$ de E, el punto q queda completamente determinado por las coordenadas de $\overrightarrow{pq}$ en esta base

$$\overrightarrow{pq} = q^1 e_1 + \ldots + q^n e_n.$$

Llamaremos *sistema de referencia cartesiano* o *sistema de coordenadas cartesiano* de (A, E) al conjunto $\{p;\ e_1, \ldots, e_n\}$ formado por un punto $p \in A$ y una base $e_1, \ldots, e_n$ de E. Llamaremos *coordenadas cartesianas* de un punto q en este sistema a las coordenadas de $\overrightarrow{pq}$ en la base $e_1, \ldots, e_n$.

Ejemplo:

Las coordenadas de p en el sistema $\{p;\ e_1, \ldots, e_n\}$ son $(0, \ldots, 0)$.

Observación:

Las coordenadas de un vector $\overrightarrow{ab}$ en la base $e_1, \ldots, e_n$ son las coordenadas de $\overrightarrow{ab} = \overrightarrow{pb} - \overrightarrow{pa}$ y, por tanto, son la diferencia de las coordenadas cartesianas de b y a en el sistema $\{p;\ e_1, \ldots, e_n\}$. Este hecho justifica la notación introducida en el §3:

$$\overrightarrow{ab} = b - a, \qquad b = a + \overrightarrow{ab}.$$

Consideremos ahora dos sistemas de referencia distintos $\{p;\ e_1, \ldots, e_n\}$, $\{q;\ v_1, \ldots, v_n\}$ y supongamos conocido el segundo en función del primero:

$$\overrightarrow{pq} = q^1 e_1 + \ldots + q^n e_n, \qquad v_i = \sum_{j=1}^{n} v_i^j e_j, \qquad i = 1, \ldots, n.$$

Dado $x \in A$, queremos estudiar la relación entre sus coordenadas en un sistema y en el otro.

$$\overrightarrow{px} = x^1 e_1 + \ldots + x^n e_n, \qquad \overrightarrow{qx} = \bar{x}^1 v_1 + \ldots + \bar{x}^n v_n.$$

Tenemos

$$\left.\begin{array}{rcl} \overrightarrow{qx} & = & \sum_{i=1}^n \bar{x}^i v_i = \sum_{i=1}^n \bar{x}^i \left(\sum_{j=1}^n v_i^j e_j\right) = \\ & = & \sum_{j=1}^n \left(\sum_{i=1}^n v_i^j \bar{x}^i\right) e_j \\ \overrightarrow{qx} & = & \overrightarrow{px} - \overrightarrow{pq} = \sum_{j=1}^n (x^j - q^j) e_j \end{array}\right\},$$

de donde $x^j = q^j + \sum_{i=1}^n v_i^j \bar{x}^i$, $j = 1, \ldots, n$.

Estas ecuaciones se pueden expresar matricialmente así: sea $V = \left(v_i^j\right)$ y x, $\bar{x}$, q las matrices de una columna formadas por las coordenadas de x, $\bar{x}$ y q, respectivamente. Entonces

$$x = q + V\bar{x};$$

o también

$$\begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{x} \\ 1 \end{pmatrix},$$

con la notación obvia.

IX.9 Ecuaciones de una variedad en coordenadas cartesianas

Sea $L = a + \langle v_1, \ldots, v_k \rangle$ una variedad lineal de dimensión k del espacio afín (A, E). Fijado un sistema de referencia cartesiano $\{p;\ e_1, \ldots, e_n\}$, nos proponemos encontrar las condiciones que deben cumplir las coordenadas de un punto x para que sea de la variedad L. Tenemos

$$x \in a + \langle v_1, \ldots, v_k \rangle \Leftrightarrow \overrightarrow{ax} \in \langle v_1, \ldots, v_k \rangle \Leftrightarrow$$

$$\Leftrightarrow \operatorname{rang} \begin{pmatrix} x^1 - a^1 & v_1^1 & \ldots & v_k^1 \\ \vdots & \vdots & & \vdots \\ x^n - a^n & v_1^n & \ldots & v_k^n \end{pmatrix} = \operatorname{rang} \begin{pmatrix} v_1^1 & \ldots & v_k^1 \\ \vdots & & \vdots \\ v_1^n & \ldots & v_k^n \end{pmatrix} = k.$$

De la segunda matriz escogemos un menor de orden k con determinante diferente de cero. Para simplificar la notación, supondremos que está formado por las k primeras filas:

$$\begin{vmatrix} v_1^1 & \ldots & v_k^1 \\ \vdots & & \vdots \\ v_1^k & \ldots & v_k^k \end{vmatrix} \neq 0.$$

Entonces, las condiciones que deben cumplir las coordenadas $(x^1, \ldots, x^n)$ de $x \in L$ son

$$\begin{vmatrix} x^1 - a^1 & v_1^1 & \ldots & v_k^1 \\ \vdots & \vdots & & \vdots \\ x^k - a^k & v_1^k & \ldots & v_k^k \\ x^i - a^i & v_1^i & \ldots & v_k^i \end{vmatrix} = 0, \qquad i = k+1, \ldots, n.$$

Observemos que estos determinantes forman un sistema de $n - k$ ecuaciones con n incógnitas.

Recíprocamente, sea

$$\left.\begin{aligned} a_1^1 x^1 + \ldots + a_n^1 x^n &= b^1 \\ \ldots\ldots\ldots\ldots & \\ a_1^d x^1 + \ldots + a_n^d x^n &= b^d \end{aligned}\right\}$$

un sistema de ecuaciones lineales cualquiera. Las soluciones del sistema homogéneo asociado, interpretadas como coordenadas de vectores de E en la base $e_1, \ldots, e_n$, forman un subespacio vectorial F de E. Sea $(a^1, \ldots, a^n)$ una solución particular del sistema dado y sea a el punto con esas coordenadas respecto al sistema de referencia $\{p;\ e_1, \ldots, e_n\}$. La solución general del sistema es la suma de $(a^1, \ldots, a^n)$ y las soluciones del sistema homogéneo; es decir, el conjunto de coordenadas de los puntos de la variedad lineal

$$a + F.$$

En estas circunstancias diremos que el sistema dado son las *ecuaciones de la variedad* $a + F$ en el sistema de referencia cartesiano $\{p;\ e_1, \ldots, e_n\}$. Observemos que la dimensión de esta variedad es $n - \operatorname{rang}\left(a_i^j\right)$.

Ejemplos:

1. Una ecuación $a_1 x^1 + \ldots + a_n x^n = b$ no trivial (es decir, con no todos los coeficientes a_i cero) representa un hiperplano.
   En general, si

$$\left.\begin{array}{rcl} a_1^1 x^1 + \ldots + a_n^1 x^n & = & b^1 \\ \ldots\ldots\ldots & & \\ a_1^d x^1 + \ldots + a_n^d x^n & = & b^d \end{array}\right\}$$

   son las ecuaciones de una variedad L, cada una de las ecuaciones representa un hiperplano H_i y la variedad L resulta ser la intersección de estos d hiperplanos. La ecuación de cualquier otro hiperplano que contenga a L será tal que, al ser añadida al sistema, éste tendrá las mismas soluciones; es decir, será una combinación lineal de las ecuaciones del sistema:

$$\lambda^1(a_1^1 x^1 + \ldots + a_n^1 x^n) + \ldots + \lambda^d(a_1^d x^1 + \ldots + a_n^d x^n) =$$
$$= \lambda^1 b^1 + \ldots + \lambda^d b^d.$$

   Este conjunto de hiperplanos se llama el *haz de hiperplanos* que pasan por L.
2. Las ecuaciones de una recta $a + (v)$ son

$$\begin{vmatrix} x^i - a^i & v^i \\ x^j - a^j & v^j \end{vmatrix} = 0 \qquad \forall i, j.$$

   Es decir, $(x^i - a^i)v^j = (x^j - a^j)v^i$ para todo i, j. Estas ecuaciones se pueden escribir como sigue:

$$\frac{x^1 - a^1}{v^1} = \ldots = \frac{x^n - a^n}{v^n},$$

con el convenio de que, si algún $v^j = 0$, el cociente correspondiente tiene numerador cero: $x^j - a^j = 0$.

El estudio de los sistemas de ecuaciones de dos variedades permite estudiar fácilmente cuestiones como intersección o paralelismo. Vamos a ver dos casos de paralelismo particularmente sencillos.

I. Dos hiperplanos

$$\left.\begin{array}{rcl} a_1x^1 + \ldots + a_nx^n & = & b \\ \bar{a}_1x^1 + \ldots + \bar{a}_nx^n & = & \bar{b} \end{array}\right\}$$

son paralelos si y sólo si tienen la misma dirección; es decir, si las ecuaciones $a_1x^1 + \ldots + a_nx^n = 0$, $\bar{a}_1x^1 + \ldots + \bar{a}_nx^n = 0$ tienen las mismas soluciones. Esto ocurre si los coeficientes son proporcionales (esto es, si existe c tal que $a_i = c\bar{a}_i$, $i = 1, \ldots, n$). Si, además, $b = c\bar{b}$, entonces los dos hiperplanos coinciden; en caso contrario, si $b \neq c\bar{b}$, no se cortan. (Comparar con (4.2).)

II. Una recta y un hiperplano de ecuaciones

$$\frac{x^1 - a^1}{v^1} = \ldots = \frac{x^n - a^n}{v^n}, \qquad b_1x^1 + \ldots + b_nx^n = b$$

son paralelos si la dirección de la recta, $\langle (v^1, \ldots, v^n) \rangle$, está contenida en la dirección del plano; es decir, si

$$b_1v^1 + \ldots + b_nv^n = 0.$$

IX.10 Razón simple

Dados tres puntos alineados a_1, a_2, a_3 de un espacio afín (A, E), se denomina *razón simple* de a_1, a_2, a_3, y se escribe $(a_1a_2a_3)$, al elemento $r \in K$ tal que

$$\overrightarrow{a_1a_3} = r\overrightarrow{a_1a_2}.$$

$(a_1a_2a_3)$ está definido siempre que $a_1 \neq a_2$, es 0 si $a_1 = a_3$ y es 1 si $a_2 = a_3$.

Proposición 10.1 *Si $a_2 \neq a_3$ y (α, β) son las coordenadas baricéntricas de a_1 en el sistema de referencia de la recta $\{a_2, a_3\}$, entonces*

$$(a_1a_2a_3) = -\frac{\alpha}{\beta}.$$

DEMOSTRACIÓN: Para cada p,

$$\overrightarrow{pa_1} = \alpha\overrightarrow{pa_2} + \beta\overrightarrow{pa_3} \text{ con } \alpha + \beta = 1.$$

En particular, $0 = \alpha\overrightarrow{a_1a_2} + \beta\overrightarrow{a_1a_3}$, de donde $\overrightarrow{a_1a_3} = -\frac{\alpha}{\beta}\overrightarrow{a_1a_2}$. □

Corolario 10.2 *Supongamos $K = \mathbf{R}$. a_1 es del segmento $\overline{a_2a_3}$ si y sólo si*

$$(a_1a_2a_3) < 0.$$

DEMOSTRACIÓN: $a_1 \in \overline{a_2a_3} \Leftrightarrow \alpha, \beta > 0$ (§6) $\Leftrightarrow (a_1a_2a_3) < 0$. □

Ejemplos:

1. Consideremos la recta afín (K, K) y sean $x_1, x_2, x_3 \in K$ tres puntos cualesquiera. Recordemos (§1) que $\overrightarrow{x_ix_j} = x_j - x_i \in K$. Por tanto, si $x_1 \neq x_2$,

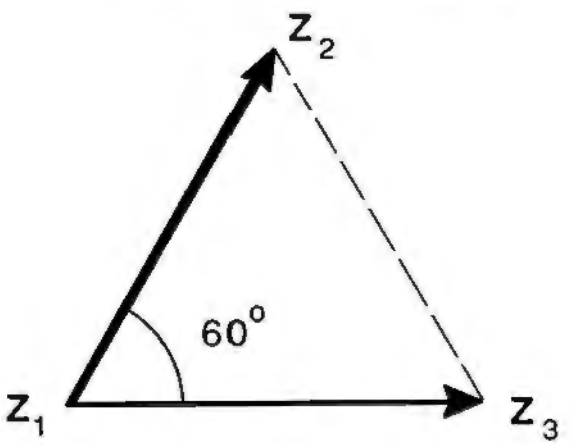
$$(x_1x_2x_3) = \frac{x_3 - x_1}{x_2 - x_1}.$$

2. En el caso particular $K = \mathbf{C}$, dados $z_1, z_2, z_3 \in \mathbf{C}$ con $z_1 \neq z_2$,

$$(z_1z_2z_3) = \frac{z_3 - z_1}{z_2 - z_1}.$$

Observemos que el módulo de la razón es el cociente de los módulos de $z_3 - z_1$ y $z_2 - z_1$, y el argumento del complejo $(z_1z_2z_3)$ es la diferencia

$$\arg(z_3 - z_1) - \arg(z_2 - z_1).$$

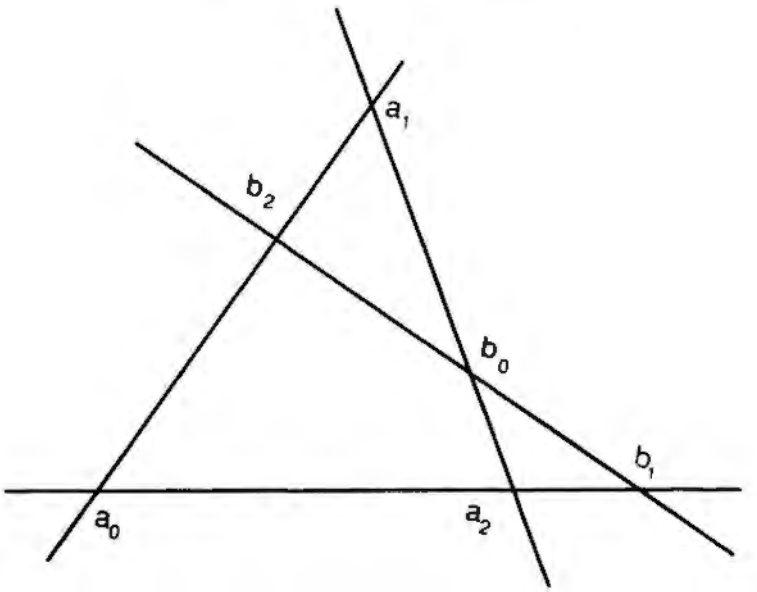


Así, por ejemplo, z_1, z_2, z_3 forman un triángulo equilátero si y sólo si

$$(z_1 z_2 z_3) = 1_{\pm 60^\circ}$$

Teorema 10.3 (de Menelao) *Sean a_0, a_1, a_2 puntos linealmente independientes de un espacio afín (A, E) y sean b_0, b_1, b_2 puntos de las rectas determinadas por a_1a_2, a_0a_2 y a_0a_1 respectivamente, alineados y diferentes de a_0, a_1, a_2. Entonces*

$$(b_0a_1a_2) \cdot (b_1a_2a_0) \cdot (b_2a_0a_1) = 1.$$



DEMOSTRACIÓN: Las coordenadas baricéntricas de los puntos b_0, b_1, b_2 en el sistema $\{a_0, a_1, a_2\}$ son del tipo

$$b_0 = (0, \alpha_0, \beta_0), \qquad b_1 = (\alpha_1, 0, \beta_1), \qquad b_2 = (\alpha_2, \beta_2, 0).$$

Entonces (10.1) nos dice que

$$(b_0a_1a_2) = -\frac{\alpha_0}{\beta_0}, \qquad (b_1a_2a_0) = -\frac{\beta_1}{\alpha_1}, \qquad (b_2a_0a_1) = -\frac{\alpha_2}{\beta_2};$$

de donde

$$(b_0a_1a_2) \cdot (b_1a_2a_0) \cdot (b_2a_0a_1) = -\frac{\alpha_0\beta_1\alpha_2}{\beta_0\alpha_1\beta_2}.$$

Pero, por otra parte, vimos en el §7 que

$$b_0, b_1, b_2 \text{ alineados} \quad \Leftrightarrow \quad 0 = \begin{vmatrix} 0 & \alpha_1 & \alpha_2 \\ \alpha_0 & 0 & \beta_2 \\ \beta_0 & \beta_1 & 0 \end{vmatrix} = \alpha_0\beta_1\alpha_2 + \alpha_1\beta_2\beta_0 \Leftrightarrow$$

$$\Leftrightarrow \quad -\frac{\alpha_0\beta_1\alpha_2}{\alpha_1\beta_2\beta_0} = 1.$$

Esto demuestra el teorema. □

Corolario 10.4 *Supongamos $K = \mathbf{R}$. Con las notaciones del teorema anterior, no es posible que*

$$b_0 \in \overline{a_1a_2}, \qquad b_1 \in \overline{a_2a_0}, \qquad b_3 \in \overline{a_0a_1}$$

simultáneamente. En otras palabras, una recta real no puede cortar simultáneamente los tres "lados" de un triángulo.

DEMOSTRACIÓN: Es consecuencia de (10.3) y (10.2). □

Teorema 10.5 (de Ceva) *Sean a_0, a_1, a_2 puntos linealmente independientes de un espacio afín (A, E) de dimensión 2. Dado $p \in A$ diferente de a_0, a_1, a_2, designemos por b_i $(i = 0, 1, 2)$ la intersección de la recta determinada por pa_i con la recta determinada por los otros dos puntos a_ka_h $(k, h \neq i)$. Entonces*

$$(b_0a_1a_2) \cdot (b_1a_2a_0) \cdot (b_2a_0a_1) = -1.$$

DEMOSTRACIÓN: Si (α, β, γ) son las coordenadas baricéntricas de p en el sistema $\{a_0, a_1, a_2\}$, resulta que

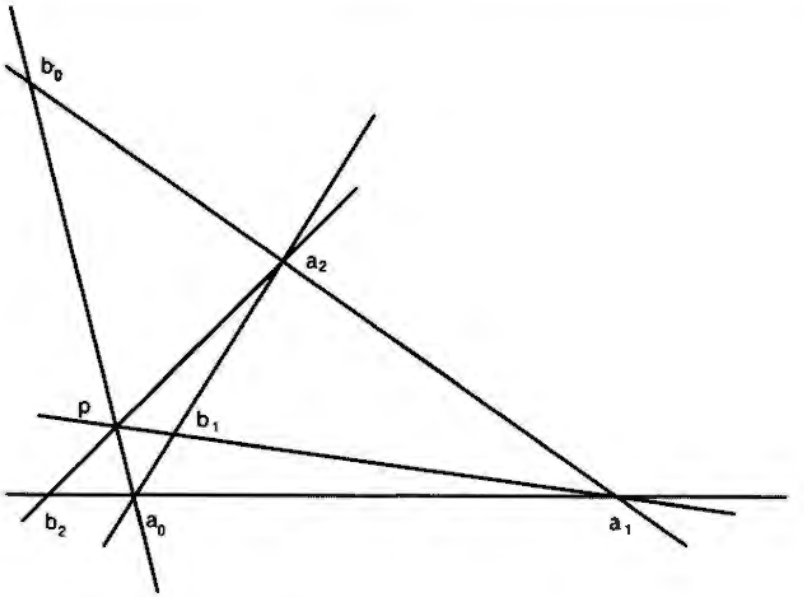
b_0 tiene por coordenadas $(0, \frac{\beta}{1-\alpha}, \frac{\gamma}{1-\alpha})$, de donde $(b_0a_1a_2) = -\frac{\beta}{\gamma}$;

b_1 tiene por coordenadas $(\frac{\alpha}{1-\beta}, 0, \frac{\gamma}{1-\beta})$, de donde $(b_1a_2a_0) = -\frac{\gamma}{\alpha}$;

b_2 tiene por coordenadas $(\frac{\alpha}{1-\gamma}, \frac{\beta}{1-\gamma}, 0)$, de donde $(b_2a_0a_1) = -\frac{\alpha}{\beta}$.

Por tanto,

$$(b_0a_1a_2) \cdot (b_1a_2a_0) \cdot (b_2a_0a_1) = -\frac{\beta\gamma\alpha}{\gamma\alpha\beta} = -1. \; \square$$



IX.11 Orientación de un espacio afín real

Sea E un espacio vectorial de dimensión n sobre $\mathbf{R}$. Diremos que dos bases $\{e_1, \ldots, e_n\}$ y $\{u_1, \ldots, u_n\}$ son *de la misma orientación* si

$$\det_{(e_i)}(u_1, \ldots, u_n) > 0.$$

En caso contrario, diremos que son *de orientaciones opuestas.* "Tener la misma orientación" es una relación de equivalencia en el conjunto de bases de E, y da lugar a dos clases de equivalencia que denominaremos *orientaciones* de E. *Orientar el espacio vectorial* E es escoger una de esas dos orientaciones. La orientación escogida se llama *orientación positiva* y la otra *orientación negativa.*

Ejemplo:

Si escogemos como orientación positiva la de e_1, e_2, e_3, entonces e_2, e_1, e_3 es de orientación negativa, pero e_2, e_3, e_1 es de orientación positiva.

Un automorfismo $f : E \longrightarrow E$ *conserva la orientación* si las bases $e_1, \ldots, e_n$ y $f(e_1), \ldots, f(e_n)$ son de la misma orientación: es decir, si

$$\det_{(e_i)}(f(e_1), \ldots, f(e_n)) = \det f > 0.$$

La definición anterior es, pues, independiente de la base $e_1, \ldots, e_n$.

Si $\det f < 0$, diremos que f *invierte la orientación.*

Orientar un espacio afín (A, E) es escoger una orientación de E. *Orientar una variedad lineal* $a + F$ es escoger una orientación de F.

IX.12 Semiespacios

En un espacio afín real (A, E) de dimensión n, cada hiperplano H divide al resto de puntos de A en dos zonas de la siguiente manera: consideremos en $A - H$ la relación

$$p \sim q \Leftrightarrow \text{ el segmento } \overline{pq} \text{ no corta a } H.$$

Las propiedades reflexiva ($p \sim p \quad \forall p \in A - H$) y simétrica ($p \sim q \Rightarrow q \sim p$) de esta relación son obvias. Antes de demostrar la propiedad transitiva, vamos a caracterizar la relación $p \sim q$ de otra manera. Sea $a_1x^1 + \ldots + a_nx^n + b = 0$ la ecuación del hiperplano H en un cierto sistema de referencia cartesiano. Para todo punto $y \in A - H$ de coordenadas $(y^1, \ldots, y^n)$, designaremos por a_y el número real

$$a_y = a_1y^1 + \ldots + a_ny^n + b \neq 0.$$

Si el segmento

$$\overline{pq} = \{x \in A,\ x^i = (1-t)p^i + tq^i,\ i = 1, \ldots, n,\ 0 \leq t \leq 1\}$$

corta a H, existe un valor de t, $0 < t < 1$, tal que

$$a_1\left((1-t)p^1 + tq^1\right) + \ldots + a_n\left((1-t)p^n + tq^n\right) + b = 0.$$

Podemos escribir esta expresión así:

$$(1-t)a_p + ta_q = 0,$$

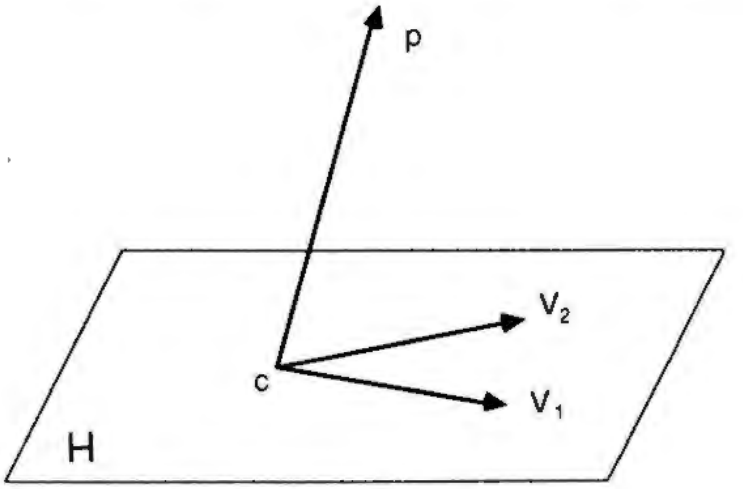
de donde $0 < t = \dfrac{a_p}{a_p - a_q} < 1$, y esto es cierto si y sólo si a_p y a_q son uno positivo y el otro negativo. Es decir,

$$p \sim q \Leftrightarrow a_p \cdot a_q > 0.$$

De ahí resulta inmediatamente que esta relación es de equivalencia y que divide a $A - H$ en dos subconjuntos que denominaremos *semiespacios* de A.

La noción de semiespacio está relacionada con el concepto de orientación definido en el §11. En efecto, sea $v_1, \ldots, v_{n-1}$ una base de la dirección de H y sea $c \in H$. Pongamos

$$a_x = \begin{vmatrix} x^1 - c^1 & v_1^1 & \ldots & v_{n-1}^1 \\ \vdots & \vdots & & \vdots \\ x^n - c^n & v_1^n & \ldots & v_{n-1}^n \end{vmatrix} = \det(\overrightarrow{cx}, v_1, \ldots, v_{n-1}).$$



$a_x = 0$ es una ecuación de H y, si $p \in A - H$,

$$a_p = \det(\overrightarrow{cp}, v_1, \ldots, v_{n-1}) \neq 0.$$

Dos puntos p, q son, por tanto, del mismo semiespacio si y sólo si las bases

$$\{\overrightarrow{cp}, v_1, \ldots, v_{n-1}\} \text{ y } \{\overrightarrow{cq}, v_1, \ldots, v_{n-1}\}$$

son de la misma orientación.

IX.13 Nota histórica

Pierre de Fermat (1601–1665) y René Descartes (1596–1650), obsesionados por la necesidad de métodos cuantitativos en la geometría e impresionados por el poder del álgebra, iniciaron la aplicabilidad de ésta al estudio de la geometría, creando los sistemas de coordenadas al asociar ecuaciones algebraicas a curvas y superficies. Esta idea ha sido una de las más ricas y fructíferas en el desarrollo de las matemáticas.

Tanto Fermat como Descartes estaban motivados por las necesidades de la ciencia y por un interés en la metodología. Especialmente Descartes (el primer gran filósofo moderno, un fundador de la biología moderna, un físico de categoría y matemático sólo incidentalmente) hizo de la metodología el objetivo principal de toda su obra.

Las primeras nociones nebulosas de un hiperespacio de dimensión $n > 3$ se pierden en la oscuridad del pasado y se mezclan con consideraciones metafísicas. El primer artículo científico que trata explícitamente del tema se debe a Arthur Cayley (1821–1895) y se remonta a 1843. Le siguen una serie de trabajos del mismo autor, de James Joseph Sylvester (1814–1897) y de William Kingdom Clifford (1845–1879) en Inglaterra y de Hermann Günther Grassmann (1809–1877) y Ludwig Schläfli (1814–1895) en el continente. La

introducción de coordenadas se lleva a cabo durante la segunda mitad del siglo 19 a través del estudio de los espacios aritméticos.

En 1818, August Ferdinand Möbius (1790–1868) ya había tenido la idea de un análisis geométrico en los espacios de dimensión 2 y 3, que desarrolló a partir de 1823 bajo el nombre de "cálculo baricéntrico", inspirado en la teoría de centros de gravedad. Eso es lo que hoy llamamos un sistema de coordenadas baricéntricas. No es, sin embargo, hasta finales de siglo que Schläfli y Camille Jordan (1838–1922) desarrollan explícitamente las nociones de la geometría afín (y de la euclídea) de dimensión n. La linealización de la geometría es un hecho.

IX.14 Ejercicios

1. En plena guerra, el servicio de contraespionaje de los buenos intercepta un mensaje de los malos que dice: "El día 23 del próximo mes de febrero, a las 6.25 horas p.m. iniciaremos un ataque contra los buenos. Comenzará con el fuego intenso de una batería de artillería de campaña que tenemos situada en el punto $(2.3, -5)$. En el punto $(-11, 0.3)$ tenemos preparada una división de infantería dispuesta a entrar en combate después del bombardeo inicial y, finalmente, en el punto $(9, 7.1)$ tenemos una batería de artillería antiaérea para defender las unidades antes citadas del ataque de los aviones de los buenos. Dirigirá la operación desde el origen de coordenadas el general Bum-Bum."

El jefe de defensa de los buenos ordena urgentemente que salgan patrullas de reconocimiento para localizar las unidades enemigas. Esas patrullas descubren que la batería de artillería de campaña está situada en el punto $(15, 7.5)$, la división de infantería está acampada en el punto $(19, -2.7)$ y la batería de artillería antiaérea se ha aposentado en el punto $(-14, 9.2)$. Todos estos datos según el sistema de referencia de los buenos, por supuesto. Después de recibir esa información, el general bueno pide que le traigan un matemático y le pregunta si es posible descubrir cuál es el origen de coordenadas de los malos, al objeto de cargarse al general Bum-Bum y desmoralizar así a los malos, que muy probablemente ya no llevarían a cabo la ofensiva. El matemático se acordaba de que cuando cursaba primero había resuelto un problema parecido en la clase de Álgebra y Geometría y en un santiamén le calculó el tan deseado punto. ¿Cómo lo hizo?

2. a) Demostrar que en $\mathbf{R}^2$ los puntos medios de cualquier cuadrilátero forman un paralelogramo.

b) Demostrar que un cuadrilátero es un paralelogramo si y sólo si las diagonales se cortan en el punto medio.

3. Sean $A_i = (a_i, b_i, c_i)$, $i = 1, 2, 3, 4$, puntos de un espacio afín de dimensión 3. Demostrar que los A_i son coplanarios si y sólo si

$$\begin{vmatrix} a_1 & b_1 & c_1 & 1 \\ a_2 & b_2 & c_2 & 1 \\ a_3 & b_3 & c_3 & 1 \\ a_4 & b_4 & c_4 & 1 \end{vmatrix} = 0.$$

4. Sean $a_1, \ldots, a_n$ puntos de un espacio afín. Demostrar que las rectas que unen cada a_i, $i = 1, \ldots, n$, con el baricentro de los puntos restantes son concurrentes (en el baricentro de $a_1, \ldots, a_n$).

5. Dado un triángulo abc del plano afín real y tres pnntos a', b', c' sobre los lados bc, ca, ab respectivamente, encontrar una condición necesaria y suficiente para que los triángulos abc y $a'b'c'$ tengan el mismo baricentro.

6. Calcular las seis razones simples que se obtienen al permutar tres puntos alineados.

7. Sea A una recta afín sobre $\mathbf{R}$ o $\mathbf{C}$. ¿En qué condiciones las seis razones simples de tres puntos de A toman como máximo tres valores diferentes?

8. (*Teorema de Desargues.*) Consideremos dos triángulos ABC y $A'B'C'$ del plano afín real y denotemos por a, b, c (resp. a', b', c') las rectas BC, AC, AB (resp. $B'C'$, $A'C'$, $A'B'$). Demostrar que las rectas AA', BB' y CC' son paralelas o concurrentes si y sólo si los puntos $a \cap a'$, $b \cap b'$ y $c \cap c'$ están alineados.

9. (*Teorema de Pappus.*) Sean A, B, C y A', B', C' dos ternas de puntos alineados y denotemos por a, b, c (resp. a', b', c') las rectas BC', AC', AB' (resp. $B'C$, $A'C$, $A'B$). Demostrar que los puntos $a \cap a'$, $b \cap b'$ y $c \cap c'$ están alineados.

10. Demostrar que un subconjunto de un espacio afín es una variedad lineal si y sólo si contiene, con cada pareja de puntos, la recta que determinan.

.1. En el plano afín $A = \mathbf{Z}/(3) \times \mathbf{Z}/(3)$ sobre el cuerpo $\mathbf{Z}/(3)$,

a) ¿Cuántos puntos hay?

b) ¿Cuántas rectas hay?

c) ¿Cuántos puntos tiene cada recta?

d) ¿Cuántas rectas hay paralelas a una dada?

e) ¿Cuántos haces diferentes de rectas paralelas hay?

12. Sea A un plano afín sobre el cuerpo $\mathbf{Z}/(5)$. Estudiar la posición relativa de las rectas

$$r : 3x + 2y + 2 = 0, \quad s : 2x + y = 0, \quad t : x - 3y + 1 = 0.$$

13. Sean $A_i x + b_i = \vec{0}$, $i = 1, 2$, las ecuaciones de dos variedades lineales. Demostrar que las ecuaciones de la variedad intersección se obtienen reduciendo por el método de Gauss la matriz

$$\begin{pmatrix} A_1 & b_1 \\ A_2 & b_2 \end{pmatrix}.$$

Si el sistema es incompatible, las dos variedades no se cortan y

a) si $\operatorname{rang}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \max(\operatorname{rang} A_1, \operatorname{rang} A_2)$, las variedades son paralelas,

b) si $\operatorname{rang}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} > \operatorname{rang} A_i$, $i = 1, 2$, las variedades se cruzan.

14. Sean $Ax + b = \vec{0}$ con $A \in M_{(n-k)\times n}(K)$, $b \in K^{n-k}$, $\operatorname{rang} A = n - k$, las ecuaciones de una variedad lineal de dimensión k en un espacio afín de dimensión n, en un sistema de referencia cartesiano.

Si $\bar{A}\bar{x} + \bar{b} = \vec{0}$ son las ecuaciones en otro sistema de referencia, demostrar que, entonces,

$$\begin{cases} \bar{A} &= AV \\ \bar{b} &= b + Aq, \end{cases}$$

donde V es la matriz del cambio de base y q las coordenadas del origen.

15. Dadas tres rectas del plano afín real de ecuaciones

$$3x + 2y = 1, \quad y = 5, \quad 6x + y = -13,$$

hallar los triángulos abc qne tienen sus medianas (XIII.11) sobre estas rectas, el vértice a sobre la primera recta y el punto $(-1, 2)$ como punto medio del lado bc.

16. Las caras de un tetraedro $abcd$ son cortadas por una recta en cuatro pnntos a', b', c', d'. Probar que los puntos medios de los segmentos aa', bb', cc', dd' son coplanarios.

17. Por los vértices de un tetraedro $abcd$ del espacio afín real se trazan cuatro rectas paralelas que cortan las caras opuestas en los puntos a', b', c', d'. Determinar $k \in \mathbf{R}$ de manera que los puntos que pertenecen a los segmentos aa', bb', cc', dd' con razón k sean coplanarios.

18. En el espacio afín real se consideran tres rectas que se cruzan dos a dos y son paralelas a un plano. Demostrar que toda recta que corte a las tres es paralela a un plano fijo. Determinar ese plano.

IX.15 Ejercicios para programar

19. Sean $p_0, \ldots, p_n$ puntos dados del espacio afín real $\mathbf{R}^n$. Hacer un programa que permita:

a) Comprobar que son linealmente independientes.

b) Dado un punto $x \in \mathbf{R}^n$ cualquiera, encontrar sus coordenadas baricéntricas respecto al sistema $\{p_0, \ldots, p_n\}$.

20. Como aplicación del ejercicio 19, preparar tres programas que permitan:

a) Decidir si un punto x dado del plano afín real es interior, exterior o está sobre un lado del triángulo determinado por tres puntos p_0, p_1, p_2. (Véase §6, ejemplo 3.)

b) Lo mismo en el espacio afín con cuatro puntos linealmente independientes.

c) Dado un polígono convexo del plano afín por la sucesión ordenada de sus vértices, decidir si un punto x dado es interior, exterior o está sobre un lado. (Indicación: descomponer el polígono en triángulos con un vértice común.)

21. Sea $\{p;\ e_1, \ldots, e_n\}$ el sistema de referencia cartesiano canónico del espacio afín $\mathbf{R}^n$ ($p = (0, \ldots, 0)$, $e_i = (0, \ldots, 1, \ldots, 0)$, $i = 1, \ldots, n$). Sea $\{q;\ v_1, \ldots, v_n\}$ otro sistema de referencia dado.

a) Hacer un programa que transporte las coordenadas de un punto dado de un sistema al otro. (Será necesario utilizar el ejercicio VII.12 para invertir la matriz $\begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix}$, donde V es la matriz de cambio de base y q las coordenadas del nuevo origen.)

b) Hacer un programa que cambie de sistema de referencia las ecuaciones cartesianas de una variedad lineal. (§9 y ejercicio 14.)

22. Dadas las ecuaciones cartesianas de dos variedades lineales de $\mathbf{R}^n$ en el sistema de referencia canónico, hacer un programa que permita decidir si se cortan, se cruzan o son paralelas. Si se cortan, dar las ecuaciones simplificadas de la intersección. (Ver ejercicio 13.)

23. Hacer un programa que calcule la razón simple de tres puntos alineados de $\mathbf{R}^n$ y estudiar el efecto de permutar los puntos.

Capítulo X

Afinidades

De la misma manera que en el capítulo de grupos considerábamos las aplicaciones entre ellos que "respetaban" sus operaciones (los homomorfismos), y que al estudiar espacios vectoriales nos interesábamos por las aplicaciones que conservaban las dos operaciones (las aplicaciones lineales), queremos ahora estudiar aplicaciones entre espacios afines que respeten su estructura afín: las aplicaciones afines o afinidades.

X.1 Definición y primeras propiedades

Sean (A_1, E_1, φ_1) y (A_2, E_2, φ_2) dos espacios afines sobre el mismo cuerpo K. Una *aplicación afín* o *afinidad* entre estos espacios es una aplicación

$$f : A_1 \longrightarrow A_2$$

junto con una aplicación lineal

$$\tilde{f} : E_1 \longrightarrow E_2$$

tales que el diagrama de aplicaciones

$$\begin{array}{ccc} A_1 \times A_1 & \xrightarrow{\varphi_1} & E_1 \\ f \times f \downarrow & & \downarrow \tilde{f} \\ A_2 \times A_2 & \xrightarrow{\varphi_2} & E_2 \end{array}$$

conmuta. Esto significa que $\tilde{f} \circ \varphi_1 = \varphi_2 \circ (f \times f)$; es decir,

$$\tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)} \qquad \forall a, b \in A_1.$$

Recordemos que la estructura de un espacio afín puede darse también por el conjunto de sus traslaciones: $T_1 : A_1 \times E_1 \longrightarrow A_1$ y $T_2 : A_2 \times E_2 \longrightarrow A_2$. Así pues, sería lógico exigir que las aplicaciones f y $\tilde{f}$ hiciesen conmutativo el diagrama

$$\begin{array}{ccc} A_1 \times E_1 & \xrightarrow{T_1} & A_1 \\ f \times \tilde{f} \downarrow & & \downarrow f \\ A_2 \times E_2 & \xrightarrow{T_2} & A_2 \end{array} ;$$

es decir, $f \circ T_1 = T_2 \circ (f \times \tilde{f})$. Esto equivale a decir que para todo $a \in A$ y todo $u \in E_1$,

$$f(a+u) = f(a) + \tilde{f}(u).$$

La proposición siguiente demuestra que esta condición equivale a la que hemos impuesto en la definición.

Proposición 1.1 *Las condiciones*

i) $\tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)} \qquad \forall a, b \in A_1$

ii) $f(a+u) = f(a) + \tilde{f}(u) \qquad \forall a \in A_1, \forall u \in E_1$

son equivalentes.

DEMOSTRACIÓN: (i) $\Rightarrow$ (ii). En efecto, dados $a \in A_1$, $u \in E_1$, sea $b \in A_1$ tal que $\overrightarrow{ab} = u$. Entonces, por (i), $\tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)}$ y, por tanto,

$$f(a+u) = f(b) = f(a) + \tilde{f}(\overrightarrow{ab}) = f(a) + \tilde{f}(u).$$

(ii) $\Rightarrow$ (i). En efecto, dados $a, b \in A_1$, de $b = a + \overrightarrow{ab}$ se deduce que $f(b) = f(a) + \tilde{f}(\overrightarrow{ab})$ y, por tanto,

$$\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}). \square$$

Por abuso de lenguaje, se dice que una aplicación $f : A_1 \longrightarrow A_2$ es una afinidad si existe una aplicación lineal $\tilde{f} : E_1 \longrightarrow E_2$ que cumple una de las condiciones de (1.1). $\tilde{f}$ se llama *aplicación lineal asociada a* f.

Vamos a demostrar ahora unos cuantos hechos que se deducen inmediatamente de las definiciones.

Proposición 1.2 *Sean* $f, g : A_1 \longrightarrow A_2$ *dos afinidades que coinciden sobre un punto* p, $f(p) = g(p)$, *y que tienen la misma aplicación lineal asociada,* $\tilde{f} = \tilde{g}$. *Entonces* $f = g$.

DEMOSTRACIÓN: Para todo $a \in A_1$,

$$f(a) = f(p + \overrightarrow{pa}) = f(p) + \tilde{f}(\overrightarrow{pa}) = g(p) + \tilde{g}(\overrightarrow{pa}) = g(p + \overrightarrow{pa}) = g(a). \square$$

Proposición 1.3 *Dada una aplicación lineal $\varphi : E_1 \longrightarrow E_2$ y un par de puntos $p \in A_1, q \in A_2$, existe una afinidad f, y sólo una, tal que $f(p) = q$ y $\tilde{f} = \varphi$.*

DEMOSTRACIÓN: Si existe, f es única por (1.2). Definamos $f : A_1 \longrightarrow A_2$ por

$$f(a) = q + \varphi(\overrightarrow{pa}), \qquad \forall a \in A_1.$$

En particular, $f(p) = q$ y $\overrightarrow{f(a)f(b)} = \overrightarrow{qf(b)} - \overrightarrow{qf(a)} = \varphi(\overrightarrow{pb}) - \varphi(\overrightarrow{pa}) = \varphi(\overrightarrow{ab})$ para todo $a, b \in A_1$. Esto demuestra que f es una afinidad con aplicación lineal asociada φ y que transforma p en q. $\square$

Proposición 1.4 *Supongamos A_1 de dimensión n. Dados $(n+1)$ puntos linealmente independientes $a_0, \ldots, a_n$ de A_1, y $(n+1)$ puntos arbitrarios $b_0, \ldots, b_n$ de A_2, existe una afinidad $f : A_1 \longrightarrow A_2$, y sólo una, tal que $f(a_i) = b_i$, $i = 0, \ldots, n$.*

DEMOSTRACIÓN: Si existe la afinidad f, su aplicación lineal asociada tiene que ser la única aplicación lineal $\tilde{f} : E_1 \longrightarrow E_2$ tal que $\tilde{f}(\overrightarrow{a_0a_i}) = \overrightarrow{b_0b_i}$ para $i = 1, \ldots, n$. Definimos pues $f : A_1 \longrightarrow A_2$ como la única aplicación afín que tiene por aplicación lineal asociada esa $\tilde{f}$ y que transforma a_0 en b_0: $f(a_0) = b_0$ (1.3). Esta f cumple

$$f(a_i) = f(a_0 + \overrightarrow{a_0a_i}) = f(a_0) + \tilde{f}(\overrightarrow{a_0a_i}) = b_0 + \overrightarrow{b_0b_i} = b_i$$

y es, por tanto, la aplicación buscada. $\square$

Proposición 1.5 *Si $f : A_1 \longrightarrow A_2$ y $g : A_2 \longrightarrow A_3$ son afinidades, entonces $g \circ f : A_1 \longrightarrow A_3$ es una afinidad, y $\widetilde{g \circ f} = \tilde{g} \circ \tilde{f}$.*

DEMOSTRACIÓN: Dados $a, b \in A_1$,

$$\tilde{g} \circ \tilde{f}(\overrightarrow{ab}) = \tilde{g}(\overrightarrow{f(a)f(b)}) = \overrightarrow{g \circ f(a)\ g \circ f(b)}.$$

La aplicación $\tilde{g} \circ \tilde{f}$ cumple, por tanto, (i) de (1.1) y es la aplicación lineal asociada a $g \circ f$. $\square$

La proposición siguiente demuestra que hay una relación muy estrecha entre una afinidad y su aplicación lineal asociada.

Proposición 1.6 *Dada una afinidad $(f, \tilde{f})$ de (A_1, E_1) en (A_2, E_2),*

a) f es inyectiva $\Leftrightarrow$ $\tilde{f}$ es inyectiva;
b) f es exhaustiva $\Leftrightarrow$ $\tilde{f}$ es exhaustiva;
c) f es biyectiva $\Leftrightarrow$ $\tilde{f}$ es biyectiva.

DEMOSTRACIÓN: Supongamos f inyectiva. Dado $v \in \operatorname{Nuc} \tilde{f}$, si $v = \overrightarrow{ab}$, tenemos

$$\vec{0} = \tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)} \quad \Rightarrow \quad f(a) = f(b) \quad \Rightarrow \quad a = b \quad \Rightarrow \quad v = \overrightarrow{ab} = \vec{0}.$$

Por tanto, $\operatorname{Nuc} \tilde{f} = \{\vec{0}\}$ y $\tilde{f}$ es inyectiva.

Supongamos $\tilde{f}$ inyectiva. Entonces,

$$\begin{aligned} f(a) = f(b) \quad &\Rightarrow \quad \vec{0} = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) \quad \Rightarrow \\ &\Rightarrow \quad \overrightarrow{ab} \in \operatorname{Nuc} \tilde{f} = \{\vec{0}\} \quad \Rightarrow \quad \overrightarrow{ab} = \vec{0} \quad \Rightarrow \quad a = b. \end{aligned}$$

Tenemos así demostrado (a).

Supongamos f exhaustiva. Dado $u \in E_2$, pongamos $u = \overrightarrow{cd}$. Sean $a, b \in A_1$, tales que $f(a) = c, f(b) = d$. Entonces $u = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab})$; por tanto, $\tilde{f}$ es exhaustiva.

Supongamos $\tilde{f}$ exhaustiva. Dado $c \in A_2$, consideremos un vector $u = \overrightarrow{f(a)c}$, donde a es un punto cualquiera de A_1. Por ser $\tilde{f}$ exhaustiva, existe $v \in E_1$ tal que $\tilde{f}(v) = u$. Escribamos $v = \overrightarrow{ab}$. Entonces $\overrightarrow{f(a)c} = u = \tilde{f}(v) = \overrightarrow{f(a)f(b)}$, de donde $f(b) = c$. Por tanto, f también es exhaustiva. Esto demuestra (b).

La afirmación (c) es consecuencia de las dos anteriores. □

Corolario 1.7 *Si f es una afinidad biyectiva con aplicación lineal asociada $\tilde{f}$, entonces f^{-1} es una afinidad con aplicación lineal asociada $\tilde{f}^{-1}$.*

DEMOSTRACIÓN: En virtud de la proposición (1.6), solamente queda demostrar que $\tilde{f}^{-1}(\overrightarrow{cd}) = \overrightarrow{f^{-1}(c)f^{-1}(d)}$; pero eso se deduce inmediatamente de

$$\tilde{f}(\overrightarrow{f^{-1}(c)f^{-1}(d)}) = \overrightarrow{ff^{-1}(c)ff^{-1}(d)} = \overrightarrow{cd}. \ \square$$

A una afinidad biyectiva la llamaremos *isomorfismo afín*. Dos espacios afines son *isomorfos* si existe un isomorfismo afín entre ellos. De (1.6)

resulta que, si A_1 y A_2 son isomorfos, sus espacios vectoriales asociados E_1, E_2 también lo son. Recíprocamente, si $\varphi : E_1 \longrightarrow E_2$ es un isomorfismo, (1.3) y (1.6) garantizan la existencia de un isomorfismo afín $f : A_1 \longrightarrow A_2$. Si los espacios son de dimensión finita obtenemos, en particular, el siguiente resultado.

Proposición 1.8 *Dos espacios afines de dimensión finita sobre el mismo cuerpo son isomorfos si y sólo si tienen la misma dimensión.* □

De (1.8) se deduce que todo espacio afín de dimensión finita es isomorfo a un espacio afín estándar K^n (IX.1).

X.2 Unos ejemplos

En este apartado estudiaremos ejemplos de aplicaciones afines de un espacio afín en sí mismo.

I Traslaciones

Sea T_v una traslación de vector $v \in E$. Si $a, b \in A$,

$$\overrightarrow{T_v(a)T_v(b)} = \overrightarrow{T_v(a)a} + \overrightarrow{ab} + \overrightarrow{bT_v(b)} = -v + \overrightarrow{ab} + v = \overrightarrow{ab} = I_E(\overrightarrow{ab}),$$

donde I_E indica la aplicación identidad de E. Esto demuestra que T_v es una afinidad con aplicación lineal asociada I_E.

Recíprocamente, si $f : A \longrightarrow A$ es una afinidad y $\tilde{f} = I_E$, entonces

$$\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) = \overrightarrow{ab} \qquad \forall a, b,$$

de donde

$$\overrightarrow{af(a)} = \overrightarrow{bf(b)}.$$

Es decir, el vector v determinado por un punto y su imagen es independiente del punto escogido: $v = \overrightarrow{af(a)}$. Así pues, para todo a, $f(a) = a + v = T_v(a)$, de donde $f = T_v$.

Proposición 2.1 *Una afinidad $f : A \longrightarrow A$ es una traslación si y sólo si $\tilde{f} = I_E$.* □

II Proyecciones

Una afinidad $f : A \longrightarrow A$ se llama una *proyección* si $f^2 = f$.

Observemos, en primer lugar, que todo punto de $\operatorname{Im} f$ es fijo: $ff(a) = f(a)$ para todo a. Además, si b es fijo, $b = f(b) \in \operatorname{Im} f$. Así pues, $\operatorname{Im} f$ es el conjunto de puntos fijos de f.

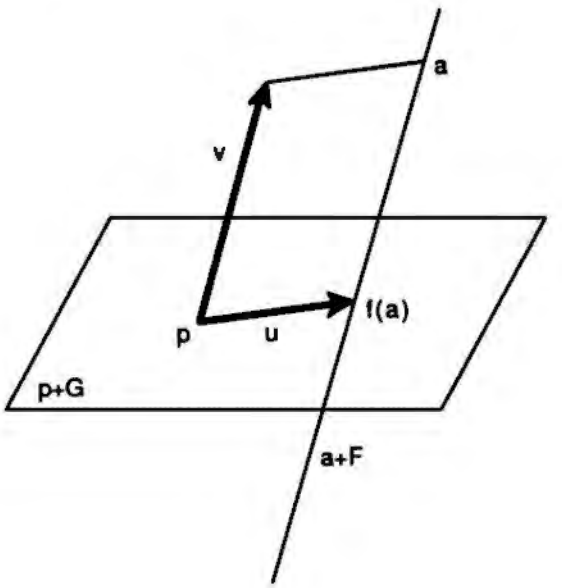
Estudiemos ahora $\tilde{f}$. Para todo $\overrightarrow{ab} \in E$,

$$\tilde{f}^2(\overrightarrow{ab}) = \tilde{f}(\overrightarrow{f(a)f(b)}) = \overrightarrow{f^2(a)f^2(b)} = \overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}).$$

Así pues, $\tilde{f}^2 = \tilde{f}$ y el polinomio mínimo de $\tilde{f}$ es un divisor de $x^2 - x$. Apliquemos los métodos de (VIII.5) al estudio de $\tilde{f}$. Pueden darse tres casos:

- Si el polinomio mínimo de $\tilde{f}$ es x, $\tilde{f} = 0$ y $\overrightarrow{f(a)f(b)} = \tilde{f}(\overrightarrow{ab}) = \vec{0}$ para todo $a, b \in A$. Así pues, $f(a) = f(b)$ para todo $a, b \in A$ y, por tanto, f aplica todo A en el mismo punto.

- Si el polinomio mínimo de $\tilde{f}$ es $x - 1$, $\tilde{f} = I_E$ y f es una traslación (2.1). Ahora bien, f tiene puntos fijos; por tanto, es una traslación de vector $\vec{0}$; es decir, f es la identidad de A.



- Si el polinomio mínimo de $\tilde{f}$ es $x(x-1)$, $E = F \oplus G$, donde F es un subespacio invariante sobre el que $\tilde{f} = 0$ y G es un subespacio invariante sobre el que $\tilde{f} = I_G$ (VIII.4.4). Entonces, si p es un punto fijo,

$$f(a) = f(p + \overrightarrow{pa}) = p + \tilde{f}(\overrightarrow{pa}) \qquad \forall a \in A.$$

Si $\overrightarrow{pa} = v + u$ con $v \in F$ y $u \in G$, $f(a) = p + u$. El conjunto de puntos fijos de f es, en este caso, $p + G$ (3.3). Por otra parte, $\overrightarrow{f(a)a} = \overrightarrow{f(a)p} + \overrightarrow{pa} = -u + (v + u) = v$, de donde $f(a) \in a + F$. El punto $f(a)$ es, por tanto, la intersección de $a + F$ e $\operatorname{Im} f = p + G$. Observemos que $E = F \oplus G$ implica que esa intersección se reduce siempre a un solo punto (IX.4.4).

III Simetrías

Una afinidad $f : A \longrightarrow A$ se llama una *simetría* si $f^2 = I_A$.

Supongamos que el cuerpo K no es de característica 2 (IX.6); los puntos medios de los pares formados por un punto a y su imagen $f(a)$ son fijos:

$$m = \frac{1}{2}a + \frac{1}{2}f(a) \Leftrightarrow \overrightarrow{am} = \frac{1}{2}\overrightarrow{af(a)} \Leftrightarrow \overrightarrow{f(a)f(m)} = \frac{1}{2}\overrightarrow{f(a)a} \Leftrightarrow$$
$$\Leftrightarrow f(m) = \frac{1}{2}f(a) + \frac{1}{2}a = m.$$

Estudiemos ahora $\tilde{f}$. Puesto que $f^2 = I_A$, también $\tilde{f}^2 = I_E$ y el polinomio mínimo de $\tilde{f}$ es un divisor de $x^2 - 1$ (VIII.4, ejemplo 2). Hay tres posibilidades:

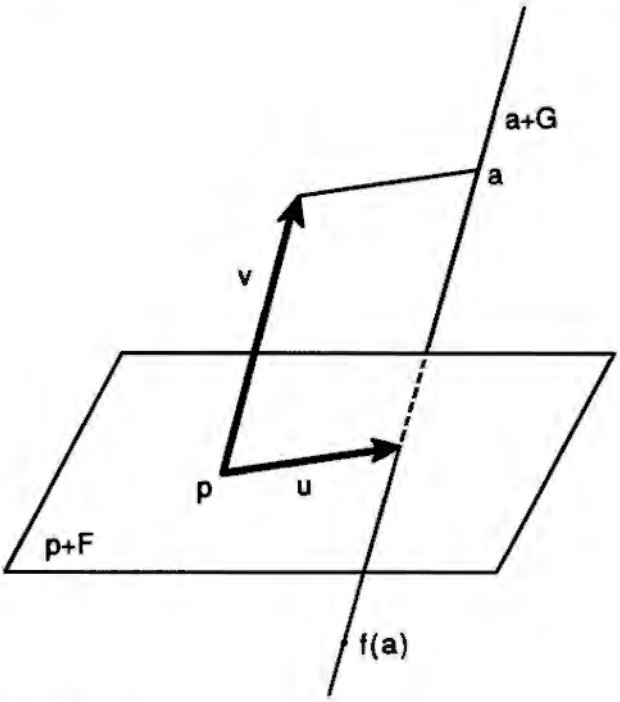
- Si el polinomio mínimo de $\tilde{f}$ es $(x - 1)$, $\tilde{f} = I_E$ y f es una traslación con puntos fijos; por tanto, $f = I_A$.
- Si el polinomio mínimo de $\tilde{f}$ es $(x + 1)$, $\tilde{f} = -I_E$. Si p es un punto fijo, entonces, para todo a, $\overrightarrow{pf(a)} = \tilde{f}(\overrightarrow{pa}) = -\overrightarrow{pa}$, y esto equivale a $p = \frac{1}{2}a + \frac{1}{2}f(a)$. Existe pues un único punto fijo p que es punto medio del par $a, f(a)$ para cada a. La afinidad f se llama entonces una *simetría central* de *centro* p.
- Si el polinomio mínimo de $\tilde{f}$ es $(x-1)(x+1)$, $E = F \oplus G$, donde F es un subespacio invariante sobre el cual $\tilde{f}$ es I_F, y G es un subespacio invariante sobre el cual $\tilde{f}$ es $-I_G$. Sea p un punto fijo; dado $a \in A$, si $\overrightarrow{pa} = u + v$ con $u \in F$ y $v \in G$,

$$f(a) = f(p + \overrightarrow{pa}) = p + \tilde{f}(u + v) = p + u - v.$$

De ahí resulta que los puntos fijos de f pertenecen a la variedad $p + F$ (proposición 3.3). Por otra parte,

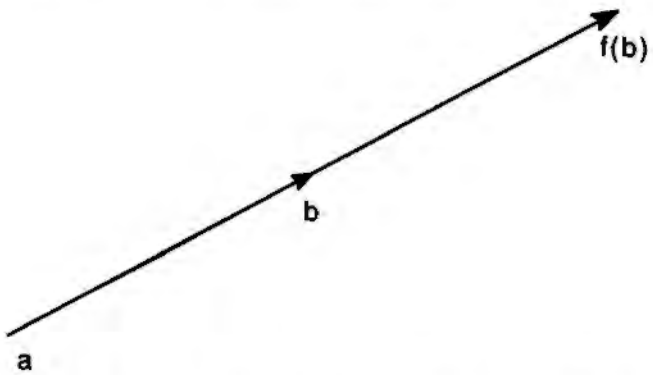
$$\overrightarrow{f(a)a} = 2v \Rightarrow f(a) \in a + G,$$
$$u = \frac{1}{2}\overrightarrow{pa} + \frac{1}{2}\overrightarrow{pf(a)} \Rightarrow \frac{1}{2}a + \frac{1}{2}f(a) = p + u \in p + F,$$

y estas dos condiciones determinan $f(a)$.



IV Homotecias

Una afinidad $f : A \longrightarrow A$ se llama una *homotecia de razón* $r \neq 0, 1$ si $\tilde{f} = rI_E$.



Para estudiar los posibles puntos fijos, escogemos un punto auxiliar $p \in A$. Entonces $a \in A$ es fijo si

$$a = f(a) = f(p + \overrightarrow{pa}) = f(p) + r\,\overrightarrow{pa} \quad \Leftrightarrow \quad \overrightarrow{pa} = \overrightarrow{pf(p)} + r\,\overrightarrow{pa} \quad \Leftrightarrow$$
$$\Leftrightarrow \quad \overrightarrow{pa} = \frac{1}{1-r}\overrightarrow{pf(p)} \quad \Leftrightarrow \quad a = p + \frac{1}{1-r}\overrightarrow{pf(p)}.$$

Este es, por tanto, el único punto fijo de f; se llama el *centro* de la homotecia. La imagen de cualquier otro punto $b \in A$ es

$$f(b) = a + r\,\overrightarrow{ab}.$$

X.3 Más propiedades de las afinidades

Proposición 3.1 *Si $f : A_1 \longrightarrow A_2$ es una afinidad y $a+F$ es una variedad lineal de A_1,*

$$f(a+F) = f(a) + \tilde{f}(F).$$

Si $b+G$ es una variedad lineal de A_2 y $a \in f^{-1}(b+G)$,

$$f^{-1}(b+G) = a + \tilde{f}^{-1}(G).$$

DEMOSTRACIÓN: La primera parte es una consecuencia inmediata de la definición de afinidad. Para demostrar la segunda parte, observemos que

$$f(a+\tilde{f}^{-1}(G)) = f(a)+G = b+G \quad \Rightarrow \quad f^{-1}(b+G) \supset a+\tilde{f}^{-1}(G).$$

Por otro lado,

$$\begin{aligned} c \in f^{-1}(b+G) \quad &\Rightarrow \quad f(c) \in b+G = f(a)+G \quad \Rightarrow \\ &\Rightarrow \quad \overrightarrow{f(a)f(c)} = \tilde{f}(\overrightarrow{ac}) \in G \quad \Rightarrow \\ &\Rightarrow \quad \overrightarrow{ca} \in \tilde{f}^{-1}(G) \quad \Rightarrow \quad c \in a+\tilde{f}^{-1}(G), \end{aligned}$$

de donde $f^{-1}(b+G) \subset a+\tilde{f}^{-1}(G)$. Esto demuestra la segunda igualdad del enunciado. □

Corolario 3.2 *a)* Im f *es una variedad lineal de dimensión* rang $\tilde{f}$.

b) f transforma variedades paralelas en variedades paralelas. En particular, f transforma puntos alineados en puntos alineados. □

Proposición 3.3 *Si el conjunto de puntos fijos de una afinidad $f : A \longrightarrow A$ no es vacío, es una variedad lineal de dirección el subespacio de vectores propios de valor propio* 1 *de $\tilde{f}$.*

DEMOSTRACIÓN: Sea a un punto fijo de f, $f(a) = a$, y E_1 el subespacio de vectores propios de valor propio $+1$. Para todo $u \in E_1$, $f(a+u) = f(a) + \tilde{f}(u) = a+u$; los puntos de $a+E_1$ son, pues, todos fijos. Recíprocamente, si b es fijo, entonces $b = f(b) = f(a+\overrightarrow{ab}) = a+\tilde{f}(\overrightarrow{ab}) \Rightarrow \tilde{f}(\overrightarrow{ab}) = \overrightarrow{ab} \Rightarrow \overrightarrow{ab} \in E_1 \Rightarrow b \in a+E_1$. □

Proposición 3.4 *Si* $f : A_1 \longrightarrow A_2$ *es una afinidad y* $x = \sum_{i=1}^{r} x^i a_i$ *con* $x^1 + \ldots + x^r = 1$, *entonces* $f(x) = \sum_{i=1}^{r} x^i f(a_i)$.

DEMOSTRACIÓN: Sea p un punto cualquiera del espacio afín A_1. Tenemos $\overrightarrow{px} = \sum_{i=1}^{r} x^i \overrightarrow{pa_i}$, de donde

$$\overrightarrow{f(p)f(x)} = \tilde{f}(\overrightarrow{px}) = \sum_{i=1}^{r} x^i \tilde{f}(\overrightarrow{pa_i}) = \sum_{i=1}^{r} x^i \overrightarrow{f(p)f(a_i)},$$

tal como queríamos demostrar. □

La proposición (3.4) nos dice, en particular, que toda afinidad transforma el baricentro de r puntos en el baricentro de sus imágenes.

Vamos a probar ahora que la propiedad (3.4) es suficientemente restrictiva como para caracterizar las afinidades.

Proposición 3.5 *Una aplicación de conjuntos* $f : A_1 \longrightarrow A_2$ *es una afinidad si y sólo si, siempre que* $x^1 + \ldots + x^r = 1$,

$$f\left(\sum_{i=1}^{r} x^i a_i\right) = \sum_{i=1}^{r} x^i f(a_i).$$

DEMOSTRACIÓN: El primer paso tiene que ser definir lo que será la aplicación lineal asociada a f. Observemos que la condición $\tilde{f}(\overrightarrow{ab}) = \overrightarrow{f(a)f(b)}$ nos determina ya $\tilde{f}$. El único problema es que cada vector $u \in E_1$ admite muchas representaciones de la forma $u = \overrightarrow{ab}$. Para evitar esta pluralidad, fijamos un punto $p \in A_1$ y tomamos todos los vectores con origen p. Así pues, definimos

$$\begin{array}{rccc} \tilde{f}: & E_1 & \longrightarrow & E_2 \\ & u = \overrightarrow{px} & \longmapsto & \tilde{f}(u) = \overrightarrow{f(p)f(x)}. \end{array}$$

Probemos que $\tilde{f}$ es lineal: dados $u = \overrightarrow{px}$, $v = \overrightarrow{py}$, sea $u + v = \overrightarrow{pz}$; entonces

$$\tilde{f}(u+v) = \tilde{f}(\overrightarrow{pz}) = \overrightarrow{f(p)f(z)}$$

$$\tilde{f}(u) + \tilde{f}(v) = \tilde{f}(\overrightarrow{px}) + \tilde{f}(\overrightarrow{py}) = \overrightarrow{f(p)f(x)} + \overrightarrow{f(p)f(y)}$$

y tendríamos que ver que $\overrightarrow{f(p)f(z)} = \overrightarrow{f(p)f(x)} + \overrightarrow{f(p)f(y)}$. Esto equivale a ver que $f(z) = -f(p) + f(x) + f(y)$. La condición del enunciado nos dice que esto será cierto si $z = -p + x + y$, es decir, si $\overrightarrow{pz} = \overrightarrow{px} + \overrightarrow{py}$, que es precisamente de donde hemos partido.

Sea ahora $u = \overrightarrow{px}$, $ku = \overrightarrow{py}$. $\tilde{f}(u) = \overrightarrow{f(p)f(x)}$, $\tilde{f}(ku) = \overrightarrow{f(p)f(y)}$ y tendríamos que ver que $\overrightarrow{f(p)f(y)} = k\,\overrightarrow{f(p)f(x)}$. Esto significa que $f(y) = (1-k)f(p) + kf(x)$. Basta con ver que $y = (1-k)p + kx$, es decir, que $\overrightarrow{py} = k\,\overrightarrow{px}$, y esto es cierto por hipótesis.

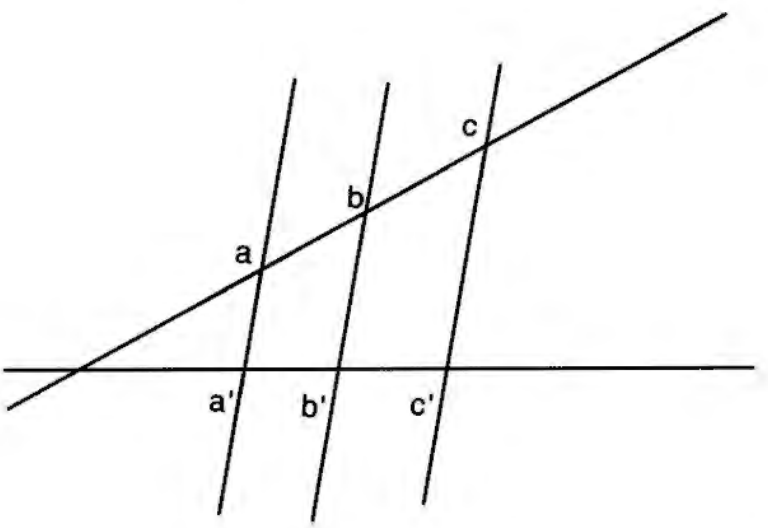
Sólo queda ahora comprobar que $\tilde{f}$ es la aplicación lineal asociada a f. En efecto, dados $a, b \in A$,

$$\tilde{f}(\overrightarrow{ab}) = \tilde{f}(\overrightarrow{pb} - \overrightarrow{pa}) = \tilde{f}(\overrightarrow{pb}) - \tilde{f}(\overrightarrow{pa}) = \overrightarrow{f(p)f(b)} - \overrightarrow{f(p)f(a)} = \overrightarrow{f(a)f(b)}. \square$$

Proposición 3.6 *Las afinidades conservan la razón simple.*

DEMOSTRACIÓN: Sea $r = (a_1a_2a_3)$. Tenemos

$$\overrightarrow{a_1a_3} = r\overrightarrow{a_1a_2} \Rightarrow \overrightarrow{f(a_1)f(a_3)} = r\overrightarrow{f(a_1)f(a_2)} \Leftrightarrow (f(a_1)f(a_2)f(a_3)) = r. \square$$



Esta proposición, en el caso particular en que f sea una proyección (§2, II), se conoce como el *teorema de Tales.*

La propiedad (3.6) también caracteriza las afinidades:

Proposición 3.7 *Si el cuerpo K no es de característica 2, una aplicación $f : A_1 \longrightarrow A_2$ es una afinidad si y sólo si conserva puntos alineados y sus razones simples.*

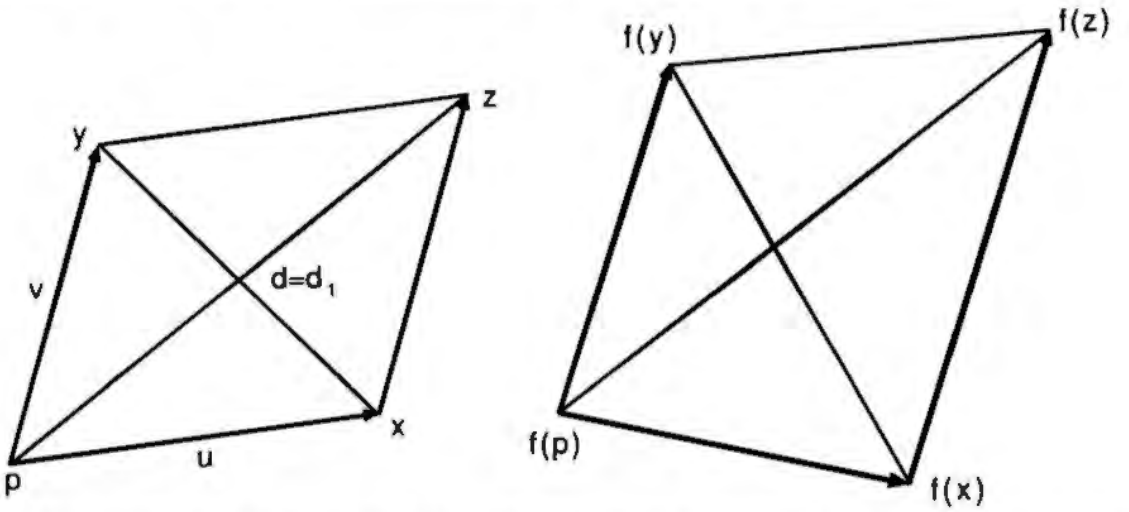
DEMOSTRACIÓN: Fijemos $p \in A_1$, y definamos

$$\begin{array}{rccc} \tilde{f}: & E_1 & \longrightarrow & E_2 \\ & u = \overrightarrow{px} & \longmapsto & \tilde{f}(u) = \overrightarrow{f(p)f(x)}. \end{array}$$

(Véase el principio de la demostración de (3.5).) Probemos que $\tilde{f}$ es lineal. Sean $u = \overrightarrow{px}$, $v = \overrightarrow{xz} = \overrightarrow{py}$ y $u + v = \overrightarrow{px} + \overrightarrow{xz} = \overrightarrow{pz}$. Designemos por d el punto medio del par p, z y por d_1 el del par y, x. Estos puntos existen siempre que en K se cumpla $2 \neq 0$ (IX.6). La condición $\overrightarrow{px} + \overrightarrow{py} = \overrightarrow{pz}$ equivale a $d = d_1$; en efecto,

$$\begin{aligned} d = \frac{1}{2}p + \frac{1}{2}z \quad &\Leftrightarrow \quad \overrightarrow{pd} = \frac{1}{2}\overrightarrow{pz} \\ d_1 = \frac{1}{2}y + \frac{1}{2}x \quad &\Leftrightarrow \quad \overrightarrow{pd_1} = \frac{1}{2}\overrightarrow{py} + \frac{1}{2}\overrightarrow{px}. \end{aligned}$$

Por tanto, si $d = d_1$, las segundas igualdades dan $\overrightarrow{pz} = \overrightarrow{py} + \overrightarrow{px}$. Recíprocamente, si $\overrightarrow{pz} = \overrightarrow{py} + \overrightarrow{px}$, obtenemos $\overrightarrow{pd} = \overrightarrow{pd_1}$; es decir, $d = d_1$.



Observemos también que las expresiones anteriores nos dicen que d es el punto medio de p, z si y sólo si $(pzd) = 1/2$. Análogamente, que d sea el punto medio de y, x equivale a $(yxd) = 1/2$. Ahora bien, por hipótesis, f conserva las razones simples; por tanto, $(f(p)f(z)f(d)) = 1/2$ y $(f(y)f(x)f(d)) = 1/2$, lo que equivale a que $f(d)$ sea el punto medio de los pares $f(p), f(z)$ y $f(y), f(x)$. Pero hemos visto ya más arriba que, si estos puntos medios coinciden, $\overrightarrow{f(p)f(z)} = \overrightarrow{f(p)f(y)} + \overrightarrow{f(p)f(x)}$. Por la definición de $\tilde{f}$, esto no es otra cosa que

$$\tilde{f}(u + v) = \tilde{f}(v) + \tilde{f}(u).$$

Sea ahora $ku = k\overrightarrow{px} = \overrightarrow{py}$. Entonces,

$$(pxy) = k \quad \Rightarrow \quad (f(p)f(x)f(y)) = k \quad \Leftrightarrow$$
$$\Leftrightarrow \quad \overrightarrow{f(p)f(y)} = k\,\overrightarrow{f(p)f(x)} \quad \Leftrightarrow \quad \tilde{f}(ku) = k\tilde{f}(u).$$

Esto termina la demostración de que $\tilde{f}$ es lineal. Para ver que $\tilde{f}$ es precisamente la aplicación lineal asociada a f, procedemos como en la demostración de (3.5). Dados $a, b \in A$,

$$\tilde{f}(\overrightarrow{ab}) = \tilde{f}(\overrightarrow{pb} - \overrightarrow{pa}) = \tilde{f}(\overrightarrow{pb}) - \tilde{f}(\overrightarrow{pa}) = \overrightarrow{f(p)f(b)} - \overrightarrow{f(p)f(a)} = \overrightarrow{f(a)f(b)}.$$

Así pues, f es una afinidad. □

X.4 Ecuaciones de una afinidad en una referencia cartesiana

Sea $f : A_1 \longrightarrow A_2$ una afinidad. Consideremos sistemas de referencia cartesianos $\{p;\ e_1, \ldots, e_n\}$, $\{q;\ u_1, \ldots, u_m\}$ de los espacios A_1, A_2 respectivamente. Sabemos que

$$f(x) = f(p) + \tilde{f}(\overrightarrow{px}) \qquad \forall x \in A_1.$$

Si las coordenadas de x en el sistema $\{p;\ e_1, \ldots, e_n\}$ son $(x^1, \ldots, x^n)$ y $M = (a_i^j)$ es la matriz de $\tilde{f}$ en las bases $\{e_1, \ldots, e_n\}$ y $\{u_1, \ldots, u_m\}$ de los espacios vectoriales asociados E_1, E_2, las coordenadas del vector $\tilde{f}(\overrightarrow{px})$ son los términos de la matriz-columna

$$Mx, \qquad \text{donde} \qquad x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

Sean ahora $(b^1, \ldots, b^m)$ las coordenadas del punto $f(p)$ en el sistema de referencia $\{q;\ u_1, \ldots, u_m\}$, e indiquemos por b la matriz-columna formada por estas coordenadas. Entonces, claramente, las coordenadas $(\bar{x}^1, \ldots, \bar{x}^m)$ de $f(x)$ son los elementos de la matriz-columna

$$\bar{x} = b + Mx.$$

Esta expresión se escribe a menudo de la siguiente manera: sean

$$N = \begin{pmatrix} a_1^1 & \ldots & a_n^1 & b^1 \\ \vdots & & \vdots & \vdots \\ a_1^m & \ldots & a_n^m & b^m \\ 0 & \ldots & 0 & 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}.$$

$$\begin{pmatrix} \bar{x}^1 \\ \vdots \\ \bar{x}^m \\ 1 \end{pmatrix} = \begin{pmatrix} \bar{x} \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

Entonces,

$$\begin{pmatrix} \bar{x} \\ 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

La matriz $\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$ se llama la *matriz de la afinidad* f en los sistemas de referencia $\{p;\ e_1, \ldots, e_n\}$, $\{q;\ u_1, \ldots, u_m\}$. La expresión desarrollada

$$\begin{cases} \bar{x}^1 = a_1^1 x^1 + \ldots + a_n^1 x^n + b^1 \\ \qquad \ldots\ldots\ldots \\ \bar{x}^m = a_1^m x^1 + \ldots + a_n^m x^n + b^m \end{cases}$$

son las *ecuaciones de la afinidad* f en los sistemas anteriores.

Observación:

Tal como acabamos de ver, fijados sistemas de referencia, toda afinidad tiene unas ecuaciones lineales que permiten calcular las coordenadas de la imagen de un punto $f(x)$ a partir de las coordenadas del punto x. Recíprocamente, toda aplicación $g : A_1 \longrightarrow A_2$ dada por unas ecuaciones de este tipo

$$\begin{cases} \bar{x}^1 = c_1^1 x^1 + \ldots + c_n^1 x^n + d^1 \\ \qquad \ldots\ldots\ldots \\ \bar{x}^m = c_1^m x^1 + \ldots + c_n^m x^n + d^m \end{cases}$$

es una afinidad. En efecto, podemos considerar la aplicación lineal $\varphi : E_1 \longrightarrow E_2$ que tiene por matriz $C = (c_i^j)$ en las bases correspondientes en los sistemas de referencia fijados. Si $(x^1, \ldots, x^n)$ son las coordenadas de $x \in A_1$, las coordenadas de $\varphi(\overrightarrow{px})$ son

$$Cx = \begin{pmatrix} c_1^1 x^1 + \cdots + c_n^1 x^n \\ \vdots \\ c_1^m x^1 + \cdots + c_n^m x^n \end{pmatrix}.$$

La imagen del punto $p = (0, \ldots, 0)$ por g es el punto de coordenadas

$$d = \begin{pmatrix} d^1 \\ \vdots \\ d^m \end{pmatrix}.$$

Por tanto, la imagen de cualquier punto x tiene por coordenadas $Cx + d$. Es decir,

$$g(x) = g(p) + \varphi(\overrightarrow{px}).$$

Esto nos dice que g es una afinidad con aplicación lineal asociada φ.

Estos argumentos prueban también que la matriz y las ecuaciones de una afinidad en unos sistemas de referencia fijados están unívocamente determinadas.

Proposición 4.1 *Sean* $N = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$, $R = \begin{pmatrix} C & d \\ 0 & 1 \end{pmatrix}$ *las matrices de las afinidades* $f : A_1 \longrightarrow A_2$, $g : A_2 \longrightarrow A_3$ *en unos ciertos sistemas de referencia de* A_1, A_2, A_3. *Entonces* RN *es la matriz de la afinidad* $g \circ f$.

DEMOSTRACIÓN: Para todo $x \in A_1$ tenemos, con las notaciones usuales,

$$\begin{pmatrix} g(f(x)) \\ 1 \end{pmatrix} = R \begin{pmatrix} f(x) \\ 1 \end{pmatrix} = RN \begin{pmatrix} x \\ 1 \end{pmatrix},$$

donde

$$RN = \begin{pmatrix} CM & Cb+d \\ 0 & 1 \end{pmatrix}.$$

La observación hecha más arriba nos dice que $g \circ f$ es una afinidad con matriz RN. □

Corolario 4.2 *Si* $f : A \longrightarrow A$ *es una afinidad biyectiva con matriz* N, *entonces* f^{-1} *es una afinidad biyectiva con matriz* N^{-1}. □

Consideremos ahora el caso particular de la aplicación identidad

$$I : A \longrightarrow A.$$

Si fijamos el mismo sistema de referencia $\{p;\ e_1, \ldots, e_n\}$ para escribir las coordenadas de los puntos $x \in A$ y de sus imágenes, la matriz de I es, claramente, la matriz identidad. Si, al contrario, consideramos las coordenadas de los puntos $x \in A$ en un sistema $\{q;\ v_1, \ldots, v_n\}$ y las de sus imágenes $f(x)$ en otro sistema $\{p;\ e_1, \ldots, e_n\}$, obtenemos una matriz

$$\begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix},$$

donde $V = (v_i^j)$ es la matriz de $\tilde{I}$:

$$\tilde{I}(v_i) = v_i = \sum_{j=1}^{n} v_i^j e_j$$

y $q = \begin{pmatrix} q^1 \\ \vdots \\ q^n \end{pmatrix}$ son las coordenadas de $I(q) = q$ en el sistema $\{p;\ e_1, \ldots, e_n\}$.
Obtenemos así

$$\begin{pmatrix} \bar{x} \\ 1 \end{pmatrix} = \begin{pmatrix} V & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix},$$

que nos da las coordenadas $\bar{x}$ de $I(x) = x$ en el sistema $\{p;\ e_1, \ldots, e_n\}$ a partir de las coordenadas x del punto $x \in A$ en el sistema $\{q;\ v_1, \ldots, v_n\}$. Este resultado lo habíamos obtenido ya en (IX.8).

Supongamos ahora que

$$N = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$$

es la matriz de $f : A_1 \longrightarrow A_2$ en los sistemas de referencia $\{p;\ e_1, \ldots, e_n\}$ y $\{q;\ u_1, \ldots, u_m\}$ de A_1 y A_2 respectivamente. ¿Cuál es, entonces, la matriz de f en unos sistemas de referencia distintos $\{p_1;\ v_1, \ldots, v_n\}$, $\{q_1;\ w_1, \ldots, w_m\}$? Consideremos f descompuesta en la forma

$$f = I_{A_2} \circ f \circ I_{A_1} : A_1 \longrightarrow A_1 \longrightarrow A_2 \longrightarrow A_2,$$

donde I_{A_i} es la identidad de A_i, $i = 1, 2$. En cada uno de esos cuatro espacios consideramos respectivamente los sistemas

$$\{p_1;\ v_1, \ldots, v_n\}, \quad \{p;\ e_1, \ldots, e_n\}, \quad \{q;\ u_1, \ldots, u_m\}, \quad \{q_1;\ w_1, \ldots, w_m\}.$$

Si $\overrightarrow{pp_1} = c^1 e_1 + \ldots + c^n e_n$ y $v_i = \sum_{j=1}^{n} r_i^j e_j$, $i = 1, \ldots, n$, la matriz de I_{A_1} es

$$\begin{pmatrix} R & c \\ 0 & 1 \end{pmatrix}.$$

Si $\overrightarrow{qq_1} = d^1 u_1 + \ldots + d^m u_m$ y $w_i = \sum_{j=1}^{m} s_i^j u_j$, $i = 1, \ldots, m$, la matriz de I_{A_2} es

$$\begin{pmatrix} S & d \\ 0 & 1 \end{pmatrix}^{-1}.$$

La matriz de f en los sistemas $\{p_1;\ v_1,\ldots,v_n\}$, $\{q_1;\ w_1,\ldots,w_n\}$ es pues

$$\begin{pmatrix} S & d \\ 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} R & c \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} S^{-1}MR & e \\ 0 & 1 \end{pmatrix},$$

donde $e = S^{-1}(Mc+b-d)$.

X.5 El grupo afín

Dado un espacio afín A de dimensión n, denotaremos por $GA(A)$ o simplemente $GA(n)$ el grupo de las aplicaciones afines biyectivas de A en A con la composición.

Ejercicio:

Demostrar que, si $A_1 \cong A_2$ como espacios afines, entonces $GA(A_1)$ y $GA(A_2)$ son isomorfos como grupos.

Este hecho justifica la notación $GA(n)$, ya que todos los espacios afines de dimensión n dan lugar a grupos isomorfos.

Consideremos la aplicación

$$\begin{array}{rccc} \Phi: & GA(n) & \longrightarrow & GL(n) \\ & f & \longmapsto & \tilde{f} \end{array}$$

donde $GL(n)$ es el *grupo lineal de orden* n, es decir, el conjunto de automorfismos de E con la composición. Por (1.3) Φ es exhaustiva y por (2.1) su núcleo está formado por las traslaciones $\mathcal{T}$. Además, Φ es un morfismo de grupos. El teorema de isomorfismo de grupos (III.5.2) nos dice entonces que

$$GA(n)/\mathcal{T} \cong GL(n).$$

En cada clase de equivalencia respecto a $\mathcal{T}$ están todas las afinidades con la misma aplicación lineal asociada. Así pues, si

$$\begin{cases} \bar{x}^1 = a_1^1 x^1 + \ldots + a_n^1 x^n + b^1 \\ \qquad\qquad \ldots\ldots\ldots \\ \bar{x}^n = a_1^n x^1 + \ldots + a_n^n x^n + b^n \end{cases}$$

son las ecuaciones de una afinidad $f : A \longrightarrow A$ en un determinado sistema de referencia, la clase de f está formada por todas las afinidades con ecuaciones

que difieren de éstas solamente en los "términos independientes" $b^1, \ldots, b^n$. En particular, una afinidad de la clase es la afinidad g dada por

$$\begin{cases} \bar{x}^1 = a_1^1 x^1 + \ldots + a_n^1 x^n \\ \qquad\qquad \ldots\ldots\ldots \\ \bar{x}^n = a_1^n x^1 + \ldots + a_n^n x^n \end{cases},$$

que transforma el punto $(0, \ldots, 0)$ en el punto $(0, \ldots, 0)$. Es decir, si hemos escogido un sistema de referencia único para escribir todos los puntos, esta afinidad g deja fijo el origen del sistema. La afinidad original f se obtiene componiendo g con una traslación de vector $v = (b^1, \ldots, b^n)$,

$$f = T_v \circ g.$$

Observemos, sin embargo, que $f \neq g \circ T_v$.

Ejemplo:

Sea A un espacio afín real tridimensional y sean

$$\begin{cases} \bar{x} = x + 2y + z + 1 \\ \bar{y} = y - z + 2 \\ \bar{z} = x + 3z - 1 \end{cases}$$

las ecuaciones de nna afinidad $f : A \longrightarrow A$ en un cierto sistema de referencia $\{p;\ e_1, e_2, e_3\}$.

Estudiemos primero $\tilde{f} : E \longrightarrow E$. Su matriz es

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 3 \end{pmatrix}.$$

$\tilde{f}$ tiene tres valores propios 0, 2, 3 y sus subespacios de vectores propios son, respectivamente,

$$E_0 = \operatorname{Nuc} \tilde{f} = \langle (-3, 1, 1) \rangle, \quad E_2 = ((1, 1, -1)\rangle, \quad E_3 = \langle (0, 1, -2) \rangle.$$

La imagen de un vector $u \in E$ se obtiene fácilmente descomponiéndolo en suma de vectores de E_0, E_2 y E_3: $u = u_0 + u_2 + u_3$. Entonces $\tilde{f}(u) = 2u_2 + 3u_3$. En particular, observemos que $\operatorname{Im} \tilde{f} = E_2 \oplus E_3$.

Supongamos ahora que $g : A \longrightarrow A$ es la afinidad de ecuaciones

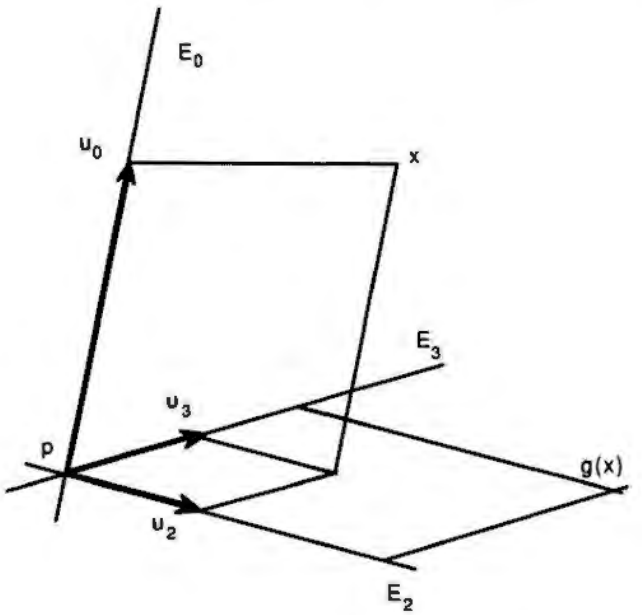
$$\begin{cases} \bar{x} = x + 2y + z \\ \bar{y} = y - z \\ \bar{z} = x + 3z. \end{cases}$$

Es decir, $\tilde{g} = \tilde{f}$ y $g(p) = p$. Entonces, para todo $x \in A$,

$$g(x) = g(p) + \tilde{g}(\overrightarrow{px}) = p + \tilde{g}(\overrightarrow{px})$$

y si $\overrightarrow{px} = u_0 + u_2 + u_3$ con $u_i \in E_i$, $i = 0, 2, 3$,

$$g(x) = p + 2u_2 + 3u_3. \qquad \text{(Ver la figura.)}$$



La afinidad inicial f se obtiene componiendo g con la traslación de vector $(1, 2, -1) = f(p)$.

Estudiemos los puntos fijos de f. Serán las soluciones del sistema

$$\begin{cases} x = x + 2y + z + 1 \\ y = y - z + 2 \\ z = x + 3z - 1. \end{cases}$$

Existe, por tanto, un único punto fijo que es $q = (-3, -\frac{3}{2}, 2)$. Entonces, para todo $x \in A$,

$$f(x) = f(q) + \tilde{f}(\overrightarrow{qx}) = q + \tilde{f}(\overrightarrow{qx})$$

y si $\overrightarrow{qx} = u_0 + u_2 + u_3$ con $u_i \in E_i$, $i = 0, 2, 3$,

$$f(x) = q + 2u_2 + 3u_3.$$

La situación es pues la misma que para g, sustituyendo p por q.

X.6 Variedades invariantes

Dada una afinidad $f : A \longrightarrow A$ de un espacio afín (A, E) en sí mismo, diremos que la variedad lineal $q + F$ es *invariante* o *doble* por f si

$$f(q + F) \subset q + F.$$

Por (3.1), $f(q + F) = f(q) + \tilde{f}(F)$. Resulta, pues, que la variedad $q + F$ es invariante si y sólo si

1. $\tilde{f}(F) \subset F$,
2. $\overrightarrow{qf(q)} \in F$.

La primera condición nos dice que las direcciones de las variedades dobles son subespacios invariantes. En particular, si la variedad es una recta y $F = \langle u \rangle$, es necesario que u sea un vector propio para que la recta pueda ser doble.

Ejemplo:

Estudiemos las rectas invariantes por la afinidad

$$\begin{cases} \bar{x} = -2y + 1 \\ \bar{y} = x + 3y - 1. \end{cases}$$

La aplicación lineal asociada tiene dos valores propios: 1 y 2. Los subespacios de vectores propios son $E_1 = \langle (2, -1) \rangle$, $E_2 = \langle (1, -1) \rangle$. Cualquier recta invariante ha de tener como dirección uno de estos dos subespacios. Además, si q es un punto de la recta, $\overrightarrow{qf(q)} \in E_i$, $i = 1$ ó 2. Si las coordenadas de q son (x_0, y_0), las de $\overrightarrow{qf(q)}$ son

$$(\bar{x}_0 - x_0, \bar{y}_0 - y_0) = (-x_0 - 2y_0 + 1, x_0 + 2y_0 - 1).$$

Este vector siempre es de E_2. Por tanto, todas las rectas con esta dirección son invariantes. En cambio, $\overrightarrow{qf(q)}$ es de E_1 sólo si se cumple $x_0 + 2y_0 - 1 = 0$. Es decir, la única recta invariante con dirección E_1 es la recta

$$x + 2y - 1 = 0.$$

Un punto es una variedad lineal de dimensión 0, y es "invariante" si y sólo si se transforma en sí mismo; es decir, si es un punto fijo. Los puntos fijos de una afinidad f de ecuación

$$\bar{x} = Mx + b$$

cumplen $x = Mx + b$; es decir,

$$(M - I)x + b = 0.$$

Este sistema de ecuaciones representa la variedad de puntos fijos.

Ejemplo:

La variedad de puntos fijos de la afinidad del ejemplo anterior está dada por el sistema

$$\begin{cases} -x - 2y + 1 = 0 \\ x + 2y - 1 = 0 \, . \end{cases}$$

Se trata, pues, de la recta $x + 2y - 1 = 0$. Observemos que, naturalmente, esta es una de las rectas invariantes de la afinidad.

El sistema $(M - I)x + b = 0$, que da los puntos fijos, tiene solución única si y sólo si $\det(M - I) \neq 0$. Esto demuestra la siguiente proposición:

Proposición 6.1 *Una afinidad tiene un único punto fijo si y sólo si la aplicación lineal asociada no tiene el valor propio* 1. □

Acabaremos este apartado dando un método para calcular los hiperplanos invariantes por una afinidad biyectiva f de ecuaciones

$$\begin{cases} \bar{x}^1 = a_1^1 x^1 + \ldots + a_n^1 x^n + b^1 \\ \qquad\qquad \ldots\ldots\ldots \\ \bar{x}^n = a_1^n x^1 + \ldots + a_n^n x^n + b^n \, . \end{cases}$$

Consideremos un hiperplano H de ecuación $c_1 x^1 + \ldots + c_n x^n + c = 0$. Un punto $(x^1, \ldots, x^n)$ se aplica en H si y sólo si

$$0 = c_1 \bar{x}^1 + \ldots + c_n \bar{x}^n + c = c_1(a_1^1 x^1 + \ldots + a_n^1 x^n + b^1) + \ldots$$
$$\ldots + c_n(a_1^n x^1 + \ldots + a_n^n x^n + b^n) + c =$$
$$= (c_1 a_1^1 + \ldots + c_n a_1^n)x^1 + \ldots + (c_1 a_n^1 + \ldots + c_n a_n^n)x^n + c_1 b^1 + \ldots + c_n b^n + c.$$

Esta es, pues, la ecuación del hiperplano $f^{-1}(H)$. H es invariante cuando $H = f^{-1}(H)$; las ecuaciones de H y $f^{-1}(H)$ han de tener, por tanto, coeficientes proporcionales:

$$\frac{c_1}{c_1 a_1^1 + \ldots + c_n a_1^n} = \ldots = \frac{c_n}{c_1 a_n^1 + \ldots + c_n a_n^n} = \frac{c}{c_1 b^1 + \ldots + c_n b^n + c} \, .$$

Son hiperplanos invariantes por f todos los que tienen ecuaciones con coeficientes que cumplen estas igualdades.

Ejemplo:

Volvamos a calcular las rectas invariantes de la afinidad del primer ejemplo de este apartado:

$$\begin{cases} \bar{x} = -2y + 1 \\ \bar{y} = x + 3y - 1. \end{cases}$$

Una recta $ax + by + c = 0$ será invariante si coincide con la de ecuación

$$\begin{aligned} 0 = a\bar{x} + b\bar{y} + c &= a(-2y + 1) + b(x + 3y - 1) + c = \\ &= bx + (-2a + 3b)y + a - b + c. \end{aligned}$$

Es decir, si

$$\frac{a}{b} = \frac{b}{-2a + 3b} = \frac{c}{a - b + c}.$$

Claramente, $a = 0$ si y sólo si $b = 0$. Pero a y b no pueden ser simultáneamente cero; por tanto, ninguna de ellas lo es. Pongamos $k = a/b$. La primera igualdad da:

$$k = \frac{1}{-2k + 3} \quad \Leftrightarrow \quad 2k^2 - 3k + 1 = 0 \quad \Leftrightarrow \quad k = 1, \frac{1}{2}.$$

Si $k = 1$, $a = b$ y las dos igualdades se cumplen para todo c. Así pues, todas las rectas $ax + ay + c = 0$ son invariantes. Si $k = 1/2$, $b = 2a$ y, por tanto, $c = -a$. La recta $ax + 2ay - a = 0$, $a \neq 0$, es invariante.

X.7 Clasificación de las afinidades de un espacio afín A en sí mismo

Hay muchas maneras de clasificar las afinidades y cada una de ellas corresponde a una relación de equivalencia (I.4). En este capítulo vamos a clasificarlas del siguiente modo: dos afinidades $f, g : A \longrightarrow A$ son *de la misma clase* si y sólo si existe un isomorfismo de espacios afines $\varphi : A \longrightarrow A$ tal que el diagrama

$$\begin{array}{ccc} A & \xrightarrow{f} & A \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ A & \xrightarrow{g} & A \end{array}$$

es conmutativo; es decir, $\varphi \circ f = g \circ \varphi$. Es fácil ver que esta relación es de equivalencia.

Las propiedades comunes a todas las afinidades de una misma clase se llaman *propiedades afines* de la afinidad. Son propiedades afines, por ejemplo, la dimensión de la variedad imagen, la dimensión de la variedad de puntos fijos, ser una simetría ($f^2 = I$), ser una proyección ($f^2 = f$), etc.. También son invariantes dentro de cada clase los valores de $\det \tilde{f}$ y $\operatorname{tr} \tilde{f}$, ya que si $\varphi \circ f = g \circ \varphi$ entonces $\tilde{g} = \tilde{\varphi} \circ \tilde{f} \circ \tilde{\varphi}^{-1}$.

Proposición 7.1 *Dos afinidades $f, g : A \longrightarrow A$ son de la misma clase si y sólo si existen dos sistemas de referencia tales que las ecuaciones de f en uno de ellos, $\{p;\ e_1, \ldots, e_n\}$, coinciden con las ecuaciones de g en el otro, $\{q;\ v_1, \ldots, v_n\}$.*

DEMOSTRACIÓN: Supongamos que f y g son de la misma clase y sea $\varphi : A \longrightarrow A$ un isomorfismo tal que $\varphi \circ f = g \circ \varphi$. Consideremos un sistema de referencia cualquiera de A, $\{p;\ e_1, \ldots, e_n\}$, y sea N la matriz de f en este sistema. Por ser $\tilde{\varphi}$ un isomorfismo de espacios vectoriales, $\tilde{\varphi}(e_1), \ldots, \tilde{\varphi}(e_n)$ es una base. La matriz de φ considerando en el primer espacio el sistema $\{p;\ e_1, \ldots, e_n\}$ y en el segundo espacio el sistema $\{\varphi(p);\ \tilde{\varphi}(e_1), \ldots, \tilde{\varphi}(e_n)\}$ es, claramente, la matriz identidad I. Entonces, si R es la matriz de g en el sistema $\{\varphi(p);\ \tilde{\varphi}(e_1), \ldots, \tilde{\varphi}(e_n)\}$, de la igualdad $\varphi \circ f = g \circ \varphi$ resulta

$$N = IN = RI = R.$$

Esto demuestra una implicación.

Supongamos ahora que N es la matriz de f en el sistema $\{p;\ e_1, \ldots, e_n\}$ y también la matriz de g en otro sistema $\{q;\ v_1, \ldots, v_n\}$. Consideremos la afinidad $\varphi : A \longrightarrow A$ tal que $\varphi(p) = q$, $\tilde{\varphi}(e_i) = v_i$, $i = 1, \ldots, n$. Claramente, φ es un isomorfismo y su matriz, tomando el sistema $\{p;\ e_1, \ldots, e_n\}$ en el primer espacio y el sistema $\{q;\ v_1, \ldots, v_n\}$ en el segundo, es la matriz identidad I. La igualdad $\varphi \circ f = g \circ \varphi$ se deduce entonces de $IN = NI$. □

Ejemplo:

Dos traslaciones T_u, T_v de vectores no nulos son siempre de la misma clase. Para demostrarlo, necesitamos un isomorfismo que cumpla $\varphi \circ T_u(a) = T_v \circ \varphi(a)$ para todo a. Esto es, tal que $\varphi(a) + \tilde{\varphi}(u) = \varphi(a) + v$. Es suficiente, pues, escoger un isomorfismo de espacios vectoriales $\tilde{\varphi}$ que transforme u en v, y cualquier afinidad φ con esa aplicación lineal asociada será un isomorfismo como el buscado.

Las ecuaciones de T_u, $u \neq \vec{0}$, en una referencia $\{p;\ u, e_2, \ldots, e_n\}$ son

$$\begin{cases} \bar{x}^1 = x^1 + 1 \\ \bar{x}^i = x^i \qquad i = 2, \ldots, n. \end{cases}$$

Cualquier traslación, en un sistema de referencia conveniente, tiene estas ecuaciones.

En los apartados siguientes clasificaremos las afinidades de la recta y del plano afín. Para hacerlo nos basaremos en la proposición (7.1): buscaremos los diferentes tipos de ecuaciones que pueden tener las afinidades en cada caso. Comprobaremos que no se puede pasar de uno de esos tipos a otro por un cambio de sistema de referencia; es decir, que corresponden a afinidades de clases diferentes. Entonces cada uno de los tipos de ecuaciones corresponderá a una clase de afinidades.

X.8 Afinidades de la recta afín

Sea A un espacio afín de dimensión 1, y $f : A \longrightarrow A$ una afinidad. $\tilde{f}$ es un endomorfismo de un espacio vectorial de dimensión 1 y, por tanto (V.5.1), es de la forma $\tilde{f} = aI$. En cualquier base la matriz de $\tilde{f}$ es (a) y $\det \tilde{f} = a$. De ahí que afinidades con parámetros a distintos sean de clases diferentes.

Examinemos los casos posibles. Si $a = 0$, $\tilde{f} = 0$ y f transforma todo A en un punto:

$$f(a) = f(p) + \tilde{f}(\overrightarrow{pa}) = f(p) \qquad \forall a.$$

Tomando como origen del sistema de referencia este punto $q = f(a)$ y una base cualquiera del espacio vectorial asociado, obtenemos la ecuación de f

$$\bar{x} = 0.$$

Si $a = 1$, $\tilde{f} = I$ y, por tanto, f es una traslación. Si el vector traslación es $u \neq \vec{0}$, en un sistema $\{p;\ u\}$ (p cualquiera), la ecuación de f es

$$\bar{x} = x + 1.$$

Si el vector traslación es $\vec{0}$, f es la identidad y, en cualquier sistema de referencia, su ecuación es

$$\bar{x} = x.$$

Si $a \neq 1$, f es una homotecia de razón a y tiene solamente un punto fijo (§2). Tomando este punto como origen y una base cualquiera del espacio vectorial asociado, la ecuación de f es

$$\bar{x} = ax.$$

Observemos que hemos obtenido una clase para cada valor de $a = \det \tilde{f}$, excepto en el caso $a = 1$, que corresponde a dos clases diferentes.

X.9 Afinidades del plano afín

Sea A un espacio afín de dimensión 2. Realizaremos el estudio de las clases de afinidades de A en tres etapas:

- Afinidades con una recta de puntos fijos.
- Afinidades sin ningún punto fijo.
- Afinidades con un único punto fijo.

Supongamos, primero, que $f : A \longrightarrow A$ tiene una recta $q + \langle u \rangle$ de puntos fijos. En un sistema de referencia $\{q;\ u, v\}$ (v cualquiera que forme base con u), las ecuaciones de f son

$$\begin{cases} \bar{x} = x + by \\ \bar{y} = ny. \end{cases}$$

Si $n \neq 1$, n es otro valor propio de $\tilde{f}$ y podemos escoger como segundo vector de la base un vector v de valor propio n. Las ecuaciones de f son

$$\begin{cases} \bar{x} = x \\ \bar{y} = ny. \end{cases}$$

La afinidad se llama entonces *homología general de razón n*. Observemos que $n = \det \tilde{f}$ y, por tanto, homologías generales de razones diferentes pertenecen a clases diferentes. Es fácil ver que una homología general f deja invariantes todas las rectas de dirección $\langle v \rangle$ y la recta de puntos fijos $q + \langle u \rangle$. Por otra parte, la imagen de un punto $a = (x, y)$ es $f(a) = (\bar{x}, \bar{y}) = (x, ny)$. La recta determinada por a y $f(a)$ se interseca con la recta de puntos fijos en un punto $b = (x, 0)$, que cumple $\overrightarrow{bf(a)} = n\,\overrightarrow{ba}$. Es decir,

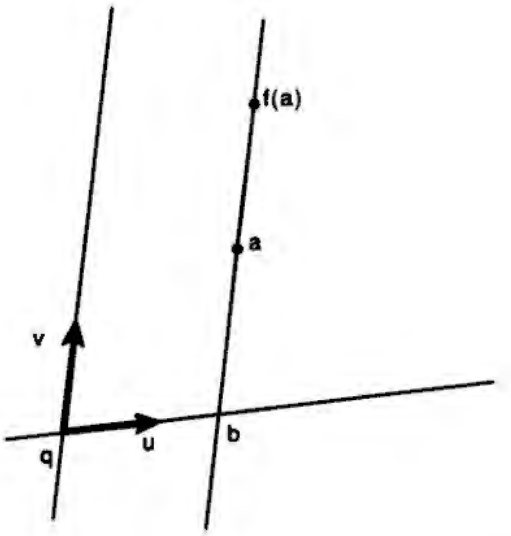
$$(b, a, f(a)) = n.$$

Consideremos ahora el caso en que $n = 1$. El único valor propio de la aplicación lineal asociada $\tilde{f}$ es 1. Si este valor propio tiene multiplicidad 2, $\tilde{f} = I$ y f es una traslación con una recta de puntos fijos; es decir, $f = I$. Sus ecuaciones en cualquier sistema son:

$$\begin{cases} \bar{x} = x \\ \bar{y} = y. \end{cases}$$

Si la multiplicidad del valor propio 1 es 1, las ecuaciones de f en el sistema $\{q;\ u, v\}$ son

$$\begin{cases} \bar{x} = x + by \\ \bar{y} = y \end{cases}$$



con $b \neq 0$. Por tanto, en la referencia $\{q;\ bu, v\}$, las ecuaciones de f son

$$\begin{cases} \bar{x} = x + y \\ \bar{y} = y. \end{cases}$$

La afinidad se llama entonces una *homología especial*. Todas las homologías especiales pertenecen a la misma clase. Sus rectas invariantes son todas las de dirección $\langle u \rangle$; una de ellas, $q + \langle u \rangle$, es recta de puntos fijos.

Supongamos ahora que $f : A \longrightarrow A$ no tiene ningún punto fijo. Entonces $\tilde{f}$ tiene el valor propio 1 (6.1). Sea $u \neq \vec{0}$ un vector propio de valor propio 1. Las ecuaciones de f en un sistema $\{p;\ u, v\}$ (p y v cualesquiera) son del tipo

$$\begin{cases} \bar{x} = x + by + c \\ \bar{y} = ny + d. \end{cases}$$

Si $n \neq 1$, n es otro valor propio de $\tilde{f}$ y podemos escoger, como segundo vector de la base, un vector de valor propio n. Las ecuaciones serán entonces

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = ny + d \end{cases}$$

con $c \neq 0$ para que no haya puntos fijos. La única recta invariante tiene ecuación

$$y = \frac{d}{1-n}.$$

Si escogemos el origen p del sistema de referencia sobre esta recta, digamos $p = (x_0, \dfrac{d}{1-n})$, resulta que $\overrightarrow{pf(p)} = (c, 0) = cu$. De ahí resulta que en el sistema de referencia $\{p;\ cu, v\}$ las ecuaciones de f son

$$\begin{cases} \bar{x} = x + 1 \\ \bar{y} = ny, \end{cases}$$

$n = \det \tilde{f}$ y, por tanto, hay tantas clases diferentes como parámetros $n \neq 1$. Esas afinidades son *homologías generales seguidas de una traslación* de dirección la de la recta de puntos fijos de la homología. Esta composición es además conmutativa.

Si $n = 1$, el único valor propio es 1. Si es de multiplicidad 2, $\tilde{f} = I$ y la afinidad es una traslación de vector $w \neq \vec{0}$. En una referencia del tipo $\{p;\ w, v\}$, las ecuaciones de f son

$$\begin{cases} \bar{x} = x + 1 \\ \bar{y} = y. \end{cases}$$

Si la multiplicidad del valor propio 1 es 1, en las ecuaciones de f en el sistema $\{p;\ u, v\}$ (u vector propio de valor propio 1)

$$\begin{cases} \bar{x} = x + by + c \\ \bar{y} = y + d \end{cases}$$

b ha de ser diferente de cero. Además, puesto que no existen puntos fijos, $d \neq 0$. En estas circunstancias, la afinidad no tiene tampoco ninguna recta invariante. Podemos escoger, sin embargo, el vector $\overrightarrow{pf(p)} = cu + dv = w$ como segundo vector de la base; entonces

$$\tilde{f}(w) = c\tilde{f}(u) + d\tilde{f}(v) = cu + d(bu + v) = dbu + w.$$

De ahí resulta que en el sistema de referencia $\{p;\ dbu, w\}$ las ecuaciones de f son

$$\begin{cases} \bar{x} = x + y \\ \bar{y} = y + 1. \end{cases}$$

Se trata, pues, de una *homología especial seguida de una traslación* de dirección diferente de la del haz de rectas invariantes.

Supongamos, finalmente, que $f : A \longrightarrow A$ tiene un único punto fijo q. Consideremos dos casos:

1. Existe un vector u tal que u, $\tilde{f}(u)$ son linealmente independientes. En el sistema de referencia $\{q;\ u, \tilde{f}(u)\}$ las ecuaciones de f son

$$\begin{cases} \bar{x} = by \\ \bar{y} = x + ny. \end{cases}$$

Observemos que $n = \operatorname{tr} \tilde{f}$, $b = -\det \tilde{f}$. Por tanto, hay tantas clases de afinidades de este tipo como parámetros b y n. Estos parámetros no pueden, sin embargo, tomar todos los valores de K: la condición de que f tenga un único punto fijo impone $b + n \neq 1$.

El polinomio característico de $\tilde{f}$ es $x^2 - nx - b$. Vamos a dar expresiones más sencillas de las ecuaciones de f en aquellos casos en que $\tilde{f}$ tiene valores propios.

Si $\tilde{f}$ tiene dos valores propios diferentes ($(n^2+4b) > 0$ cuando $K = \mathbf{R}$), es diagonalizable en una cierta base $\{w, v\}$. Las ecuaciones de f en un sistema $\{q;\ w, v\}$ son del tipo

$$\begin{cases} \bar{x} = ax \\ \bar{y} = cy \qquad\qquad a \neq c. \end{cases}$$

Si $n^2 + 4b = 0$, $\tilde{f}$ tiene un único valor propio a. Sea w un vector propio y $\{w, v\}$ una base. En el sistema $\{q;\ w, v\}$ las ecuaciones de f son

$$\begin{cases} \bar{x} = ax + b'y \\ \bar{y} = n'y \end{cases}$$

con $n' = a$ (en caso contrario n' sería otro valor propio). Además, $b' \neq 0$, ya que en caso contrario $\tilde{f} = aI$ y todos los pares u, $\tilde{f}(u)$ serían linealmente dependientes. Podemos, pues, tomar el sistema de referencia $\{q;\ b'w, v\}$ y obtenemos como ecuaciones de f

$$\begin{cases} \bar{x} = ax + y \\ \bar{y} = ay \qquad\qquad a \neq 1. \end{cases}$$

2. Si para todo vector u $\{u, \tilde{f}(u)\}$ son linealmente dependientes, $\tilde{f}$ ha de ser una homotecia vectorial. En efecto, sea $\{u, v\}$ una base y $\tilde{f}(u) = au$, $\tilde{f}(v) = bv$. Es necesario, pues, que $\tilde{f}(u + v) = au + bv = c(u + v)$, de donde $a = c = b$. Todos los vectores son, por tanto, vectores propios del mismo valor propio y $\tilde{f} = aI$ ($a \neq 1$ porque f tiene un único punto fijo). La afinidad f es una *homotecia de razón* a (§2), y sus ecuaciones en el sistema $\{q;\ u, v\}$ son

$$\begin{cases} \bar{x} = ax \\ \bar{y} = ay. \end{cases}$$

En este caso, $a^2 = \det \tilde{f}$ y, por tanto, afinidades con parámetros a distintos son de clases diferentes.

X.10 Nota histórica

El estudio de las transformaciones adecuadas entre ciertas estructuras adquiere toda su importancia a raíz de la conferencia que Felix Klein (1849–1925) dio en 1872, con motivo de su admisión en la Universidad de Erlangen, con el título "Vergleichende Betrachtungen über neuere geometrische

Forschungen" (Una revisión comparativa de investigaciones recientes en geometría). Los puntos de vista expresados en esa conferencia se conocen hoy como el "Programa de Erlangen".

La idea básica de Klein es que toda geometría puede caracterizarse por un grupo de transformaciones y que la geometría trata esencialmente de los invariantes por ese grupo de transformaciones. La geometría afín queda caracterizada por el grupo de las afinidades (el grupo afín) y no es más que el estudio de los invariantes por este grupo.

X.11 Ejercicios

1. Sea A un plano afín. Demostrar que dadas dos rectas que se cortan y un punto que no pertenece a ninguna de las dos rectas, y dada otra configuración análoga, existen dos afinidades de A que transforman una configuración en la otra. Hallar esas afinidades en el caso

$$r : x - y = 2, \quad s : x - 2y = -1, \quad p = (0,0)$$
$$r' : x = 1, \quad s' : x - y = 1, \quad p' = (2,2).$$

2. Escribir la ecuación de todas las homologías generales de $\mathbf{R}^2$ que tienen el eje de homología paralelo al eje de abscisas.

3. Ecuaciones de las afinidades del plano que transforman las rectas r_1, r_2, r_3 en r_2, r_3, r_1 respectivamente, donde

$$r_1 : x + y = 1, \quad r_2 : x + 2y = 0, \quad r_3 : 4x - y = 2.$$

   Clasificar esas afinidades.

4. Estudiar las afinidades de $\mathbf{R}^2$ que dejan fija la hipérbola $xy = 1$.

5. Sea Z la unión de dos rectas del plano afín A que se cortan. Describir el grupo de las afinidades biyectivas que dejan Z fijo. Explicitar ese grupo en el caso de las rectas del ejercicio 1.

6. Estudiar el grupo de afinidades del plano que dejan fijo un triángulo dado.

7. Dibujar la imagen de los puntos $(0,0), (1,0), (1,1), (0,1)$ para cada una de las formas simplificadas de las afinidades del plano afín que se han obtenido, tomando $K = \mathbf{R}$. Discutir diferentes valores de los parámetros cuando los haya.

8. Determinar el lugar geométrico de las imágenes de un punto dado x por todas las afinidades que tienen una recta dada r de puntos fijos y una recta dada s, que se cruza con r, fija.

9. Demostrar que hay una única afinidad del plano que transforma cada uno de los vértices de un triángulo dado en el punto medio del lado opuesto. Estudiar esa afinidad.

10. Sea f una afinidad de un espacio afín real. Demostrar:

a) Si f^2 tiene algún punto fijo, f también.

b) Si existe un $n \in \mathbf{N}$ tal que f^n tiene algún punto fijo, entonces f también.

11. Sea f una afinidad y $\tilde{f}$ el endomorfismo asociado. Demostrar:

a) $e \in \operatorname{Nuc} \tilde{f}$ si y sólo si todas las rectas de dirección (e) se transforman por f en un punto.

b) e es un vector propio de $\tilde{f}$ de valor propio diferente de cero si y sólo si todas las rectas de dirección (e) se transforman por f en una recta paralela.

12. Consideremos tres rectas concurrentes r, s, t del plano afín ordinario y dos rectas paralelas l, l' que cortan a r, s, t en los puntos a, b, c y a', b', c' respectivamente. Demostrar que $(abc) = (a'b'c')$.

13. En la familia de afinidades del plano de ecuaciones

$$\begin{cases} \bar{x} &= ax + ay + b \\ \bar{y} &= ax + 6y + b^2 \end{cases}$$

hay cuatro homologías, cuyos ejes son los lados de un paralelogramo. Determinar los vértices de ese paralelogramo.

14. Estudiar según los valores del parámetro a las afinidades dadas por las ecuaciones

$$\begin{cases} \bar{x} &= ax + y + z + 1 \\ \bar{y} &= x + ay + z + 1 \\ \bar{z} &= x + y + az + 1. \end{cases}$$

15. Estudiar la afinidad de ecuaciones

$$\begin{cases} \bar{x} &= x - \dfrac{1}{8}y - \dfrac{1}{8} \\ \bar{y} &= 2x - \dfrac{1}{8} \end{cases}$$

y expresarla como producto de una homotecia y una homología.

16. Estudiar todas las afinidades de la forma

$$\begin{cases} \bar{x} &= ax + y + a \\ \bar{y} &= x + ay + a. \end{cases}$$

Determinar el lugar geométrico de las imágenes de un punto dado por todas esas afinidades.

17. Estudiar las afinidades de ecuaciones

$$\begin{cases} \bar{x} &= (1+a)x - ay + 1 \\ \bar{y} &= a^2x + (1 + 2a - 4a^2 + a^3)y \end{cases}$$

que no tienen puntos fijos.

18. Demostrar que un subconjunto del espacio afín K^n es una variedad lineal de dimensión r si y sólo si es el conjunto de ceros de una afinidad exhaustiva $f : K^n \longrightarrow K^{n-r}$.

X.12 Ejercicios para programar

19. Hacer un programa que cambie de sistema de referencia las ecuaciones cartesianas de una afinidad $f : \mathbf{R}^n \longrightarrow \mathbf{R}^m$. (Indicación: seguir el método explicado al final del §4.)

20. Aplicar el ejercicio 19 para cambiar de sistema de referencia las coordenadas de los puntos de $\mathbf{R}^n$. (Indicación: considerar la aplicación identidad $I : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ y sus ecuaciones cartesianas tomando sistemas de referencia diferentes a derecha e izquierda.)

21. Aplicar el ejercicio 19 para cambiar de sistema de referencia las ecuaciones cartesianas de una variedad lineal de $\mathbf{R}^n$. (Indicación: utilizar los ejercicios IX.14 y X.18.)

22. Hacer un programa que permita encontrar las ecuaciones de la variedad de puntos fijos de una afinidad $f : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ dada por sus ecuaciones cartesianas en una cierta referencia. (Indicación: si $f(x) = Ax + b$, reducir por el método de Gauss el sistema de ecuaciones $(A - I)x + b = \vec{0}$.)

23. Sea $f : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ una afinidad dada por $f(x) = Ax + b$. Hacer un programa que calcule:

a) La variedad de puntos fijos de f (ejercicio 22).

b) Los valores propios y los vectores propios reales de A (ejercicio VIII.23).

Siguiendo la clasificación hecha en el §9, se pueden obtener las ecuaciones simplificadas de f y la referencia en que se obtienen.

Capítulo XI

Espacios vectoriales euclídeos y unitarios

XI.1 Formas bilineales y sesquilineales

Sea E un espacio vectorial sobre $\mathbf{R}$. Una aplicación del producto cartesiano $E \times E$ en $\mathbf{R}$

$$\phi : E \times E \longrightarrow \mathbf{R}$$

se llama una *forma bilineal* si cumple

i) $\phi(u_1 + u_2, v) = \phi(u_1, v) + \phi(u_2, v) \quad \forall u_1, u_2, v \in E,$
$\phi(ku, v) = k\phi(u, v) \quad \forall u, v \in E \quad \forall k \in \mathbf{R};$

ii) $\phi(u, v_1 + v_2) = \phi(u, v_1) + \phi(u, v_2) \quad \forall u, v_1, v_2 \in E,$
$\phi(u, kv) = k\phi(u, v) \quad \forall u, v \in E \quad \forall k \in \mathbf{R}.$

Proposición 1.1 *Sea $e_1, \ldots, e_n$ una base del espacio vectorial real E.*

1. *La matriz $B = (b_i^j)$, donde $b_i^j = \phi(e_j, e_i)$, determina ϕ. Más concretamente, si $u = \sum_{i=1}^n u^i e_i$, $v = \sum_{i=1}^n v^i e_i$*

$$\phi(u, v) = u^t B v,$$

donde $v = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$ *y* $u^t = (u^1 \ldots u^n)$ *es la traspuesta de* $u = \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$.

2. *Dada una matriz $B = (b_i^j)$ cualquiera, existe siempre una forma bilineal ϕ tal que $\phi(e_j, e_i) = b_i^j$.*

DEMOSTRACIÓN:

1. Aplicando las condiciones (i) y (ii) de forma bilineal, obtenemos

$$\begin{aligned}\phi(u,v) &= \phi\left(\sum_{i=1}^n u^i e_i, \sum_{j=1}^n v^j e_j\right) = \sum_{i=1}^n u^i \phi\left(e_i, \sum_{j=1}^n v^j e_j\right) = \\ &= \sum_{i=1}^n u^i \left(\sum_{j=1}^n v^j \phi(e_i, e_j)\right) = \sum_{i,j=1}^n u^i b_j^i v^j = u^t B v.\end{aligned}$$

2. Dada B, definimos $\phi : E \times E \longrightarrow \mathbf{R}$ por

$$\phi(u,v) = u^t B v$$

(con las notaciones del enunciado). Es fácil ver que ϕ es bilineal. □

La matriz B de (1.1) se llama *matriz de ϕ en la base $e_1, \ldots, e_n$*. Sea C la matriz de ϕ en otra base $u_1, \ldots, u_n$ y sea $A = (a_i^j)$ la matriz del cambio de base:

$$u_i = \sum_{j=1}^n a_i^j e_j, \qquad i = 1, \ldots, n.$$

Por (1.1), $c_j^i = \phi(u_i, u_j) = a_i^t B a_j$, donde a_i representa la matriz formada por la i-ésima columna de A (es decir, las coordenadas de u_i en la base $e_1, \ldots, e_n$). De ahí que

$$C = A^t B A.$$

Sea E un espacio vectorial sobre los complejos $\mathbf{C}$. Una aplicación $\phi : E \times E \longrightarrow \mathbf{C}$ se llama una *forma sesquilineal* si cumple

i) $\phi(u_1 + u_2, v) = \phi(u_1, v) + \phi(u_2, v) \quad \forall u_1, u_2, v \in E,$
$\phi(ku, v) = k\phi(u, v) \quad \forall u, v \in E \quad \forall k \in \mathbf{C};$

ii) $\phi(u, v_1 + v_2) = \phi(u, v_1) + \phi(u, v_2) \quad \forall u, v_1, v_2 \in E,$
$\phi(u, kv) = \bar{k}\phi(u, v) \quad \forall u, v \in E \quad \forall k \in \mathbf{C}$, donde $\bar{k}$ indica el conjugado de $k \in \mathbf{C}$.

Proposición 1.2 *Sea $e_1, \ldots, e_n$ una base de E (espacio vectorial sobre $\mathbf{C}$).*

1. *La matriz $B = (b_i^j)$, donde $b_i^j = \phi(e_j, e_i)$, determina ϕ. Más concretamente, si $u = \sum_{i=1}^n u^i e_i$, $v = \sum_{i=1}^n v^j e_i$*

$$\phi(u,v) = u^t B \bar{v},$$

donde $\bar{v} = \begin{pmatrix} \bar{v}^1 \\ \vdots \\ \bar{v}^n \end{pmatrix}$ *y* $u^t = (u^1 \ldots u^n)$ *es la traspuesta de* $u = \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$.

2. *Dada una matriz compleja* $B = (b_i^j)$ *cualquiera, existe siempre una forma sesquilineal* ϕ *tal que* $\phi(e_j, e_i) = b_i^j$.

La demostración es una adaptación fácil de la de (1.1). □

La matriz B de (1.2) se llama *matriz de* ϕ *en la base* $e_1, \ldots, e_n$. Sea C la matriz de ϕ en otra base $u_1, \ldots, u_n$ y sea $A = (a_i^j)$ la matriz del cambio de base:

$$u_i = \sum_{j=1}^{n} a_i^j e_j, \qquad i = 1, \ldots, n.$$

Igual que en el caso real, se deduce que

$$C = A^t B \bar{A}.$$

Una forma bilineal en un espacio vectorial real E, $\phi : E \times E \longrightarrow \mathbf{R}$, se llama *simétrica* si $\phi(u, v) = \phi(v, u)$ para todo $u, v \in E$.

Proposición 1.3 *Una forma bilineal es simétrica si y sólo si su matriz es simétrica (una matriz* B *se llama* simétrica *si* $B^t = B$*).*

DEMOSTRACIÓN: Si ϕ es simétrica, $b_j^i = \phi(e_i, e_j) = \phi(e_j, e_i) = b_i^j$. Si B es simétrica, $\phi(u, v) = \sum_{i,j=1}^{n} u^i b_j^i v^j = \sum_{i,j=1}^{n} v^j b_i^j u^i = \phi(v, u)$. □

Una forma sesquilineal $\phi : E \times E \longrightarrow \mathbf{C}$ sobre un espacio vectorial complejo E se llama *hermítica* si $\phi(u, v) = \overline{\phi(v, u)}$ para todo $u, v \in E$.

Proposición 1.4 *Una forma sesquilineal es hermítica si y sólo si su matriz* B *es hermítica (*B *se llama* hermítica *si* $B^t = \bar{B}$*).*

DEMOSTRACIÓN: Se procede como en el caso real. □

XI.2 Producto escalar

Sea E un espacio vectorial real. Un *producto escalar* en E es una forma bilineal simétrica

$$\phi : E \times E \longrightarrow \mathbf{R}$$

que cumple

$$\phi(u, u) \geq 0 \quad \forall u \in E;$$
$$\phi(u, u) = 0 \Leftrightarrow u = \vec{0}.$$

Una forma que cumple estas dos propiedades se llama *definida positiva*.

Sea E un espacio vectorial complejo. Un *producto escalar* en E es una forma sesquilineal hermítica $\phi : E \times E \longrightarrow \mathbf{C}$ definida positiva; es decir, tal que

$$\phi(u,u) \geq 0 \text{ (real positivo)} \qquad \forall u \in E;$$
$$\phi(u,u) = 0 \Leftrightarrow u = \vec{0}.$$

Sea E un espacio vectorial real o complejo con un producto escalar ϕ. Un vector u se llama *unitario* si

$$\phi(u,u) = 1.$$

Dos vectores $u, v \in E$ se llaman *ortogonales* si

$$\phi(u,v) = 0.$$

Observaciones:

1. Si $u \neq 0$, $\dfrac{u}{\sqrt{\phi(u,u)}}$ es unitario. (Indicamos por $\sqrt{\phi(u,u)}$ la determinación positiva de la raíz.)
2. Si S es un conjunto de vectores diferentes de $\vec{0}$ y ortogonales dos a dos, S es linealmente independiente. En efecto, si $\sum \lambda^i v_i = \vec{0}$ con $v_i \in S$, para cada v_k tenemos

$$0 = \phi\left(\sum \lambda^i v_i, v_k\right) = \sum \lambda^i \phi(v_i, v_k) = \lambda^k \phi(v_k, v_k).$$

Pero $\phi(v_k, v_k) \neq 0$, ya que $v_k \neq \vec{0}$. Por tanto, $\lambda^k = 0$ para todo k. Nuestro objetivo inmediato es demostrar que siempre existe una base $u_1, \ldots, u_n$ en la cual la matriz del producto escalar es la matriz identidad; es decir,

$$\phi(u_i, u_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$$

Los vectores $u_1, \ldots, u_n$ son, pues, unitarios y ortogonales dos a dos. Diremos entonces que $u_1, \ldots, u_n$ es una *base ortonormal*.

3. Si $w = w^1 u_1 + \ldots + w^n u_n$ y $v = v^1 u_1 + \ldots + v^n u_n$, donde $u_1, \ldots, u_n$ es una base ortonormal,

$$\phi(w,v) = w^1 v^1 + \ldots + w^n v^n \quad \text{en el caso real;}$$
$$\phi(w,v) = w^1 \bar{v}^1 + \ldots + w^n \bar{v}^n \quad \text{en el caso complejo.}$$

4. Las coordenadas de un vector $v = v^1 u_1 + \ldots + v^n u_n$ en una base ortonormal $u_1, \ldots, u_n$ son

$$v^i = \phi(v, u_i).$$

Proposición 2.1 *Sea E un espacio vectorial de dimensión finita n sobre* $\mathbf{R}$ *o* $\mathbf{C}$, *con un producto escalar ϕ. Existe siempre una base ortonormal de E.*

DEMOSTRACIÓN: Sea $e_1, \ldots, e_n$ una base cualquiera de E. Consideremos los subespacios

$$E_1 = \langle e_1 \rangle \subset E_2 = \langle e_1, e_2 \rangle \subset \ldots \subset E_n = \langle e_1, \ldots, e_n \rangle = E.$$

- E_1 tiene una base ortonormal que es $u_1 = \dfrac{e_1}{\sqrt{\phi(e_1, e_1)}}$.

- Supongamos que $u_1, \ldots, u_r$ es una base ortonormal de E_r. Construyamos una base ortonormal de $E_{r+1} = \langle u_1, \ldots, u_r, e_{r+1} \rangle$ de la siguiente manera: consideremos un vector de la forma

$$u'_{r+1} = e_{r+1} - (k^1 u_1 + \ldots + k^r u_r),$$

ortogonal a cada u_i, $i = 1, \ldots, r$: $0 = \phi(u'_{r+1}, u_i) = \phi(e_{r+1}, u_i) - k^i$. Tenemos que tomar, pues, $k^i = \phi(e_{r+1}, u_i)$, $i = 1, \ldots, r$. La observación 2 nos dice que $u_1, \ldots, u_r, u'_{r+1}$ son linealmente independientes y forman, por tanto, una base de E_{r+1}. Pongamos

$$u_{r+1} = \frac{u'_{r+1}}{\sqrt{\phi(u'_{r+1}, u'_{r+1})}}$$

y $u_1, \ldots, u_r, u_{r+1}$ será una base ortonormal de E_{r+1}. Por inducción obtenemos, así, que $E_n = E$ tiene una base ortonormal. □

Este proceso de construcción de una base ortonormal se conoce con el nombre de *método de Gram-Schmidt.* La proposición siguiente es el recíproco de (2.1):

Proposición 2.2 *Si una forma bilineal o sesquilineal ϕ (sobre* $\mathbf{R}$ *o* $\mathbf{C}$ *respectivamente) tiene la matriz identidad en una base $u_1, \ldots, u_n$, entonces ϕ es un producto escalar.*

DEMOSTRACIÓN: Si $w = w^1u_1 + \ldots + w^nu_n$ y $v = v^1u_1 + \ldots + v^nu_n$, resulta

$$\phi(w,v) = w^1v^1 + \ldots + w^nv^n \quad \text{en el caso real;}$$
$$\phi(w,v) = w^1\bar{v}^1 + \ldots + w^n\bar{v}^n \quad \text{en el caso complejo.}$$

De ahí resultan fácilmente las propiedades que debe cumplir un producto escalar. □

Nota:

A partir de ahora, si no indicamos lo contrario, por comodidad de notación, pensaremos siempre que los escalares son complejos (teniendo en cuenta que $\mathbf{R} \subset \mathbf{C}$), y escribiremos $\bar{k}$ donde sea necesario, bien entendido que $\bar{k} = k$ cuando $k \in \mathbf{R}$.

Proposición 2.3 *Sea ϕ una forma bilineal (o sesquilineal) con matriz B en la base $e_1, \ldots, e_n$. Designemos por B_r el menor formado por las r primeras filas y las r primeras columnas de B. Entonces ϕ es un producto escalar si y sólo si B es simétrica (o hermítica) y $\det B_r > 0$ para todo r.*

DEMOSTRACIÓN: Designemos por E_r el subespacio $\langle e_1, \ldots, e_r \rangle$ y por ϕ_r la restricción de ϕ a $E_r \times E_r$:

$$\begin{array}{rccc} \phi_r : E_r \times E_r & \longrightarrow & \mathbf{R} & (\text{o } \mathbf{C}) \\ (u,v) & \longmapsto & \phi(u,v). & \end{array}$$

La matriz de ϕ_r en la base $e_1, \ldots, e_r$ es precisamente B_r. Supongamos que ϕ es un producto escalar en E. Entonces ϕ_r es un producto escalar en E_r y, por (2.1), existe una base ortonormal de E_r. Si P es la matriz de cambio de la base $e_1, \ldots, e_r$ a la base ortonormal, tenemos

$$I = P^t B_r \bar{P},$$

de donde $1 = \det B_r \cdot |\det P|^2$ y $\det B_r > 0$.

Supongamos ahora que $\det B_r > 0$ para todo r. Vamos a construir una base $u_1, \ldots, u_n$ tal que $\phi(u_i, u_j) = 0$ si $i \neq j$, $\phi(u_i, u_i) = 1$; entonces (2.2) nos asegurará que ϕ es un producto escalar.

Para construir la base "ortonormal", usaremos el mismo método de Gram-Schmidt utilizado en (2.1).

- $\phi(e_1, e_1) = b_1^1 = \det B_1 > 0$. Por tanto, existe $\sqrt{\phi(e_1, e_1)}$ y podemos construir
$$u_1 = \frac{e_1}{\sqrt{\phi(e_1, e_1)}},$$
que es un vector unitario, base de $E_1 = \langle e_1 \rangle$.

- Supongamos que $u_1, \ldots, u_r$ es una base de E_r tal que

$$\phi(u_i, u_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$$

Igual que en (2.1), resulta que el vector

$$u'_{r+1} = e_{r+1} - (k^1 u_1 + \ldots + k^r u_r),$$

con $k^i = \phi(e_{r+1}, u_i)$, es ortogonal a cada uno de los vectores $u_1, \ldots, u_r$. Si demostramos que $\phi(u'_{r+1}, u'_{r+1}) > 0$, el vector unitario

$$u_{r+1} = \frac{u'_{r+1}}{\sqrt{\phi(u'_{r+1}, u'_{r+1})}}$$

será tal que $u_1, \ldots, u_r, u_{r+1}$ formarán una base ortonormal de E_{r+1}. El resultado se obtiene, entonces, por inducción. Calculemos, pues,

$$\begin{aligned}
&\phi(u'_{r+1}, u'_{r+1}) = \\
&= \phi(e_{r+1}, e_{r+1}) - \phi(e_{r+1}, \sum_i k^i u_i) - \\
&\quad -\phi(\sum_i k^i u_i, e_{r+1}) + \phi(\sum_i k^i u_i, \sum_j k^j u_j) = \\
&= \phi(e_{r+1}, e_{r+1}) - \sum_i \bar{k}^i \phi(e_{r+1}, u_i) - \sum_i k^i \phi(u_i, e_{r+1}) + \sum_i k^i \bar{k}^i = \\
&= \phi(e_{r+1}, e_{r+1}) - \sum_i \overline{\phi(e_{r+1}, u_i)} \phi(e_{r+1}, u_i) = \\
&\quad - \sum_i \phi(e_{r+1}, u_i)\phi(u_i, e_{r+1}) + \sum_i \phi(e_{r+1}, u_i)\overline{\phi(e_{r+1}, u_i)} = \\
&= \phi(e_{r+1}, e_{r+1}) - \sum_i \phi(e_{r+1}, u_i)\phi(u_i, e_{r+1}) = \\
&= \begin{vmatrix} 1 & \cdots & 0 & \phi(u_1, e_{r+1}) \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & \phi(u_r, e_{r+1}) \\ \phi(e_{r+1}, u_1) & \cdots & \phi(e_{r+1}, u_r) & \phi(e_{r+1}, e_{r+1}) \end{vmatrix}.
\end{aligned}$$

(Esto se puede comprobar, por ejemplo, desarrollando por la última fila.)

La matriz que aparece aquí es la de ϕ_{r+1} en la base $u_1, \ldots, u_r, e_{r+1}$ y se obtiene, por tanto, de la matriz B_{r+1} por un cambio de base. Tiene, pues, la forma

$$P^t B_{r+1} \bar{P}$$

y su determinante es

$$\det B_{r+1} \, |\det P|^2 > 0,$$

por ser $\det B_{r+1} > 0$. Esto termina la demostración. □

Un *espacio vectorial euclídeo* es un espacio vectorial sobre $\mathbf{R}$ con un producto escalar. Un *espacio vectorial unitario* es un espacio vectorial sobre $\mathbf{C}$ con un producto escalar. En ambos casos se acostumbra a designar $\phi(u, v)$ por $\langle u, v \rangle$ o por $u \cdot v$.

Observación:

Sea E un espacio vectorial sobre $\mathbf{R}$ o $\mathbf{C}$ y sea $u_1, \ldots, u_n$ una base cualquiera de E. Por (2.2), existe siempre un producto escalar ϕ en E con el que $u_1, \ldots, u_n$ es una base ortonormal:

$$\phi(w, v) = w^1 \bar{v}^1 + \ldots + w^n \bar{v}^n,$$

donde $(w^1, \ldots, w^n)$, $(v^1, \ldots, v^n)$ son las coordenadas de w, v en la base $u_1, \ldots, u_n$.

En $\mathbf{R}^n$ y $\mathbf{C}^n$ consideraremos como producto escalar estándar aquel con el que la base $(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$ es ortonormal.

XI.3 Norma

Sea E un espacio vectorial sobre $\mathbf{R}$ o $\mathbf{C}$. Una *norma* en E es una aplicación

$$\begin{array}{rccl} \| \ \| : & E & \longrightarrow & \mathbf{R} \quad (\text{¡siempre en } \mathbf{R}!) \\ & v & \longmapsto & \|v\| \end{array}$$

que cumple

1. $\|v\| = 0 \Leftrightarrow v = \vec{0}$;
2. $\|kv\| = |k| \cdot \|v\|$;
3. $\|u + v\| \leq \|u\| + \|v\|$ (*desigualdad triangular*).

$|k|$ indica el valor absoluto si $k \in \mathbf{R}$ y el módulo si $k \in \mathbf{C}$.

Sea E un espacio vectorial con un producto escalar ϕ. Utilizaremos la notación

$$\phi(u, v) = u \cdot v.$$

Sabemos que $u \cdot u$ es siempre real positivo; designemos por $\sqrt{u \cdot u}$ la determinación positiva de la raíz cuadrada de $u \cdot u$. La aplicación $u \mapsto \sqrt{u \cdot u}$ es una norma en E. Para demostrarlo necesitamos un lema.

Lema 3.1 (Desigualdad de Cauchy-Schwarz)

$$|u \cdot v|^2 \leq (u \cdot u)(v \cdot v) \qquad \forall u, v \in E.$$

DEMOSTRACIÓN: Si $v = \vec{0}$, la desigualdad es cierta. Supongamos $v \neq \vec{0}$ y consideremos

$$k = \frac{u \cdot v}{v \cdot v}.$$

Entonces

$$\begin{aligned} 0 \;\leq\; & (u - kv) \cdot (u - kv) = u \cdot u - k(v \cdot u) - \bar{k}(u \cdot v) + k\bar{k}(v \cdot v) = \\ = \; & u \cdot u - \frac{(u \cdot v)(v \cdot u)}{(v \cdot v)} - \frac{\overline{(u \cdot v)}(u \cdot v)}{(v \cdot v)} + \frac{(u \cdot v)\overline{(u \cdot v)}}{(v \cdot v)} = \\ = \; & u \cdot u - \frac{(u \cdot v)\overline{(u \cdot v)}}{(v \cdot v)} = u \cdot u - \frac{|u \cdot v|^2}{(v \cdot v)}, \end{aligned}$$

de donde $|u \cdot v|^2 \leq (u \cdot u)(v \cdot v)$. □

Proposición 3.2 *Sea E un espacio vectorial con un producto escalar. La aplicación*

$$\begin{array}{rccc} \| \; \| : & E & \longrightarrow & \mathbf{R} \\ & v & \longmapsto & \sqrt{v \cdot v} \end{array}$$

es una norma.

DEMOSTRACIÓN: La condición 1 de norma resulta de que el producto escalar es definido positivo. Para probar 2, observemos que $(kv) \cdot (kv) = k\bar{k}(v \cdot v) = |k|^2 (v \cdot v)$.

Para demostrar 3, hacemos

$$\begin{aligned} (u+v) \cdot (u+v) \; & = u \cdot u + u \cdot v + v \cdot u + v \cdot v = \\ & = u \cdot u + v \cdot v + (u \cdot v + \overline{u \cdot v}) = \\ & \leq u \cdot u + v \cdot v + 2\,|u \cdot v| \leq \text{ por (3.1)} \\ & \leq u \cdot u + v \cdot v + 2\sqrt{u \cdot u}\sqrt{v \cdot v} = (\sqrt{u \cdot u} + \sqrt{v \cdot v})^2, \end{aligned}$$

de donde $\sqrt{(u+v) \cdot (u+v)} \leq \sqrt{u \cdot u} + \sqrt{v \cdot v}$. □

XI.4 Producto escalar y espacio dual

Sea E un espacio vectorial euclídeo o unitario de dimensión finita. Para todo $v \in E$, la aplicación

$$\begin{array}{rccc} v^* : & E & \longrightarrow & \mathbf{R} \text{ (o } \mathbf{C}) \\ & u & \longmapsto & u \cdot v \end{array}$$

es lineal. Podemos definir, pues,

$$\begin{array}{rccc} \varphi : & E & \longrightarrow & E' \\ & v & \longmapsto & v^*, \end{array}$$

que cumple las siguientes propiedades:

a) φ es inyectiva, ya que $u^* = v^* \Rightarrow u^*(w) = v^*(w)\ \forall w \Rightarrow w \cdot u = w \cdot v\ \forall w \Rightarrow w \cdot (u - v) = 0\ \forall w \Rightarrow u - v = \vec{0} \Rightarrow u = v$.

b) φ es exhaustiva. En efecto, dado $w \in E'$, consideremos una base ortonormal $u_1, \ldots, u_n$ y el vector

$$u = \overline{w(u_1)}u_1 + \ldots + \overline{w(u_n)}u_n.$$

El vector u es una antiimagen de w, ya que $u^*(u_i) = u_i \cdot u = w(u_i)$ para todo i; es decir,

$$u^* = w.$$

c) $\varphi(v+u) = \varphi(v) + \varphi(u)$, ya que $(v+u)^*(w) = w(v+u) = w \cdot v + w \cdot u = v^*(w) + u^*(w) = (v^* + u^*)(w)\ \forall w$ y, por tanto, $(v+u)^* = v^* + u^*$.

d) En el caso euclídeo, $\varphi(kv) = k\varphi(v)$, ya que $(kv)^*(w) = w \cdot (kv) = k(w \cdot v) = k(v^*(w)) = (kv^*)(w)\ \forall w$, de donde $(kv)^* = kv^*$.

d') En el caso unitario, $\varphi(kv) = \bar{k}\varphi(v)$, ya que $(kv)^*(w) = w \cdot (kv) = \bar{k}(w \cdot v) = \bar{k}(v^*(w)) = (\bar{k}v^*)(w)\ \forall w$, de donde $(kv)^* = \bar{k}v^*$.

Hemos demostrado, en particular, el

Teorema 4.1 *Si E es un espacio euclídeo, la aplicación*

$$\begin{array}{rccc} \varphi : & E & \longrightarrow & E' \\ & v & \longmapsto & v^* \end{array}$$

es un isomorfismo canónico. □

Ejercicio:

Demostrar que, si $u_1, \ldots, u_n$ es una base ortonormal, $u_1^*, \ldots, u_n^*$ es su base dual en E'.

XI.5 Subespacios ortogonales

Sea E un espacio vectorial euclídeo o unitario de dimensión finita y S un subconjunto de E. Denominaremos *subespacio ortogonal* de S a

$$S^{\perp} = \{v \in E \mid u \cdot v = 0 \quad \forall u \in S\}.$$

Proposición 5.1 *Se cumple*

1. $S^{\perp}$ *es un subespacio vectorial de* E;
2. $S \subset R \Rightarrow R^{\perp} \subset S^{\perp}$;
3. $S^{\perp} = \langle S \rangle^{\perp}$;
4. $\langle S \rangle \cap S^{\perp} = \{\vec{0}\}$;
5. $\langle S \rangle \subset (S^{\perp})^{\perp}$.

Ejercicio:

Demostrar (5.1).

Proposición 5.2 *Si* F *es un subespacio vectorial de* E, *entonces*

$$E = F \oplus F^{\perp}.$$

DEMOSTRACIÓN: De la propiedad 4 de (5.1) se deduce que $F \cap F^{\perp} = \{\vec{0}\}$.

Sea $u_1, \ldots, u_n$ una base ortonormal de F. Completémosla hasta obtener una base $u_1, \ldots, u_r, e_{r+1}, \ldots, e_n$ de E y apliquemos el método de Gram-Schmidt para conseguir una base ortonormal $u_1, \ldots, u_r, u_{r+1}, \ldots, u_n$ de E. Observemos que $u_j \in F^{\perp}$ si $j = r+1, \ldots, n$. Entonces, para todo $v \in E$,

$$v = (v^1 u_1 + \ldots + v^r u_r) + (v^{r+1} u_{r+1} + \ldots + v^n u_n) \in F + F^{\perp},$$

de donde resulta que $E = F \oplus F^{\perp}$. □

Corolario 5.3 $\dim F^{\perp} = \dim E - \dim F$. □

Corolario 5.4 *Si* F *es un subespacio vectorial de* E, $F^{\perp\perp} = F$.

DEMOSTRACIÓN: Por la propiedad 5 de (5.1), $F \subset F^{\perp\perp}$. Por (5.3), F y $F^{\perp\perp}$ tienen la misma dimensión. Por tanto, $F = F^{\perp\perp}$. □

Observación:

La biyección $\varphi : E \longrightarrow E'$ definida en el apartado 4 aplica el ortogonal de un subespacio F, tal como lo acabamos de definir, sobre el ortogonal de F en E' definido en (V.7). En efecto:

$$\{v \in E \mid u \cdot v = 0 \ \ \forall u \in F\} \xrightarrow{\varphi} \{v^* \in E' \mid v^*(u) = u \cdot v = 0 \ \ \forall u \in F\}.$$

Observemos que las propiedades de $F^\perp$ demostradas en (5.1) y (5.2) son consecuencia inmediata de esta biyección.

XI.6 Aplicaciones adjuntas y autoadjuntas

Sea E un espacio vectorial euclídeo o unitario. Un endomorfismo g se llama la *aplicación adjunta* de $f \in \mathrm{End}(E)$ si

$$v \cdot g(u) = f(v) \cdot u \quad \forall u, v \in E.$$

La adjunta, si existe, es única. En efecto, si g_1 también es una adjunta de f tenemos $v \cdot g(u) = f(v) \cdot u = v \cdot g_1(u)$ para todo v, u, de donde $g(u) = g_1(u)$ para todo u; por tanto, $g = g_1$.

¿Existe siempre la adjunta de f? Vamos a responder a esta cuestión demostrando que la adjunta de f no es más que la dual $f' : E' \longrightarrow E'$ considerada como aplicación de E en E vía la biyección φ del apartado 4. Es decir, veremos que la aplicación

$$g = \varphi^{-1} f' \varphi : E \longrightarrow E' \longrightarrow E' \longrightarrow E$$

es lineal y $g(u) \cdot v = u \cdot f(v)$ para todo u, v. En el caso real, la linealidad de g es consecuencia de la linealidad de φ y de f'. En el caso complejo, la linealidad respecto a la suma de φ y f' implica la linealidad respecto a la suma de g y, dado $k \in \mathbf{C}$,

$$\begin{aligned} g(ku) &= \varphi^{-1} f' \varphi(ku) = \varphi^{-1} f'(\bar{k}\varphi(u)) \\ &= \varphi^{-1}(\bar{k} f' \varphi(u)) = k(\varphi^{-1} f' \varphi(u)) = kg(u). \end{aligned}$$

Observemos ahora cuál es la imagen de un vector $u \in E$ por g:

$$u \xmapsto{\varphi} u^* \xmapsto{f'} u^* \circ f \xmapsto{\varphi} g(u).$$

$g(u) \in E$ cumple $g(u)^* = u^* \circ f$, de donde $g(u)^*(v) = u^* f(v)$, que equivale a

$$v \cdot g(u) = f(v) \cdot u.$$

Proposición 6.1 *Si g es la aplicación adjunta de $f \in \mathrm{End}(E)$ y $\dim E$ es finita, entonces*

$$\mathrm{Nuc}\, g = (\mathrm{Im}\, f)^{\perp} \quad e \quad \mathrm{Im}\, g = (\mathrm{Nuc}\, f)^{\perp}.$$

DEMOSTRACIÓN: $\mathrm{Nuc}\, g = \{u \in E \mid g(u) = \vec{0}\} = \{u \in E; v \cdot g(u) = 0\ \forall v\}$ $= \{u \in E \mid f(v) \cdot u = 0\ \forall v\} = (\mathrm{Im}\, f)^{\perp}$. De ahí, tomando ortogonales, obtenemos $\mathrm{Im}\, f = (\mathrm{Nuc}\, g)^{\perp}$. Ahora bien, si g es la adjunta de f, entonces f es la adjunta de g y, en particular, se cumple $\mathrm{Im}\, g = (\mathrm{Nuc}\, f)^{\perp}$. □

Proposición 6.2 *Sea $A = (a_i^j)$ la matriz de $f : E \longrightarrow E$ en una base ortonormal $u_1, \ldots, u_n$. La matriz de la adjunta de f en la base $u_1, \ldots, u_n$ es A^t en el caso real y $\bar{A}^t$ en el caso complejo.*

DEMOSTRACIÓN: Sea g la adjunta de f y $B = (b_i^j)$ su matriz. En el caso complejo tenemos

$$g(u_i) = \sum_{j=1}^{n} b_i^j u_j,$$

de donde

$$b_i^j = g(u_i) \cdot u_j = u_i \cdot f(u_j) = u_i \cdot \left(\sum_{k=1}^{n} a_j^k u_k \right) = \bar{a}_j^i.$$

Por tanto, $B = \bar{A}^t$. El mismo razonamiento vale en el caso real y se obtiene $B = A^t$. □

Una *aplicación autoadjunta* $f \in \mathrm{End}(E)$ es una aplicación lineal que coincide con su adjunta; es decir, tal que

$$v \cdot f(u) = f(v) \cdot u \qquad \forall u, v \in E.$$

De (6.2) resulta inmediatamente la

Proposición 6.3 *Si A es la matriz de f en una base ortonormal, f es autoadjunta si y sólo si*

$A = A^t$ *(A simétrica) en el caso real;*
$A = \bar{A}^t$ *(A hermítica) en el caso complejo.* □

XI.7 Diagonalización de matrices simétricas y hermíticas

Toda matriz simétrica real o hermítica compleja es la matriz de una aplicación autoadjunta en una base ortonormal. Esto se deduce de (6.3) y de una observación hecha al final del apartado 2. El problema de diagonalizar esas matrices equivale, pues, al de encontrar una base de vectores propios de una aplicación autoadjunta.

Teorema 7.1 *Si E es un espacio vectorial unitario y $f : E \longrightarrow E$ es autoadjunta, existe una base de vectores propios ortonormal.*

DEMOSTRACIÓN: Procederemos por inducción sobre la dimensión de E. Si $\dim E = 1$, todo vector es propio y no hay nada que demostrar. Si $\dim E = n$, el polinomio característico de f, $p(x) = \det(f - xI) \in \mathbf{C}[x]$, tiene siempre una raíz. Sea v un vector unitario de valor propio esa raíz λ: $f(v) = \lambda v$. El subespacio

$$F = \langle v \rangle^{\perp} = \{u \in E \mid u \cdot v = 0\}$$

es invariante por f. En efecto, si $u \in F$,

$$f(u) \cdot v = u \cdot f(v) = u \cdot (\lambda v) = \bar{\lambda}(u \cdot v) = \bar{\lambda} \cdot 0 = 0,$$

de donde $f(u) \in F$. Por hipótesis de inducción, existe una base ortonormal y de vectores propios de F: $u_2, \ldots, u_n$. Entonces $u_1 = v, u_2, \ldots, u_n$ es una base ortonormal de vectores propios de f (por (5.2)). □

Teorema 7.2 *Si E es un espacio euclídeo y $f : E \longrightarrow E$ es autoadjunta, existe una base de vectores propios ortonormal.*

DEMOSTRACIÓN: La misma que en el caso unitario vale siempre que podamos demostrar que el polinomio característico, $p(x) \in \mathbf{R}[x]$, tiene una raíz en $\mathbf{R}$. Esto y más nos lo demuestra el lema siguiente.

Lema 7.3 *Sea E un espacio vectorial euclídeo o unitario y sea $f : E \longrightarrow E$ una aplicación autoadjunta. Entonces el polinomio característico de f es de la forma*

$$p(x) = \pm(x - \lambda_1) \cdots (x - \lambda_n),$$

con $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$ (tanto en el caso euclídeo como en el unitario).

DEMOSTRACIÓN: En el caso unitario, la demostración se reduce a ver que todos los valores propios son reales. En efecto, si $f(v) = \lambda v$ con $v \neq \vec{0}$, tenemos $\lambda(v \cdot v) = (\lambda v) \cdot v = f(v) \cdot v = v \cdot f(v) = v \cdot (\lambda v) = \bar{\lambda}(v \cdot v)$. Puesto que $v \cdot v \neq 0$, es necesario que $\lambda = \bar{\lambda}$. Es decir, $\lambda \in \mathbf{R}$.

En el caso euclídeo, haremos la demostración mediante una "complejificación" del problema. Sea A la matriz simétrica correspondiente a f en una cierta base ortonormal. Si consideramos A como una matriz compleja, A es hermítica y, por tanto, es la matriz de una cierta aplicación lineal $\bar{f} : \bar{E} \longrightarrow \bar{E}$ de un espacio vectorial unitario. Puesto que f y $\bar{f}$ tienen la misma matriz, su polinomio característico será el mismo y, por la primera parte de la demostración, tendrá todas las raíces reales, tal como queríamos demostrar. Esto acaba también la demostración de 7.2. □

XI.8 Producto vectorial

Sea E un espacio vectorial euclídeo de dimensión 3 y e_1, e_2, e_3 una base ortonormal de E. Fijados $u, v \in E$, la aplicación

$$\begin{array}{ccc} E & \longrightarrow & \mathbf{R} \\ w & \longmapsto & \det_{(e_i)}(u, v, w) \end{array}$$

es lineal y, por tanto, un elemento del dual E'. Sea x el vector correspondiente a este elemento a través del isomorfismo de (4.1):

$$\begin{array}{ccc} E & \longrightarrow & E' \\ x & \longmapsto & x^* = \det_{(e_i)}(u, v, \quad); \end{array}$$

es decir, para todo w,

$$w \cdot x = \det_{(e_i)}(u, v, w).$$

El vector x es, por definición, el *producto vectorial* de u y v, y lo denotaremos por

$$u \wedge v.$$

Proposición 8.1 *El producto vectorial cumple*

1. $w \cdot (u \wedge v) = \det_{(e_i)}(u, v, w)$;
2. $u \wedge v = -v \wedge u$;
3. $(ku) \wedge v = k(u \wedge v)$;
4. $(u + u') \wedge v = u \wedge v + u' \wedge v$;

5. $u \wedge v$ *es ortogonal a* u *y a* v;
6. $u \wedge v = \vec{0}$ *si y sólo si* u, v *son linealmente dependientes;*
7. *Si* $u \wedge v \neq \vec{0}$, u, v, $u \wedge v$ *es una base de la misma orientación que* e_1, e_2, e_3.

DEMOSTRACIÓN: 1 es la misma definición. Para demostrar 2, observemos que para todo w se cumple $w \cdot (u \wedge v) = \det_{(e_i)}(u, v, w) = \det_{(e_i)}(v, u, -w) = (-w) \cdot (v \wedge u) = w \cdot (-v \wedge u)$, de donde resulta que $u \wedge v = -v \wedge u$.

De manera análoga se demuestran 3 y 4; 5 resulta de 1:

$$u \cdot (u \wedge v) = \det_{(e_i)}(u, v, u) = 0$$

y, de la misma manera, $v \cdot (u \wedge v) = 0$. Si $u \wedge v = \vec{0}$, entonces, para todo w, $\det_{(e_i)}(u, v, w) = 0$. Si u y v fuesen linealmente independientes, existiría un w tal que u, v, w sería una base y $\det_{(e_i)}(u, v, w) \neq 0$. Por tanto, u y v son linealmente dependientes. El recíproco es claro; así tenemos 6. Por último, si $u \wedge v \neq \vec{0}$, por 1,

$$\det_{(e_i)}(u, v, u \wedge v) = (u \wedge v) \cdot (u \wedge v) = \|u \wedge v\|^2 > 0,$$

lo que demuestra 7. □

Proposición 8.2 *a)* $e_1 \wedge e_2 = e_3$, $e_2 \wedge e_3 = e_1$, $e_3 \wedge e_1 = e_2$.

b) Si $u = u^1 e_1 + u^2 e_2 + u^3 e_3$ *y* $v = v^1 e_1 + v^2 e_2 + v^3 e_3$, *entonces*

$$u \wedge v = (u^2 v^3 - u^3 v^2) e_1 + (u^3 v^1 - v^3 u^1) e_2 + (u^1 v^2 - u^2 v^1) e_3.$$

Este resultado justifica la notación

$$u \wedge v = \begin{vmatrix} e_1 & u^1 & v^1 \\ e_2 & u^2 & v^2 \\ e_3 & u^3 & v^3 \end{vmatrix}$$

(¡desarrollar formalmente por la primera columna!).

DEMOSTRACIÓN: De (8.1.1) resulta fácilmente que $e_i \cdot (e_1 \wedge e_2) = e_i \cdot e_3$ para todo i. Por tanto, $e_1 \wedge e_2 = e_3$. Análogamente se demuestra que $e_2 \wedge e_3 = e_1$ y $e_3 \wedge e_1 = e_2$. Entonces, por (8.1),

$$\begin{aligned} u \wedge v &= \textstyle\sum_{i,j} u^i v^j e_i \wedge e_j \\ &= (u^2 v^3 - u^3 v^2) e_1 + (u^3 v^1 - u^1 v^3) e_2 + (u^1 v^2 - u^2 v^1) e_3. \end{aligned}$$ □

Proposición 8.3 $(u \wedge v) \wedge w = (u \cdot w)v - (v \cdot w)u$.

DEMOSTRACIÓN: Si $u \wedge v = 0$, $u = kv$, de donde $(u \cdot w)v - (v \cdot w)u = k(v \cdot w)v - k(v \cdot w)v = 0$ y la igualdad del enunciado es cierta. Si $u \wedge v \neq \vec{0}$, $u \wedge v \in (u, v)^{\perp}$, de donde $(u \wedge v) \wedge w \in \langle u, v \rangle$. Por tanto, $(u \wedge v) \wedge w = kv - hu$. Este vector ha de ser también ortogonal a w, de donde resulta que

$$0 = (kv - hu) \cdot w; \quad \text{es decir,} \quad k(v \cdot w) = h(u \cdot w).$$

Si $v \cdot w = 0$ y $u \cdot w = 0$, $w \in \langle u \wedge v \rangle$, de donde $(u \wedge v) \wedge w = \vec{0}$ y la igualdad del enunciado es cierta. Si uno de los dos productos es $\neq 0$, $k = a(u \cdot w)$ y $h = a(v \cdot w)$ para un cierto número real a. Es decir,

$$(u \wedge v) \wedge w = a((u \cdot w)v - (v \cdot w)u).$$

Para ver que $a = 1$, calculemos la primera coordenada de este vector:

$$\begin{aligned}&(u^3v^1 - u^1v^3)w^3 - (u^1v^2 - u^2v^1)w^2 = \\ &\quad = a[(u^1w^1 + u^2w^2 + u^3w^3)v^1 - (v^1w^1 + v^2w^2 + v^3w^3)u^1].\end{aligned}$$

De ahí resulta que $a = 1$. □

El producto vectorial no es asociativo (tal como se puede deducir tanto de (8.2) como de (8.3)); en lugar de la asociatividad, cumple la siguiente propiedad:

Proposición 8.4 (Identidad de Jacobi)

$$(u \wedge v) \wedge w + (v \wedge w) \wedge u + (w \wedge u) \wedge v = \vec{0}.$$

DEMOSTRACIÓN: Es consecuencia de (8.3). □

Proposición 8.5 $(u_1 \wedge u_2) \cdot (v_1 \wedge v_2) = (u_1 \cdot v_1)(u_2 \cdot v_2) - (u_1 \cdot v_2)(u_2 \cdot v_1)$.

DEMOSTRACIÓN:

$$\begin{aligned}&(u_1 \wedge u_2) \cdot (v_1 \wedge v_2) = \\ &\quad = \det_{(e_i)}(v_1, v_2, u_1 \wedge u_2) = \det_{(e_i)}(u_1 \wedge u_2, v_1, v_2) = v_2 \cdot ((u_1 \wedge u_2) \wedge v_1) = \\ &\quad = v_2 \cdot ((u_1 \cdot v_1)u_2 - (u_2 \cdot v_1)u_1) = (u_1 \cdot v_1)(u_2 \cdot v_2) - (u_1 \cdot v_2)(u_2 \cdot v_1).\end{aligned}$$ □

Corolario 8.6 $\|u \wedge v\|^2 = \|u\|^2\|v\|^2 - (u \cdot v)^2$. □

(Esto nos proporciona el término que falta en (3.1) en el caso particular de un espacio euclídeo de dimensión 3.)

Véase la observación 2 de (XII.5) para otra expresión de $\|u \wedge v\|$.

Observaciones:

1. Las propiedades 5 y 7 de (8.1) y el corolario (8.6) determinan el producto vectorial $u \wedge v$ de dos vectores linealmente independientes.
2. El producto vectorial depende de la base ortonormal e_1, e_2, e_3. Consideremos otra base ortonormal u_1, u_2, u_3;

$$\det_{(u_i)}(u, v, w) = \det_{(u_i)}(e_1, e_2, e_3) \cdot \det_{(e_i)}(u, v, w).$$

Si las dos bases son de la misma orientación, $\det_{(u_i)}(e_1, e_2, e_3) = +1$ y los productos vectoriales definidos utilizando una base y la otra coinciden. Si las bases son de orientaciones opuestas, $\det_{(u_i)}(e_1, e_2, e_3) = -1$ y los productos vectoriales definidos utilizando una base y la otra son vectores opuestos.

XI.9 Nota histórica

Los primeros pasos en la teoría de operadores lineales y en la introducción de productos escalares se deben a Erhard Schmidt (1876–1959) en 1907, aunque se encuentran antecedentes en los trabajos de David Hilbert sobre las ecuaciones integrales. Schmidt introdujo el concepto de norma, de producto escalar y de ortogonalidad y demostró una generalización del teorema de Pitágoras y del hecho de que vectores ortogonales dos a dos son linealmente independientes.

Schmidt desarrolló la teoría utilizando métodos introducidos por Hermann Amandus Schwarz (1843–1921), quien demostró, en particular, la desigualdad que lleva su nombre.

XI.10 Ejercicios

1. Demostrar que si $\|u\| = \|v\|$ entonces $(u + v) \cdot (u - v) = 0$. ($u + v$ y $u - v$ nos dan las bisectrices de u y v.)

2. Demostrar la *ley del paralelogramo*

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

3. Sea G la matriz de un producto escalar en una base e_1, e_2, e_3. Demostrar que $\det G = \det(e_1, e_2, e_3)^2$.

4. Sea E el espacio vectorial de los polinomios reales de grado ≤ 2. Para todo par $p(x), q(x) \in E$, definimos

$$\phi(p,q) = \int_0^1 p(x) \cdot q(x)\, dx.$$

a) Demostrar que ϕ es un producto escalar en E.

b) Encontrar una base ortonormal de E.

c) Encontrar una base del subespacio ortogonal al polinomio $2x+1$.

5. Calcular $\det(u \wedge v, v \wedge w, w \wedge u)$ en función de $\det(u,v,w)$ (los determinantes referidos a una misma base).

6. Dados dos vectores u, v del espacio vectorial euclídeo ordinario, consideramos el endomorfismo definido por

$$f(x) = (u \wedge x) \wedge v.$$

Comprobar que

a) si $u = \vec{0}$ o $v = \vec{0}$, entonces $f = 0$;

b) si $u \neq \vec{0}$, $v \neq \vec{0}$ y $u \cdot v = 0$, entonces $\operatorname{Nuc} f = \langle v \rangle^{\perp}$, $\operatorname{Im} f = \langle u \rangle$ y $f^2 = 0$;

c) si $u \cdot v \neq 0$, entonces $\operatorname{Nuc} f = (u\rangle$ e $\operatorname{Im} f = \langle v\rangle^{\perp}$;

d) si $u \cdot v = 1$, f es una proyección.

7. Dados dos vectores u, v linealmente independientes del espacio vectorial euclídeo ordinario, definimos un endomorfismo f por

$$f(x) = u \wedge (v \wedge x) - v \wedge (u \wedge x).$$

a) Demostrar que f es lineal y que $\det f = 2(u \cdot v)\|u \wedge v\|^2$.

b) Determinar los valores y vectores propios de f.

8. Sean $a_1, \ldots, a_k$ puntos dados del espacio afín $\mathbf{R}^n$, $F = \langle a_1\vec{a}_2, \ldots, a_1\vec{a}_k\rangle$ y $u_{k+1}, \ldots, u_n$ una base de $F^{\perp}$. Si $U \in M_{n\times(n-k)}(\mathbf{R})$ tiene por columnas los vectores $u_{k+1}, \ldots, u_n$, demostrar que la ecuación cartesiana de la variedad lineal determinada por $a_1, \ldots, a_k$ es $U^t x + b = \vec{0}$, donde $b = -U^t a_1$.

9. Sea E un espacio vectorial euclídeo, f un endomorfismo de E tal que $\|f(x)\| \leq \|x\|$ para todo $x \in E$ y g su adjunta.

Demostrar que

a) $\|g(x)\| \leq \|x\| \quad \forall x \in E$;

b) $\text{Nuc}(g - I) = \text{Nuc}(f - I)$;

c) $E = \text{Nuc}(f - I) + \text{Im}(f - I)$.

10. Sea E un espacio vectorial euclídeo y f un endomorfismo de E tal que $f(x) \cdot y = -x \cdot f(y)$ para todo $x, y \in E$. Demostrar que

a) Nuc f e Im f son subespacios ortogonales de E;

b) $E = \text{Nuc}\, f \oplus \text{Im}\, f$;

c) Si (a^i_j) es la matriz de f en una base ortonormal, entonces $a^i_j = -a^j_i$ para todo i, j.

11. Sea $f : E \longrightarrow F$ una aplicación lineal entre espacios vectoriales con producto escalar. Se llama *adjunta de* f a una aplicación lineal $g : F \longrightarrow E$ tal que $v \cdot g(u) = f(v) \cdot u$ para todo $u \in F$, $v \in E$. Demostrar que g existe y es única.

12. Sea $f : E \longrightarrow F$ una aplicación lineal entre espacios vectoriales euclídeos y g su adjunta (ejercicio 11). Demostrar que

a) $g \circ f$ es diagonalizable en una base ortonormal.

b) Todos los valores propios de $g \circ f$ son positivos. Designémoslos por $a_1, \ldots, a_n$, $n = \dim E$.

c) Existen bases ortonormales $e_1, \ldots, e_n$ de E y $u_1, \ldots, u_n$ de F tales que $f(e_i) = \sqrt{a_i}\, u_i$, $i = 1, \ldots, n$.

XI.11 Ejercicios para programar

13. Hacer un programa que, dada una matriz simétrica, estudie si es o no definida positiva.

14. Sea $e_1, \ldots, e_n$ una base dada de $\mathbf{R}^n$ y ϕ un producto escalar dado. Supongamos que tenemos las componentes de los e_i y la matriz de ϕ en la base canónica de $\mathbf{R}^n$. Hacer un programa que permita construir una base ortonormal siguiendo el método de Gram-Schmidt.

15. Hacer un programa que, dada una familia de vectores $e_1, \ldots, e_k$ de $\mathbf{R}^n$ y un producto escalar ϕ, encuentre una base del subespacio ortogonal a $F = \langle e_1, \ldots, e_k \rangle$. (Para su posterior utilización, preparar este programa como un subprograma.)

(Indicación: hay dos métodos;

a) Resolver el sistema homogéneo

$$\begin{cases} (e_1)^t Bx &= 0 \\ &\vdots \\ (e_k)^t Bx &= 0, \end{cases}$$

donde B es la matriz de ϕ.

b) Siguiendo la demostración de (5.2), extraer una base de F, completarla a una base de $\mathbf{R}^n$ (ejercicio IV.18) y ortonormalizar (ejercicio 14).)

16. Hacer un programa que

a) Dé las ecuaciones cartesianas $Ax + b = \vec{0}$, en la referencia canónica, de una variedad lineal de $\mathbf{R}^n$ a partir de la expresión vectorial $a + F$.

b) Recíprocamente, a partir de las ecuaciones cartesianas, encuentre un punto a y una base de F.

(Indicación: para (a), utilizar el ejercicio 8. Para (b), el punto a es una solución del sistema $Ax + b = \vec{0}$ y una base de F es una base de soluciones del sistema homogéneo $Ax = \vec{0}$ (ejercicio VII.12).)

17. Hacer un programa que, dados $a_1, \ldots, a_k$ puntos de $\mathbf{R}^n$, calcule las ecuaciones cartesianas de la variedad lineal que generan.

(Indicación: usar los ejercicios 8 y 16.)

18. Preparar un programa que, dados $u, v \in \mathbf{R}^3$, calcule $u \wedge v$. Comprobar en ejemplos concretos las identidades del §8.

Capítulo XII

Aplicaciones ortogonales. Aplicaciones unitarias

El capítulo anterior ha estado dedicado al estudio de una nueva estructura algebraica: los espacios vectoriales con un producto escalar. Ahora, tal como hemos hecho siempre que hemos introducido una estructura algebraica, queremos estudiar las aplicaciones que conservan esa estructura: las aplicaciones ortogonales y unitarias.

XII.1 Definiciones

Sea E un espacio vectorial con un producto escalar (euclídeo o unitario). Una aplicación $f : E \longrightarrow E$ se llama *ortogonal*, en el caso real, o *unitaria*, en el caso complejo, si

$$f(u) \cdot f(v) = u \cdot v \qquad \forall u, v \in E.$$

Proposición 1.1 *Toda aplicación que conserve el producto escalar es lineal.*

DEMOSTRACIÓN: Para probar la linealidad respecto a la suma, vamos a ver que, para todo u, v, $f(u+v) - f(u) - f(v)$ tiene norma cero y, por tanto, es el vector $\vec{0}$. En efecto,

$$\begin{aligned}
& (f(u+v) - f(u) - f(v)) \cdot (f(u+v) - f(u) - f(v)) = \\
= \; & f(u+v) \cdot f(u+v) - f(u+v) \cdot f(u) - f(u+v) \cdot f(v) - \\
& -f(u) \cdot f(u+v) + f(u) \cdot f(u) + f(u) \cdot f(v) - \\
& -f(v) \cdot f(u+v) + f(v) \cdot f(u) + f(v) \cdot f(v) = \\
= \; & (u+v) \cdot (u+v) - (u+v) \cdot u - (u+v) \cdot v - u \cdot (u+v) + \\
& +u \cdot u + +u \cdot v - v \cdot (u+v) + v \cdot u + v \cdot v = \\
= \; & ((u+v) - u - v) \cdot ((u+v) - u - v) = \vec{0} \cdot \vec{0} = 0.
\end{aligned}$$

Análogamente, tenemos

$$\begin{aligned}(kf(u) - f(ku)) \cdot (kf(u) - f(ku)) &= \\ &= k\bar{k}f(u) \cdot f(u) - kf(u) \cdot f(ku) - \bar{k}f(ku) \cdot f(u) + f(ku) \cdot f(ku) = \\ &= k\bar{k}u \cdot u - ku \cdot (ku) - \bar{k}(ku) \cdot u + (ku) \cdot (ku) = \\ &= (ku - ku) \cdot (ku - ku) = 0,\end{aligned}$$

de donde $kf(u) - f(ku) = \vec{0}$ y $f(ku) = kf(u)$. □

Proposición 1.2 *Si f es ortogonal o unitaria, se cumple*

1. $\|f(u)\| = \|u\|$ *para todo* $u \in E$;
2. v, u *son ortogonales si y sólo si* $f(v)$, $f(u)$ *lo son;*
3. f *es biyectiva;*
4. *si* k *es un valor propio de* f, $|k| = 1$, *donde* $|k|$ *indica el valor absoluto o el módulo según sea real o complejo;*
5. *si* u, v *son dos vectores propios de valores propios* $k \neq h$, *entonces* u, v *son ortogonales.*

DEMOSTRACIÓN: 1 y 2 son consecuencia inmediata de la conservación del producto escalar. 3 se deduce de 1. Para demostrar 4, supongamos $f(v) = kv$ con $v \neq \vec{0}$. Entonces

$$v \cdot v = f(v) \cdot f(v) = k\bar{k}(v \cdot v) = |k|^2(v \cdot v),$$

de donde $|k| = 1$.

Para probar 5, hacemos $u \cdot v = f(u) \cdot f(v) = k\bar{h}(u \cdot v)$. Si $u \cdot v \neq 0$, $k\bar{h} = 1$ y, puesto que $|h| = 1$, $\bar{h} = h^{-1}$. Por tanto, $k = h$, en contra de la hipótesis. Así pues, $u \cdot v = 0$ y u, v son ortogonales. □

Proposición 1.3 *Sea A la matriz de $f \in \mathrm{End}(E)$ en la base $e_1, \ldots, e_n$ de E. Entonces*

$$f \text{ es } \left.\begin{matrix}\textit{ortogonal} \\ \textit{unitaria}\end{matrix}\right\} \Leftrightarrow (a)\ \ f(e_i) \cdot f(e_j) = e_i \cdot e_j\ \ \forall i, j$$

$$\Leftrightarrow (b)\ \begin{cases} A^t G A = G \\ A^t G \bar{A} = G. \end{cases}$$

(G indica la matriz del producto escalar en la base $e_1, \ldots, e_n$.)

DEMOSTRACIÓN: de (a): la implicación ($\Rightarrow$) es consecuencia de la conservación del producto escalar.

$\Leftarrow$) Calculemos $f(u) \cdot f(v)$. Si $u = \sum u^i e_i$ y $v = \sum v^i e_i$, entonces

$$f(u) = \sum u^i f(e_i) \quad \text{y} \quad f(v) = \sum v^i f(e_i)$$

y, por tanto,

$$f(u) \cdot f(v) = \sum_{i,j} u^i \bar{v}^j f(e_i) \cdot f(ej) = \sum_{i,j} u^i \bar{v}^j e_i \cdot e_j = u \cdot v.$$

Análogamente en el caso real.

Para demostrar (b), observemos que $f(e_1), \ldots, f(e_n)$ es base por (1.2.(3)) y que la matriz del producto escalar en esa base es $A^t G A$ en el caso real y $A^t G \bar{A}$ en el caso complejo. Si $f(e_i) \cdot f(e_j) = e_i \cdot e_j$ para todo i, j, entonces esa matriz coincide con la matriz del producto escalar en la base $e_1, \ldots, e_n$, que es G. □

Corolario 1.4 *Sea A la matriz de $f \in \mathrm{End}(E)$ en una base ortonormal. Entonces,*

f es ortogonal $\Leftrightarrow$ $A^t A = I$;

f es unitaria $\Leftrightarrow$ $A^t \bar{A} = I$. □

Corolario 1.5 *Sea A la matriz de cambio de una base ortonormal a otra también ortonormal. Entonces $A^t A = I$ (caso real) y $A^t \bar{A} = I$ (caso complejo).* □

Una matriz real A se llama *ortogonal* si $A^t A = I$ (es decir, si $A^{-1} = A^t$); entonces $\det A = \pm 1$. Una matriz compleja A se llama *unitaria* si $A^t \bar{A} = I$ (es decir, si $A^{-1} = \bar{A}^t$); entonces $|\det A| = 1$.

Las aplicaciones ortogonales de un espacio vectorial euclídeo E de dimensión n forman un grupo que denominaremos *grupo ortogonal de orden n* y denotaremos por $O(n)$. Las matrices reales $n \times n$ ortogonales forman un grupo que, por (1.4), es isomorfo al grupo de las aplicaciones ortogonales.

Las aplicaciones unitarias de un espacio unitario E de dimensión n forman un grupo que denominaremos *grupo unitario de orden n* y denotaremos por $U(n)$. Las matrices complejas $n \times n$ unitarias forman un grupo que, por (1.4), es isomorfo al grupo de las aplicaciones unitarias.

Ejercicio:

Estudiar los grupos $O(1)$ y $U(1)$.

XII.2 Diagonalización de matrices unitarias

Teorema 2.1 *Sea E un espacio vectorial unitario y sea $f \in \mathrm{End}(E)$ una aplicación unitaria. Existe siempre una base ortonormal de E formada por vectores propios de f.*

DEMOSTRACIÓN: Por (VIII.2.5), sabemos que existe una base $v_1, \ldots, v_n$ en la que f tiene una matriz triangular; es decir, tal que

$$f(v_i) \in \langle v_1, \ldots, v_i \rangle.$$

Sea $w_1, \ldots, w_n$ una base ortonormal obtenida a partir de la anterior por el método de Gram-Schmidt (XI.2.1). Como $w_i \in \langle v_1, \ldots, v_i \rangle$, resulta que

$$f(w_i) \in \langle f(v_1), \ldots, f(v_i) \rangle = \langle v_1, \ldots, v_i \rangle = \langle w_1, \ldots, w_i \rangle$$

y, por tanto, la matriz de f en la base $w_1, \ldots, w_n$ también es triangular.

Vamos a ver que $w_1, \ldots, w_n$ son vectores propios. w_1 es, claramente, un vector propio. Supongamos que $w_1, \ldots, w_k$ son vectores propios y sea

$$f(w_{k+1}) = a^1 w_1 + \ldots + a^k w_k + a^{k+1} w_{k+1}.$$

Entonces tenemos que, si $f(w_i) = b^i w_i$ para $i = 1, \ldots, k$,

$$0 = w_{k+1} \cdot w_i = f(w_{k+1}) \cdot f(w_i) = f(w_{k+1}) \cdot (b^i w_i) = a^i b^i.$$

Pero $|b^i| = 1$ y, por tanto, $b^i \neq 0$, de donde $a^i = 0$ para $i = 1, \ldots, k$, y

$$f(w_{k+1}) = a^{k+1} w_{k+1}.$$

Es decir, w_{k+1} es también un vector propio. □

XII.3 Forma canónica de una matriz ortogonal

Teorema 3.1 *Sea E un espacio vectorial euclídeo y $f : E \longrightarrow E$ una aplicación ortogonal. Existe una base ortonormal de E tal que la matriz correspondiente a f es de la forma*

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & 0 & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & A_1 & \\ & & 0 & & & & \ddots \\ & & & & & & & A_r \end{pmatrix},$$

donde las matrices A_i son del tipo

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \text{ con } a^2 + b^2 = 1.$$

DEMOSTRACIÓN: Sean E_1 y E_{-1} los subespacios de vectores propios de valor propio 1 y -1 respectivamente. Los vectores de E_1 y E_{-1} son ortogonales entre sí (1.2.(5)) y, por tanto, $E_1 \cap E_{-1} = \{\vec{0}\}$. Designemos por F el subespacio ortogonal a $E_1 \oplus E_{-1}$. Es fácil ver que F es invariante por f. La base buscada de

$$E = E_1 \oplus E_{-1} \oplus F$$

es unión de bases ortonormales de cada uno de esos subespacios. Puesto que F no contiene vectores propios, el polinomio característico de la restricción de f a F es producto de polinomios irreducibles de grado dos. En particular, la dimensión de F es par. Designemos por A la matriz de la restricción de f a F en una cierta base ortonormal $e_1, \ldots, e_{2r}$.

Sea $\widehat{F}$ un espacio vectorial unitario y $\widehat{e}_1, \ldots, \widehat{e}_{2r}$ una base ortonormal de $\widehat{F}$. Consideremos la aplicación

$$\widehat{f} : \widehat{F} \longrightarrow \widehat{F}$$

que en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$ tiene la matriz A. Por (1.4) $\widehat{f}$ es unitaria y por (2.1) existe una base ortonormal de vectores propios. El polinomio característico de $\widehat{f}$ es el mismo que el de f y sus raíces son, por tanto, conjugadas dos a dos. Sea v un vector propio unitario de valor propio $z = a+bi$. Si las coordenadas de v en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$ son $(x^1 + iy^1, \ldots, x^n + iy^n)$, designaremos por $\bar{v}$ el vector de coordenadas $(x^1 - iy^1, \ldots, x^n - iy^n)$. Entonces, $\bar{v}$ es un vector unitario de valor propio $\bar{z}$. En efecto, con la notación usual, tenemos

$$A\bar{v} = \overline{Av} = \overline{zv} = \bar{z} \cdot \bar{v}.$$

Podemos escoger, por tanto, una base ortonormal $v_1, \ldots, v_{2r}$ de vectores propios de $\widehat{f}$ formada por pares de vectores $v, \bar{v}$ con valores propios conjugados, $z, \bar{z}$.

Consideremos ahora, para cada par $v, \bar{v}$, vectores

$$w = \frac{1}{\sqrt{2}}(v + \bar{v}), \qquad w' = \frac{-i}{\sqrt{2}}(v - \bar{v}).$$

Observemos que w y w' tienen coordenadas reales:

$$\begin{aligned} w &= \sqrt{2}\,(x^1, \ldots, x^n) \\ w' &= \sqrt{2}\,(y^1, \ldots, y^n). \end{aligned}$$

Teniendo en cuenta que $v \cdot \bar{v} = 0$ (por 1.2 (5)) y que $v \cdot v = \bar{v} \cdot \bar{v} = 1$, resulta que

$$w \cdot w' = 0 \ \text{ y } \ \|w\| = 1 = \|w'\|.$$

Además,

$$\begin{aligned} \widehat{f}(w) &= \frac{1}{\sqrt{2}}\widehat{f}(v+\bar{v}) = \frac{1}{\sqrt{2}}(zv + \bar{z}\bar{v}) = \\ &= \frac{1}{\sqrt{2}}\left((a+bi)\frac{1}{\sqrt{2}}(w+iw') + (a-bi)\frac{1}{\sqrt{2}}(w-iw')\right) = \\ &= aw - bw'. \end{aligned}$$

Análogamente, $\widehat{f}(w') = bw + aw'$. Es decir, en lenguaje de matrices,

$$\begin{aligned} Aw &= aw - bw' \\ Aw' &= bw + aw'. \end{aligned}$$

Sustituyendo en la base $v_1, \ldots, v_{2r}$ cada par $v, \bar{v}$ por los correspondientes w, w', obtenemos una nueva base ortonormal de $\widehat{F}$, $w_1, \ldots, w_{2r}$, formada por vectores de coordenadas reales (en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$).

Designemos por u_i el vector de F que en la base $e_1, \ldots, e_{2r}$ tiene las mismas coordenadas que w_i en la base $\widehat{e}_1, \ldots, \widehat{e}_{2r}$. Claramente, $u_1, \ldots, u_{2r}$ es una base ortonormal de F formada por pares u, u' tales que

$$\begin{aligned} Au &= au - bu' \\ Au' &= bu + au'. \end{aligned}$$

La matriz de f en esta base es de la forma

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & A_r \end{pmatrix},$$

donde las cajas A_i son del tipo

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

con $a^2 + b^2 = 1$. □

XII.4 Los grupos $O(2)$ y $SO(2)$

En este apartado, E es un espacio vectorial euclídeo de dimensión 2. El conjunto de aplicaciones ortogonales de E con la composición es $O(2)$, grupo ortogonal de orden 2. Por (1.4), si $f \in O(2)$, $\det f = \pm 1$. Consideremos el morfismo

$$\begin{array}{ccc} O(2) & \longrightarrow & \{+1,-1\} \\ f & \longmapsto & \det f. \end{array}$$

El núcleo de este morfismo,

$$SO(2) = \{f \in O(2) \mid \det f = +1\},$$

es un subgrupo normal de $O(2)$ que se llama el *grupo ortogonal especial* de orden 2. Empezaremos por estudiar este subgrupo.

Fijemos una base ortonormal e_1, e_2 de E. La matriz de $f \in SO(2)$

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

debe cumplir $AA^t = I$ (o $A^{-1} = A^t$) (por (1.4)). Es decir,

$$A^{-1} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A^t$$

(ya que $\det A = +1$), o sea, $a = d$, $c = -b$. Así pues,

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{con } a^2 + b^2 = 1.$$

Proposición 4.1 *El grupo $SO(2)$ es conmutativo.*

DEMOSTRACIÓN: Sean $f, g \in SO(2)$ con matrices en la base ortonormal e_1, e_2

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \begin{pmatrix} c & -d \\ d & c \end{pmatrix};$$

entonces

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -ad-bc \\ ad+bc & ac-bd \end{pmatrix} = \begin{pmatrix} c & -d \\ d & c \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

de donde $f \circ g = g \circ f$. □

Proposición 4.2 *Sea $f \in SO(2)$ y*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

su matriz en la base ortonormal e_1, e_2. Entonces, si u es un vector unitario cualquiera,

$$a = u \cdot f(u) \quad y \quad b = \det_{(e_i)}(u, f(u)).$$

DEMOSTRACIÓN: Sea $u = \alpha e_1 + \beta e_2$ con $\alpha^2 + \beta^2 = 1$. Entonces

$$f(u) = (a\alpha - b\beta)e_1 + (b\alpha + a\beta)e_2,$$

de donde

$$u \cdot f(u) = \alpha(a\alpha - b\beta) + \beta(b\alpha + a\beta) = a(\alpha^2 + \beta^2) = a,$$

y

$$\det_{(e_i)}(u, f(u)) = \begin{vmatrix} \alpha & a\alpha - b\beta \\ \beta & b\alpha + a\beta \end{vmatrix} = b(\alpha^2 + \beta^2) = b. \ \square$$

Esta proposición nos dice, en particular, que a es independiente de la base ortonormal escogida y que b varía, como mucho, en el signo. En efecto, si u_1, u_2 es otra base ortonormal,

$$\det_{(u_i)}(u, f(u)) = \det_{(u_i)}(e_1, e_2) \cdot \det_{(e_i)}(u, f(u)) = \pm b,$$

ya que $\det_{(u_i)}(e_1, e_2) = \pm 1$ (1.5). Este determinante es $+1$ cuando e_1, e_2 y u_1, u_2 son de la misma orientación y -1 en caso contrario (IX.11).

Proposición 4.3 *Dados dos vectores $u, v \in E$ que tengan la misma norma, $\|u\| = \|v\| \neq 0$, existe una $f \in SO(2)$ y sólo una tal que $f(u) = v$.*

DEMOSTRACIÓN: Sea e_1, e_2 una base ortonormal. Si $f \in SO(2)$ cumple $f(u) = v$, su matriz en esa base ha de ser

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

donde

$$a = \frac{u}{\|u\|} \cdot \frac{v}{\|v\|}, \qquad b = \det_{(e_i)}\left(\frac{u}{\|u\|}, \frac{v}{\|v\|}\right).$$

Definimos, pues, f como la aplicación lineal que tiene esa matriz. Para probar que f es ortogonal, debemos ver que $a^2 + b^2 = 1$. Sean

$$u = \alpha e_1 + \beta e_2, \qquad v = \gamma e_1 + \delta e_2.$$

Tenemos

$$a^2+b^2=\frac{(u\cdot v)^2+\det(u,v)^2}{\|u\|^2\|v\|^2}=\frac{(\alpha\gamma+\beta\delta)^2+(\alpha\delta-\gamma\beta)^2}{(\alpha^2+\beta^2)(\gamma^2+\delta^2)}=1.$$

Demostremos ahora que $f(u)=v$. Por (4.2),

$$\det_{(e_i)}\left(\frac{u}{\|u\|},\frac{f(u)}{\|u\|}\right)=b=\det_{(e_i)}\left(\frac{u}{\|u\|},\frac{v}{\|v\|}\right),$$

de donde $\det_{(e_i)}(u,f(u)-v)=0$ y, por tanto, $f(u)-v=ku$. Por otra parte,

$$\frac{u}{\|u\|}\cdot\frac{f(u)}{\|u\|}=a=\frac{u}{\|u\|}\cdot\frac{v}{\|v\|}$$

implica $u\cdot(f(u)-v)=0$, de donde $u\cdot(ku)=0$.

Por tanto, $k=0$ y $f(u)-v=0$. □

Corolario 4.4 *Si $f\in SO(2)$ deja un vector fijo, entonces $f=I$.* □

Estudiemos ahora las aplicaciones $f\in O(2)$ con $\det f=-1$.

Esas aplicaciones forman una clase del cociente de $O(2)$ por el núcleo $SO(2)$ de la aplicación $\det: O(2)\longrightarrow\{+1,-1\}$ (III.5). Consideremos, en particular, la que en la base ortonormal e_1,e_2 tiene por matriz

$$\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

Cualquier otra se obtiene de ésta, componiendo con una aplicación de $SO(2)$, y su matriz será, por tanto, de la forma

$$\begin{pmatrix}a&b\\b&-a\end{pmatrix}\qquad\text{con } a^2+b^2=1.$$

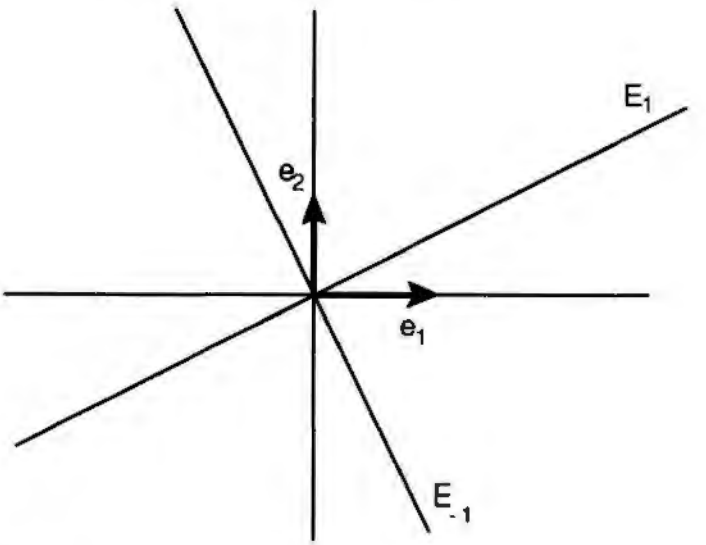
Observemos que

$$\begin{pmatrix}a&b\\b&-a\end{pmatrix}^2=\begin{pmatrix}1&0\\0&1\end{pmatrix}.$$

Se trata, por tanto, de simetrías (X.2.III). Los subespacios de vectores propios de valor propio $+1$ y -1 son, respectivamente,

$$E_1=\langle(b,1-a)\rangle\quad\text{y}\quad E_{-1}=\langle(-b,1+a)\rangle.$$

Estos subespacios son ortogonales y, por ello, esas simetrías se llaman *simetrías ortogonales.*



XII.5 Ángulos

Sea E un espacio vectorial euclídeo de dimensión 2. En el conjunto de pares de vectores unitarios definimos una relación de equivalencia de la siguiente manera:

$$(u,u') \sim (v,v') \Leftrightarrow \text{existe } f \in SO(2) \text{ tal que } f(u)=v, f(u')=v' \Leftrightarrow$$
$$\Leftrightarrow \text{existe } g \in SO(2) \text{ tal que } g(u)=u', g(v)=v'.$$

Demostremos en primer lugar que estas dos condiciones son equivalentes.

$\Rightarrow$) Supongamos que existe la aplicación f y sea $g \in SO(2)$ tal que $g(u)=u'$ (4.3). Entonces, por la conmutatividad de $SO(2)$,

$$g(v) = gf(u) = fg(u) = f(u') = v'.$$

$\Leftarrow$) Supongamos ahora que existe g y sea $f \in SO(2)$ tal que $f(u)=v$. Entonces

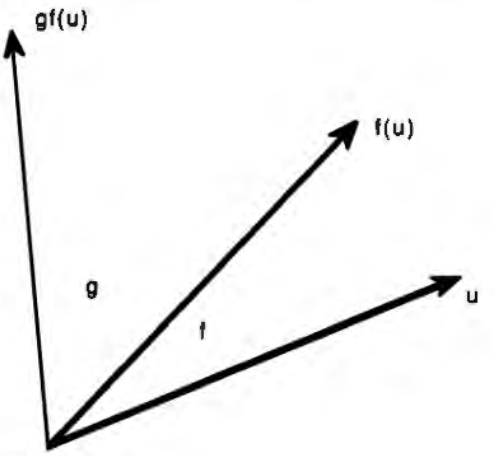
$$f(u') = fg(u) = gf(u) = g(v) = v'.$$

Llamaremos *ángulo* a cada una de las clases de equivalencia por esa relación. El ángulo determinado por un par (u,u') será denotado por $\widehat{uu'}$ o simplemente por (u,u'). El ángulo de dos vectores u, v no necesariamente unitarios, $\widehat{uv}$, es el ángulo de los vectores $\dfrac{u}{\|u\|}, \dfrac{v}{\|v\|}$.

Designaremos por A el conjunto de ángulos. La aplicación

$$\begin{array}{ccc} SO(2) & \longrightarrow & A \\ f & \longmapsto & \widehat{uf(u)} \end{array}$$

donde u es cualquier vector unitario, es una biyección que nos permite transportar la operación de $SO(2)$ a A: dados $\alpha, \beta \in A$ con antiimágenes f y g,



respectivamente, definimos la *suma* $\alpha+\beta$ como el ángulo correspondiente a $g \circ f$.

Así, si u es un vector unitario,

$$\left.\begin{array}{rcl} \alpha & = & (u, f(u)) \\ \beta & = & (f(u), gf(u)) \end{array}\right\} \Rightarrow \alpha+\beta = (u, gf(u)).$$

Naturalmente, la suma definida en A tiene las mismas propiedades que la operación de $SO(2)$. A es, pues, un grupo conmutativo con elemento neutro $0 = \widehat{uu}$. El opuesto de un ángulo $\widehat{uf(u)}$ es $\widehat{f(u)u}$.

Fijada una orientación en E, a cada $f \in SO(2)$ le corresponde por (4.2) una matriz bien determinada

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{con } a^2+b^2 = 1.$$

Sea α el ángulo correspondiente a f. Definimos el *coseno* de α ($\cos\alpha$) y el *seno* de α ($\operatorname{sen}\alpha$) como los valores a y b en esa matriz:

$$\cos\alpha = a, \qquad \operatorname{sen}\alpha = b.$$

Observemos que al cambiar la orientación del espacio E cambia el signo de $\operatorname{sen}\alpha$, pero no el de $\cos\alpha$. Además, se tiene

$$\cos^2\alpha + \operatorname{sen}^2\alpha = 1.$$

El ángulo 0 corresponde a la aplicación identidad; por tanto,

$$\cos 0 = 1, \qquad \operatorname{sen} 0 = 0.$$

El ángulo opuesto de α corresponde a la aplicación inversa f^{-1}, que tiene por matriz la traspuesta de la matriz de f; por tanto,

$$\cos(-\alpha) = \cos\alpha, \qquad \operatorname{sen}(-\alpha) = -\operatorname{sen}(\alpha).$$

El ángulo $\alpha+\beta$ corresponde a la composición de las aplicaciones correspondientes a α y β; se obtiene

$$\begin{aligned}\cos(\alpha+\beta) &= \cos\alpha\cdot\cos\beta-\operatorname{sen}\alpha\cdot\operatorname{sen}\beta\\ \operatorname{sen}(\alpha+\beta) &= \operatorname{sen}\alpha\cdot\cos\beta+\cos\alpha\cdot\operatorname{sen}\beta.\end{aligned}$$

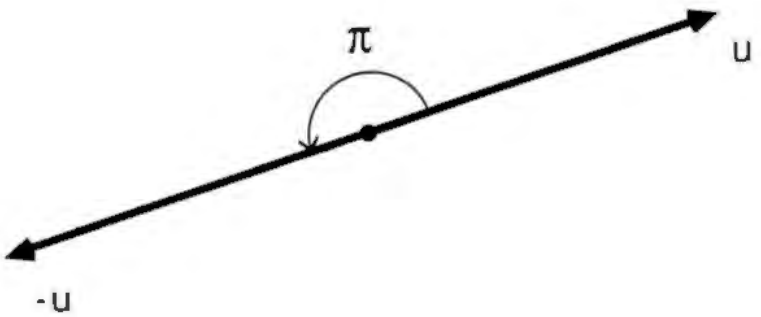
Proposición 5.1 *Existe un ángulo $\pi\neq 0$, y sólo uno, tal que $2\pi=0$. π es el ángulo tal que*

$$\cos\pi=-1 \qquad y \qquad \operatorname{sen}\pi=0.$$

DEMOSTRACIÓN: Sea $\begin{pmatrix} a & -b\\ b & a\end{pmatrix}=A$ la matriz correspondiente a π en una orientación prefijada. La condición $2\pi=0$ equivale a $AA=I$; es decir, $A^{-1}=A$. Ahora bien, puesto que A es ortogonal ($A^{-1}=A^t$), $A=A^t$, de donde resulta que $b=0$ y $a=\pm 1$. En el caso $a=1$, $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ corresponde al ángulo 0.

En el caso $a=-1$, $\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$ corresponde al ángulo buscado. □

Observemos que $\pi=\widehat{u(-u)}$, donde u es un vector cualquiera. π se llama el *ángulo llano.*



Proposición 5.2 *Existen dos ángulos δ_1,δ_2, y sólo dos, tales que $2\delta_i=\pi$. δ_1 y δ_2 son los ángulos con*

$$\cos\delta_1=\cos\delta_2=0, \qquad \operatorname{sen}\delta_1=-\operatorname{sen}\delta_2=1.$$

DEMOSTRACIÓN: Sea

$$A=\begin{pmatrix}a&-b\\b&a\end{pmatrix}$$

la matriz del ángulo δ buscado. La condición $2\delta=\pi$ equivale a $AA=-I$; es decir, $A^{-1}=-A$. Por ser A ortogonal, $A^{-1}=A^t$, de donde $A^t=-A$ y, por tanto, $a=0$, $b=\pm 1$. □

Los ángulos δ_1, δ_2 se llaman los *ángulos rectos.*

Ejercicio:

$\widehat{uv}$ es un ángulo recto si y sólo si u y v son ortogonales.

Proposición 5.3 *Dado un ángulo α, existen dos ángulos ϕ_1, ϕ_2, y sólo dos, tales que $2\phi = \alpha$. Además,*

$$\cos\alpha = \cos^2\phi - \operatorname{sen}^2\phi \qquad y \qquad \operatorname{sen}\alpha = 2\cos\phi \cdot \operatorname{sen}\phi.$$

DEMOSTRACIÓN: Sean

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \qquad B = \begin{pmatrix} c & -d \\ d & c \end{pmatrix}$$

las matrices de α y ϕ respectivamente. La condición $2\phi = \alpha$ equivale a $BB = A$:

$$\begin{pmatrix} c^2 - d^2 & -cd \\ 2cd & c^2 - d^2 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

De $c^2 - d^2 = a$, $c^2 + d^2 = 1$ resulta que $c = \pm\sqrt{\dfrac{1+a}{2}}$, $d = \pm\sqrt{\dfrac{1-a}{2}}$, de donde

$$2cd = \pm\sqrt{1-a^2} = \pm\sqrt{b^2} = \pm|b|.$$

Pero $2cd = b$. Por tanto, si $b > 0$, c y d tienen que ser ambos positivos o ambos negativos; si $b < 0$, c y d tienen que ser uno positivo y el otro negativo. En ambos casos hay dos soluciones, tal como se trataba de ver. El caso $b = 0$ ya ha sido considerado en (5.1) y (5.2). □

Observaciones:

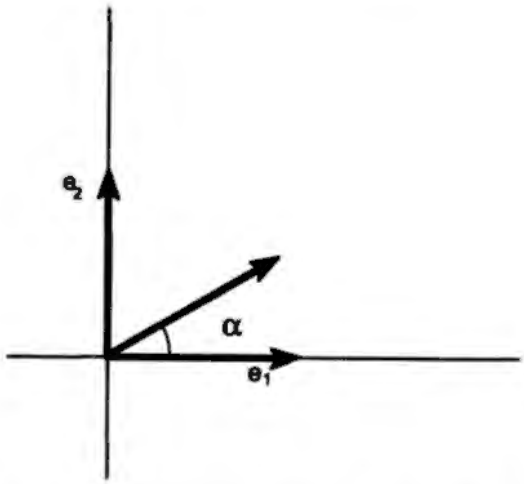
1. Sea $\alpha = \widehat{uv}$. Si u y v no son unitarios, recordemos que $\widehat{uv}$ es, por definición, el ángulo de $\dfrac{u}{\|u\|}$ y $\dfrac{v}{\|v\|}$. Por (4.2) y la definición de $\cos\alpha$, tenemos que

$$\cos\alpha = \frac{u}{\|u\|} \cdot \frac{v}{\|v\|},$$

de donde $u \cdot v = \|u\|\|v\|\cos\alpha$.

2. De la observación anterior y de (XI.8.6) resulta que

$$\|u \wedge v\| = \|u\|\|v\|\,|\operatorname{sen}\widehat{uv}|.$$



3. Sea e_1, e_2 una base ortonormal y $v = ae_1 + be_2$ un vector unitario ($a^2 + b^2 = 1$).
El ángulo $\alpha = \widehat{e_1 v}$ corresponde a una aplicación $f \in SO(2)$ tal que $f(e_1) = v$. Su matriz será

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Por tanto, $a = \cos\alpha$ y $b = \operatorname{sen}\alpha$. Es decir,

$$v = \cos\alpha \cdot e_1 + \operatorname{sen}\alpha \cdot e_2.$$

La aplicación $f \in SO(2)$ correspondiente a un ángulo α se llama una *rotación (vectorial) de ángulo* α. $SO(2)$ es el *grupo de las rotaciones* de E. Observemos que la traza de cualquier matriz de f es $2\cos\alpha$. Así pues, α es el ángulo tal que

$$\cos\alpha = \frac{1}{2}\operatorname{tr} f.$$

El signo de $\operatorname{sen}\alpha$ no queda determinado; depende de la orientación de la base en la que trabajemos.

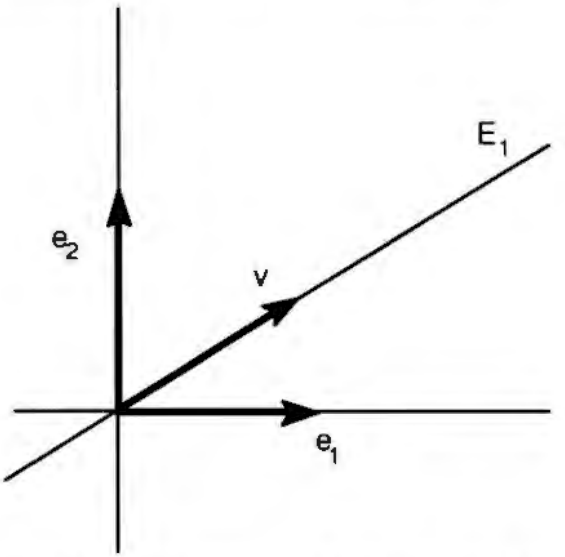
Sea ahora $f \in O(2)$ con $\det f = -1$ y matriz, en una base ortonormal e_1, e_2,

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \qquad a^2 + b^2 = 1.$$

f es una simetría ortogonal de eje $E_1 = \langle (b, 1-a) \rangle$.

Supongamos $a \neq 1$. Consideremos el vector unitario

$$v = \left(\frac{b}{\sqrt{2(1-a)}}, \frac{1-a}{\sqrt{2(1-a)}} \right).$$



Observemos que su segunda coordenada es positiva ($a < 1$). Pongamos $\alpha = \widehat{e_1 v}$. Entonces

$$\left.\begin{array}{l} \cos\alpha = \dfrac{b}{\sqrt{2(1-a)}} \\ \operatorname{sen}\alpha = \dfrac{1-a}{\sqrt{2(1-a)}} \end{array}\right\} \Rightarrow \begin{array}{rcccl} \cos 2\alpha & = & \cos^2\alpha - \operatorname{sen}^2\alpha & = & a \\ \operatorname{sen} 2\alpha & = & 2\operatorname{sen}\alpha\cdot\cos\alpha & = & b. \end{array}$$

La matriz de f es, por tanto, de la forma

$$\begin{pmatrix} \cos 2\alpha & \operatorname{sen} 2\alpha \\ \operatorname{sen} 2\alpha & -\cos 2\alpha \end{pmatrix}.$$

Si $a = 1$, entonces $v = e_1$, $\alpha = 0$, y vale el mismo resultado.

Dado un ángulo $\alpha = \widehat{uv}$ y una aplicación $f \in \operatorname{End}(E)$, designaremos por $f\alpha$ el ángulo $\widehat{f(u)f(v)}$.

Si $f \in SO(2)$, y $g \in SO(2)$ es la aplicación correspondiente al ángulo α, tenemos

$$f\alpha = \widehat{f(u)fg(u)} = \widehat{f(u)gf(u)} = \alpha.$$

Si f es una simetría ortogonal, es muy fácil ver que

$$fg = g^{-1}f.$$

Entonces

$$f\alpha = \widehat{f(u)fg(u)} = \widehat{f(u)g^{-1}f(u)} = -\alpha.$$

La proposición siguiente resume estos dos hechos.

Proposición 5.4 *Las rotaciones vectoriales conservan los ángulos. Las simetrías ortogonales los invierten.* □

XII.6 El grupo $O(3)$

En este apartado, E es un espacio vectorial euclídeo de dimensión 3. Igual que en el caso de dimensión 2 (§4),

$$SO(3) = \{f \in O(3) \mid \det f = +1\}.$$

Estudiemos primero las aplicaciones de $SO(3)$. El polinomio característico de $f \in SO(3)$ es de grado 3 y tiene, por tanto, una raíz real. Sea $v \neq \vec{0}$ un vector propio. El subespacio $F = \langle v \rangle^{\perp}$ es invariante por f y la restricción de f a F es ortogonal. Si v es de valor propio -1, el determinante de esa restricción es -1; será, pues, una simetría ortogonal y, en particular, tendrá el valor propio $+1$. Resulta, por tanto, que $+1$ es siempre valor propio de f.

Sea u_3 un vector propio unitario de valor propio $+1$ y sea u_1, u_2, u_3 una base ortonormal. La restricción

$$f : \langle u_1, u_2 \rangle = \langle u_3 \rangle^{\perp} \longrightarrow \langle u_3 \rangle^{\perp}$$

tiene determinante $+1$ y es una rotación vectorial (de ángulo α). La matriz de f en la base u_1, u_2, u_3 es

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha & 0 \\ \operatorname{sen}\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Diremos entonces que f es una *rotación vectorial* de *eje* $\langle u_3 \rangle$ y *ángulo* α. Observemos que $\cos\alpha$ queda determinado por la traza:

$$\cos\alpha = \frac{1}{2}(\operatorname{tr} f - 1).$$

El signo de $\operatorname{sen}\alpha$ depende de la orientación de la base u_1, u_2 de $\langle u_3 \rangle^{\perp}$.

¡Atención!:

El signo de $\operatorname{sen}\alpha$ no depende de la orientación de u_1, u_2, u_3.

El eje de f está contenido en el núcleo de la aplicación

$$\phi = f - f^{-1}.$$

Sea $A = (a_i^j)$ la matriz de f en una base ortonormal e_1, e_2, e_3. La matriz de f^{-1} es $A^{-1} = A^t$ y la de ϕ

$$A - A^t = \begin{pmatrix} 0 & a_2^1 - a_1^2 & a_3^1 - a_1^3 \\ a_1^2 - a_2^1 & 0 & a_3^2 - a_2^3 \\ a_1^3 - a_3^1 & a_2^3 - a_3^2 & 0 \end{pmatrix}.$$

- Si $\phi \neq 0$, $\operatorname{Nuc}\phi = \langle (a_2^3 - a_3^2), (a_3^1 - a_1^3), (a_1^2 - a_2^1) \rangle$ es el eje de f.
- Si $\phi = 0$ (es decir, si $A = A^t$ es simétrica), entonces $f = f^{-1}$ y $f^2 = I$. Las únicas aplicaciones de $SO(3)$ que cumplen estas condiciones son la identidad y una rotación de ángulo π o *simetría axial*. Para todo v,

$$f(v + f(v)) = f(v) + f^2(v) = f(v) + v$$

es del eje. En particular,

$$e_1 + f(e_1) = (a_1^1 + 1, a_1^2, a_1^3)$$

$$e_2 + f(e_2) = (a_2^1, a_2^2 + 1, a_2^3)$$

$$e_3 + f(e_3) = (a_3^1, a_3^2, a_3^3 + 1)$$

son del eje. Observemos que estos tres vectores no pueden ser simultáneamente $\vec{0}$, ya que entonces $f = -I \notin SO(3)$.

Observación:

Es fácil ver que una aplicación lineal $\Phi : E \longrightarrow E$ tiene una matriz de la forma

$$\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

en una base ortonormal si y sólo si, para todo $v \in E$, $v \cdot \phi(v) = 0$. Existe, entonces, un vector $u \in E$ ($u = (-c, b, -a) \in \operatorname{Nuc}\Phi$) tal que

$$\Phi = u \wedge - \ ;$$

es decir,

$$\Phi(v) = u \wedge v \qquad \forall v \in E.$$

Pasemos ahora a estudiar el conjunto de $f \in O(3)$ con $\det f = -1$. El mismo razonamiento hecho para demostrar que, si $f \in SO(3)$, f tiene el valor propio $+1$, prueba que si $f \in O(3)$ con $\det f = -1$, f tiene el valor propio -1. Sea u_3 un vector propio unitario de valor propio -1 y sea u_1, u_2, u_3 una base ortonormal. La restricción

$$f : \langle u_1, u_2 \rangle = \langle u_3 \rangle^{\perp} \longrightarrow \langle u_3 \rangle^{\perp}$$

es una aplicación ortogonal con determinante $+1$; es decir, una rotación en $\langle u_3\rangle^\perp$. La matriz de f en la base u_1, u_2, u_3 es de la forma

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha & 0 \\ \operatorname{sen}\alpha & \cos\alpha & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

En particular, si $\alpha = 0$, diremos que se trata de una *simetría especular* (vectorial) respecto al subespacio $\langle u_3\rangle^\perp$. En general, f es composición de una simetría especular y una rotación de eje ortogonal al plano de simetría.

La matriz de f que acabamos de encontrar nos dice que

$$\cos\alpha = \frac{1}{2}(\operatorname{tr} f + 1).$$

Tanto el plano de simetría $\langle u_3\rangle^\perp$ como el eje de rotación están determinados por el subespacio de vectores propios de valor propio -1. Este subespacio está contenido en el núcleo de

$$\phi = f - f^{-1}.$$

- Si $\phi \neq 0$, obtenemos, como en el caso del grupo ortogonal especial,

$$\operatorname{Nuc}\phi = \langle (a_2^3 - a_3^2), (a_3^1 - a_1^3), (a_1^2 - a_2^1)\rangle = \langle u_3\rangle.$$

- Si $\phi = 0$ (A simétrica: $A = A^t = A^{-1}$), $f^2 = I$. Las únicas simetrías ortogonales con $\det f = -1$ son las simetrías especulares y la *simetría central* $f = -I$. Para todo v,

$$f(f(v) - v) = f^2(v) - f(v) = -(f(v) - v).$$

En particular,

$$f(e_1) - e_1 = (a_1^1 - 1, a_1^2, a_1^3)$$

$$f(e_2) - e_2 = (a_2^1, a_2^2 - 1, a_2^3)$$

$$f(e_3) - e_3 = (a_3^1, a_3^2, a_3^3 - 1)$$

son ortogonales al plano de simetría.

XII.7 Otra determinación de las rotaciones

Una rotación $f \in SO(3)$ queda determinada por el eje $\langle u \rangle$ y el ángulo α. Éste, por su parte, está determinado por una orientación de $\langle u \rangle^\perp$ y por $\operatorname{sen}\alpha$ y $\cos\alpha$.

Supongamos fijada una orientación en el espacio E. Entonces, toda orientación en $\langle u \rangle$ induce una orientación en $\langle u \rangle^\perp$ y viceversa, de la siguiente manera: si $v \neq 0$ tiene orientación positiva en $\langle u \rangle$, una base $v_1, v_2 \in \langle u \rangle^\perp$ es de orientación positiva en $\langle u \rangle^\perp$ si y sólo si v_1, v_2, v es de orientación positiva en E.

Consideremos el eje orientado de forma que $\operatorname{sen}\alpha \geq 0$. Sea $v \in \langle u \rangle$ con

$$\|v\| = \left|\operatorname{sen}\frac{\alpha}{2}\right|$$

y orientación positiva. v nos da, por tanto, la orientación de $\langle u \rangle^\perp$ y el valor de $\cos\alpha$ y $\operatorname{sen}\alpha$:

$$\begin{aligned} \cos\alpha &= \cos^2\frac{\alpha}{2} - \operatorname{sen}^2\frac{\alpha}{2} = 1 - 2\|v\|^2 \\ \operatorname{sen}\alpha &= 2\cos\frac{\alpha}{2}\cdot\operatorname{sen}\frac{\alpha}{2} = 2\|v\|\sqrt{1-\|v\|^2}. \end{aligned}$$

La rotación queda, pues, completamente determinada por v y la denotaremos por g_v. Observemos que cualquier vector v con $\|v\| \leq 1$ determina una rotación.

Cuando $\operatorname{sen}\alpha = 0$, cualquiera de las dos orientaciones del eje cumple $\operatorname{sen}\alpha \geq 0$. Este caso se presenta si $\alpha = 0$ o $\alpha = \pi$.

- $\alpha = 0 \Rightarrow \|v\| = 0 \Rightarrow v = \vec{0}$.
- $\alpha = \pi \Rightarrow \|v\| = 1$ y v puede ser cualquiera de los dos vectores unitarios de $\langle u \rangle$. Entonces $g_v = g_{-v}$ es una simetría axial, que denotaremos por s_v.

XII.8 Composición de rotaciones

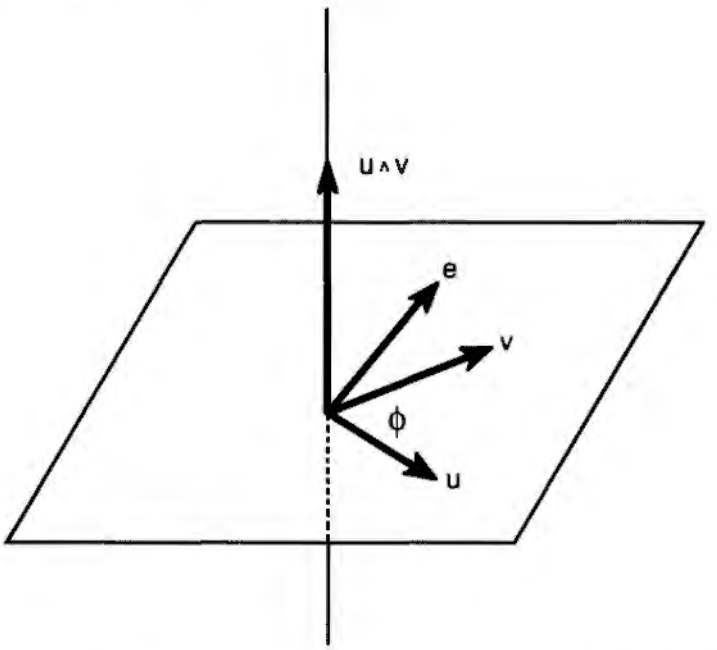
Si dos rotaciones tienen el mismo eje, es muy fácil ver que su composición es una rotación con ese eje y ángulo la suma de los ángulos de las dos rotaciones. La determinación de los elementos que caracterizan la rotación es mucho más compleja en el caso de rotaciones con ejes distintos. El objetivo de este apartado es, dadas dos rotaciones g_{v_1}, g_{v_2} (con la notación del §7), calcular el vector w tal que $g_w = g_{v_2} \circ g_{v_1}$.

Empezaremos por estudiar la composición de dos simetrías axiales.

Proposición 8.1

$$\begin{aligned} s_v \circ s_u &= g_{u\wedge v} \quad si \quad u\cdot v \geq 0 \\ &= g_{v\wedge u} \quad si \quad u\cdot v \leq 0. \end{aligned}$$

DEMOSTRACIÓN: $u\wedge v$ ortogonal a u y a $v \Rightarrow s_v \circ s_u(u\wedge v) = s_v(-u\wedge v) = u\wedge v$ $\Rightarrow \langle u \wedge v\rangle$ es del eje de $s_v \circ s_u$. $u \wedge v = \vec{0} \Leftrightarrow u = \pm v$ (ya que los dos son unitarios) $\Leftrightarrow s_u = s_v \Leftrightarrow s_v \circ s_u = I = g_{\vec{0}}$.



Supongamos $u \wedge v \neq \vec{0}$. El ángulo de la rotación $s_v \circ s_u$ es el ángulo que forman un vector cualquiera de $\langle u, v\rangle$, por ejemplo u, y su imagen: $s_v s_u(u) = s_v(u)$. Ese ángulo es precisamente $\alpha = 2\widehat{uv}$; en efecto, la matriz de s_v en una base ortonormal $\{u, e\}$ es, por el §5,

$$\begin{pmatrix} \cos 2\varphi & \operatorname{sen} 2\varphi \\ \operatorname{sen} 2\varphi & -\cos 2\varphi \end{pmatrix},$$

donde φ es el ángulo de v con el eje: $\varphi = \widehat{uv}$. Por tanto,

$$s_v(u) = \cos 2\varphi \cdot u + \operatorname{sen} 2\varphi \cdot e,$$

de donde $\alpha = \widehat{us_v(u)} = 2\varphi$.

Por otra parte, $\|u \wedge v\| = |\operatorname{sen} \widehat{uv}| = \left|\operatorname{sen} \dfrac{\alpha}{2}\right|$, de donde resulta que el vector que determina la rotación $s_v \circ s_u$ es $u \wedge v$ o $v \wedge u$.

Observemos que, con la notación anteriormente utilizada,

$$\det_{(u,e)}(u, v) = \operatorname{sen} \varphi$$

y, por tanto, $\operatorname{sen}\varphi$ es positivo o negativo según que la orientación de $\langle u, v\rangle$ sea la dada por u, v o la opuesta. Entonces tenemos que

- $u \cdot v > 0 \Leftrightarrow \cos\varphi > 0 \Rightarrow \operatorname{sen}\alpha = 2\operatorname{sen}\varphi\cos\varphi \geq 0$ sólo si $\operatorname{sen}\varphi \geq 0 \Rightarrow$ la orientación en $\langle u, v\rangle$ tiene que ser la dada por $u, v \Rightarrow$ la orientación del eje tiene que ser la de $u \wedge v \Rightarrow$

$$s_v \circ s_u = g_{u\wedge v}.$$

- $u \cdot v < 0 \Leftrightarrow \cos\varphi < 0 \Rightarrow \operatorname{sen}\alpha \geq 0$ sólo si $\operatorname{sen}\varphi \leq 0 \Rightarrow$ la orientación de $\langle u, v\rangle$ tiene que ser la de $v, u \Rightarrow$ la orientación del eje es la de $v \wedge u \Rightarrow$

$$s_v \circ s_u = g_{v\wedge u}.$$

- si $u \cdot v = 0$, entonces $\|u \wedge v\| = 1$ y por tanto $g_{u\wedge v} = g_{v\wedge u}$ vuelve a ser una simetría axial. □

La proposición (8.1) sugiere que toda rotación g_v se puede descomponer en producto de dos simetrías axiales de ejes perpendiculares al eje de la rotación. De hecho, una de las simetrías puede escogerse arbitrariamente.

Proposición 8.2 *Dada una simetría axial s_e ($\|e\| = 1$) y una rotación g_v con $v \cdot e = 0$, existen simetrías axiales s_w y s_u tales que*

$$g_v = s_e \circ s_w, \qquad g_v = s_u \circ s_e.$$

DEMOSTRACIÓN: $v = \vec{0} \Rightarrow g_v = I \Rightarrow s_w = s_e,\ s_u = s_e$.

Supongamos $v \neq \vec{0}$. Por (8.1), debemos buscar $w \in \langle v\rangle^{\perp} = \langle e, v \wedge e\rangle$. Escribamos

$$w = ae + b(v \wedge e).$$

Dado que $s_w = s_{-w}$, podemos suponer que $a = e \cdot w \geq 0$. Entonces, por (8.1),

$$v = w \wedge e = b(v \wedge e) \wedge e = b((v \cdot e)e - (e \cdot e)v) = -bv,$$

de donde $b = -1$. Además,

$$1 = \|w\|^2 = a^2 + \|v \wedge e\|^2 = a^2 + \|v\|^2 \Rightarrow a = \sqrt{1 - \|v\|^2}.$$

Pero $\|v\| = \left|\operatorname{sen}\dfrac{\alpha}{2}\right|$, donde α es el ángulo de la rotación g_v. Por tanto, $a = \left|\cos\dfrac{\alpha}{2}\right|$; es decir, si w existe, tiene que ser:

$$w = \left|\cos\frac{\alpha}{2}\right| e - v \wedge e.$$

Es inmediato comprobar que este w cumple efectivamente $s_e \circ s_w = g_v$. Por un razonamiento análogo a éste se obtiene

$$u = \left|\cos\frac{\alpha}{2}\right| e + v \wedge e. \ \square$$

El problema de hallar la rotación composición de dos g_{v_1}, g_{v_2} con ejes no paralelos ($v_1 \wedge v_2 \neq \vec{0}$) se reduce, por (8.3), a un problema de composición de simetrías axiales que (8.2) resuelve. En efecto, si e es un vector unitario ortogonal a v_1 y a v_2, (8.3) nos da vectores u_1, u_2 tales que

$$g_{v_1} = s_e \circ s_{u_1}, \qquad g_{v_2} = s_{u_2} \circ s_e.$$

Entonces, por (8.2),

$$\begin{aligned} g_{v_2} \circ g_{v_1} = s_{u_2} \circ s_{u_1} &= g_{u_1 \wedge u_2} \quad \text{si} \quad u_1 \cdot u_2 \geq 0 \\ &= g_{u_2 \wedge u_1} \quad \text{si} \quad u_1 \cdot u_2 \leq 0. \end{aligned}$$

Vamos a calcular estos vectores. Para ello, elijamos un vector e concreto $e = \dfrac{v_1 \wedge v_2}{\|v_1 \wedge v_2\|}$. Entonces, si α_1, α_2 son los ángulos de las rotaciones g_{v_1}, g_{v_2},

$$u_1 = \left|\cos\frac{\alpha_1}{2}\right| e + e \wedge v_1 \qquad \text{y} \qquad u_2 = \left|\cos\frac{\alpha_2}{2}\right| e - e \wedge v_2,$$

de donde

$$\begin{aligned} u_1 \wedge u_2 &= \left|\cos\frac{\alpha_1}{2}\right| v_2 + \left|\cos\frac{\alpha_2}{2}\right| v_1 - (e \wedge v_1) \wedge (e \wedge v_2) = \\ &= \left|\cos\frac{\alpha_1}{2}\right| v_2 + \left|\cos\frac{\alpha_2}{2}\right| v_1 - (v_1 \wedge v_2), \\ u_1 \cdot u_2 &= \left|\cos\frac{\alpha_1}{2}\right| \left|\cos\frac{\alpha_2}{2}\right| - (e \wedge v_1)(e \wedge v_2) = \\ &= \left|\cos\frac{\alpha_1}{2}\right| \left|\cos\frac{\alpha_2}{2}\right| - v_1 \cdot v_2 = \\ &= \left|\cos\frac{\alpha_1}{2}\right| \left|\cos\frac{\alpha_2}{2}\right| - \left|\operatorname{sen}\frac{\alpha_1}{2}\right| \left|\operatorname{sen}\frac{\alpha_2}{2}\right| \cos\widehat{v_1 v_2}. \end{aligned}$$

Estas expresiones de $u_1 \wedge u_2$ y $u_1 \cdot u_2$ son válidas también si las dos rotaciones tienen el mismo eje: $v_1 \wedge v_2 = \vec{0}$. En ese caso, en los cálculos anteriores, e es un vector unitario cualquiera ortogonal al eje y se cumple $(e \wedge v_1) \wedge (e \wedge v_2) = \vec{0} = v_1 \wedge v_2$. Tenemos, pues, el siguiente resultado:

Proposición 8.3 *Sean g_{v_1}, g_{v_2} dos rotaciones de ángulos α_1 y α_2 respectivamente. Si el valor de*

$$\left|\cos\frac{\alpha_1}{2}\right|\left|\cos\frac{\alpha_2}{2}\right| - \left|\operatorname{sen}\frac{\alpha_1}{2}\right|\left|\operatorname{sen}\frac{\alpha_2}{2}\right|\cos\widehat{v_1 v_2}$$

es positivo, entonces

$$g_{v_2} \circ g_{v_1} = g_w,$$

donde

$$w = \left|\cos\frac{\alpha_1}{2}\right| v_2 + \left|\cos\frac{\alpha_2}{2}\right| v_1 - (v_1 \wedge v_2).$$

Si aquel valor es negativo,

$$g_{v_2} \circ g_{v_1} = g_{-w}$$

para el mismo w. □

XII.9 Nota histórica

Pese a que el nombre "matriz ortogonal" ya fue utilizado en 1854 por Charles Hermite (1822–1901), no fue hasta 1878 que Georg Ferdinand Frobenius (1849–1917) dio una definición formal de este concepto, demostrando sus primeras propiedades. Hermite había considerado ya en 1855 las matrices hermíticas, demostrando que los valores propios son reales, resultado demostrado por Arthur Buckheim (1859–1888) en 1885 para las matrices simétricas.

XII.10 Ejercicios

1. Demostrar que si F es un subespacio vectorial de E invariante por $f \in O(E)$, $F^\perp$ también es invariante por f.

2. Demostrar que para toda $A \in M_{n\times n}(\mathbf{R})$, $A = (a_i^j)$, $\sum_{i,j}(a_i^j)^2$ es invariante por cambios de base ortogonales.

   (Indicación: probar que $\sum_{i,j}(a_i^j)^2 = \operatorname{tr}(AA^t)$.)

3. Estudiar el subgrupo de $O(2)$ que deja invariante el conjunto formado por dos subespacios vectoriales de dimensión 1.

4. Estudiar los elementos de $O(3)$ que dejan invariante el conjunto formado por dos planos ortogonales.

5. Dados dos vectores $u \neq v$ de igual norma de un espacio vectorial euclídeo de dimensión 2, estudiar las aplicaciones ortogonales f tales que $f(u) = v$ y $f(v) = u$.

6. En una base u, v, una rotación vectorial tiene por matriz

$$\begin{pmatrix} 1 & -3/2 \\ 2/3 & 0 \end{pmatrix}.$$

Determinar el ángulo de la rotación, el ángulo de los vectores u, v y la razón $\dfrac{\|v\|}{\|u\|}$.

7. Sea f una simetría axial de $\mathbf{R}^2$ de eje $\langle u \rangle$ y v un vector cualquiera. Demostrar que $\widehat{vf(v)} = 2\widehat{vu}$.

8. Sea

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

una matriz simétrica y $D = Q^{-1}AQ$ su forma diagonal. Demostrar que Q es una rotación de ángulo

$$\phi = \frac{1}{2} \arctan\left(\frac{2b}{d-a}\right).$$

9. Sea $f \in \mathrm{End}(E)$ tal que, para todo $u, v \in E$,

$$f(u) \cdot f(v) = \lambda(u \cdot v) \qquad (\lambda > 0 \text{ fijo}).$$

Demostrar que:

a) $\|f(v)\| = \mu\|v\|$ con $\mu = +\sqrt{\lambda}$, para todo $v \in E$;
b) f es biyectiva;
c) los únicos valores propios posibles de f son $\pm\mu$;
d) $f = g \circ h$, donde $g \in O(n)$ y h es una homotecia.

10. Demostrar que si una aplicación $f \in \mathrm{End}(\mathbf{R}^2)$ tiene, en una base ortonormal, una matriz de la forma

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} a & b \\ b & -a \end{pmatrix},$$

entonces existe un $\lambda \in \mathbf{R}$ tal que

$$f(u) \cdot f(v) = \lambda(u \cdot v) \quad \forall u, v \in \mathbf{R}^2.$$

Calcular λ y determinar la descomposición del apartado (d) del ejercicio anterior.

11. Demostrar que toda matriz cuadrada real A admite una descomposición $A = QR$ con Q ortogonal y R triangular superior.

(Indicación: demostrar primero que, si $B \in M_{2\times 2}(\mathbf{R})$, existe siempre una rotación P tal que la matriz PB es triangular superior. Entonces, dada $A \in M_{n\times n}(\mathbf{R})$, se pueden ir eliminando todos los elementos subdiagonales con rotaciones bidimensionales apropiadas.)

Demostrar que, si $\det A \neq 0$, la descomposición $A = QR$ es única.

12. Demostrar que todo endomorfismo de un espacio vectorial euclídeo de dimensión finita se puede descomponer en producto de un endomorfismo autoadjunto y uno ortogonal. Estudiar la unicidad de la descomposición.

(Indicación: utilizar el ejercicio XI.12.)

XII.11 Ejercicios para programar

13. Hacer un programa que, dados dos vectores u, v de $\mathbf{R}^2$, calcule el ángulo $\widehat{uv}$.

14. Utilizando el ejercicio XI.14, preparar una lista de matrices de $O(3)$ para utilizarlas como ejemplos donde convenga.

15. Preparar un programa que permita

a) Dada $A \in O(3)$, encontrar el eje y el ángulo salvo el signo (§6).

b) Dada $A \in SO(3)$, encontrar el vector $v \in \mathbf{R}^3$ tal que $A = g_v$ (§7).

c) Dado $v \in \mathbf{R}^3$ unitario, encontrar la matriz de la rotación g_v en la base canónica.

16. Dadas dos rotaciones g_u, g_v de $\mathbf{R}^3$, calcular su composición $g_u \circ g_v$ (8.3).

Proponemos a continuación cuatro métodos iterativos para calcular valores propios reales de matrices (véase el ejercicio VIII.26).

17. Método QR

Escribir $A = QR$ donde Q es ortogonal y R es triangular superior (ejercicio 11). Sea $A_1 = RQ$. El algoritmo consiste en ir descomponiendo $A_k = Q_k R_k$ y haciendo $A_{k+1} = R_k Q_k$, $k = 1, 2, \ldots$. Las matrices A_k son todas ellas equivalentes a A y tienden hacia una matriz triangular superior, bajo hipótesis más generales que en el método LU.

18. Método de Jacobi (Se aplica solamente cuando la matriz es simétrica.)

Dada A simétrica, cada rotación bidimensional Q como la del ejercicio 11 anula un par de elementos de A. La rotación siguiente no respeta los ceros conseguidos, pero es fácil deducir del ejercicio 2 que el valor $\sum_{i\neq j}(a_i^j)^2$ disminuye en cada rotación.

Preparar el programa de la siguiente manera: sea a_i^j el elemento de A de módulo máximo entre los que no están en la diagonal principal. Pongamos $A_1 = {Q_1}^{-1} A Q_1$, donde Q_1 es una rotación bidimensional que anula a_i^j. Repitiendo el proceso se obtiene una sucesión $A_k = {Q_k}^{-1} A Q_k$ que converge hacia una matriz diagonal.

Observemos que el método nos da también una base ortonormal de vectores propios.

19. Método de la potencia

Dada $A \in M_{n\times n}(\mathbf{R})$ diagonalizable en $\mathbf{C}$ y tal que sus valores propios satisfacen $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$, este algoritmo permite aproximar el valor propio dominante λ_1 y el vector propio u_1 correspondiente.

Escoger un vector $x_0 \in \mathbf{R}^n$ cualquiera ($x_0 \neq \vec{0}$). Considerar la sucesión

$$x_{k+1} = A\left(\frac{x_k}{\|x_k\|}\right), \quad k = 0, 1, 2, \ldots.$$

Es sencillo ver que $\{x_k\}$ converge hacia el vector propio u_1. El valor propio λ_1 se puede obtener a partir de una componente cualquiera no nula de x_k, ya que, para k grande, Ax_k aproxima $\lambda_1 u_1$.

Este método también es válido si λ_1 tiene multiplicidad mayor que 1. En ese caso, $\{x_k\}$ converge hacia algún vector del subespacio de vectores propios de valor propio λ_1 (dependiendo del x_0 inicial).

20. Método de iteración inversa

Una sencilla variante del método anterior permite aproximar todos los valores propios reales de A y los vectores propios correspondientes.

Sólo hay que observar que, si $\alpha \in \mathbf{R}$ es un parámetro cualquiera, entonces, si $\alpha \neq \lambda_i$,

$$Au_i = \lambda_i u_i \iff (A - \alpha I)^{-1} u_i = \frac{1}{\lambda_i - \alpha} u_i.$$

Así pues, si λ es el valor propio de A más próximo a α, resulta que $1/(\lambda - \alpha)$ es valor propio dominante de $(A - \alpha I)^{-1}$. Por tanto, la

sucesión

$$\begin{cases} x_0 \neq \vec{0} \qquad \text{cualquiera} \\ x_{k+1} = (A - \alpha I)^{-1} \left(\dfrac{x_k}{\|x_k\|} \right), \quad k = 0, 1, 2, \ldots \end{cases}$$

tiende hacia el vector propio u tal que $Au = \lambda u$. Entonces, los cocientes x_{k+1}^i / x_k^i tienden hacia $1/(\lambda - \alpha)$.

Así, dando valores a α, podemos ir "cazando" uno a uno los valores propios reales de A.

Capítulo XIII

Espacios afines euclídeos

En el estudio de los espacios afines llevado a cabo en el capítulo IX no aparecen conceptos tan usuales en la geometría clásica como distancias, perpendicularidad y ángulos. El motivo es que estas nociones están relacionadas con la existencia de un producto escalar en el espacio vectorial asociado al espacio afín. En este capítulo estudiaremos los espacios afines reales que tienen asociado un espacio vectorial euclídeo: los espacios afines euclídeos.

XIII.1 Espacios afines euclídeos

Un espacio afín real (A, E) se llama un *espacio afín euclídeo* si en E hay un producto escalar, es decir, si E es un espacio vectorial euclídeo.

Dados dos puntos de un espacio afín euclídeo, $p, q \in A$, se llama *distancia entre p y q* al número real

$$\mathrm{d}(p,q) = \|\overrightarrow{pq}\|.$$

La aplicación $\mathrm{d} : A \times A \longrightarrow \mathbf{R}$ que asigna a cada par $(p,q) \in A \times A$ el número real $\mathrm{d}(p,q)$ se llama *aplicación distancia* y cumple las siguientes propiedades para todo $p, q, r \in A$:

1. $\mathrm{d}(p,q) \geq 0; \quad \mathrm{d}(p,q) = 0 \Leftrightarrow p = q.$
2. $\mathrm{d}(p,q) = \mathrm{d}(q,p).$
3. $\mathrm{d}(p,q) \leq \mathrm{d}(p,r) + \mathrm{d}(r,q) \quad$ (*desigualdad triangular*).
4. $\mathrm{d}(p,q) \geq |\mathrm{d}(p,r) - \mathrm{d}(r,q)|.$

Las tres primeras propiedades son consecuencia de propiedades de la norma. Para demostrar la cuarta, observemos que, por 3 y 2,

$$\begin{cases} \mathrm{d}(p,q)+\mathrm{d}(q,r)\geq \mathrm{d}(p,r) \Rightarrow \mathrm{d}(p,q)\geq \mathrm{d}(p,r)-\mathrm{d}(r,q) \\ \mathrm{d}(r,p)+\mathrm{d}(p,q)\geq \mathrm{d}(r,q) \Rightarrow \mathrm{d}(p,q)\geq \mathrm{d}(r,q)-\mathrm{d}(p,r). \end{cases}$$

Ejercicio:

Probar que $\mathrm{d}(p,q)=\mathrm{d}(p,x)+\mathrm{d}(x,q)$ si y sólo si el punto x es del segmento $\overline{pq}$.

Proposición 1.1 (Teorema de Pitágoras) *Sean p,q,r tres puntos de un espacio afín euclídeo (A,E). Si $\overrightarrow{pq}\cdot\overrightarrow{pr}=0$, se cumple*

$$\mathrm{d}(p,q)^2+\mathrm{d}(p,r)^2=\mathrm{d}(q,r)^2.$$

DEMOSTRACIÓN: $\mathrm{d}(q,r)^2=\|\overrightarrow{qr}\|^2=\overrightarrow{qr}\cdot\overrightarrow{qr}=(\overrightarrow{pr}-\overrightarrow{pq})\cdot(\overrightarrow{pr}-\overrightarrow{pq})=\|\overrightarrow{pr}\|^2+\|\overrightarrow{pq}\|^2=\mathrm{d}(p,r)^2+\mathrm{d}(p,q)^2$. □

Sea (A,E) un espacio afín euclídeo de dimensión n.

Dos variedades $a+F$ y $b+G$ tales que cualquier vector $u\in F$ es ortogonal a cualquier vector $v\in G$ $(u\cdot v=0)$ se llaman *ortogonales* o *perpendiculares*. Entonces $F\subset G^\perp$ y, por tanto, la suma de las dimensiones de esas variedades es $\leq n$. Dos variedades $a+F$ y $b+G$ tales que $\dim F+\dim G\geq n$ diremos que son *ortogonales* o *perpendiculares* si las variedades $a+F^\perp$ y $b+G^\perp$ son ortogonales, es decir, si $F^\perp\subset G$.

Ejemplos:

1. Sea $a_1x^1+\ldots+a_nx^n=b$ la ecuación de un hiperplano en un sistema de referencia $\{p;\ e_1,\ldots,e_n\}$ ortonormal (tal que la base $e_1,\ldots,e_n$ es ortonormal). El vector de coordenadas $(a_1,\ldots,a_n)$ es ortogonal a cualquier vector $(x^1,\ldots,x^n)$ de la dirección de H, ya que estos vectores cumplen $a_1x^1+\ldots+a_nx^n=0$.

2. Sea A el espacio afín euclídeo de dimensión 3. Dos rectas $a+\langle w\rangle$ y $b+\langle v\rangle$ se cruzan si no son paralelas ni se cortan. Entonces w,v son linealmente independientes y $\overrightarrow{ab}\notin\langle w,v\rangle$ (IX.4.1).

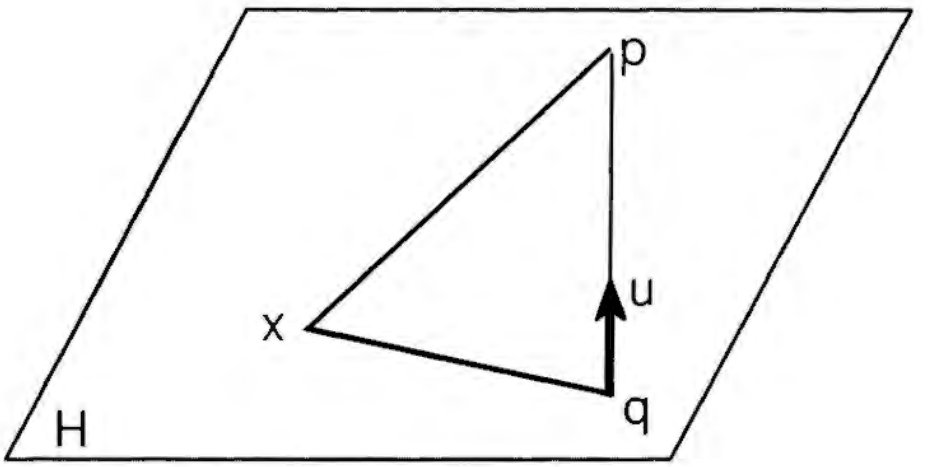
   Consideremos los planos que contienen una de las rectas y la dirección perpendicular a ambas. Si $\vec{0}\neq u\in\langle w,v\rangle^\perp$, estos planos son

$$a+\langle w,u\rangle, \qquad b+\langle v,u\rangle.$$

Puesto que w, v, u forman una base, $\overrightarrow{ab} \in \langle w, u\rangle + \langle v, u\rangle$ y los planos se cortan en una recta $c + \langle u\rangle$. Esa recta corta a $a + \langle w\rangle$, ya que $\overrightarrow{ac} \in \langle w, u\rangle$, y corta a $b + \langle v\rangle$, ya que $\overrightarrow{bc} \in \langle v, u\rangle$. La recta $c + \langle u\rangle$ se llama la *perpendicular común* a $a + \langle w\rangle$ y $b + \langle v\rangle$; los puntos de intersección con esas rectas se llaman los *pies de la perpendicular común.*

3. Sea $H = a + F$ un hiperplano y p un punto tal que $p \notin H$. Dado que $F \oplus F^{\perp} = E$, la recta $p + F^{\perp}$ corta a H exactamente en un punto (IX, §4). Este punto se llama la *proyección ortogonal de p sobre H.*

   Si lo designamos por q, entonces para todo $x \in H$ se tiene $\overrightarrow{pq} \cdot \overrightarrow{qx} = 0$, ya que $\overrightarrow{pq} \in F^{\perp}$ y $\overrightarrow{qx} \in F$.



XIII.2 Distancia entre dos variedades lineales

Dadas dos variedades lineales L_1, L_2, el conjunto de las distancias entre sus puntos tiene un mínimo que denominaremos *distancia entre* L_1 *y* L_2:

$$\mathrm{d}(L_1, L_2) = \mín\{\mathrm{d}(p, q) \mid p \in L_1,\ q \in L_2\}.$$

La existencia de este mínimo es consecuencia de la continuidad de la aplicación distancia y del hecho de que el conjunto de distancias está acotado inferiormente por 0. Además, se puede ver también que siempre existen puntos $a \in L_1$, $b \in L_2$ tales que $\mathrm{d}(a, b) = \mathrm{d}(L_1, L_2)$. Nosotros no vamos a demostrar aquí ninguno de estos hechos en general; nos limitaremos a un

par de casos particulares interesantes. El tratamiento de esos casos es, por otra parte, un buen modelo para hallar la distancia entre otros tipos de variedades.

I Distancia de un punto p a un hiperplano $H = a + F$

Sea q la proyección ortogonal de p sobre H. Entonces, para todo $x \in H$,

$$\mathrm{d}(p,x)^2 = \mathrm{d}(p,q)^2 + \mathrm{d}(q,x)^2 \geq \mathrm{d}(p,q)^2 \quad \text{(por (1.1))},$$

ya que $\overrightarrow{pq} \cdot \overrightarrow{qx} = 0$. Tenemos, pues, que $\mathrm{d}(p,q)$ es el mínimo de las distancias del punto p a los puntos de H: $\mathrm{d}(p,H) = \mathrm{d}(p,q)$. Para hallar concretamente el valor de esta distancia, consideremos un vector unitario u ortogonal a la dirección de H. Entonces $\overrightarrow{pq} = ku$ y $\mathrm{d}(p,q) = |k|$. Sea $x \in H$ cualquiera;

$$\overrightarrow{px} \cdot u = (\overrightarrow{pq} + \overrightarrow{qx}) \cdot u = \overrightarrow{pq} \cdot u + \overrightarrow{qx} \cdot u = ku \cdot u = k,$$

ya que $\overrightarrow{qx} \cdot u = 0$. Por tanto,

$$\mathrm{d}(p,H) = \left|\overrightarrow{px} \cdot u\right|,$$

donde $x \in H$ es arbitrario y u es un vector unitario ortogonal a la dirección de H.

Si $a_1x^1 + \ldots + a_nx^n + b = 0$ es la ecuación de H en un sistema de referencia ortonormal, elegimos

$$u = \frac{1}{r}(a_1, \ldots, a_n),$$

donde $r = \|(a_1, \ldots, a_n)\| = \sqrt{{a_1}^2 + \ldots + {a_n}^2}$. Así pues,

$$\begin{aligned}
\overrightarrow{px} \cdot u &= \frac{1}{r}\left((x^1 - p^1)a_1 + \ldots + (x^n - p^n)a_n\right) = \\
&= \frac{1}{r}\left(x^1a_1 + \ldots + x^na_n - p^1a_1 - \ldots - p^na_n\right) = \\
&= -\frac{1}{r}(b + p^1a_1 + \ldots + p^na_n),
\end{aligned}$$

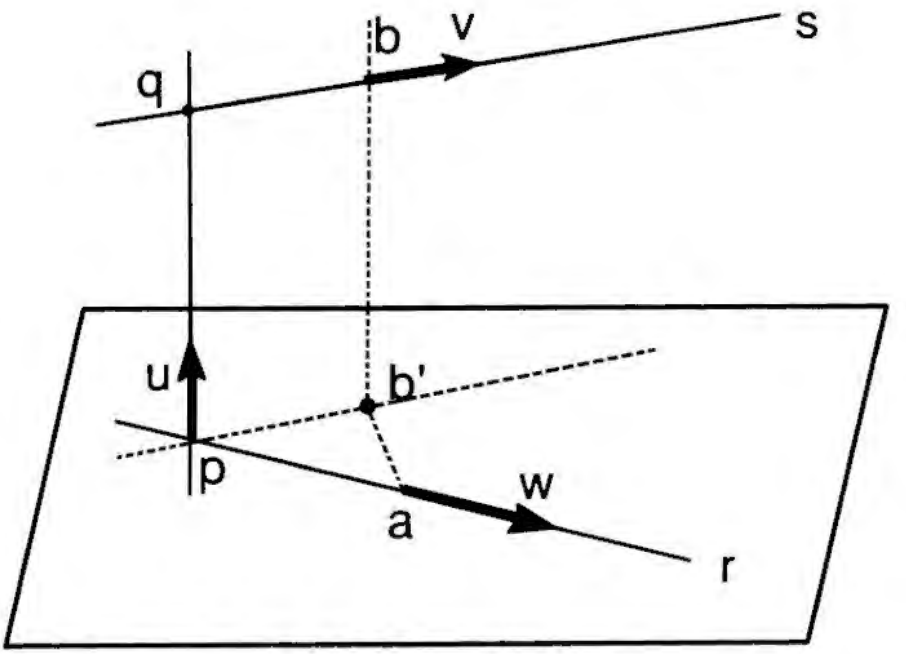
y, por tanto,

$$\mathrm{d}(p,H) = \frac{|a_1p^1 + \ldots + a_np^n + b|}{\sqrt{{a_1}^2 + \ldots + {a_n}^2}}.$$

Observemos que el numerador de esta expresión es el valor de la ecuación de H en el punto $p = (p^1, \ldots, p^n)$. En particular, si $p \in H$, entonces obviamente $\mathrm{d}(p, H) = 0$.

II Distancia entre dos rectas que se cruzan, en un espacio afín euclídeo de dimensión 3

Sean $r = a + \langle w \rangle$ y $s = b + \langle v \rangle$ esas dos rectas, p, q los pies de la perpendicular común y $p + \langle u \rangle$ esa perpendicular (§1). Supongamos u unitario.



Consideremos el plano $a + \langle w, v \rangle$ y la proyección ortogonal b' de b sobre ese plano: b' es la intersección de $b + \langle u \rangle$ con $a + \langle w, v \rangle$. Entonces $\overrightarrow{bb'}$ y $\overrightarrow{b'a}$ son ortogonales y $\mathrm{d}(a, b) \geq \mathrm{d}(b', b)$ por (1.1). Ahora bien, el punto $c = b + \overrightarrow{qp}$ es de $b + \langle u \rangle$ y de $p + \langle w, v \rangle = a + \langle w, v \rangle$, ya que $\overrightarrow{pc} = \overrightarrow{pb} + \overrightarrow{qp} = \overrightarrow{qb} \in \langle v \rangle$. Por tanto, $b' = c = b + \overrightarrow{qp}$, de donde $\overrightarrow{b'b} = \overrightarrow{qp}$ y $\mathrm{d}(b', b) = \mathrm{d}(p, q)$. Así pues, $\mathrm{d}(p, q)$ es la mínima distancia entre los puntos de r y s:

$$\mathrm{d}(r, s) = \mathrm{d}(p, q).$$

Si $\overrightarrow{pq} = ku$, $\mathrm{d}(p, q) = |k| = \left|\overrightarrow{pq} \cdot u\right|$. De hecho, este producto escalar no depende de los puntos p y q sobre las rectas r y s; en efecto, de $\overrightarrow{ap} \cdot u = 0$ y $\overrightarrow{qb} \cdot u = 0$ resulta que

$$\overrightarrow{ab} \cdot u = (\overrightarrow{ap} + \overrightarrow{pq} + \overrightarrow{qb}) \cdot u = \overrightarrow{pq} \cdot u$$

y

$$\mathrm{d}(r,s) = \left|\overrightarrow{ab}\cdot u\right|,$$

donde a y b son puntos arbitrarios de r y s respectivamente, y u es un vector unitario perpendicular a las direcciones de r y s.

Cuando el sistema de referencia $\{O;\ e_1,e_2,e_3\}$ es ortonormal, tomando $u = \dfrac{w\wedge v}{\|w\wedge v\|}$, obtenemos

$$\mathrm{d}(r,s) = \frac{\det_{(e_i)}(w,v,\overrightarrow{ab})}{\|w\wedge v\|}.$$

XIII.3 Isometrías

Una aplicación $f : A_1 \longrightarrow A_2$ entre dos espacios afines euclídeos se llama una *isometría* si

$$\mathrm{d}(f(a),f(b)) = \mathrm{d}(a,b) \qquad \forall a,b\in A_1.$$

Las isometrías son siempre inyectivas, ya que si $f(a) = f(b)$,

$$0 = \mathrm{d}(f(a),f(b)) = \mathrm{d}(a,b)$$

y, por tanto, $a = b$.

Las aplicaciones que conservan la estructura de espacio afín euclídeo deberían ser las afinidades tales que la aplicación lineal asociada conservase el producto escalar. La siguiente proposición nos asegura que estas aplicaciones son, precisamente, las isometrías.

Proposición 3.1 *Una aplicación $f : A_1 \longrightarrow A_2$ entre espacios afines euclídeos es una isometría si y sólo si es una afinidad y su aplicación lineal asociada conserva el producto escalar.*

DEMOSTRACIÓN: Supongamos que f es una isometría y definamos una aplicación entre los espacios vectoriales asociados de la siguiente manera: fijado un punto $p\in A$,

$$\begin{array}{rccc} \phi: & E_1 & \longrightarrow & E_2 \\ & u=\overrightarrow{pa} & \longmapsto & \phi(u)=\overrightarrow{f(p)f(a)} \end{array}$$

(ver demostración de (X.3.5)). Esta aplicación conserva el producto escalar, ya que, dados $u = \overrightarrow{pa}$ y $v = \overrightarrow{pb}$,

$$\begin{aligned} \mathrm{d}(a,b)^2 &= \|\overrightarrow{ab}\|^2 = \|\overrightarrow{ap}+\overrightarrow{pb}\|^2 = \|\overrightarrow{ap}\|^2 + \|\overrightarrow{pb}\|^2 + 2\overrightarrow{ap}\cdot\overrightarrow{pb} = \\ &= \mathrm{d}(a,p)^2 + \mathrm{d}(p,b)^2 - 2u\cdot v. \end{aligned}$$

Análogamente se obtiene

$$\mathrm{d}(f(a),f(b))^2 = \mathrm{d}(f(a),f(p))^2 + \mathrm{d}(f(p),f(b))^2 - 2\,\overrightarrow{f(p)f(a)}\cdot\overrightarrow{f(p)f(b)};$$

de donde

$$u\cdot v = \overrightarrow{f(p)f(a)}\cdot\overrightarrow{f(p)f(b)} = \phi(u)\cdot\phi(v).$$

Por conservar el producto escalar, ϕ es lineal (XII.1.1). Finalmente,

$$\phi(\overrightarrow{ab}) = \phi(\overrightarrow{pb}-\overrightarrow{pa}) = \phi(\overrightarrow{pb}) - \phi(\overrightarrow{pa}) = \overrightarrow{f(p)f(b)} - \overrightarrow{f(p)f(a)} = \overrightarrow{f(a)f(b)}$$

y, por tanto, f es una afinidad con aplicación lineal asociada ϕ.

Supongamos ahora que f es una afinidad y que $\tilde{f}$ conserva el producto escalar. Entonces

$$\mathrm{d}(f(a),f(b)) = \|\overrightarrow{f(a)f(b)}\| = \|\tilde{f}(\overrightarrow{ab})\| = \|\overrightarrow{ab}\| = \mathrm{d}(a,b)$$

y f es una isometría. □

Las isometrías tienen, pues, todas las propiedades de las afinidades. En particular, conservan el paralelismo y la razón simple. Además, por (3.1) conservan también la perpendicularidad.

Dos espacios afines euclídeos A_1, A_2 se llaman *isomorfos* si existe una isometría biyectiva $f : A_1 \longrightarrow A_2$. Escribiremos entonces $A_1 \cong A_2$. Dos espacios afines euclídeos isomorfos son de la misma dimensión (X.1.6). Recíprocamente, si $\dim A_1 = \dim A_2 = n < \infty$, $A_1 \cong A_2$. En efecto, sean $\{e_1,\ldots,e_n\}$ y $\{u_1,\ldots,u_n\}$ bases ortonormales de E_1 y E_2 respectivamente. Designemos por $\phi : E_1 \longrightarrow E_2$ el isomorfismo de espacios vectoriales que aplica una base en la otra: $\phi(e_i) = u_i$, $i = 1,\ldots,n$. ϕ conserva el producto escalar y cualquier afinidad f con aplicación lineal asociada ϕ es un isomorfismo de espacios afines euclídeos. Hemos probado así el siguiente resultado:

Proposición 3.2 *Dos espacios afines euclídeos de dimensión finita son isomorfos si y sólo si tienen la misma dimensión.* □

Así, salvo isomorfismos, existe un único espacio afín euclídeo para cada dimensión n. Un modelo particular es el espacio afín estándar $\mathbf{R}^n$ con el producto escalar para el que la base $(1,0,\ldots,0),\ldots,(0,\ldots,0,1)$ es ortonormal.

XIII.4 Clasificación de los desplazamientos

Un *desplazamiento* o *movimiento* es una isometría de un espacio afín euclídeo en sí mismo.

Dos desplazamientos $f, g : A \longrightarrow A$ son *de la misma clase* si existe un isomorfismo de espacios afines euclídeos $\phi : A \longrightarrow A$ que hace conmutativo el diagrama

$$\begin{array}{ccc} A & \xrightarrow{f} & A \\ \phi\downarrow & & \downarrow\phi \\ A & \xrightarrow{g} & A; \end{array}$$

es decir, si $\phi \circ f = g \circ \phi$. Esta relación es claramente de equivalencia. Compararla con la relación definida en (X.7) para afinidades. La proposición siguiente corresponde a la proposición 7.1 de aquel capítulo.

Proposición 4.1 *Dos desplazamientos $f, g : A \longrightarrow A$ son de la misma clase si y sólo si existen sistemas de referencia ortonormales tales que las ecuaciones de f en uno de ellos coinciden con las ecuaciones de g en el otro.*

DEMOSTRACIÓN: Vale la misma demostración de (X.7.1), observando simplemente que $\tilde{\phi}$ conserva el producto escalar si y sólo si transforma bases ortonormales en bases ortonormales. □

Ejemplo:

Dos traslaciones T_u y T_v son de la misma clase si existe un isomorfismo de espacios afines euclídeos ϕ tal que $\phi \circ T_u(a) = T_v \circ \phi(a)$ para todo a; es decir, $\phi(a) + \tilde{\phi}(u) = \phi(a) + v$. Por tanto, $\tilde{\phi}(u) = v$. Esto es posible si y sólo si u y v tienen la misma norma. Recordemos, sin embargo, que, como afinidades y según la clasificación establecida en (X.7), T_u y T_v son de la misma clase siempre que $u \neq \vec{0} \neq v$.

Las ecuaciones de T_u, $u \neq \vec{0}$, en una referencia $\{p; \dfrac{u}{\|u\|}, e_2, \ldots, e_n\}$, son

$$\begin{cases} \bar{x}^1 = x^1 + \|u\| \\ \bar{x}^i = x^i \qquad\qquad i = 2, \ldots, n. \end{cases}$$

La aplicación lineal asociada a un desplazamiento es una aplicación ortogonal; su determinante es, pues, ± 1. Diremos que un desplazamiento $f : A \longrightarrow A$ es *propio* (o *directo*) si $\det \tilde{f} = +1$, y que es *impropio* (o *inverso*) si $\det \tilde{f} = -1$.

XIII.5 Desplazamientos de la recta euclídea

Sea $f : A \longrightarrow A$ un desplazamiento de un espacio afín euclídeo A de dimensión 1. Entonces $\tilde{f}$ es ortogonal y, por tanto, es la identidad I o $-I$. En el primer caso, o bien $f = I$ ($\bar{x} = x$) o bien $f = T_u$ con $u \neq \vec{0}$ y, en un sistema de referencia como el del ejemplo de (XIII.4),

$$\bar{x} = x + \|u\|.$$

Si $\tilde{f} = -I$, f es una homotecia de razón -1 y, tal como vimos en (X.2), tiene un único punto fijo. En cualquier referencia que tenga este punto como origen, la ecuación de f es

$$\bar{x} = -x.$$

f se llama entonces una *simetría central* de *centro* el punto fijo.

XIII.6 Desplazamientos del plano euclídeo

Sea $f : A \longrightarrow A$ un desplazamiento de un espacio afín euclídeo A de dimensión 2. Si f es un desplazamiento propio, $\tilde{f} \in SO(2)$ y en cualquier base ortonormal $\{e_1, e_2\}$ su matriz es del tipo

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Si $\cos\theta \neq 1$, $\tilde{f}$ no tiene el valor propio 1 y, por tanto, f tiene un único punto fijo q (X.6.1). Diremos, en este caso, que f es un *giro* o *rotación* de *ángulo* θ y *centro* q. Sus ecuaciones en la referencia $\{q;\ e_1, e_2\}$ son

$$\begin{cases} \bar{x} = x\cos\theta - y\operatorname{sen}\theta \\ \bar{y} = x\operatorname{sen}\theta + y\cos\theta. \end{cases}$$

En el caso particular en que $\cos\theta = -1$, f es una *simetría central* de *centro* q.

Si $\cos\theta = 1$, $\tilde{f} = I$ y f es una traslación de vector diferente de $\vec{0}$ o bien la identidad.

Cuando f es un desplazamiento impropio, $\tilde{f} \in O(2)$ con $\det\tilde{f} = -1$. Existe, entonces, una base ortonormal $\{e_1, e_2\}$ en la que la matriz de $\tilde{f}$ es

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Si f tiene un punto fijo q, sus ecuaciones en la referencia $\{q;\ e_1, e_2\}$ son

$$\begin{cases} \bar{x} = x \\ \bar{y} = -y \end{cases}$$

y se llama una *simetría axial.* La recta $q + \langle e_1 \rangle$ es de puntos fijos y se llama el *eje* de la simetría.

Si no existe ningún punto fijo, las ecuaciones de f en una referencia $\{p;\ e_1, e_2\}$, con $p \in A$ cualquiera, son

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = -y + d \end{cases}$$

con $c \neq 0$. La recta $y = d/2$ es invariante. Si escogemos el origen sobre esa recta, $q = (x_0, d/2)$, obtenemos $\overrightarrow{qf(q)} = (c, 0) = c\,e_1$. Las ecuaciones de f en la referencia $\{q;\ e_1, e_2\}$ son, por tanto,

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = -y. \end{cases}$$

Este desplazamiento es la *composición de una simetría axial* de eje $q + \langle e_1 \rangle$ *y una traslación* paralela al eje, de vector $c\,e_1$. Esta composición es conmutativa.

Observación:

Supongamos dada una afinidad $f : A \longrightarrow A$, A espacio afín euclídeo de dimensión 2, en una referencia no necesariamente ortonormal:

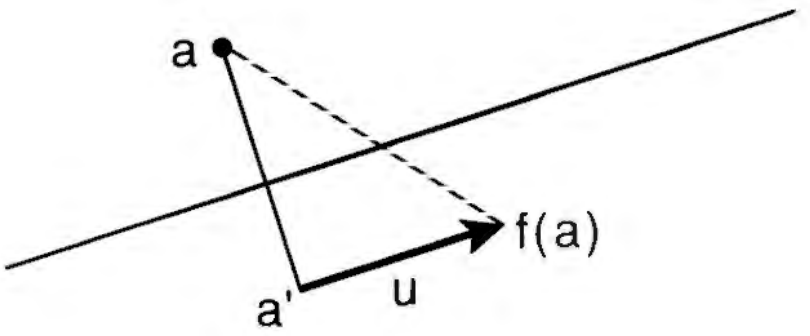
$$\bar{x} = Mx + b.$$

Entonces f es un desplazamiento si y sólo si $\tilde{f}$ es ortogonal o, equivalentemente, si $M^t G M = G$, donde G es la matriz del producto escalar en esa referencia. En caso de ser desplazamiento, f es propio o impropio según que $\det M$ sea $+1$ o -1. En el primer caso, f es la identidad ($M = I$, $b = 0$), o una traslación ($M = I$, $b \neq 0$), o un giro de centro el único punto fijo y ángulo θ tal que $\operatorname{tr} M = 2\cos\theta$.

Si f es un desplazamiento impropio, se trata de una simetría axial si existe una recta de puntos fijos. Si no hay puntos fijos, $f = T_u \circ s$, donde s es una simetría axial y T_u una traslación paralela al eje de s. Este eje es la única recta invariante por f y su dirección es el subespacio E_1 de vectores propios de valor propio $+1$:

$$\text{Eje} = \{x \in A \mid \overrightarrow{xf(x)} \in E_1\}.$$

El vector u de la traslación es precisamente $u = \overrightarrow{xf(x)}$, para cualquier punto x del eje.



Ejercicio:

Si f es un desplazamiento impropio, entonces para todo $a \in A$ el punto medio del segmento $\overline{af(a)}$ es del eje. Tenemos así otra manera de calcular el eje.

Resumen de los tipos de desplazamientos del plano euclídeo

Sea $\bar{x} = Mx + b$ un desplazamiento del plano euclídeo expresado en una referencia cualquiera. Entonces:

$$\det M = 1 \quad (\cos\theta = \tfrac{1}{2}\operatorname{tr} M) \begin{cases} \cos\theta = 1\ (M = I) & \begin{cases} b = 0 & \ldots. \text{ identidad} \\ b \neq 0 & \ldots. \text{ traslación} \end{cases} \\ \cos\theta = -1\ (M = -I) & \ldots. \text{ simetría central} \\ |\cos\theta| \neq 1 & \ldots. \text{ rotación} \end{cases}$$

$$\det M = -1 \begin{cases} \nexists \text{ puntos fijos} & \ldots. \text{ simetría axial compuesta con una traslación paralela al eje} \\ \exists \text{ puntos fijos} & \ldots. \text{ simetría axial.} \end{cases}$$

XIII.7 Desplazamientos del espacio euclídeo tridimensional

Sea $f : A \longrightarrow A$ un desplazamiento de un espacio afín euclídeo A de dimensión 3. Si f es propio, $\vec{f} \in SO(3)$ y, en una base ortonormal conveniente $\{e_1, e_2, e_3\}$, su matriz es de la forma

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si hay un punto fijo q, hay toda una recta de puntos fijos: $q + \langle e_3 \rangle$. El desplazamiento se llama un *giro* o *rotación* de *ángulo* θ y *eje* la recta de

puntos fijos. Sus ecuaciones en la referencia $\{q;\ e_1, e_2, e_3\}$ son

$$\begin{cases} \bar{x} = x\cos\theta - y\,\mathrm{sen}\,\theta \\ \bar{y} = x\,\mathrm{sen}\,\theta + y\cos\theta \\ \bar{z} = z. \end{cases}$$

Si $\cos\theta = -1$, se dice que f es una *simetría axial.*

Si f no tiene ningún punto fijo, escojamos de momento un origen p cualquiera. Sean

$$\begin{cases} \bar{x} = x\cos\theta - y\,\mathrm{sen}\,\theta + c \\ \bar{y} = x\,\mathrm{sen}\,\theta + y\cos\theta + d \\ \bar{z} = z + n \end{cases}$$

las ecuaciones de f en la referencia $\{p;\ e_1, e_2, e_3\}$. Si $\cos\theta = 1$, f es una traslación. Si $\cos\theta \neq 1$, en el sistema de puntos fijos

$$\begin{cases} x(\cos\theta - 1) - y\,\mathrm{sen}\,\theta + c = 0 \\ x\,\mathrm{sen}\,\theta + y(\cos\theta - 1) + d = 0 \\ n = 0 \end{cases}$$

las dos primeras ecuaciones tienen siempre una solución única (x_0, y_0). Como estamos suponiendo que no hay puntos fijos, este sistema ha de ser incompatible y, por tanto, $n \neq 0$. La recta $x = x_0$, $y = y_0$ es invariante y, si escogemos el origen q sobre ella, $q = (x_0, y_0, z)$, obtenemos $\overrightarrow{qf(q)} = (0, 0, n)$. Las ecuaciones de f en la referencia $\{q;\ e_1, e_2, e_3\}$ son, por tanto,

$$\begin{cases} \bar{x} = x\cos\theta - y\,\mathrm{sen}\,\theta \\ \bar{y} = x\,\mathrm{sen}\,\theta + y\cos\theta \\ \bar{z} = z + n. \end{cases}$$

Esto es uua *rotación* de ángulo θ y eje $q + \langle e_3\rangle$ *compuesta con una traslación* de vector ne_3 paralelo al eje. Esta composición es conmutativa. Este tipo de desplazamiento se llama *movimiento helicoidal.* Observemos que los restantes desplazamientos propios son un caso particular de movimiento helicoidal para $\cos\theta = 1$ y/o $n = 0$.

Si el desplazamiento f es impropio, hay una base ortonormal $\{e_1, e_2, e_3\}$ en la cual la matriz de $\vec{f}$ es

$$\begin{pmatrix} \cos\theta & -\mathrm{sen}\,\theta & 0 \\ \mathrm{sen}\,\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Consideraremos tres casos:

1. $\cos\theta = 1$ y existe un punto fijo q. Entonces el plano $q + \langle e_1, e_2\rangle$ es de puntos fijos y f se llama una *simetría especular* respecto a ese plano (*plano de simetría*). En la referencia $\{q;\ e_1, e_2, e_3\}$ las ecuaciones de f son

$$\begin{cases} \bar{x} = x \\ \bar{y} = y \\ \bar{z} = -z. \end{cases}$$

2. $\cos\theta = 1$ y no hay ningún punto fijo. Las ecuaciones de f en una referencia $\{p;\ e_1, e_2, e_3\}$ son

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = y + d \\ \bar{z} = -z + n, \end{cases}$$

con $c^2 + d^2 \neq 0$ para que no haya puntos fijos. El plano $z = n/2$ es invariante y, tomando el origen sobre él, $q = (x_0, y_0, n/2)$, tenemos $\overrightarrow{qf(q)} = (c, d, 0)$. De ahí resulta que las ecuaciones de f en la referencia $\{q;\ e_1, e_2, e_3\}$ son

$$\begin{cases} \bar{x} = x + c \\ \bar{y} = y + d \\ \bar{z} = -z. \end{cases}$$

f es, por tanto, la *composición de una simetría especular* respecto a $q + \langle e_1, e_2\rangle$ *y una traslación* de vector paralelo al plano de simetría. Esta composición es conmutativa.

3. $\cos\theta \neq 1$. Entonces $\tilde{f}$ no tiene el valor propio 1 y, por tanto, f tiene un único punto fijo q (X.6.1). En la referencia $\{q;\ e_1, e_2, e_3\}$ las ecuaciones de f son

$$\begin{cases} \bar{x} = x\cos\theta - y\,\operatorname{sen}\theta \\ \bar{y} = x\,\operatorname{sen}\theta + y\cos\theta \\ \bar{z} = -z. \end{cases}$$

f es la *composición de una simetría especular* respecto al plano $q + \langle e_1, e_2\rangle$ *y una rotación* de eje $q + \langle e_3\rangle$ perpendicular al plano de simetría. Esta composición es conmutativa. En el caso particular $\cos\theta = -1$, f se llama una *simetría central* de *centro* q.

Observación:

Supongamos que A es un espacio afín euclídeo de dimensión 3 y $f : A \longrightarrow A$ una afinidad que en una referencia no necesariamente ortonormal tiene por ecuaciones

$$\bar{x} = Mx + b.$$

f es un desplazamiento si y sólo si $M^t G M = G$, donde G es la matriz del producto escalar en esa referencia. Si f es un desplazamiento, es propio o impropio según que $\det M$ sea $+1$ o -1. En el primer caso f puede ser una traslación ($M = I$) o un giro, compuesto o no con una traslación. El ángulo de giro θ cumple

$$2\cos\theta + 1 = \operatorname{tr} M.$$

Si f es un giro y tiene puntos fijos, el eje es la recta de puntos fijos. Si no hay puntos fijos, el eje es una recta invariante, de dirección el subespacio E_1 de vectores propios de valor propio $+1$:

$$\text{Eje} = \{x \in A \mid \overrightarrow{xf(x)} \in E_1\}.$$

El vector de la traslación es $u = \overrightarrow{xf(x)}$, donde x es un punto cualquiera del eje.

En el segundo caso, f es una simetría especular, compuesta o no con una traslación de vector paralelo al plano de simetría, o con una rotación de eje perpendicular a ese plano.

El plano de simetría es un plano invariante. Para todo $a \in A$, el punto medio del segmento $\overline{af(a)}$ es de ese plano.

El ángulo de la rotación viene dado por $2\cos\theta - 1 = \operatorname{tr} M$.

El eje es una recta invariante con un punto fijo, la dirección de la cual está contenida en el subespacio E_{-1} de vectores propios de valor propio -1. Si este subespacio no tiene dimensión 1, tiene dimensión 3 y f es una simetría central, de centro el punto fijo.

Si $\cos\theta = 1$, se trata de una simetría especular seguida de una traslación y el vector de la traslación es $u = \overrightarrow{xf(x)}$, donde x es un punto cualquiera del plano de simetría.

Ejercicio:

Si f es una simetría axial, entonces para todo $a \in A$ el punto medio del segmento $\overline{af(a)}$ es del eje.

Resumen de los tipos de desplazamientos del espacio euclídeo

Sea $\bar{x} = Mx + b$ un desplazamiento del espacio euclídeo tridimensional expresado en una referencia cualquiera. Entonces:

$\det M = 1$ $(\cos\theta = \frac{1}{2}(\operatorname{tr} M - 1))$

- $\exists$ puntos fijos
  - $\cos\theta = 1$ $(M = I)$ ... identidad
  - $\cos\theta = -1$ ... simetría axial
  - $|\cos\theta| \neq 1$ ... rotación
- $\nexists$ puntos fijos
  - $\cos\theta = 1$ ... traslación
  - $\cos\theta \neq 1$ ... movimiento helicoidal

$\det M = -1$ $(\cos\theta = \frac{1}{2}(\operatorname{tr} M + 1))$

- $\exists$ puntos fijos
  - $\cos\theta = 1$ ... simetría especular
  - $\cos\theta = -1$ $(M = -I)$ ... simetría central
  - $|\cos\theta| \neq 1$ ... simetría especular compuesta con una rotación perpendicular al plano de simetría.
- $\nexists$ puntos fijos ... simetría especular compuesta con una traslación paralela al plano de simetría.

XIII.8 Semejanzas

Una *semejanza* es una aplicación $f : A \longrightarrow A$ de un espacio afín euclídeo en sí mismo que cumple

$$\mathrm{d}(f(a), f(b)) = r\,\mathrm{d}(a, b) \qquad \forall a, b \in A,$$

donde r es un número real fijo $(r > 0)$ que se llama la *razón* de la semejanza. En particular, los desplazamientos son semejanzas de razón 1. Igual que en el caso de los desplazamientos, las semejanzas son siempre inyectivas. También es fácil ver que la composición de dos semejanzas de razones r, r' es una semejanza de razón rr'.

Proposición 8.1 *Una aplicación $f : A \longrightarrow A$, donde A es un espacio afín euclídeo, es una semejanza si y sólo si f es una afinidad y $\tilde{f}$ se descompone en producto de una aplicación ortogonal y una homotecia vectorial.*

DEMOSTRACIÓN: Supongamos que f es una semejanza de razón r y fijemos un punto $p \in A$ cualquiera. Sea h la homotecia de centro p y razón r^{-1}.

$$\mathrm{d}(h(a),h(b)) = \|\overrightarrow{h(a)h(b)}\| = \|\tilde{h}(\overrightarrow{ab})\| = \|r^{-1}\overrightarrow{ab}\| = r^{-1}\mathrm{d}(a,b) \qquad \forall a,b$$

nos dice que h es una semejanza de razón r^{-1} y, por tanto, $g = h \circ f$ es una semejanza de razón $r^{-1}r = 1$, es decir, un desplazamiento. Así pues, $f = h^{-1} \circ g$ es una afinidad (producto de dos afinidades) y $\tilde{f} = \tilde{h}^{-1} \circ \tilde{g}$ con $\tilde{g}$ ortogonal y $\tilde{h}^{-1} = rI$ una homotecia vectorial. Esto demuestra la primera parte.

Observemos que $g' = f \circ h$ también es un desplazamiento y $f = g' \circ h^{-1}$. En general $g' \neq g$, pero $\tilde{g}' = r^{-1}\tilde{f} = \tilde{g}$.

Supongamos ahora que f es una afinidad y $f = hog$, con h una homotecia de razón r y g un desplazamiento. Para todo par a, b,

$$\mathrm{d}(f(a),f(b)) = \mathrm{d}(hg(a),hg(b)) = r\,\mathrm{d}(g(a),g(b)) = r\,\mathrm{d}(a,b).$$

Por tanto, f es una semejanza. Un razonamiento análogo puede hacerse si $f = g \circ h$ con g un desplazamiento y h una homotecia. □

Proposición 8.2 *Toda semejanza de razón $r \neq 1$ tiene un punto fijo y sólo uno.*

DEMOSTRACIÓN: Sea $f = h \circ g$ una descomposición de la semejanza como producto de una homotecia h de razón r y un desplazamiento g. Para todo vector u,

$$\tilde{f}(u) = r\tilde{g}(u),$$

de donde resulta que los únicos valores propios posibles de $\tilde{f}$ son $\pm r \neq 1$. Entonces (X.6.1) asegura que existe un único punto fijo. □

El punto fijo q de una semejanza f de razón $r \neq 1$ se llama el *centro* de la semejanza. Si en la primera parte de la demostración de (8.1) tomamos como centro de la homotecia h el punto q, los desplazamientos $g = h \circ f$ y $g' = f \circ h$ dejarán ambos fijo el punto q. Puesto que $\tilde{g} = \tilde{g}'$, resulta que $g = g'$ y $f = h^{-1} \circ g = g \circ h^{-1}$.

Observación:

Consideremos una afinidad $f : A \longrightarrow A$ de ecuaciones

$$\bar{x} = Mx + b$$

en una base no necesariamente ortonormal. f es una semejanza de razón r si y sólo si $M = (rI)N$, N ortogonal, o, equivalentemente, si $N^tGN = G$, $M = (rI)N$, donde G es la matriz del producto escalar. Es decir, si y sólo si

$$M^tGM = r^2G.$$

Esta igualdad nos permite estudiar si f es una semejanza y calcular su razón en caso afirmativo. (Recordemos que siempre $r > 0$.)

Observemos también que $\det M = r^n \det N = \pm r^n$, donde n es la dimensión del espacio A.

Una semejanza f se llama *directa* si $\det \tilde{f} > 0$, y se llama *inversa* si $\det \tilde{f} < 0$.

Ejercicio:

Probar que las únicas semejanzas de razón $\neq 1$ de la recta afín euclídea son las homotecias.

XIII.9 Semejanzas del espacio afín euclídeo tridimensional

Sea $f : A \longrightarrow A$ una semejanza de razón $r \neq 1$, $\dim A = 3$. Denotemos por q el centro de f. Entonces

$$f = h \circ g = g \circ h,$$

donde g es un desplazamiento que deja fijo q y h es la homotecia de razón r y centro q.

Si f es una semejanza directa, $\det \tilde{f} = r^3 \det \tilde{g} > 0$, de donde $\det \tilde{g} > 0$ y g es un desplazamiento propio con un punto fijo q. Así pues, g es una rotación cuyo eje pasa por q.

Si f es una semejanza inversa, $\det \tilde{f} = r^3 \det \tilde{g} < 0$, de donde $\det \tilde{g} < 0$ y g es un desplazamiento impropio con un punto fijo. Así pues, g es una simetría especular seguida (o no) de una rotación de eje perpendicular al plano de simetría e intersección el punto q. Podemos, sin embargo, obtener

una interpretación geométrica de f más clara de la manera siguiente. Consideremos la homotecia h' de centro g y razón $-r$. Tenemos

$$h' = h \circ s,$$

donde s es la simetría central de centro q. Luego,

$$f = h \circ g = h \circ s \circ s \circ g = h' \circ g',$$

donde $g' = s \circ g$ es un desplazamiento propio con un punto q fijo. Así pues, g' es una rotación cuyo eje pasa por q y f se obtiene componiendo g' con una homotecia de razón negativa $-r$. Hemos demostrado así la siguiente proposición:

Proposición 9.1 *Las semejanzas de razón $r \neq 1$ del espacio afín euclídeo tridimensional son las rotaciones seguidas de una homotecia de centro sobre el eje de la rotación y razón $\pm r$.* □

XIII.10 Semejanzas del plano afín euclídeo

Sea $f : A \longrightarrow A$ una semejanza de razón $r \neq 1$, $\dim A = 2$. Designemos por q el centro de f. Entonces

$$f = h \circ g = g \circ h,$$

donde g es un desplazamiento que deja fijo q y h es la homotecia de centro q y razón r.

Si f es directa, $\det \bar{f} = r^2 \det \tilde{g} > 0$, de donde $\det \tilde{g} > 0$ y g es propio con $g(q) = q$. Así pues, g es un giro de centro q.

Si f es inversa, $\det \bar{f} = r^2 \det \tilde{g} < 0$, de donde $\det \tilde{g} < 0$ y g es impropio con $g(q) = q$. En este caso, g es una simetría axial respecto a un eje que pasa por q.

Estudiaremos ahora las semejanzas en un modelo de espacio afín euclídeo de dimensión 2 concreto: los complejos. Este espacio afín está formado por el conjunto $A = \mathbf{C}$, el espacio vectorial $E = \mathbf{C}$ de dimensión 2 sobre $\mathbf{R}$ y la aplicación

$$\begin{array}{rccc} \phi: & \mathbf{C}\times\mathbf{C} & \longrightarrow & \mathbf{C} \\ & (z_1, z_2) & \longmapsto & z_2 - z_1, \end{array}$$

donde $z_j = a_j + ib_j \in \mathbf{C}$, $j = 1, 2$. Como producto escalar en $\mathbf{C}$ consideraremos el que hace que la base $\{1, i\}$ sea ortonormal. Es decir,

$$\langle z_1, z_2 \rangle = a_1 a_2 + b_1 b_2.$$

En particular, $\|z_1\|^2 = {a_1}^2 + {b_1}^2$.

En el conjunto $\mathbf{C}$ tenemos, además de estas estructuras, un producto y una noción de conjugado de un $z = a + ib \in \mathbf{C}$, que escribiremos $\bar{z} = a - ib$. En particular, $\|z\|^2 = z \cdot \bar{z}$.

Una aplicación $f : \mathbf{C} \longrightarrow \mathbf{C}$ es una semejanza de razón $r > 0$ si y sólo si es una afinidad y $\vec{f}$ tiene en la base $\{1, i\}$ una matriz de la forma

$$\begin{pmatrix} ra & -rb \\ rb & ra \end{pmatrix}, \qquad a^2 + b^2 = 1, \qquad \text{si } f \text{ es directa;}$$

$$\begin{pmatrix} ra & rb \\ rb & -ra \end{pmatrix}, \qquad a^2 + b^2 = 1, \qquad \text{si } f \text{ es inversa.}$$

Dado $z = x + iy \in \mathbf{C}$, sus coordenadas en el sistema de referencia ortonormal $\{0;\ 1, i\}$ son (x, y). Las ecuaciones de f en el caso directo son, por tanto, del tipo

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} ra & -rb \\ rb & ra \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} r(ax - by) + c \\ r(bx + ay) + d \end{pmatrix}.$$

Así pues,

$$\begin{aligned} z^* = x^* + iy^* &= (r(ax - by) + c) + i(r(bx + ay) + d) = \\ &= r(a + ib)(x + iy) + c + id = \\ &= \alpha z + \beta, \end{aligned}$$

donde $\alpha = r(a + ib)$, $\beta = (c + id)$. Obsérvese que $|\alpha| = r$.

Recíprocamente, toda aplicación $f : \mathbf{C} \longrightarrow \mathbf{C}$ dada por $f(z) = z^* = \alpha z + \beta$, $\alpha, \beta \in \mathbf{C}$, es una semejanza de razón el módulo de α, $|\alpha|$. En efecto, si $\alpha = |\alpha|(a + ib)$, $\beta = c + id$, $z = x + iy$, $z^* = x^* + iy^*$, la ecuación $z^* = \alpha z + \beta$ equivale a

$$\begin{cases} x^* = |\alpha| ax - |\alpha| by + c \\ y^* = |\alpha| bx + |\alpha| ay + d \end{cases}$$

y esto es, claramente, una semejanza directa de razón $|\alpha|$.

Si f es inversa, sus ecuaciones en la referencia $\{0;\ 1, i\}$ son del tipo

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} ra & rb \\ rb & -ra \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} r(ax + by) + c \\ r(bx - ay) + d \end{pmatrix}.$$

Así pues,

$$\begin{aligned} z^* = x^* + iy &= (r(ax + by) + c) + i(r(bx - ay) + d) = \\ &= r(a + ib)(x - iy) + (c + id) = \\ &= \alpha \bar{z} + \beta, \end{aligned}$$

donde $\alpha = r(a+ib)$, $\beta = c+id$, $|\alpha| = r$.

Recíprocamente, dada $f : \mathbf{C} \longrightarrow \mathbf{C}$ por $f(z) = z^* = \alpha\bar{z}+\beta$, $\alpha, \beta \in \mathbf{C}$, si ponemos $\alpha = |\alpha|(a+ib)$, $\beta = c+id$, $z = x+iy$, $z^* = x^*+iy^*$, obtenemos

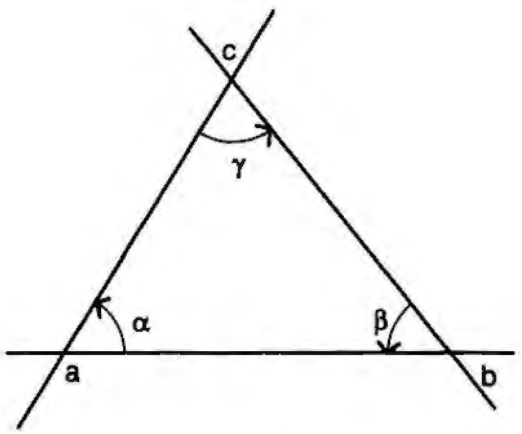
$$\begin{cases} x^* = |\alpha| ax + |\alpha| by + c \\ y^* = |\alpha| bx - |\alpha| ay + d \,, \end{cases}$$

de donde resulta fácilmente que f es una semejanza inversa de razón $|\alpha|$.

Observemos que en estas expresiones de las semejanzas del plano complejo están incluidos también los desplazamientos.

XIII.11 Algunos ejemplos y aplicaciones

En este apartado vamos a deducir unos cuantos resultados bien conocidos de la geometría elemental. Nuestro objetivo es poner de manifiesto que la "geometría lineal" que hemos estado estudiando es la "geometría ordinaria" que ya conocíamos en parte, aunque quizás con otro lenguaje. Estos ejemplos pueden servir también de modelo para "traducir" otros resultados de un lenguaje al otro.



Triángulos

Situémonos en el plano afín euclídeo real.

Un *triángulo* es un conjunto de tres puntos linealmente independientes $\{a, b, c\}$ que llamaremos *vértices*. Denominaremos *ángulos* de un triángulo $\{a, b, c\}$ a los ángulos

$$\alpha = \widehat{\overrightarrow{ac}\,\overrightarrow{ab}}, \qquad \beta = \widehat{\overrightarrow{ba}\,\overrightarrow{bc}}, \qquad \gamma = \widehat{\overrightarrow{cb}\,\overrightarrow{ca}}.$$

La composición de las aplicaciones ortogonales correspondientes a estos án-

gulos (XII.5) actúa así:

$$\frac{\overrightarrow{ac}}{\|\overrightarrow{ac}\|} \overset{\alpha}{\longmapsto} \frac{\overrightarrow{ab}}{\|\overrightarrow{ab}\|} \overset{\beta}{\longmapsto} \frac{\overrightarrow{cb}}{\|\overrightarrow{cb}\|} \overset{\gamma}{\longmapsto} \frac{\overrightarrow{ca}}{\|\overrightarrow{ca}\|} = -\frac{\overrightarrow{ac}}{\|\overrightarrow{ac}\|}.$$

Esta composición es, por tanto, $-I$ y

$$\alpha + \beta + \gamma = \pi.$$

Así pues,

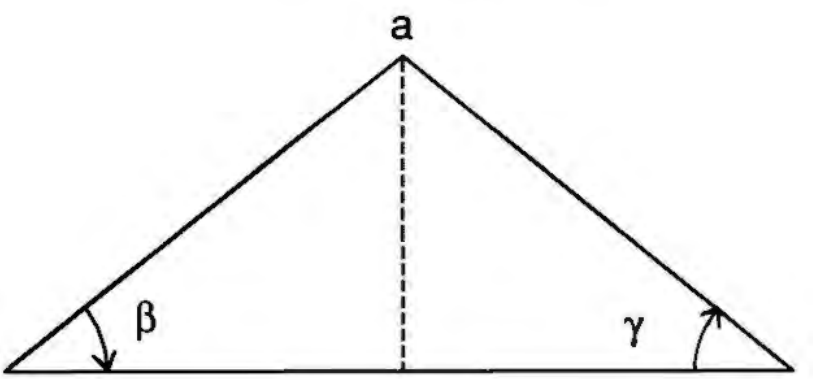
Proposición 11.1 *La suma de los ángulos de un triángulo es* π. □

Llamaremos *lados* de un triángulo a, b, c a los segmentos determinados por sus vértices. Llamaremos *longitud* de un lado a la distancia entre los vértices correspondientes.

Proposición 11.2 *Un triángulo tiene dos ángulos iguales si y sólo si los dos lados opuestos a esos ángulos tienen la misma longitud. Diremos entonces que el triángulo es* isósceles.

DEMOSTRACIÓN: Supongamos que $\{a, b, c\}$ es un triángulo con $\mathrm{d}(a, b) = \mathrm{d}(a, c)$. Designemos por f la simetría axial de eje $a + \langle \overrightarrow{ab} + \overrightarrow{ac} \rangle$. Entonces

$$\vec{f}(\overrightarrow{ab} + \overrightarrow{ac}) = \overrightarrow{ab} + \overrightarrow{ac}.$$



Además,

$$(\overrightarrow{ab} - \overrightarrow{ac}) \cdot (\overrightarrow{ab} + \overrightarrow{ac}) = \|\overrightarrow{ab}\|^2 - \|\overrightarrow{ac}\|^2 = 0,$$

de donde $\vec{f}(\overrightarrow{ab} - \overrightarrow{ac}) = -\overrightarrow{ab} + \overrightarrow{ac}$. Esta igualdad, junto con la de más arriba, nos da

$$\vec{f}(\overrightarrow{ab}) = \overrightarrow{ac}\,, \quad \vec{f}(\overrightarrow{ac}) = \overrightarrow{ab}\,, \quad \vec{f}(\overrightarrow{cb}) = \overrightarrow{bc}.$$

Por tanto, $\tilde{f}$ transforma el ángulo $\beta = \widehat{\overrightarrow{ba}\,\overrightarrow{bc}}$ en el ángulo $\widehat{\overrightarrow{cd}\,\overrightarrow{cb}} = -\gamma$. Entonces, por (XII.5.4), $\beta = \gamma$.

Supongamos ahora que $\beta = \gamma$. En una base ortonormal positiva (XII.4.2) tenemos

$$\det\left(\frac{\overrightarrow{ba}}{\|\overrightarrow{ba}\|}, \frac{\overrightarrow{bc}}{\|\overrightarrow{bc}\|}\right) = \operatorname{sen}\beta = \operatorname{sen}\gamma = \det\left(\frac{\overrightarrow{cb}}{\|\overrightarrow{cb}\|}, \frac{\overrightarrow{cd}}{\|\overrightarrow{cd}\|}\right),$$

de donde

$$\det\left(\frac{\overrightarrow{ba}}{\|\overrightarrow{ba}\|} - \frac{\overrightarrow{cd}}{\|\overrightarrow{cd}\|}, \frac{\overrightarrow{bc}}{\|\overrightarrow{bc}\|}\right) = 0$$

y, por tanto, existe un k tal que

$$\frac{\overrightarrow{ba}}{\|\overrightarrow{ba}\|} - \frac{\overrightarrow{cd}}{\|\overrightarrow{cd}\|} = k\frac{\overrightarrow{bc}}{\|\overrightarrow{bc}\|} = \frac{k}{\|\overrightarrow{bc}\|}(\overrightarrow{ba} - \overrightarrow{cd}).$$

Es decir, $\|\overrightarrow{ba}\| = \|\overrightarrow{cd}\|$ y $\mathrm{d}(a,b) = \mathrm{d}(a,c)$. □

Puntos notables del triángulo

Se llaman *medianas* de un triángulo las rectas que pasan por un vértice y el punto medio del lado opuesto. Las tres medianas de un triángulo pasan por el *baricentro* de los vértices (ejercicio IX.4).

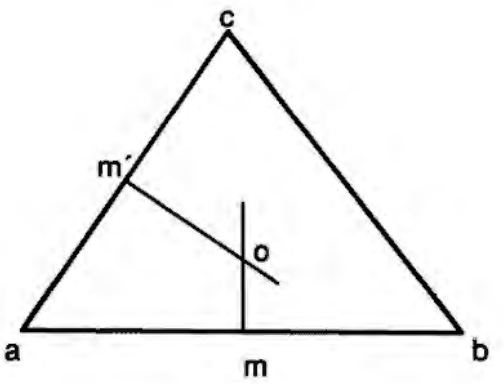
Se llaman *alturas* de un triángulo las rectas que pasan por un vértice y son perpendiculares a la recta determinada por los otros dos vértices. En la situación de la proposición (11.2), la mediana $a + \langle\overrightarrow{am}\rangle$, donde $m = \frac{1}{2}b + \frac{1}{2}c$, es precisamente la altura por a, ya que

$$\overrightarrow{am}\cdot\overrightarrow{bc} = \left(\frac{1}{2}\overrightarrow{ab} + \frac{1}{2}\overrightarrow{ac}\right)\cdot(\overrightarrow{ac} - \overrightarrow{ab}) = \frac{1}{2}(\|\overrightarrow{ac}\|^2 - \|\overrightarrow{ab}\|^2) = 0.$$

Las *mediatrices* de un triángulo $\{a,b,c\}$ son las rectas perpendiculares a los lados que pasan por su punto medio. Sean $m = \frac{1}{2}a + \frac{1}{2}b$, $m' = \frac{1}{2}a + \frac{1}{2}c$ y sea O el punto de intersección de las mediatrices por m y m'. Observemos que

$$\mathrm{d}(O,a)^2 = \|\overrightarrow{Oa}\|^2 = \|\overrightarrow{Om}\|^2 + \|\overrightarrow{ma}\|^2 = \|\overrightarrow{Om}\|^2 + \|\overrightarrow{mb}\|^2 = \|\overrightarrow{Ob}\|^2 = \mathrm{d}(O,b)^2.$$

Análogamente, $\mathrm{d}(O,a) = \mathrm{d}(O,c)$. El triángulo $\{O,b,c\}$ es, pues, isósceles y su altura corta a $\overline{cb}$ en el punto medio. El punto O es, por tanto, la intersección de las tres mediatrices y se llama *circuncentro*. O es centro de una circunferencia que pasa por a, b y c: la circunferencia circunscrita al triángulo.

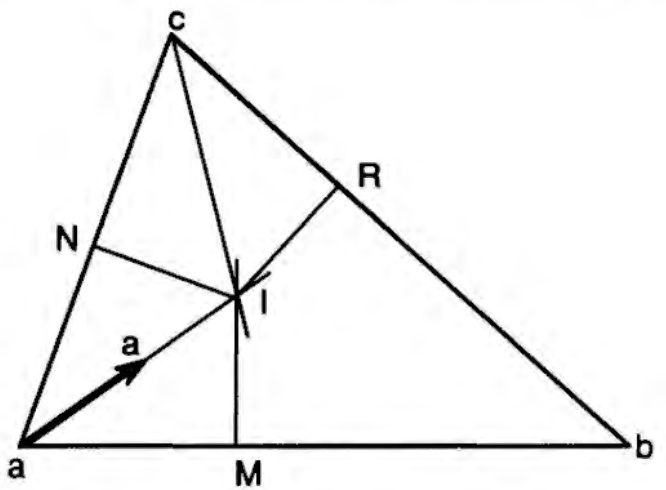


La *bisectriz* del ángulo α de un triángulo $\{a, b, c\}$ es la recta $a + \langle u \rangle$ tal que $\widehat{\overrightarrow{ac}u} = \widehat{u\overrightarrow{ab}}$.

Ejercicio:

Demostrar que en el triángulo isósceles de (11.2) la mediana am es también la bisectriz de α.

Sea I el punto de intersección de las bisectrices de α y de γ. Sean M, N, R los pies de las perpendiculares por I a las rectas ab, ac, cb. I está



en la bisectriz de β si y sólo si $\mathrm{d}(I, R) = \mathrm{d}(I, M)$. Por el mismo motivo, puesto que I está en las bisectrices de α y γ, $\mathrm{d}(I, M) = \mathrm{d}(I, N) = \mathrm{d}(I, R)$. Por tanto, I es la intersección de las bisectrices y se llama *incentro* del triángulo. El incentro es el centro de una circunferencia tangente a los lados del triángulo en M, N y R: la circunferencia inscrita.

Triángulos semejantes

Diremos que dos triángulos son *congruentes* o *iguales* si existe un desplazamiento que transforma uno de ellos en el otro. Diremos que son *semejantes* si existe una semejanza que transforma uno de ellos en el otro.

Proposición 11.3 *Dos triángulos son semejantes si y sólo si se cumple una de las dos condiciones equivalentes:*

1. *Las longitudes de los lados son proporcionales.*
2. *Los tres ángulos de un triángulo son iguales a los ángulos del otro o son iguales a sus opuestos.*

DEMOSTRACIÓN: Las homotecias de razón positiva conservan los ángulos. Entonces (8.1) y (XII.5.4) nos dicen que las semejanzas directas conservan los ángulos y las inversas los invierten. Por tanto, si dos triángulos son semejantes, se cumple 2. Por otra parte, al aplicar una semejanza, las longitudes de los lados quedan multiplicadas por la razón, y también se cumple 1.

Supongamos ahora que se cumple 1. Si $\{a,b,c\}$ y $\{a',b',c'\}$ son los triángulos, pongamos

$$\begin{aligned} u &= \overrightarrow{ab}, \quad v = \overrightarrow{ac}, \quad v-u = \overrightarrow{bc}, \\ u' &= \overrightarrow{a'b'}, \quad v' = \overrightarrow{a'c'}, \quad v'-u' = \overrightarrow{b'c'}. \end{aligned}$$

Por hipótesis, $\|u\| = k\,\|u'\|$, $\|v\| = k\,\|v'\|$, $\|v-u\| = k\,\|v'-u'\|$. Designemos por f la afinidad que transforma $\{a,b,c\}$ en $\{a',b',c'\}$ y por h la homotecia de razón k y centro un punto p cualquiera. Sea $g = h \circ f$. Entonces

$$\tilde{g}(u) = ku', \quad \tilde{g}(v) = kv' \qquad \Rightarrow \qquad \|\tilde{g}(u)\| = \|u\|, \quad \|\tilde{g}(v)\| = \|v\|.$$

$$\begin{aligned} \tilde{g}(u)\cdot\tilde{g}(v) = k^2 u'\cdot v' &= k^2\frac{1}{2}(\|u'\|^2 + \|v'\|^2 - \|u'-v'\|^2) = \\ &= \frac{1}{2}(\|u\|^2 + \|v\|^2 - \|u-v\|^2) = u\cdot v. \end{aligned}$$

$\tilde{g}$ es, pues, ortogonal y g un desplazamiento. Por tanto, f es una semejanza y $\{a,b,c\}$, $\{a',b',c'\}$ son semejantes.

Demostremos, por último, que 2 $\Rightarrow$ 1. Con las notaciones anteriores,

$$\begin{aligned} \widehat{uv} = \pm\widehat{u'v'} \quad &\Rightarrow \quad \det\left(\frac{u}{\|u\|},\frac{v}{\|v\|}\right) = \pm\det\left(\frac{u'}{\|u'\|},\frac{v'}{\|v'\|}\right) \quad \Rightarrow \\ &\Rightarrow \quad \det(u,v) = \pm kk'\det(u',v'), \end{aligned}$$

donde $k = \dfrac{\|u\|}{\|u'\|}$, $k' = \dfrac{\|v\|}{\|v'\|}$. Análogamente,

$$\widehat{u(u-v)} = \pm\widehat{u'(u'-v')} \Rightarrow \det(u, u-v) = \pm k'k'' \det(u', u'-v') \Leftrightarrow$$
$$\Leftrightarrow \det(u, v) = \pm k'k'' \det(u', v'),$$

donde $k'' = \dfrac{\|u-v\|}{\|u'-v'\|}$. Las dos igualdades obtenidas implican que $k' = k''$. Por el mismo motivo, resulta que $k = k'$ y, por tanto,

$$\frac{\|u\|}{\|u'\|} = \frac{\|v\|}{\|v'\|} = \frac{\|u-v\|}{\|u'-v'\|}. \square$$

Proposición 11.4 *Si dos triángulos $\{a, b, c\}$, $\{a', b', c'\}$ tienen un ángulo igual salvo el signo, $\alpha = \pm\alpha'$, y los dos lados que lo forman proporcionales, $\dfrac{d(a,b)}{d(a',b')} = \dfrac{d(a,c)}{d(a',c')}$, entonces los dos triángulos son semejantes.*

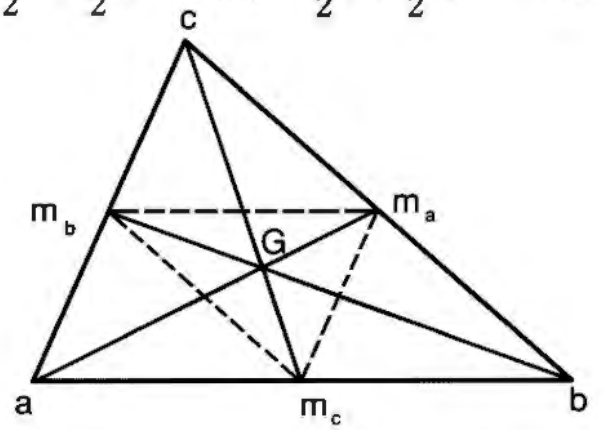
DEMOSTRACIÓN: Sea f la afinidad que aplica $\{a, b, c\}$ en $\{a', b', c'\}$. Con la notación de (11.3) tenemos $\|u\| = k\|u'\|$, $\|v\| = k\|v'\|$. La aplicación lineal $\phi = k\tilde{f}$ es ortogonal, ya que

$$\|\phi(u)\| = \|ku'\| = \|u\|, \qquad \|\phi(v)\| = \|kv'\| = \|v\|,$$
$$\phi(u) \cdot \phi(v) = k^2 u' \cdot v' = k^2 \|u'\| \, \|v'\| \cos\widehat{u'v'} = \|u\| \, \|v\| \cos\widehat{uv} = u \cdot v.$$

Por tanto, $\tilde{f} = k^{-1}\phi$ y f es una semejanza. $\square$

Consideremos un triángulo $\{a, b, c\}$ y los puntos medios de los lados:

$$m_a = \frac{1}{2}b + \frac{1}{2}c, \qquad m_b = \frac{1}{2}a + \frac{1}{2}c, \qquad m_c = \frac{1}{2}a + \frac{1}{2}b.$$



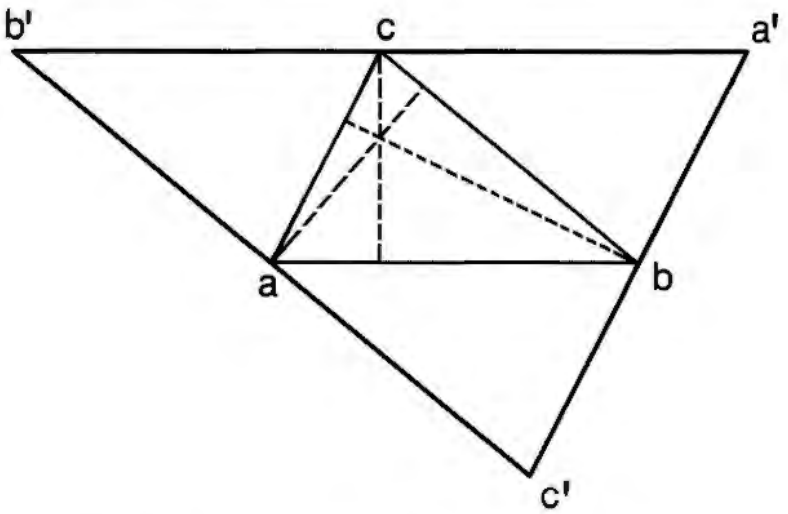
La homotecia de centro a y razón $1/2$ aplica $\{a,b,c\}$ en $\{a,m_c,m_b\}$. De $\overrightarrow{m_c m_b} = \overrightarrow{h(b)h(c)} = \tilde{h}(\overrightarrow{bc}) = \frac{1}{2}\overrightarrow{bc}$ se deduce que las rectas $m_c m_b$ y bc son paralelas. Análogamente, las rectas $m_b m_a$ y $m_a m_c$ son paralelas a ab y ca respectivamente. De ahí resulta que los ángulos de los triángulos $\{a,b,c\}$ y $\{m_a,m_b,m_c\}$ son iguales y, por (11.3), los triángulos son semejantes. Sea f la semejanza que pasa del uno al otro. Estos dos triángulos tienen las mismas medianas y baricentro G, que será un punto fijo de f. Además, $\tilde{f}$ tiene vectores propios: $\overrightarrow{Ga}$, $\overrightarrow{Gb}$, $\overrightarrow{Gc}$; por tanto, es una homotecia de razón $-1/2$, ya que $\|\overrightarrow{m_a m_b}\| = \frac{1}{2}\|\overrightarrow{ab}\|$. En particular, obtenemos

$$\mathrm{d}(G,m_a) = 2\,\mathrm{d}(G,a)$$

y lo mismo para los otros vértices.

Vamos a usar la construcción anterior para demostrar que las alturas de un triángulo se cortan en un punto: el *ortocentro*.

Dado un triángulo $\{a,b,c\}$, consideremos las rectas que pasan por un vértice y son paralelas al lado opuesto. Los puntos de intersección de esas rectas forman un triángulo $\{a',b',c'\}$ y los puntos medios de sus lados son precisamente a,b,c. (En efecto, sean $m_{a'},m_{b'},m_{c'}$ los puntos medios. Entonces las rectas ab y $m_{a'}m_{b'}$ son paralelas porque ambas lo son a $a'b'$. Lo mismo pasa con los otros dos pares. Entonces $m_{a'} \in \overline{ac'} \Leftrightarrow m_{b'} \in \overline{bc'} \Leftrightarrow m_{c'} \in \overline{cb'} \Leftrightarrow m_{a'} \in \overline{ab'}$. Por tanto, $m_{a'} = a$, y análogamente $m_{b'} = b$, $m_{c'} = c$.)



Las alturas de $\{a,b,c\}$ son, pues, las mediatrices de $\{a',b',c'\}$ y, en particular, se cortan en un punto, tal como queríamos probar.

Un triángulo con los tres lados iguales se llama *equilátero*. Por (11.2) sus ángulos también son iguales: $\alpha = \beta = \gamma$. De $\alpha+\beta+\gamma = 3\alpha = \pi$ resulta fácilmente que $\cos\alpha = \dfrac{1}{2}$ y $\operatorname{sen}\alpha = \pm\dfrac{\sqrt{3}}{2}$ (el signo depende de la orientación

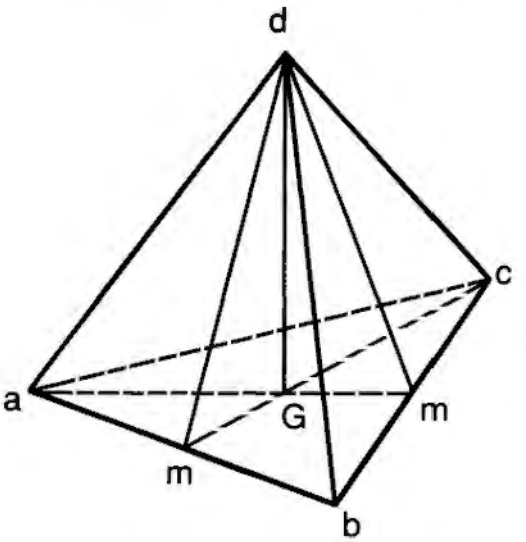
del plano).

En un triángulo equilátero las alturas, mediatrices, medianas y bisectrices coinciden y sus puntos de intersección también.

El tetraedro

Situémonos ahora en el espacio afín euclídeo tridimensional.

Un *tetraedro* es un conjunto de cuatro puntos linealmente independientes. Si la distancia entre sus vértices es constante, se dice que el tetraedro es *regular*. Se denominan *caras* de un tetraedro a sus subconjuntos de tres elementos. Si un tetraedro es regular, sus caras son triángulos equiláteros.



Sea $\{a, b, c, d\}$ un tetraedro regular. Si m es el punto medio de $\overline{ab}$, la recta dm es la altura de la cara $\{a, b, d\}$ y, por tanto, perpendicular a $\overline{ab}$. Análogamente, $\overline{cm}$ es perpendicular a $\overline{ab}$, y por ello el plano dmc es perpendicular a $\overline{ab}$. De la misma manera, si m' es el punto medio de $\overline{bc}$, el plano $dm'a$ es perpendicular a $\overline{cb}$. La intersección de estos dos planos es la recta que pasa por d y el baricentro G del triángulo $\{a, b, c\}$. Por estar en los dos planos, $\overline{dG}$ es perpendicular al plano abc. La recta dG pasa, claramente, por el baricentro de $\{a, b, c, d\}$ y, por tanto, este es el punto de intersección de las alturas del tetraedro.

Observemos por último que el triángulo $\{d, m, c\}$ es isósceles. Su altura es la perpendicular común a los lados $\overline{ab}$ y $\overline{dc}$ y los corta en sus puntos medios.

XIII.12 Nota histórica

Según las ideas del "Programa de Erlangen", la geometría métrica es el estudio de los invariantes por el grupo de las isometrías; es decir, es una

subgeometría de la geometría afín. Este punto de vista, que Felix Klein (1849–1925) introdujo en 1872, ha perdurado y es el que hemos adoptado.

En los capítulos de Geometría de su *Introductio*, Leonhard Euler (1707–1783) se ocupó de buscar las curvas invariantes por una isometría del plano, y los razonamientos que utilizó conducen a la conclusión de que una tal isometría es una traslación, o una rotación, o una traslación seguida de una simetría axial de eje la dirección de la traslación. En 1776 demostró que toda semejanza directa del plano tiene un punto fijo. El año anterior, en conexión con sus trabajos de Mecánica, había intentado demostrar que un desplazamiento propio del espacio tenía una recta fija, buscando una demostración de que el endomorfismo ortogonal asociado tenía el valor propio 1, pero no lo había conseguido. Lexell lo demostró un año más tarde, pero no es hasta los trabajos de Michel Charles (1793–1880) en 1830 donde se halla finalmente una exposición completa y coherente de la clasificación de los desplazamientos del espacio y de sus composiciones.

XIII.13 Ejercicios

1. Dados tres puntos linealmente independientes de $\mathbf{R}^3$, demostrar que la intersección de los planos que pasan por un punto y son perpendiculares a la recta determinada por los otros dos es una recta perpendicular al plano que contiene a los tres puntos.

2. En un cubo de arista 1 calcular el ángulo que forman las diagonales de dos caras contiguas que concurren en un mismo vértice.

3. Dados tres números complejos de módulo 1 que sumen 0, demostrar que forman un triángulo equilátero.

4. Un número complejo y sus raíces cuadradas son vértices de un triángulo equilátero. Determinar su área.

5. Se define la *razón doble* de cuatro números complejos a_1, a_2, a_3, a_4 como el cociente, si existe,

$$(a_1\, a_2\, a_3\, a_4) = \frac{(a_1\, a_3\, a_4)}{(a_2\, a_3\, a_4)} \in \mathbf{C}.$$

   Demostrar que $(a_1\, a_2\, a_3\, a_4) \in \mathbf{R}$ si y sólo si los cuatro puntos están en una circunferencia.

6. Sean $(a_1, b_1), (a_2, b_2), (a_3, b_3)$ tres puntos linealmente independientes del plano en una referencia ortonormal. Demostrar que el área del

triángulo que forman es

$$A = \frac{1}{2}\begin{vmatrix} a_1 & b_1 & 1 \\ a_2 & b_2 & 1 \\ a_3 & b_3 & 1 \end{vmatrix}.$$

Enunciar y demostrar una fórmula análoga para el volumen de un tetraedro.

7. Si g es el baricentro de un triángulo de vértices a, b, c, demostrar que los triángulos abg, bcg y cag tienen la misma área.

Enunciar y demostrar un ejercicio análogo para un tetraedro.

8. Sea $\{p;\ e_1, e_2, e_3\}$ una referencia ortonormal del espacio afín euclídeo (A, E). Designemos por s la simetría respecto al eje $p + \langle (a, b, c) \rangle$ y por $\tilde{s}$ su endomorfismo asociado.

a) Demostrar que, para todo $v \in E$, $\tilde{s}(v) + v$ es un vector propio de valor propio $+1$ (o bien es $\vec{0}$).

b) Deducir de (a) la matriz de $\tilde{s}$ en función de a, b, c.

c) Hallar las ecuaciones de la simetría axial respecto a la recta

$$\begin{cases} 3x - 4y - 25 = 0 \\ z = 2. \end{cases}$$

9. Sea $\{p;\ e_1, e_2, e_3\}$ una referencia ortonormal del espacio afín euclídeo (A, E), s la simetría especular respecto al plano $ax + by + cz + d = 0$ y $\tilde{s}$ el endomorfismo asociado a s.

a) Demostrar que, para todo $v \in E$, $\tilde{s}(v) - v$ es ortogonal al plano de simetría.

b) Deducir de (a) la matriz de $\tilde{s}$ en función de a, b, c.

c) Hallar las ecuaciones de la simetría especular respecto al plano $x + 2y - 3z + 2 = 0$.

10. Lugar geométrico de las imágenes del punto $(1, 1)$ por todos los giros de $\mathbf{R}^2$ de ángulo $\frac{\pi}{2}$ y centro sobre la recta $x + y = 1$.

11. Dadas dos rectas del plano no paralelas, hacemos corresponder a cada punto M el punto medio M^* de las proyecciones ortogonales de M sobre cada una de las rectas dadas.

a) Estudiar la correspondencia $M \longrightarrow M^*$.

b) ¿Cómo han de ser las dos rectas dadas para que esta correspondencia sea una homotecia?

12. Estudiar las semejanzas y los desplazamientos de la siguiente familia de afinidades:

$$\begin{cases} \bar{x} &= \left(\frac{a}{3}+b\right)x+\frac{a}{3}y+\frac{a}{3}z+a \\ \bar{y} &= \frac{a}{3}x+\left(\frac{a}{3}+b\right)y+\frac{a}{3}z+a \\ \bar{z} &= \frac{a}{3}x+\frac{a}{3}y+\left(\frac{a}{3}+b\right)z+a. \end{cases}$$

13. Determinar el valor del parámetro a para que la afinidad de ecuaciones

$$\begin{cases} \bar{x} &= ax-20y-15z+46 \\ \bar{y} &= 20x+9y-12z+16 \\ \bar{z} &= 15x-12y+16z-60 \end{cases}$$

sea una semejanza. Determinar su centro, eje, ángulo y razón.

14. Consideremos la familia de afinidades de $\mathbf{R}^2$

$$\begin{cases} \bar{x} &= cy+a \\ \bar{y} &= x+b. \end{cases}$$

a) Determinar los valores de los parámetros para los cuales estas afinidades son desplazamientos y estudiarlos.

b) ¿Para qué valores de los parámetros se cumple

$$\min \mathrm{d}(p,\bar{p}) = 4\ ?$$

c) Determinar el lugar geométrico de las imágenes de un punto p dado por todos los desplazamientos del apartado (b).

15. ¿Puede ser

$$\begin{pmatrix} 1 & -1 & -2 \\ 1 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

la matriz de una semejanza en alguna referencia? En caso afirmativo, estudiarla.

16. a) Demostrar que las simetrías axiales generan el grupo de los desplazamientos del plano. En particular, todo giro g es producto de dos simetrías axiales, $g = s_1 \circ s_2$, donde s_1 (o s_2) tiene por eje una recta prefijada que pasa por el centro de g. Deducir de este hecho un procedimiento geométrico para hallar el centro del giro producto de dos dados.

b) Enunciar y demostrar un ejercicio análogo para el espacio (sustituyendo simetría axial por simetría especular, recta por plano y centro del giro por eje del giro).

17. Sean $\{r_1, r_2\}, \{r_3, r_4\}, \{r_5, r_6\}$ los pares de aristas opuestas de un tetraedro regular. Si s_i denota la simetría axial de eje r_i, estudiar la composición $s_6 \circ s_5 \circ s_4 \circ s_3 \circ s_2 \circ s_1$.

18. ¿Qué condiciones debe cumplir un cuadrilátero del plano para que exista una afinidad que lo transforme en un cuadrado? Si existe alguno que satisface esas condiciones, ¿hay alguna homología que lo transforme en un cuadrado?

19. Estudiar el grupo de los desplazamientos de $\mathbf{R}^2$ que dejan fijo cada uno de los siguientes conjuntos:

i) un triángulo equilátero;

ii) un cuadrado;

iii) dos rectas que se cortan en un punto;

iv) una recta y un punto no contenido en ella.

20. Estudiar el grupo de los desplazamientos de $\mathbf{R}^3$ que dejan fijo cada uno de los siguientes conjuntos:

i) un tetraedro regular;

ii) dos rectas que se cortan en un punto;

iii) dos rectas que se cruzan;

iv) un cuadrado y un punto que no están en un mismo plano;

v) un triángulo equilátero y una recta que pasa por su baricentro.

21. Sea G el conjunto de todas las traslaciones y todas las homotecias.

a) Demostrar que G es un grupo generado por las homotecias.

b) Demostrar que las traslaciones forman un subgrupo normal de G.

c) Estudiar si el conjunto H de las homotecias de centro un punto p fijado es un subgrupo normal de G.

d) Determinar el lugar geométrico de los centros de las homotecias de los conjuntos

i) hHh^{-1}, donde h es una homotecia dada;

ii) THT^{-1}, donde T es una traslación dada.

22. Demostrar que una afinidad es una semejanza si y sólo si conserva la ortogonalidad y si y sólo si conserva los ángulos.

XIII.14 Ejercicios para programar

23. Hacer un programa que, dadas dos rectas de $\mathbf{R}^3$ que se cruzan, calcule

a) las ecuaciones de la perpendicular común (§1, ejemplo 2);

b) los pies de esa perpendicular;

c) la distancia entre las dos rectas (§2, II).

24. Preparar un programa que, dado un hiperplano H de $\mathbf{R}^n$ y un punto p cualquiera, permita calcular

a) la proyección ortogonal de p sobre H (§1, ejemplo 3);

b) la distancia de p a H (§2, I).

25. Sea $f : \mathbf{R}^3 \longrightarrow \mathbf{R}^3$ dado en la referencia canónica por $f(x) = Mx + b$ con $f \in O(3)$. Hacer un programa que

a) calcule $\det M$ y $\operatorname{tr} M$;

b) decida si hay puntos fijos (ejercicio X.22);

c) observando la tabla del §7, clasifique f;

d) calcule (cuando existan):

- plano de simetría;
- eje de rotación;
- vector de traslación.

Nota: se pueden utilizar las matrices preparadas en el ejercicio XII.14 para ensayar el programa.

26. (Programa gráfico)

Hacer un programa que permita ver desde cualquier ángulo una cierta figura del espacio. Se puede seguir el método siguiente:

a) Representar la figura en una referencia apropiada y guardar en una matriz $A \in M_{3\times n}(\mathbf{R})$ las coordenadas de sus puntos.

b) Elegir la rotación que se desee efectuar: el ángulo φ y el eje (X, Y o Z). Si R es la matriz de esa rotación, calcular RA.

c) Para hacer una traslación de vector b, sumar este vector a cada columna de la matriz. Se pueden hacer diversas rotaciones y traslaciones sucesivas repitiendo los mismos pasos.

d) Supongamos que el plano de la pantalla es el YZ. Para proyectar en él la matriz resultante hay que hacer lo siguiente:

   1) Eliminar la primera coordenada de todos los puntos.

   2) Para que el dibujo quede centrado en la pantalla, se debe aplicar a cada punto (y_i, z_i) una afinidad

$$\begin{cases} \bar{y}_i &= ry_i + c \\ \bar{z}_i &= rz_i + d, \end{cases}$$

   donde (c, d) es el centro de la pantalla según las coordenadas del ordenador y la razón r debe escogerse de manera que la figura quede dentro de los límites de esas coordenadas.

e) Dibujar los puntos $(\bar{y}_i, \bar{z}_i)$ obtenidos. Si en la figura inicial había un segmento que unía (x_i, y_i, z_i) con (x_j, y_j, z_j), dibujar ahora un segmento que una $(\bar{y}_i, \bar{z}_i)$ con $(\bar{y}_j, \bar{z}_j)$.

Observación: este mismo programa sirve para observar el efecto de una afinidad $f(x) = Mx + b$ sobre la figura. Solamente hace falta, en el segundo paso, usar la matriz M en lugar de R.

Índice alfabético

abelianizado de un grupo 65
adjunto 126
afinidad 217
álgebra 104
    de endomorfismos 103
altura 320
ángulo 280
    de un triángulo 318
    de una rotación 284, 307
    llano 282
    recto 283
anillo 17
    conmutativo 17
aplicación adjunta 260, 268
    afín 217
    autoadjunta 261
    canónica 100
    dual 106
    lineal 89
    ortogonal 271
    unitaria 271
automorfismo 91

baricentro 196, 320
base de un espacio vectorial 72
    de un ideal 10
    dual 105
    ortonormal 252
Bézout, identidad de 11
bisectriz 321

característica de un cuerpo 196
caras de un tetraedro 325
Cauchy-Schwarz, desigualdad 257
Cayley-Hamilton, teorema de 166
centro de un grupo 65
    de una homotecia 224
    de una rotación 307
    de una semejanza 314
    de simetría 223, 311
cero de un polinomio 32
Ceva, teorema de 208
ciclo 44
circuncentro 320
clases de equivalencia 16
    de restos 15
clausura algebraica 168
cociente 9, 25
coeficientes de un polinomio 23
combinación lineal 70
conjunto cociente 16
conmutador de un grupo 65
coordenadas baricéntricas 195
    cartesianas 202
    de un vector 82
coseno 281
Cramer, regla de 137
cuerpo 18

dependencia lineal 72
Desargues, teorema de 213
desigualdad triangular 256, 299
desplazamiento 306
    directo 306
    impropio 306
    inverso 306
    propio 306
determinante 118
    de un endomorfismo 122
    de una matriz 121

dimensión de un espacio afín 184
de un espacio vectorial 76
de una variedad lineal 188
dirección 187
distancia 299
entre variedades lineales 301
divisor 9, 26
de cero 18

ecuación diofántica 18
ecuaciones de una afinidad 230
de una variedad 201, 204
eje de simetría 308
de una rotación 286, 309
endomorfismo 91
diagonalizable 152
simultáneamente 179
nilpotente 180
triangulable 155
epimorfismo 50, 91
escalar 68
espacio
afín 184
estándar 185
euclídeo 299
bidual 107
de aplicaciones lineales 102
dual 105
vectorial 67
cociente 80
euclídeo 256
unitario 256
Euclides, algoritmo de 12, 29
teorema de 13, 30
extensión algebraica 36

Fermat, pequeño teorema de 20
Fitting, descomposición de 180
forma 105
bilineal 249
simétrica 251
definida positiva 251
multilineal alternada 116
sesquilineal 250
hermítica 251

Gauss, método de 140
generador de un grupo cíclico 57
generadores 49, 71
giro 307, 309
grado de un polinomio 23
Gram-Schmidt, método de 253
Grassmann, fórmula de 77, 191
grupo 41
abeliano 41
afín 233
alternado 54
cíclico 57
conmutativo 41
de las rotaciones 284
lineal 233
ortogonal 273
especial 277
unitario 273

haz de hiperplanos 204
homología especial 242
general 241
homomorfismo de grupos 49
homotecia 104, 224, 244

ideal 10, 27
imagen 50, 90
incentro 321
independencia lineal 72, 193
índice de un subgrupo 59
involución 162
isometría 304
isomorfismo 50, 91
afín 220

Jacobi, identidad de 265

Laplace, regla de 125

matriz 69
adjunta 133

canónica de Jordan 173
de cambio de base 83
de una afinidad 230
de una aplicación lineal 95
de una forma bilineal 250
sesquilineal 251
diagonal 152
hemisimétrica 86, 133
hermítica 251
identidad 84
inversa 84
ortogonal 273
simétrica 86, 251
traspuesta 107
triangular 155
unitaria 273
matrices equivalentes 152
máximo común divisor 11, 28
mediana 320
mediatriz 320
Menelao, teorema de 207
menor de una matriz 125
mínimo común múltiplo 10, 28
monomorfismo 50, 91
morfismo de grupos 49
movimiento 306
helicoidal 310
multiplicidad de un cero 32
de un valor propio 150
múltiplo 9, 26

norma 256
núcleo 50, 90
número primo 13
números congruentes 15
primos entre sí 13

orden de un elemento 58
de un grupo finito 58
orientación 209
ortocentro 324

Pappus, teorema de 213
paralelismo 189
partición 16
permutación 43
impar 46
par 46
regular 64
perpendicular común 301
Pitágoras, teorema de 300
plano de simetría 311
polinomio 23
anulador 157
característico 150
irreducible 30
mínimo 157, 158
primo 30
polinomios congruentes 35
primos entre sí 30
producto de matrices 83
directo de grupos 55
de subgrupos 55
escalar 251
vectorial 263
proyección 222
ortogonal 301
proyector 113
punto medio 197

raíz de un polinomio 32
rango 91, 128
razón de una homología 241
de una homotecia 104, 224
de una semejanza 313
doble 326
simple 205
referencia baricéntrica 195
cartesiana 202
relación 15
de equivalencia 16
resto 9, 25
rotación 284, 286, 307, 309

segmento 197
semejanza 313
directa 315

inversa 315
semiespacio 210
seno 281
signo 47
simetría 223
axial 287, 308, 310
central 223, 288, 307, 311
especular 288, 311
ortogonal 279
Steinitz, teorema de 74
subespacio cíclico 168
complementario 79
invariante 159
ortogonal 109, 259
vectorial 70
subgrupo 48
normal 53
sucesión 23
suma de ángulos 281
de aplicaciones lineales 102
de subespacios vectoriales 77
de variedades lineales 191
directa 78, 79

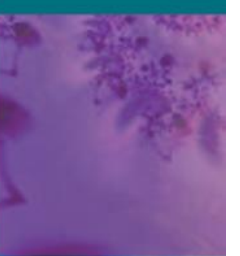
Tales, teorema de 227
Taylor, fórmula de 112
teorema chino del resto 21
de descomposición 160, 169
de diagonalización 155, 165
de isomorfismo 54, 99
de triangulación 155
fundamental del Álgebra 34
terna pitagórica 21
tetraedro 325
traslación 186
trasposición 45
traza 152
triángulo 318
equilátero 324
isósceles 319
triángulos congruentes 322
semejantes 322

valor de un polinomio 32
propio 150
variedad lineal 187
invariante 236
variedades lineales
ortogonales 300
paralelas 189
perpendiculares 300
que se cruzan 190
vector 68
propio 149
unitario 252
vectores ortogonales 252
vértices de un triángulo 318

Wilson, congruencia de 20

Manuel Castellet \ Irene Llerena

Álgebra Lineal y Geometría



EDITORIAL REVERTÉ
www.reverte.com

UAB
Universitat Autònoma de Barcelona