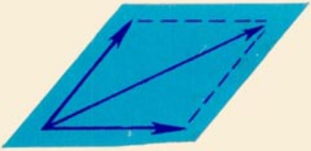
L. I. GOLOVINA

ALGEBRA LINEAL Y ALGUNAS DE SUS APLICACIONES



ИЗДАТЕЛЬСТВО · МИР ·

Л. И. ГОЛОВИНА

ЛИНЕЙНАЯ АЛГЕБРА
И НЕКОТОРЫЕ ЕЕ ПРИЛОЖЕНИЯ

ИЗДАТЕЛЬСТВО
НАУКА

L. I. GOLOVINÁ

ALGEBRA LINEAL
Y ALGUNAS DE SUS APLICACIONES

Traducido por Carlos Vega,
catedrático de Matemáticas Superiores

Segunda edición

EDITORIAL MIR
MOSCÚ

Primera edición 1974
Segunda edición 1980

На испанском языке

© Traducción al español. Editorial Mir. 1980

Impreso en la URSS 1980

ÍNDICE

CAPÍTULO IX

NOCIÓN DE TENSORES

CAPÍTULO X

CONCEPTOS PRINCIPALES DE LA TEORÍA DE GRUPOS

PREFACIO

Este libro es un manual de Algebra lineal para los estudiantes de los Institutos Politécnicos y de las Facultades de Ciencias Naturales de las Universidades. Será también útil para el lector que desee, de un modo individual, estudiar los conceptos principales del Algebra lineal en una fuente que no requiere conocimientos previos de las Matemáticas superiores. Sólo se supone que el lector, además del curso elemental de Matemáticas, está familiarizado con los elementos de la Geometría analítica. Los conceptos del Análisis matemático que se emplean en el libro (la derivada y la integral) sólo aparecen en ejemplos que pueden ser omitidos en la lectura sin perjuicio para la comprensión del material.

El capítulo I es de carácter introductorio; contiene todos los elementos de la teoría de determinantes y de la teoría de sistemas de ecuaciones lineales necesarios en lo sucesivo. Los capítulos del II al VI son los principales: exponen un curso breve del Álgebra lineal propiamente dicha.

Los cuatro últimos capítulos no se refieren, realmente, al Algebra lineal, pero sus resultados se basan en el material anterior («. . . y algunas de sus aplicaciones»). Estos capítulos son independientes y pueden ser leídos en cualquier orden (véase el esquema de dependencia de los capítulos).

El capítulo VII está dedicado a la teoría general de curvas y de superficies de segundo grado. Completa y profundiza la parte correspondiente del curso de Geometría analítica sin pretender a sustituirla.

Poco común para un manual de Algebra lineal, el capítulo VIII, dedicado a la teoría especial de la relatividad, ha sido inspirado, en gran medida, por el curso que P. K. Rashevski dictó en la Universidad de Moscú en 1940. Este capítulo puede ser omitido al estudiar el Algebra lineal, pero la experiencia muestra que despierta generalmente gran interés en los oyentes.

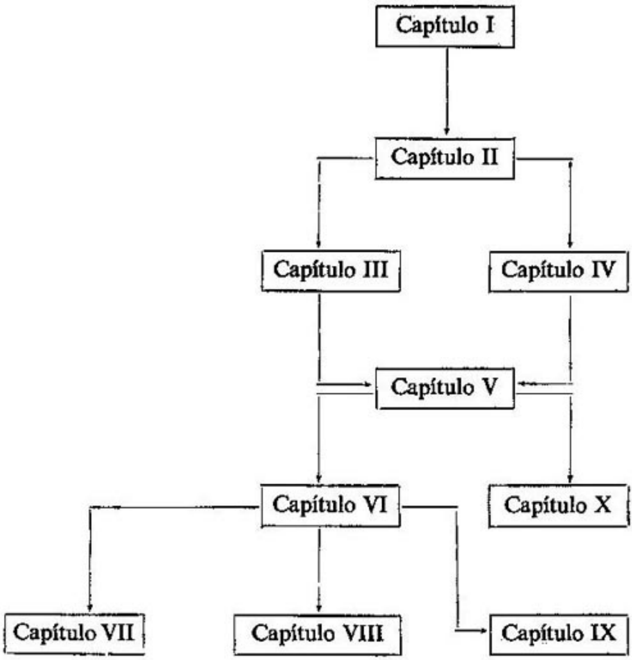
El capítulo IX está dedicado a los conceptos principales del Algebra tensorial. Tiene más bien carácter de resumen y puede servir de introducción a exposiciones más detalladas sobre el mismo tema.

El último capítulo contiene solamente las nociones más generales de la teoría de grupos, cuyo conocimiento es hoy día indispensable no sólo para el matemático, sino también para el ingeniero. Sin embargo, la teoría de grupos no es, ni mucho menos, una parte del Algebra lineal; por lo tanto, para estudiarla con cierto detalle, el lector debe recurrir a la literatura especial.

El presente libro equivale, por su contenido, al curso algo ampliado que la autora dictó varias veces en la Sección de Química Física de la Facultad de Química de la Universidad de Moscú.

L. I. Goloviná

ESQUEMA DE DEPENDENCIA DE LOS CAPÍTULOS



CAPÍTULO I

DETERMINANTES Y SISTEMAS DE ECUACIONES LINEALES

Este capítulo contiene el material auxiliar relacionado con la solución de sistemas de ecuaciones lineales (es decir, de ecuaciones de primer grado). Para el estudio de estos sistemas se introduce el importante concepto de determinante. Los resultados de este capítulo, que tienen interés tanto por sí mismos como por sus aplicaciones en la Geometría analítica, son indispensables para la comprensión de los capítulos sucesivos del libro.

§ 1. Sistemas de ecuaciones con dos y con tres incógnitas

En la solución de *una ecuación de primer grado con una incógnita*

$$ax = b$$

se pueden presentar tres casos:

1. Si $a \neq 0$, la ecuación tiene una solución *única* $x = \frac{b}{a}$.

2. Si $a = 0$ y $b = 0$, la ecuación tiene un *conjunto infinito* de soluciones: cualquier número real x satisface la ecuación $ax = b$ (ya que $0 \cdot x = 0$) y es, por consiguiente, solución de la misma.

3. Si $a = 0$, pero $b \neq 0$, la ecuación *no tiene* soluciones, ya que al sustituir x por cualquier número real se obtiene en el primer miembro el cero, mientras que el segundo miembro es diferente de cero.

Estos mismos tres casos se pueden presentar al resolver un sistema arbitrario de ecuaciones lineales.

Consideremos un *sistema de dos ecuaciones con dos incógnitas:*

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2. \end{cases} \tag{1}$$

Se llama *solución* de este sistema a todo par de valores $x = \alpha$ e $y = \beta$ que, al sustitur por ellos a x y a y, convierten ambas ecuaciones en identidades. Para resolver este sistema, multipliquemos la primera ecuación por b_2, la segunda por $-b_1$ y sumémoslas; obtendremos

$$x(a_1b_2-a_2b_1) = c_1b_2-c_2b_1.$$

De aquí, siendo $a_1b_2-a_2b_1 \neq 0$, tendremos

$$x = \frac{c_1b_2-c_2b_1}{a_1b_2-a_2b_1}. \tag{2}$$

Análogamente encontramos que

$$y = \frac{a_1c_2-a_2c_1}{a_1b_2-a_2b_1}. \tag{3}$$

Por consiguiente, en el caso en que $a_1b_2-a_2b_1 \neq 0$, el sistema (1) tiene una solución *única*.

Las expresiones que figuran en los numeradores y en los denominadores de los segundos miembros de las igualdades (2) y (3) tienen la misma estructura. A saber, consideremos la tabla cuadrada de números

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}.$$

Las tablas de este tipo se llaman ***matrices***. Las líneas horizontales de los números que componen la matriz se denominan *filas* de la misma y las líneas verticales se denominan *columnas*. Los números a_1, b_1, a_2 y b_2 que componen la matriz se denominan *elementos* de la misma. En nuestro ejemplo nos hemos encontrado con una matriz cuadrada de *segundo orden*. La diagonal que va del ángulo izquierdo superior al ángulo derecho inferior de la matriz se llama *diagonal principal* de ésta. Los denominadores de las fracciones que aparecen en los segundos miembros de las igualdades (2) y (3) tienen la siguiente estructura: del producto de los elementos que se hallan en la diagonal principal de la matriz A se resta el producto de los elementos que figuran en su otra diagonal, llamada *secundaria*:

$$a_1b_2-a_2b_1.$$

La expresión así obtenida lleva el nombre de ***determinante*** de la matriz A (determinante de segundo orden) y se indica como sigue:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Luego, por definición,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1.$$

En esta notación el numerador de la fracción que figura en el segundo miembro de la igualdad (2) representa el determinante

$$\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} = c_1b_2 - c_2b_1$$

que se obtiene del denominador al sustituir la primera columna por la columna de los términos independientes, mientras el numerador de la fracción que aparece en el segundo miembro de la igualdad (3) representa el determinante

$$\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} = a_1c_2 - a_2c_1$$

que se obtiene del denominador sustituyendo la segunda columna por la columna de los términos independientes de las ecuaciones del sistema (1).

Hemos encontrado, pues, que

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \quad \text{e} \quad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}.$$

Estas son las *fórmulas de Cramer* para la solución de un sistema de dos ecuaciones con dos incógnitas.

Ejemplo 1. *Resuélvase, empleando las fórmulas de Cramer, el sistema de ecuaciones*

$$\begin{cases} 2x+5y = 8, \\ 3x+y = -1. \end{cases}$$

Solución.

$$x = \frac{\begin{vmatrix} 8 & 5 \\ -1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 5 \\ 3 & 1 \end{vmatrix}} = \frac{8+5}{2-15} = -1; \quad y = \frac{\begin{vmatrix} 2 & 8 \\ 3 & -1 \end{vmatrix}}{-13} = \frac{-2-24}{-13} = 2.$$

Consideremos ahora el caso en que

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1 = 0. \tag{4}$$

La igualdad (4) puede ser representada así[1]:

$$\frac{a_1}{a_2} = \frac{b_1}{b_2},$$

es decir, los coeficientes de las incógnitas son en este caso proporcionales. Si además se tiene que

$$\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} = 0, \text{ es decir, } \quad \frac{c_1}{c_2} = \frac{b_1}{b_2},$$

también los términos independientes son proporcionales a los coeficientes de las incógnitas y tenemos, en realidad, una ecuación con dos incógnitas que admite un conjunto *infinito* de soluciones.

Finalmente, si

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = 0, \text{ pero } \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \neq 0,$$

es decir, si

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2},$$

las ecuaciones son, evidentemente, contradictorias y el sistema *no tiene solución alguna.*

La solución $x = \alpha$, $y = \beta$ del sistema de ecuaciones (1) determina el punto de intersección de las rectas

$$a_1x + b_1y = c_1 \quad \text{y} \quad a_2x + b_2y = c_2.$$

Si $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$, son dos rectas distintas no paralelas que tienen, por consiguiente, un punto común *único*. En el caso en que $\frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2}$, las rectas son paralelas y, por consiguiente, no tienen *ningún* punto común. Finalmente, si es $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$, ambas

[1] Aquí y en lo que sigue aceptamos que los denominadores son diferentes de cero; el caso en que esto no sea así, considérelo usted mismo.

ecuaciones determinan *una misma* recta y *todos* sus puntos son precisamente «puntos de intersección» de las rectas dadas.

Consideremos ahora un sistema de *tres ecuaciones lineales con tres incógnitas*

$$\begin{cases} a_1x+b_1y+c_1z = d_1, \\ a_2x+b_2y+c_2z = d_2, \\ a_3x+b_3y+c_3z = d_3. \end{cases} \tag{5}$$

Se llama *solución* de este sistema a toda terna de números $x = \alpha, y = \beta, z = \gamma$ que convierten las tres ecuaciones en identidades. Multiplicando la primera ecuación por $\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} = b_2c_3-b_3c_2$, la segunda por $-\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} = b_3c_1-b_1c_3$, la tercera por $\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} =$ $= b_1c_2-b_2c_1$ y sumándolas todas, obtenemos

$$x(a_1b_2c_3-a_1b_3c_2+a_2b_3c_1-a_2b_1c_3+a_3b_1c_2-a_3b_2c_1) =$$
$$= d_1b_2c_3-d_1b_3c_2+d_2b_3c_1-d_2b_1c_3+d_3b_1c_2-d_3b_2c_1$$

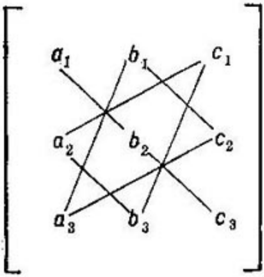
(es fácil ver que los coeficientes de y y z serán iguales a cero). De aquí, siendo el coeficiente de x diferente de cero, obtenemos

$$x = \frac{d_1b_2c_3-d_1b_3c_2+d_2b_3c_1-d_2b_1c_3+d_3b_1c_2-d_3b_2c_1}{a_1b_2c_3-a_1b_3c_2+a_2b_3c_1-a_2b_1c_3+a_3b_1c_2-a_3b_2c_1}. \tag{6}$$

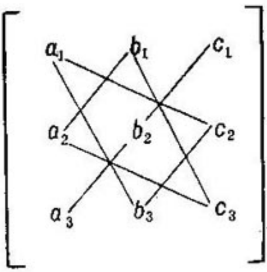
Veamos la estructura que tiene la expresión que aparece en el denominador del segundo miembro de la igualdad (6). Consideremos con este fin la tabla cuadrada *(la matriz de tercer orden)*

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}.$$

El denominador de la fórmula (6) representa la suma algebraica de seis términos cada uno de los cuales es un producto de tres elementos tomados uno a uno de cada fila y de cada columna de la matriz A, con la particularidad de que el signo *más* lo tienen el producto de los elementos que pertenecen a la diagonal principal y los dos productos de los elementos que forman en la matriz triángulos (isósceles) de base paralela a la diagonal principal:



mientras que el signo *menos* lo tienen el producto de los elementos pertenecientes a la diagonal secundaria y los dos productos de los elementos que forman triángulos de base paralela a la diagonal secundaria:



Una expresión de este tipo se llama *determinante* de la matriz A (determinante de *tercer orden*) y se indica como sigue

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Es decir, por definición,

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - \\ - a_3b_2c_1 - a_2b_1c_3 - a_1b_3c_2 .$$

La expresión que figura en el numerador del segundo miembro de la fórmula (6) se obtiene del denominador sustituyendo cada una de las letras a por la letra d con el mismo índice, es decir,

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}} = \frac{D_1}{D}. \tag{6a}$$

Se puede demostrar análogamente que, siendo $D \neq 0$, del sistema (5) se deducen las igualdades

$$y = \frac{D_2}{D} \quad \text{y} \quad z = \frac{D_3}{D}, \tag{7}$$

donde D_i, $i = 1, 2, 3$, es el determinante que se obtiene del determinante D sustituyendo su i-ésima columna por la columna de los términos independientes. Estas son *las fórmulas de Cramer* para un sistema de tres ecuaciones con tres incógnitas.

Ejemplo 2. *Resuélvase empleando las fórmulas de Cramer el sistema de ecuaciones*

$$\begin{cases} x+2y+3z = 7, \\ x-3y+2z = 5, \\ x+ y+ z = 3. \end{cases}$$

Solución.

$$D = \begin{vmatrix} 1 & 2 & 3 \\ 1 & -3 & 2 \\ 1 & 1 & 1 \end{vmatrix} = -3+3+4+9-2-2 = 9 \neq 0;$$

$$D_1 = \begin{vmatrix} 7 & 2 & 3 \\ 5 & -3 & 2 \\ 3 & 1 & 1 \end{vmatrix} = -21+15+12+27-10-14 = 9,$$

$$D_2 = \begin{vmatrix} 1 & 7 & 3 \\ 1 & 5 & 2 \\ 1 & 3 & 1 \end{vmatrix} = 5+14+9-15-7-6 = 0,$$

$$D_3 = \begin{vmatrix} 1 & 2 & 7 \\ 1 & -3 & 5 \\ 1 & 1 & 3 \end{vmatrix} = -9+7+10+21-6-5 = 18.$$

Por consiguiente,

$$x=\frac{D_1}{D}=1,\quad y=\frac{D_2}{D}=0 \quad \text{y} \quad z=\frac{D_3}{D}=2.$$

Con el fin de introducir el concepto de *determinante de orden n*, consideremos de nuevo los determinantes de segundo y de tercer órdenes:

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2-a_2b_1 \tag{8}$$

y

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3+a_2b_3c_1+a_3b_1c_2-a_3b_2c_1- \\ -a_2b_1c_3-a_1b_3c_2. \tag{9}$$

Vemos que el determinante es la suma algebraica de todos los productos posibles de sus elementos formados de modo que haya uno de cada fila y uno de cada columna. Cada uno de estos productos se llama *término* del determinante. Coloquemos los factores de todo término del determinante de segundo orden en el orden de secuencia de las *columnas* del determinante:

$$a_1b_2-a_2b_1.$$

Consideremos las ordenaciones correspondientes *(permutaciones)* de los índices inferiores (que indican el número de las filas):

$$1,2 \quad \text{y} \quad 2,1.$$

En el primer producto los índices aparecen en el orden de crecimiento y el producto correspondiente figura en el determinante con el signo *más;* en el segundo producto los índices forman, como suele decirse, un desorden o una *inversión* 2, 1 y el término correspondiente aparece en el determinante con el signo *menos*.

El determinante de orden tres tiene seis términos. Si colocamos los factores de cada uno de ellos en el orden de secuencia de las columnas, los índices inferiores de los términos que aparecen con el signo *más* formarán las permutaciones

$$1, 2, 3; \quad 2, 3, 1 \quad \text{y} \quad 3, 1, 2.$$

Consideremos tres pares de índices 1, 2; 1, 3 y 2, 3 correspondientes a la primera permutación 1, 2, 3; en cada uno de estos pares los números aparecen en el orden de crecimiento; se dice que esta permutación tiene *cero* inversiones. La segunda permutación 2, 3, 1 origina tres pares de índices: 2, 3; 2, 1 y 3, 1 y *dos* de ellos (a saber, 2, 1 y 3, 1) forman inversiones. La tercera permutación 3, 1, 2 origina tres pares de índices — 3,1; 1, 2 y 3, 2 — y *dos* de ellos — 3, 1 y 3, 2 — forman inversiones.

A los productos que aparecen con el signo *menos* les corresponden tres permutaciones de los índices inferiores:

3, 2, 1;   2, 1, 3   y   1, 3, 2;

es fácil ver que la primera contiene *tres* inversiones: 3, 2; 3, 1 y 2, 1, mientras que la segunda y tercera, una inversión: 2, 1 y 3, 2, respectivamente. Por consiguiente, con el signo más aparecen aquellos términos que contienen un número *par* de inversiones en la permutación de los índices inferiores y con el signo menos aparecen aquellos para los cuales este número es *impar*.

Es conveniente para lo sucesivo introducir unas nuevas notaciones para los determinantes de segundo y tercer órdenes:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \text{y} \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

donde todos los elementos del determinante se designan por una misma letra a con dos índices, de los cuales el primero corresponde al número de la fila y el segundo, al número de la columna. Entonces, se tiene

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

y

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum \pm a_{i_1 1} a_{i_2 2} a_{i_3 3},$$

donde el signo más lo llevan aquellos productos en los cuales la permutación i_1, i_2, i_3 es *par* (es decir, tiene un número par de inversiones) y el signo menos lo llevan aquellos, en los cuales ésta es *impar*. Esto mismo se puede escribir así:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum (-1)^{\alpha} a_{i_1 1} a_{i_2 2} a_{i_3 3},$$

donde α es el número de inversiones en la permutación i_1, i_2, i_3 de los primeros índices (los segundos índices están en su orden natural) y la suma se extiende a todas las seis permutaciones i_1, i_2, i_3 de los tres números 1, 2, 3.

§ 2. Permutaciones y trasposiciones. Determinante de orden n

Sean dados n elementos $a_1, a_2, \ldots, a_n$ (que pueden ser, por ejemplo, los números $1, 2, 3, \ldots, n$). Como se sabe, se llaman *permutaciones* de n elementos a todas las ordenaciones posibles de estos elementos. Con n elementos se puede formar un total de $n!$ permutaciones (demuéstrese esto).

Si un par a_i, a_k de elementos de una permutación aparece en ésta de modo que el elemento con mayor índice antecede al elemento con menor índice, se dice que estos elementos forman una *inversión*. Supongamos que debemos encontrar el número de inversiones que presenta una permutación cualquiera formada por los números naturales $1, 2, 3, \ldots, n$ (éstos pueden ser los índices de los elementos $a_1, a_2, \ldots, a_n$). Para ello se puede proceder así. Calculemos primero el número de elementos que *anteceden a la unidad;* todos estos elementos, y sólo ellos, forman inversiones con la unidad. Tachemos después la unidad y calculemos el número de elementos que *anteceden al dos;* éstos serán todos aquellos elementos que forman inversión con el dos (sin tomar en consideración la unidad tachada que también puede formar una inversión con el dos; pero en dicho caso esta inversión ha sido ya fijada anteriormente). Tachemos luego el dos y calculemos el número de elementos que *anteceden al tres*, etc. Sumemos todos los números obtenidos; esta suma será igual precisamente al número total de inversiones. El número de inversiones en la permutación $i_1, i_2, \ldots, i_n$ se indica así: $[i_1, i_2, \ldots, i_n]$. Por ejemplo,

$$[2, 5, 1, 4, 7, 3, 6] = 2+0+3+1+0+1 = 7.$$

Las permutaciones que presentan un número *par* de inversiones se llaman **pares** y las permutaciones que presentan un número *impar* de inversiones se llaman permutaciones **impares**.

Sea $a_1, a_2, \ldots, a_i, \ldots, a_k, \ldots, a_n$ una permutación dada de n elementos. Pongamos cada uno de los elementos a_i y a_k en el lugar del otro; obtendremos así la permutación $a_1, a_2, \ldots, a_k, \ldots, a_i, \ldots, a_n$. Esta operación de traslado de dos elementos de una permutación se llama *trasposición*.

Teorema 1. *Una trasposición altera la paridad de la permutación* (es decir, una permutación par se hace impar y una permutación impar se hace par).

Demostración. Consideremos primero el caso en que se trasponen dos elementos *consecutivos* α y β de una permutación

$$a_1, a_2, \ldots, a_l, \alpha, \beta, b_1, b_2, \ldots, b_m. \tag{10}$$

Después de la trasposición de los elementos α y β obtendremos la permutación

$$a_1, a_2, \ldots, a_l, \beta, \alpha, b_1, b_2, \ldots, b_m. \tag{11}$$

Puesto que las permutaciones (10) y (11) difieren sólo en la posición recíproca de los elementos α y β (y como quiera que la posición recíproca de cada uno de estos elementos con otro cualquiera, así como la posición recíproca de dos cualesquiera de los restantes elementos, es la misma), resulta que el número de inversiones en la permutación (11) es mayor o menor en una unidad que el número de inversiones en la permutación (10) y, por consiguiente, una de estas permutaciones es par, mientras que la otra es impar.

Consideremos ahora el caso general. Supongamos que en la permutación $a_1, \ldots, a_l, \alpha, c_1, \ldots, c_k, \beta, b_1, \ldots, b_m$ se trasponen los elementos α y β entre los cuales figuran otros k elementos $c_1, c_2, \ldots, c_k$. Podemos realizar la trasposición de los elementos α y β mediante varias trasposiciones de elementos consecutivos: traspongamos primero α con c_1, después con c_2, etc. y, finalmente, con c_k (realizando así k trasposiciones de elementos consecutivos); traspongamos después α y β (una trasposición más) y, finalmente, traspongamos β consecutivamente con c_k, c_{k-1}, etc. y c_1 (otras k trasposiciones de elementos consecutivos). En definitiva β ocupará el lugar de α (y viceversa). Cada una de estas trasposiciones altera, como hemos visto, la paridad de la permutación. Puesto que ésta se altera $2k+1$ veces, es decir, un número impar de veces, resulta que una permutación impar se hace par y una par se hace impar.

Corolario. *El número de permutaciones impares de n elementos es igual al número de permutaciones pares* (y, por consiguiente, es igual a $n!/2$).

Demostración. Supongamos que entre las $n!$ permutaciones de n elementos hay p pares y q impares. Realicemos en cada una de las permutaciones pares *una misma trasposición*, trasponiendo, por ejemplo, los dos primeros elementos. Toda permutación par se convertirá en una impar y es evidente, además, que todas las p permutaciones impares obtenidas de esta forma serán diferentes. Puesto que, por hipótesis, el número total de permutaciones impares de n elementos es igual a q, resulta que $p \leqslant q$. Podemos comprobar de la misma forma que, al contrario, $p \geqslant q$. Por consiguiente, $p = q$.

Demos ahora la definición general de un determinante. Sea dada una tabla cuadrada *(una matriz de orden n)*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Los números a_{ik} se llaman *elementos* de la misma; las líneas horizontales de elementos se llaman *filas* y las verticales, *columnas* de la misma. Se llama *determinante* de esta matriz (determinante de *orden n*) a la suma algebraica de todos los productos posibles de elementos formados de modo que haya un factor de cada fila y uno de cada columna de la matriz A. Si en cada uno de estos productos (*términos* del determinante) los factores se colocan en el orden de secuencia de las columnas, se toman con el signo *más* aquellos productos para los cuales es *par* la permutación formada por los primeros índices y con el signo *menos* aquellos para los cuales esta permutación es *impar*. Resumiendo:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (-1)^{[i_1, i_2, \dots, i_n]} a_{i_1 1} a_{i_2 2} \dots, a_{i_n n},$$

donde la suma se extiende a todas las permutaciones posibles i_1, $i_2, \dots, i_n$ de los n números $1, 2, 3, \dots, n$. El determinante de orden n contiene $n!$ términos, ya que el número de permutaciones de n

elementos es igual a $n!$ Debido al corolario del teorema 1, justamente la mitad de éstos, es decir, $n!/2$ términos, aparecen en el determinante con el signo más y la misma cantidad, con el signo menos.

§ 3. Propiedades de los determinantes

Con el crecimiento del orden del determinante aumenta rápidamente el número de sus elementos. Así, el determinante de orden cuatro consta de 24 términos, el determinante de orden cinco, de 120, el determinante de orden seis, de 720 términos, etc. Por lo tanto, resulta prácticamente imposible calcular los determinantes de orden superior a tres valiéndose sólo de la definición. Para poder calcular estos determinantes, deberemos estudiar sus propiedades. Demostraremos ante todo una proposición auxiliar.

Lema (del signo del término de un determinante). *El producto* $a_{i_1k_1}\, a_{i_2k_2}\, \ldots\, a_{i_nk_n}$ *aparece en el determinante de orden n con el signo determinado por la expresión*

$$(-1)^{[i_1,\, i_2,\, \ldots,\, i_n]+[k_1,\, k_2,\, \ldots,\, k_n]};$$

diremos simplemente que aparece con el signo

$$(-1)^{[i_1,\, i_2,\, \ldots,\, i_n]+[k_1,\, k_2,\, \ldots,\, k_n]}$$

Demostración. Observemos, ante todo, que colocando dos factores del producto $a_{i_1k_1}\, a_{i_2k_2}\, \ldots\, a_{i_nk_n}$ uno en lugar del otro, se provoca una trasposición tanto en los primeros como en sus segundos índices y, por consiguiente, se altera la paridad de cada uno de los números

$$[i_1, i_2, \ldots, i_n] \quad \text{y} \quad [k_1, k_2, \ldots, k_n],$$

mientras que la paridad de la suma de éstos $[i_1, i_2, \ldots, i_n] + [k_1\ k_2, \ldots, k_n]$ se conserva.

Sea dado el producto $a_{i_1k_1}a_{i_2k_2} \ldots a_{i_nk_n}$. Coloquemos sus factores, mediante una serie de trasposiciones, de modo que los segundos índices aparezcan en su orden natural. Con este fin realicemos primero una trasposición tal que en la primera posición aparezca el elemento de la primera columna, después una trasposición tal que en el segundo lugar aparezca el elemento de la segunda columna, etc. (Así, por ejemplo, el producto $a_{45}a_{14}a_{52}a_{21}a_{33}$ se transforma consecutivamente en $a_{21}a_{14}a_{52}a_{45}a_{33}$, después en $a_{21}a_{52}a_{14}a_{45}a_{33}$, en $a_{21}a_{52}a_{33}a_{45}a_{14}$ y, finalmente, en $a_{21}a_{52}a_{33}a_{14}a_{45}$.) Si después de colocar los segundos índices en orden de crecimiento los primeros forman la permutación $[m_1, m_2, \ldots, m_n]$, esto significará, por definición, que el término considerado aparece en el determinante con el signo $(-1)^{[m_1, m_2, \ldots, m_n]}$. Pero como durante la trasposición de los factores no ha cambiado la paridad de la suma $[i_1, i_2, \ldots, i_n] + [k_1, k_2, \ldots, k_n]$ del número de inversiones en los primeros índices y del número de inversiones en los segundos índices, resulta que la paridad de esta suma en la ordenación inicial de los factores coincide con la paridad del número $[m_1, m_2, \ldots, m_n]$ que representa el número de inversiones correspondiente a la permutación de los primeros índices en la ordenación definitiva, cuando los segundos índices forman cero inversiones. Por consiguiente,

$$(-1)^{[m_1, m_2, \ldots, m_n]} = (-1)^{[i_1, i_2, \ldots, i_n] + [k_1, k_2, \ldots, k_n]}$$

y esto demuestra nuestra afirmación.

Ejemplo 3. *Determínese el signo que corresponde al producto* $a_{32}a_{43}a_{51}a_{15}a_{24}$ *en el determinante de quinto orden*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}. \tag{12}$$

Solución.

$[3, 4, 5, 1, 2] = 3+3 = 6,$

$[2, 3, 1, 5, 4] = 2+1 = 3;$

$(-1)^{6+3} = (-1)^9 = -1.$

El producto considerado aparece en el determinante (12) con el signo *menos*.

Propiedad 1 (paridad de las columnas y de las filas de un determinante). *El valor de un determinante no varíe si éste se* ***traspone****, es decir, si se cambia cada una de sus filas por la columna del mismo número.*

Demostración. Consideremos los determinantes

$$D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

y

$$D' = \begin{vmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix}.$$

Debemos demostrar que $D' = D$.

Todo término del determinante D es un término del determinante D', ya que sus factores siguen permaneciendo en el determinante D' en distintas filas y en distintas columnas; recíprocamente, todo término del determinante D' es al mismo tiempo un término del determinante D. Por consiguiente, ambos determinantes representan la «suma algebraica» (es decir, una suma en la que algunos sumandos se toman con el signo menos) de los mismos términos de tipo $a_{i_1k_1} a_{i_2k_2} \dots a_{i_nk_n}$. La diferencia estriba solamente en que en el determinante D los primeros índices corresponden a los números de las filas y los segundos a los números de las columnas, mientras que en el determinante D' sucede al revés. Pero, según el lema del signo del término de un determinante, el signo de este producto tanto en el primero como en el segundo determinante será el mismo:

$$(-1)^{[i_1, i_2, \dots, i_n]+[k_1, k_2, \dots, k_n]};$$

luego, $D' = D$.

Propiedad 2. *Si se cambian entre sí dos filas o dos columnas de un determinante, el determinante cambia sólo de signo, pero su valor absoluto no varía.*

Demostremos esta afirmación, por ejemplo, en el caso de las columnas. Cambiando entre sí la p-ésima y la q-ésima columnas del determinante

$$D_1 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1p} & \dots & a_{1q} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2p} & \dots & a_{2q} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{np} & \dots & a_{nq} & \dots & a_{nn} \end{vmatrix},$$

obtendremos el determinante

$$D_2 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1q} & \dots & a_{1p} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2q} & \dots & a_{2p} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nq} & \dots & a_{np} & \dots & a_{nn} \end{vmatrix}.$$

Todo término del determinante D_1 es también un término del determinante D_2, ya que sus factores siguen permaneciendo en el determinante D_2 en distintas filas y en distintas columnas, y viceversa. Tomemos un término cualquiera

$$a_{i_1 1}a_{i_2 2}\ \dots\ a_{i_p p}\ \dots\ a_{i_q q}\ \dots\ a_{i_n n}$$

del determinante D_1. Puesto que sus factores aparecen en el orden de secuencia de las columnas de D_1, resulta que este término figura en el determinante D_1 con el signo $(-1)^{[i_1, i_2, \dots, i_p, \dots, i_q, \dots, i_n]}$. Para determinar el signo de este término en el determinante D_2, coloquemos sus factores en el orden de secuencia de las columnas de D_2:

$$a_{i_1 1}a_{i_2 2}\ \dots\ a_{i_q q}\ \dots\ a_{i_p p}\ \dots\ a_{i_n n}$$

(el elemento $a_{i_q q}$ pertenece a la p-ésima columna del determinante D_2 y el elemento $a_{i_p p}$ a la q-ésima columna). Los primeros índices, tanto en el determinante D_2 como en el determinante D_1, corresponden a los números de las filas; por lo tanto, el producto considerado aparece en el determinante D_2 con el signo $(-1)^{[i_1, i_2, \dots, i_q, \dots, i_p, \dots, i_n]}$. Pero la permutación $i_1, i_2, \dots, i_q, \dots, i_p, \dots, i_n$ se obtiene de la permutación $i_1, i_2, \dots, i_p, \dots, i_q, \dots, i_n$ mediante *una trasposición* y, por consiguiente, los números

$$[i_1, i_2, \dots, i_p, \dots, i_q, \dots, i_n] \quad \text{y} \quad [i_1, i_2, \dots, i_q, \dots, i_p, \dots, i_n]$$

son de paridad *distinta*. Luego, todo término del determinante D_1 aparece en el determinante D_2 con el signo opuesto y, por consiguiente, $D_2 = -D_1$.

Para demostrar la proposición correspondiente a las filas, pasemos a los determinantes traspuestos D_1' (el determinante traspuesto de D_1) y D_2' (el determinante traspuesto de D_2). Si el determinante D_2 se obtiene de D_1 cambiando entre sí la p-ésima y la q-ésima filas, resulta que D_2' se obtiene de D_1' cambiando entre sí la p-ésima y la q-ésima columnas y, por consiguiente, $D_2' = -D_1'$. Pero $D_1' = D_1$ y $D_2' = D_2$; luego, $D_2 = -D_1$.

Corolario. *Un determinante que tiene dos filas o dos columnas idénticas es igual a cero.*

Para demostrar esto, cambiemos entre sí las filas (o las columnas) idénticas del determinante D; claro está que el determinante en este caso no varía. Pero, como según la propiedad 2 debe cambiar de signo, resulta que $D = -D$, de donde se tiene que $D = 0$.

Propiedad 3. *Si se multiplican todos los elementos de una fila o de una columna de un determinante por un mismo número, el valor del determinante queda multiplicado por este mismo número.*

Realicemos la demostración, por ejemplo, para el caso de las columnas. Si todos los elementos de la k-ésima columna del determinante

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nk} & \dots & a_{nn} \end{vmatrix}$$

se multiplican por c, se obtiene el determinante

$$D_1 = \begin{vmatrix} a_{11} & a_{12} & \dots & ca_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & ca_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & ca_{nk} & \dots & a_{nn} \end{vmatrix}$$

igual a

$$\sum (-1)^{[i_1, i_2, \dots, i_n]} a_{i_1 1} a_{i_2 2} \dots (ca_{i_k k}) \dots a_{i_n n} =$$
$$= c \sum (-1)^{[i_1, i_2, \dots, i_n]} a_{i_1 1} a_{i_2 2} \dots a_{i_k k} \dots a_{i_n n} = cD.$$

La propiedad correspondiente a las filas se demuestra fácilmente pasando a los determinantes traspuestos.

Por consiguiente, *un factor común de todos los elementos de una fila o de una columna de un determinante puede ser separado como el factor del determinante.*

Corolario. *Un determinante con dos filas o dos columnas proporcionales es igual a cero.*

En efecto, separando el «factor de proporcionalidad» de la fila o de la columna como factor del determinante, obtenemos un determinante con dos filas o dos columnas idénticas que es igual a cero debido al corolario de la propiedad 2.

Propiedad 4. *Si todo elemento de la k-ésima columna de un determinante viene dado como la suma de dos sumandos:* $a_{ik} = b_{ik} + c_{ik}$, *es decir, si*

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & b_{1k}+c_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_{2k}+c_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & b_{nk}+c_{nk} & \dots & a_{nn} \end{vmatrix},$$

se puede representar D como la suma de dos determinantes en la forma siguiente:

$$D=\begin{vmatrix} a_{11} & a_{12} & \dots & b_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & b_{nk} & \dots & a_{nn} \end{vmatrix}+$$

$$+\begin{vmatrix} a_{11} & a_{12} & \dots & c_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & c_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & c_{nk} & \dots & a_{nn} \end{vmatrix}=D_1+D_2.$$

Una afirmación análoga es válida también para las filas.

La demostración resulta de la igualdad:

$$\begin{aligned} D &= \sum (-1)^{[i_1, i_2, \dots, i_n]} a_{i_1 1} a_{i_2 2} \dots (b_{i_k k}+c_{i_k k}) \dots a_{i_n n} = \\ &= \sum (-1)^{[i_1, i_2, \dots, i_n]} a_{i_1 1} a_{i_2 2} \dots b_{i_k k} \dots a_{i_n n} + \\ &+ \sum (-1)^{[i_1, i_2, \dots, i_n]} a_{i_1 1} a_{i_2 2} \dots c_{i_k k} \dots a_{i_n n} = D_1+D_2. \end{aligned}$$

Observación. Es fácil ver que tiene lugar incluso la siguiente afirmación más general: *Si todo elemento de la k-ésima columna de un determinante D viene dado como la suma de p sumandos:* $a_{ik}=a_{ik}^1+a_{ik}^2+\dots+a_{ik}^p$, *el determinante D puede ser representado como la suma de p determinantes:*

$$D=\sum_{j=1,2,\dots,p}\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k}^j & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k}^j & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nk}^j & \dots & a_{nn} \end{vmatrix}.$$

Corolario. *El valor de un determinante no varía si a todos los elementos de una de sus filas o una de sus columnas cualquiera se suman los elementos correspondientes de la línea paralela multiplicados por un mismo número.*

Efectivamente, sea dado el determinante

$$D=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1p} & \dots & a_{1q} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2p} & \dots & a_{2q} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{np} & \dots & a_{nq} & \dots & a_{nn} \end{vmatrix}.$$

Sumando a los elementos de su p-ésima columna los elementos correspondientes de la q-ésima columna multiplicados por un mismo número c, obtenemos el determinante

$$D_1 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1p}+ca_{1q} & \dots & a_{1q} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2p}+ca_{2q} & \dots & a_{2q} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{np}+ca_{nq} & \dots & a_{nq} & \dots & a_{nn} \end{vmatrix}.$$

En virtud de la propiedad 4, el determinante D_1 es igual a

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1p} & \dots & a_{1q} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2p} & \dots & a_{2q} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{np} & \dots & a_{nq} & \dots & a_{nn} \end{vmatrix} +$$

$$+ \begin{vmatrix} a_{11} & a_{12} & \dots & ca_{1q} & \dots & a_{1q} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & ca_{2q} & \dots & a_{2q} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & ca_{nq} & \dots & a_{nq} & \dots & a_{nn} \end{vmatrix} = D$$

(el segundo sumando es igual a cero, ya que representa un determinante con dos columnas proporcionales).

§ 4. Menores y complementos algebraicos

Se llama *menor* M_{ik} del elemento a_{ik} de un determinante D de orden n al determinante de orden $n-1$ que se obtiene suprimiendo en D la i-ésima fila y la k-ésima columna.

Se llama *complemento algebraico* A_{ik} del elemento a_{ik} a su menor tomado con el signo $(-1)^{i+k}$:

$$A_{ik} = (-1)^{i+k} M_{ik}.$$

Teorema 2. *Si en un determinante D de orden n todos los elementos de la k-ésima columna (fila), a excepción de uno, son iguales a cero, el determinante es igual al producto de este elemento diferente de cero por su complemento algebraico.*

Demostración. Consideremos primero el caso en que todos los elementos de la *primera* columna del determinante D, a excepción de a_{11}, son iguales a cero:

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Todo término del determinante D contiene exactamente un elemento de la primera columna; pero como todos estos elementos, a excepción de a_{11}, son iguales a cero, resulta que se anulan todos aquellos términos del determinante D en los cuales figura no el elemento a_{11} sino otro elemento cualquiera de la primera columna. Por consiguiente,

$$D = \sum (-1)^{[1, i_2, \dots, i_n]} a_{11} a_{i_2 2} \dots a_{i_n n},$$

donde $i_2, \dots, i_n$ toman los valores $2, 3, \dots, n$. El elemento a_{11} puede ser separado como factor y puesto que la unidad, que aparece en la posición primera, no forma inversión alguna, se tiene $[1, i_2, \dots, i_n] = [i_2, \dots, i_n]$; luego,

$$D = a_{11} \sum (-1)^{[i_2, \dots, i_n]} a_{i_2 2} \dots a_{i_n n},$$

donde la suma se extiende a todas las permutaciones posibles $i_2, i_3, \dots, i_n$ de los números $2, 3, \dots, n$. Puesto que la suma

$$\sum (-1)^{[i_2, \dots, i_n]} a_{i_2 2} \dots a_{i_n n}$$

es igual al determinante de orden $n-1$ que se obtiene de D suprimiendo la primera fila y la primera columna, es decir, es igual a M_{11}, y puesto que $A_{11} = (-1)^{1+1} M_{11} = M_{11}$, resulta que

$$D = a_{11} M_{11} = a_{11} A_{11}.$$

Consideremos ahora el caso general en que todos los elementos de la k-ésima columna del determinante D, a excepción de a_{ik}, son iguales a cero, es decir, el caso en que el determinante es de la forma

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & 0 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & 0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ik} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & 0 & \dots & a_{nn} \end{vmatrix};$$

Traslademos la i-ésima fila del determinante D a la primera posición trasponiéndola sucesivamente con la fila $(i-1)$-ésima, con la fila $(i-2)$-ésima, etc. y, finalmente, con la primera fila. Esto requiere $i-1$ trasposiciones de filas y en cada una de estas trasposiciones el signo del determinante varía. Traslademos después la k-ésima columna del determinante D a la primera posición trasponiéndola sucesivamente con la columna $(k-1)$-ésima, con la columna $(k-2)$-ésima, etc. y, finalmente, con la primera columna. Esto requiere $k-1$ trasposiciones de columnas y en cada una de estas trasposiciones el signo del determinante también varía. En definitiva, obtendremos el determinante

$$D_1 = \begin{vmatrix} a_{ik} & a_{i1} & \dots & a_{in} \\ 0 & a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n1} & \dots & a_{nn} \end{vmatrix},$$

que difiere del determinante D en el signo $(-1)^{i+k-2}$. Pero, como acabamos de ver, el determinante D_1 es igual al producto de a_{ik} por el determinante de orden $n-1$ que se obtiene de D_1 suprimiendo la primera columna y la primera fila o, que es lo mismo, por el determinante que se obtiene de D suprimiendo la k-ésima columna y la i-ésima fila, es decir,

$$D_1 = a_{ik}M_{ik}$$

y, por consiguiente,

$$D = (-1)^{i+k-2}D_1 = (-1)^{i+k}a_{ik}M_{ik} = a_{ik}A_{ik}.$$

El teorema demostrado permite, empleando además el corolario de la propiedad 4, calcular los determinantes de cualquier orden.

Ejemplo 4. *Calcúlese el determinante de quinto orden*

$$D = \begin{vmatrix} 2 & 0 & 1 & 3 & 1 \\ -1 & 1 & 2 & 2 & 3 \\ 1 & 4 & 0 & -1 & 5 \\ 2 & 1 & 3 & 1 & 2 \\ 1 & 2 & -1 & 3 & 1 \end{vmatrix}.$$

Solución. Restando de la primera columna del determinante D la tercera duplicada, de la cuarta la tercera triplicada y de la quinta la tercera, obtenemos

$$D = \begin{vmatrix} 0 & 0 & 1 & 0 & 0 \\ -5 & 1 & 2 & -4 & 1 \\ 1 & 4 & 0 & -1 & 5 \\ -4 & 1 & 3 & -8 & -1 \\ 3 & 2 & -1 & 6 & 2 \end{vmatrix} = a_{13}A_{13} = 1\cdot(-1)^{1+3} \begin{vmatrix} -5 & 1 & -4 & 1 \\ 1 & 4 & -1 & 5 \\ -4 & 1 & -8 & -1 \\ 3 & 2 & 6 & 2 \end{vmatrix}.$$

En el determinante obtenido de orden cuatro procederemos de la misma forma continuando haciendo ceros: agreguemos a la primera columna la cuarta multiplicada por 5, restemos de la segunda columna la cuarta y agreguemos a la tercera columna la cuarta multiplicada por 4:

$$D = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 26 & -1 & 19 & 5 \\ -9 & 2 & -12 & -1 \\ 13 & 0 & 14 & 2 \end{vmatrix} = a_{14}A_{14} = 1\cdot(-1)^{1+4} \begin{vmatrix} 26 & -1 & 19 \\ -9 & 2 & -12 \\ 13 & 0 & 14 \end{vmatrix}.$$

Hemos llegado a un determinante de tercer orden que puede ser calculado o bien directamente o bien reduciéndolo a un determinante de segundo orden: agregando a la segunda fila la primera duplicada, obtenemos

$$\begin{vmatrix} 26 & -1 & 19 \\ 43 & 0 & 26 \\ 13 & 0 & 14 \end{vmatrix} = a_{12}A_{12} = (-1)(-1)^{1+2} \begin{vmatrix} 43 & 26 \\ 13 & 14 \end{vmatrix} = 2 \begin{vmatrix} 43 & 13 \\ 13 & 7 \end{vmatrix} =$$
$$= 2(301-169) = 264;$$

por consiguiente, $D = -264$.

§ 5. Desarrollo de un determinante por los elementos de una fila o de una columna

Teorema 3. *Todo determinante es igual a la suma de los productos de los elementos de una de sus filas (columnas) cualquiera por los correspondientes complementos algebraicos.*

Demostraremos que para cualesquiera i, k se tiene

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}$$

(desarrollo por los elementos de la i-ésima fila) y

$$D = a_{1k}A_{1k} + a_{2k}A_{2k} + \dots + a_{nk}A_{nk}$$

(desarrollo por los elementos de la k-ésima columna).

Para realizar la **demostración** observemos, ante todo, que, dados dos determinantes que difieren uno del otro sólo en los elementos de una columna (fila), los complementos algebraicos de los elementos de esta columna (fila) son los mismos en ambos determinantes, ya que al calcular estos complementos se suprime precisamente la columna (fila) en la que difieren los determinantes.

Demostremos ahora que para el determinante D es válido, por ejemplo, el desarrollo por los elementos de la k-ésima columna. Con este fin representemos el determinante en la forma siguiente:

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k}+0+\dots+0 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & 0+a_{2k}+\dots+0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & 0+0+\dots+a_{nk} & \dots & a_{nn} \end{vmatrix}$$

(aquí todo elemento de la k-ésima columna viene dado como la suma de n sumandos y $n-1$ de estos sumandos son iguales a cero). Según la propiedad 4 (véase la observación de la pág. 30), se tiene

$$D = D_1+D_2+\dots+D_n,$$

donde

$$D_1 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & 0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & 0 & \dots & a_{nn} \end{vmatrix},$$

$$D_2 = \begin{vmatrix} a_{11} & a_{12} & \dots & 0 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & 0 & \dots & a_{nn} \end{vmatrix}, \dots$$

$$\dots, D_n = \begin{vmatrix} a_{11} & a_{12} & \dots & 0 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & 0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nk} & \dots & a_{nn} \end{vmatrix}.$$

El determinante D_1 es igual al producto del elemento a_{1k} por su complemento algebraico *en este determinante*. Sin embargo, puesto que el determinante D_1 difiere del determinante D sólo en la

k-ésima columna, este complemento algebraico coincide con el complemento algebraico A_{1k} del elemento a_{1k} *en el determinante* D:

$$D_1 = a_{1k}A_{1k}.$$

Análogamente,

$$D_2 = a_{2k}A_{2k}, \ldots, D_n = a_{nk}A_{nk}.$$

Hemos demostrado que

$$D = a_{1k}A_{1k} + a_{2k}A_{2k} + \ldots + a_{nk}A_{nk}.$$

Este teorema puede ser también empleado para el cálculo de los determinantes ya que permite reducirlos a determinantes de órdenes inferiores.

Ejemplo 5. *Calcúlese el determinante de orden cuatro*

$$D = \begin{vmatrix} -5 & 1 & -4 & 1 \\ 1 & 4 & -1 & 5 \\ -4 & 1 & -8 & -1 \\ 3 & 2 & 6 & 2 \end{vmatrix}.$$

Solución. Desarrollemos el determinante, por ejemplo, por los elementos de la primera fila:

$$D = (-5)(-1)^{1+1}\begin{vmatrix} 4 & -1 & 5 \\ 1 & -8 & -1 \\ 2 & 6 & 2 \end{vmatrix} +$$

$$+1(-1)^{1+2}\begin{vmatrix} 1 & -1 & 5 \\ -4 & -8 & -1 \\ 3 & 6 & 2 \end{vmatrix} + (-4)(-1)^{1+3}\begin{vmatrix} 1 & 4 & 5 \\ -4 & 1 & -1 \\ 3 & 2 & 2 \end{vmatrix} +$$

$$+1(-1)^{1+4}\begin{vmatrix} 1 & 4 & -1 \\ -4 & 1 & -8 \\ 3 & 2 & 6 \end{vmatrix} = -5\cdot 74 - (-15) - 4(-31) - 33 =$$

$$= -370 + 15 + 124 - 33 = -264.$$

Teorema 4. *La suma de los productos de los elementos de cualquier fila (o de cualquier columna) de un determinante por los complementos algebraicos de los elementos correspondientes de una línea paralela es igual a cero.*

Demostración. Sea dado un determinante

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{nk} & \dots & a_{nn} \end{vmatrix}.$$

Consideremos otro determinante D_1 que difiere del determinante D sólo en que su k-ésima columna repite la i-ésima

$$D_1 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{ni} & \dots & a_{nn} \end{vmatrix}.$$

El determinante D_1 es igual a cero (según el corolario de la propiedad 2). Desarrollémoslo por los elementos de la k-ésima columna; tendremos

$$D_1 = a_{1i}A_{1k} + a_{2i}A_{2k} + \dots + a_{ni}A_{nk},$$

donde A_{jk} son los complementos algebraicos de los elementos de la k-ésima columna del determinante D_1 que al mismo tiempo son los complementos algebraicos de los elementos de la k-ésima columna del determinante D, ya que el determinante D_1 difiere del determinante D sólo en la k-ésima columna. Luego, para todo i y todo $k \neq i$ se tiene

$$a_{1i}A_{1k} + a_{2i}A_{2k} + \dots + a_{ni}A_{nk} = 0.$$

Análogamente, para todo i y todo $k \neq i$ se tiene

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \dots + a_{in}A_{kn} = 0.$$

§ 6. Sistemas de *n* ecuaciones lineales con *n* incógnitas

Consideremos un *sistema de n ecuaciones lineales con n incógnitas*

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{array} \right. \qquad (13)$$

Se llama *solución* del sistema (13) todo conjunto de valores de las incógnitas $x_1 = \alpha_1, x_2 = \alpha_2, \ldots, x_n = \alpha_n$ que al ser introducido en las ecuaciones convierte todas éstas en identidades. Supongamos que el determinante formado por los coeficientes de las incógnitas del sistema (13) es diferente de cero:

$$D = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \neq 0.$$

Multipliquemos por A_{11} la primera ecuación del sistema, por A_{21} la segunda, etc., por A_{n1} la última y sumemos todas las ecuaciones. Obtendremos la ecuación

$$\begin{aligned} &x_1(a_{11}A_{11}+a_{21}A_{21}+\ldots+a_{n1}A_{n1})+ \\ &\quad +x_2(a_{12}A_{11}+a_{22}A_{21}+\ldots+a_{n2}A_{n1})+\ldots \\ &\quad\quad \ldots+x_n(a_{1n}A_{11}+a_{2n}A_{21}+\ldots+a_{nn}A_{n1}) = \\ &\quad\quad\quad = b_1A_{11}+b_2A_{21}+\ldots+b_nA_{n1}, \end{aligned} \tag{14}$$

es decir,

$$x_1D = b_1A_{11}+b_2A_{21}+\ldots+b_nA_{n1}, \tag{15}$$

ya que los coeficientes de las incógnitas $x_2, x_3, \ldots, x_n$, comprendidos en los paréntesis de la ecuación (14), son iguales a cero en virtud del teorema 4 y el coeficiente de x_1 es igual a D debido al teorema 3. Además, para el segundo miembro se tiene

$$b_1A_{11}+b_2A_{21}+\ldots+b_nA_{n1} = D_1,$$

donde D_1 es el determinante que se obtiene de D sustituyendo su primera columna por la columna de los términos independientes. (En el segundo miembro de las ignaldades (14) y (15) aparece el desarrollo del determinante D_1 por la primera columna.) Análogamente a la ecuación (15) obtenemos

$$x_2D = D_2, \ldots, x_nD = D_n, \tag{15a}$$

donde D_i es el determinante que se obtiene de D sustituyendo la i-ésima columna por la columna de los términos independientes.

El sistema formado por las ecuaciones (15) y (15a) es un corolario del sistema (13). Hemos demostrado, por consiguiente, que el sistema (13) tiene solución, que ésta será también solución del sistema (15) y (15a) y que

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \ \ldots, \ x_n = \frac{D_n}{D}. \tag{16}$$

Las fórmulas (16) se llaman *fórmulas de Cramer*.

Realizando la sustitución directa de estos valores de las incógnitas en todas las ecuaciones del sistema (13), se puede comprobar que forman efectivamente una solución del mismo.

Teorema 5. *En el caso en que* $D \neq 0$, *el sistema* (13) *tiene una solución única determinada por las fórmulas de Cramer.*

§ 7. Rango de una matriz

Consideremos de nuevo las tablas de números (matrices) pero sin exigir ahora que el número de filas de la matriz sea igual al número de sus columnas. Introduciremos para estas *matrices (rectangulares)* el importante concepto de **rango.**

Consideremos una matriz rectangular formada por m filas y n columnas (una $[m \times n]$-matriz). Sea $k \leqslant m$ y $k \leqslant n$. Fijemos en esta matriz cualesquiera k filas y k columnas. Con los elementos que figuran en los cruces de las filas y de las columnas fijadas se puede formar un determinante de orden k. Todos los determinantes de este tipo se llaman *menores* de esta matriz. Es obvio que de una $[m \times n]$-matriz se puede formar $C_m^k \cdot C_n^k$ menores de orden k. Por ejemplo, dada la matriz

$$A = \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ 0 & 4 & 5 & 1 \end{bmatrix}$$

se puede formar $C_4^1 \cdot C_3^1 = 12$ menores de orden uno que vienen a ser los propios elementos de la matriz A, se puede formar $C_4^2 \cdot C_3^2 = 6 \cdot 3 = 18$ menores de segundo orden:

$$\begin{vmatrix} 3 & 2 \\ 2 & 0 \end{vmatrix}, \ \begin{vmatrix} 3 & 1 \\ 2 & -1 \end{vmatrix}, \ \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix}, \ \begin{vmatrix} 2 & 1 \\ 0 & -1 \end{vmatrix}, \ \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix}, \ \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix},$$

$$\begin{vmatrix} 3 & 2 \\ 0 & 4 \end{vmatrix}, \ \begin{vmatrix} 3 & 1 \\ 0 & 5 \end{vmatrix}, \ \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix}, \ \begin{vmatrix} 2 & 1 \\ 4 & 5 \end{vmatrix}, \ \begin{vmatrix} 2 & 2 \\ 4 & 1 \end{vmatrix}, \ \begin{vmatrix} 1 & 2 \\ 5 & 1 \end{vmatrix},$$

$$\begin{vmatrix} 2 & 0 \\ 0 & 4 \end{vmatrix}, \ \begin{vmatrix} 2 & -1 \\ 0 & 5 \end{vmatrix}, \ \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix}, \ \begin{vmatrix} 0 & -1 \\ 4 & 5 \end{vmatrix}, \ \begin{vmatrix} 0 & 1 \\ 4 & 1 \end{vmatrix}, \ \begin{vmatrix} -1 & 1 \\ 5 & 1 \end{vmatrix},$$

y se puede formar $C_4^3 \cdot C_3^3 = 4$ menores de tercer orden:

$$\begin{vmatrix} 3 & 2 & 1 \\ 2 & 0 & -1 \\ 0 & 4 & 5 \end{vmatrix}, \quad \begin{vmatrix} 3 & 2 & 2 \\ 2 & 0 & 1 \\ 0 & 4 & 1 \end{vmatrix}, \quad \begin{vmatrix} 3 & 1 & 2 \\ 2 & -1 & 1 \\ 0 & 5 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2 & 1 & 2 \\ 0 & -1 & 1 \\ 4 & 5 & 1 \end{vmatrix}.$$

Es fácil comprobar que todos los menores de tercer orden de la matriz A son iguales a cero y que no todos los menores de segundo orden son iguales a cero (el primero de los menores escritos de orden dos ya es diferente de cero). En este caso diremos que **el rango de la matriz A es igual a 2.**

Se llama ***rango*** de una matriz *el orden máximo de sus menores diferentes de cero.*

Por consiguiente, si el rango de una matriz es igual a r, entre los menores de esta matriz existe al menos un menor de orden r diferente de cero mientras que todos sus menores de orden $r+1$ o de orden superior son iguales a cero. Indicaremos por $r(A)$ el rango de la matriz A.

Se llaman *transformaciones elementales* de una matriz las siguientes transformaciones de la misma:

1. La trasposición, es decir, la sustitución de toda fila por la columna del mismo número y viceversa.
2. El cambio entre sí de dos filas o de dos columnas.
3. La multiplicación de todos los elementos de una fila o de una columna por cualquier número c diferente de cero.
4. La adición a todos los elementos de una fila o de una columna de los elementos correspondientes de una línea paralela multiplicados por un mismo número.

Teorema 6 (sobre las transformaciones elementales). *El rango de una matriz no varía en las transformaciones elementales.*

Demostración. Consideremos por separado cada una de las transformaciones. En los tres primeros casos nuestra afirmación es casi evidente:

1. Según la propiedad 1 de los determinantes, todo menor de la matriz traspuesta es igual a uno de los menores de la matriz dada y viceversa.
2. Cambiando entre sí dos filas o dos columnas de la matriz A, obtenemos una matriz nueva en la que todo menor es o bien igual a cierto menor de la matriz A o bien difiere sólo en el signo de un menor de la matriz A.

3. Al multiplicar todos los elementos de una fila o de una columna de una matriz por un número c, una parte de sus menores no varía, mientras que los demás se multiplican por c; pero como $c \neq 0$, el orden máximo de los menores diferentes de cero de esta matriz continuará siendo el mismo.

4. Consideremos la matriz B que se obtiene de la matriz A agregando a todos los elementos de su i-ésima columna los elementos correspondientes de la k-ésima columna, multiplicados por c:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1i} & \ldots & a_{1k} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2i} & \ldots & a_{2k} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \ldots & a_{mi} & \ldots & a_{mk} & \ldots & a_{mn} \end{bmatrix},$$

$$B = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1i}+ca_{1k} & \ldots & a_{1k} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2i}+ca_{2k} & \ldots & a_{2k} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \ldots & a_{mi}+ca_{mk} & \ldots & a_{mk} & \ldots & a_{mn} \end{bmatrix}.$$

Supongamos que el rango $r(A)$ de la matriz A es igual a r. Demostremos que el rango de la matriz B es *no mayor* que r. Para ello es suficiente probar que todo menor de orden mayor que r de la matriz B es igual a 0. Sea D un menor de orden mayor que r de la matriz B. Si D no contiene elementos de la i-ésima columna, es igual exactamente al menor correspondiente de la matriz A, es decir, es igual a 0 por ser un menor de orden mayor que r formado a partir de una matriz de rango r.

Si D contiene elementos de la i-ésima y de la k-ésima columnas de la matriz B, también coincide, en virtud de la propiedad 4, con el menor correspondiente de la matriz A y, por consiguiente, es igual a 0.

Finalmente, si el determinante D contiene elementos de la i-ésima columna de la matriz B, pero no contiene elementos de su k-ésima columna, resulta, en virtud de la propiedad 4, que este determinante puede ser representado como la suma de dos determinantes: $D = D_1 + D_2$, con la particularidad de que uno de ellos es igual al correspondiente menor de la matriz A y el otro difiere del correspondiente menor de la matriz A en el factor $\pm c$. (El signo menos aparece aquí debido a que la columna con los elementos

a_{ik} puede «no estar en su sitio».) Así, por ejemplo,

$$D = \begin{vmatrix} a_{21} & a_{22}+ca_{24} & a_{23} \\ a_{41} & a_{42}+ca_{44} & a_{43} \\ a_{51} & a_{52}+ca_{54} & a_{53} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{vmatrix} + \begin{vmatrix} a_{21} & ca_{24} & a_{23} \\ a_{41} & ca_{44} & a_{43} \\ a_{51} & ca_{54} & a_{53} \end{vmatrix}.$$

Por consiguiente, ambos determinantes D_1 y D_2 son iguales a 0 y

$$D = 0.$$

Luego, todo menor de orden mayor que r de la matriz B es igual a cero y, por lo tanto,

$$r(B) \leqslant r(A).$$

Pero la matriz A se obtiene, a su vez, de la matriz B mediante una transformación elemental del cuarto tipo: para obtener la matriz A es necesario agregar a la i-ésima columna de la matriz B su k-ésima columna multiplicada por $-c$. Según hemos demostrado, el rango de la matriz en este caso no aumenta, es decir,

$$r(A) \leqslant r(B).$$

Por consiguiente,

$$r(A) = r(B).$$

Es fácil comprobar que *toda matriz puede ser reducida, mediante transformaciones elementales, a la forma*

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix},$$

donde en la «diagonal principal» aparecen r unidades, mientras que todos los demás elementos de la matriz son iguales a cero. Está claro que el rango de esta matriz y, por consiguiente, el rango de la matriz inicial son iguales a r.

Ejemplo 6. *Calcúlese, empleando las transformaciones elementales, el rango de la matriz*

$$A = \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ 0 & 4 & 5 & 1 \end{bmatrix}$$

(considerada al principio del parágrafo).

Solución. Restando de la tercera fila la primera duplicada, dividiendo la segunda columna por 2 y restando después de la primera columna la segunda triplicada, de la tercera columna la segunda y de la cuarta columna la segunda duplicada, obtenemos consecutivamente:

$$A \sim \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 1 & 2 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{bmatrix} \sim$$

$$\sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 1 \\ -6 & 0 & 3 & -3 \end{bmatrix},$$

donde el signo $\sim$ significa que las matrices que él une se obtienen una de la otra mediante transformaciones elementales y tienen, por consiguiente, el mismo rango.

Agregando ahora a la tercera fila la segunda triplicada, dividiendo la primera columna por 2, agregándola a la tercera columna, restándola de la cuarta columna y, finalmente, cambiando entre sí las dos primeras columnas, obtendremos

$$A \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Encontramos de nuevo que el rango de la matriz A es igual a 2.

§ 8. Noción de dependencia lineal

Si indicamos por

$$e_1 = (3, 2, 1, 2), \quad e_2 = (2, 0, -1, 1), \quad e_3 = (0, 4, 5, 1)$$

las filas de la matriz A (véase el § 7), es evidente que tiene lugar la igualdad

$$e_3 = 2e_1 - 3e_2,$$

comprendida en el sentido de *adición término por término:* todo elemento de la fila e_3 es igual al correspondiente elemento de la fila e_1 multiplicado por 2 menos el correspondiente elemento de la fila e_2 multiplicado por 3:

En general, siendo $e_1, e_2, \ldots, e_m$ unas filas de una matriz A y siendo, por ejemplo,

$$e_m = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_{m-1} e_{m-1}, \tag{17}$$

donde $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}$ son unos números cualesquiera, diremos que la m-ésima fila de esta matriz *se expresa linealmente* en términos de sus primeras $m-1$ filas o que la fila e_m es *una combinación lineal* de las filas $e_1, e_2, \ldots, e_{m-1}$. De la igualdad (17) se deduce que

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_{m-1} e_{m-1} - e_m = 0,$$

donde el cero del segundo miembro es comprendido como *la fila nula* (es decir, como la fila formada por n ceros).

Diremos que unas filas $e_1, e_2, \ldots, e_m$ de una matriz A son *linealmente dependientes*, si se pueden encontrar unos números $\gamma_1, \gamma_2, \ldots, \gamma_m$, no todos iguales a cero, tales que

$$\gamma_1 e_1 + \gamma_2 e_2 + \ldots + \gamma_m e_m = 0. \tag{18}$$

Si tales números γ_i no existen, es decir, si la igualdad (18) tiene lugar sólo en el caso en que todos los $\gamma_i = 0$, se dice que las filas $e_1, e_2, \ldots, e_m$ son *linealmente independientes*.

Está claro que si una fila de una matriz se expresa linealmente en términos de las demás, las filas de esta matriz son linealmente dependientes. Recíprocamente, supongamos que entre las filas de la matriz A existe una relación de dependencia lineal (18). Puesto que al menos uno de los números γ_i, por ejemplo, γ_m, es diferente de cero, tenemos

$$e_m = -\frac{\gamma_1}{\gamma_m} e_1 - \frac{\gamma_2}{\gamma_m} e_2 - \ldots - \frac{\gamma_{m-1}}{\gamma_m} e_{m-1},$$

es decir, al menos una de las filas de la matriz se expresa en este caso linealmente en términos de las restantes.

Se puede también introducir un concepto análogo de dependencia lineal para las columnas de una matriz.

Teorema 7 (sobre el rango de una matriz). *Siendo r el rango de una matriz, se puede encontrar en esta matriz r filas (columnas) linealmente independientes en términos de las cuales se expresan linealmente todas las demás filas (columnas).*

Demostración. Sea dada una $[m\times n]$-matriz de rango r. Supongamos, para concretar, que el menor de orden r diferente de cero de esta matriz aparece en el ángulo *superior de la izquierda*, es decir, supongamos que

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad D=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{vmatrix}\neq 0.$$

Demostremos que en estas condiciones *las primeras r filas de esta matriz serán linealmente independientes.* (Si es diferente de cero no éste sino otro menor de orden r de la matriz A, serán linealmente independientes precisamente las filas que forman este menor.) Supongamos, por el contrario, que estas filas son linealmente dependientes; entonces, una de ellas —digamos la e_r para concretar— se expresa linealmente en términos de las restantes:

$$e_r=\alpha_1e_1+\alpha_2e_2+\dots+\alpha_{r-1}e_{r-1}.$$

Restemos de la r-ésima fila de la matriz A la primera fila multiplicada por α_1, la segunda multiplicada por α_2, etc., y, finalmente, la $(r-1)$-ésima multiplicada por α_{r-1}. Después de estas transformaciones la r-ésima fila de la matriz A quedará compuesta solamente de ceros. Además, el determinante D, que en virtud del corolario de la propiedad 4 no debe variar, será igual a cero. La contradicción obtenida demuestra la independencia lineal de las r primeras filas de la matriz A.

Demostremos ahora la segunda parte del teorema, es decir, que *todas las demás filas de la matriz A se expresan linealmente en términos de sus r primeras filas*. Sea $r<k\leqslant m$ y $1\leqslant l\leqslant n$; consideremos el determinante

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2l} \\ \dots & \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{rl} \\ a_{k1} & a_{k2} & \dots & a_{kr} & a_{kl} \end{vmatrix}$$

de orden $r+1$ que es igual a cero cualesquiera que sean k y l: si $l\leqslant r$, por tener dos columnas idénticas, y si $l>r$, por ser un menor de orden $r+1$ de una matriz de rango r.

Desarrollemos el determinante Δ por los elementos de la última columna:

$$\Delta = a_{1l}A_1 + a_{2l}A_2 + \ldots + a_{rl}A_r + a_{kl}A_{r+1} = 0. \tag{19}$$

Los complementos algebraicos $A_1, A_2, \ldots, A_r, A_{r+1}$ de los elementos de la última columna dependen de k, pero no dependen de l, ya que al calcularlos suprimimos la última columna. Además, $A_{r+1} = D \neq 0$ y, por consiguiente, podemos dividir la igualdad (19) por A_{r+1}; así resulta

$$a_{kl} = \alpha_1 a_{1l} + \alpha_2 a_{2l} + \ldots + \alpha_r a_{rl},$$

donde los coeficientes $\alpha_i = -\frac{A_i}{D}$ no dependen de l. Tomando $l = 1, 2, \ldots, n$, tendremos

$$\begin{aligned} a_{k1} &= \alpha_1 a_{11} + \alpha_2 a_{21} + \ldots + \alpha_r a_{r1}, \\ a_{k2} &= \alpha_1 a_{12} + \alpha_2 a_{22} + \ldots + \alpha_r a_{r2}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{kn} &= \alpha_1 a_{1n} + \alpha_2 a_{2n} + \ldots + \alpha_r a_{rn}. \end{aligned}$$

Pero esto significa que la k-ésima columna de la matriz A se expresa linealmente en términos de sus r primeras filas:

$$e_k = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r.$$

Corolario 1. *El número máximo de columnas linealmente independientes de una matriz es igual al número máximo de filas linealmente independientes*, ya que al trasponer la matriz sus filas se convierten en columnas y el rango de la matriz no varía.

Corolario 2. *Para que un determinante sea igual a cero es necesario y suficiente que sus filas (columnas) sean linealmente dependientes.*

§ 9. Sistemas arbitrarios de ecuaciones lineales

Consideremos ahora *un sistema de m ecuaciones lineales con n incógnitas:*

$$\left\{ \begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ &\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \right. \tag{20}$$

donde *no se supone* que el número de las incógnitas es igual al número de las ecuaciones.

Se llama *solución* del sistema (20) todo conjunto de n valores de las incógnitas $x_1 = \alpha_1$, $x_2 = \alpha_2$, ..., $x_n = \alpha_n$ que al ser introducido en las ecuaciones convierte todas éstas en identidades. Un sistema que tiene al menos una solución se llama *compatible* y un sistema que no tiene solución alguna se llama *incompatible*. Un sistema que posee una solución única se llama *determinado* y un sistema que tiene más de una solución se llama *indeterminado*.

Consideremos dos matrices: *la matriz A formada por los coeficientes de las incógnitas del sistema* (20) y la matriz

$$B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix},$$

que se obtiene de A agregando la columna de los términos independientes y que se denomina *matriz ampliada*. Está claro que $r(B) \geqslant r(A)$, ya que todo menor de la matriz A es un menor de la matriz B pero no viceversa.

Teorema 8 (criterio de compatibilidad de un sistema de ecuaciones lineales). *Para que el sistema* (20) *sea compatible es necesario y suficiente que el rango de la matriz ampliada B sea igual al rango de la matriz de los coeficientes.*

Demostración de la necesidad. Supongamos que el sistema (20) es compatible, es decir, que existen unos números $x_1 = \alpha_1$, $x_2 = \alpha_2$, ..., $x_n = \alpha_n$ tales que

$$\begin{aligned} a_{11}\alpha_1 + a_{12}\alpha_2 + \dots + a_{1n}\alpha_n &= b_1, \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \dots + a_{2n}\alpha_n &= b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \dots + a_{mn}\alpha_n &= b_m. \end{aligned}$$

Restando de la última columna de la matriz B su primera columna multiplicada por α_1, su segunda columna multiplicada por α_2, etc., y, finalmente, su n-ésima columna multiplicada por α_n, obtenemos la matriz

$$B_1 = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 \end{bmatrix},$$

cuyo rango, en virtud del teorema sobre las transformaciones elementales, es igual al rango de la matriz B:

$$r(B_1) = r(B).$$

Por otra parte, está claro que $r(B_1) = r(A)$, porque todos los menores no nulos de la matriz B_1 son iguales a los correspondientes menores de la matriz A y viceversa. Por consiguiente, $r(B) = r(A)$.

Demostración de la suficiencia. Sea

$$r(B) = r(A) = r$$

y supongamos, para concretar, que el determinante diferente de cero de orden r de la matriz A aparece en el ángulo *superior de la izquierda* de la misma:

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

Las r primeras filas de la matriz B son entonces linealmente independientes y puesto que su rango es igual exactamente a r, resulta que las restantes filas de la matriz B se expresan linealmente en términos de sus r primeras filas. Pero esto significa que *las r primeras ecuaciones del sistema* (20) *son independientes y que las demás $m-r$ ecuaciones de este sistema* son «combinaciones lineales» de aquéllas, es decir, son simplemente *corolarios de ésas*. De hecho el sistema contiene en este caso solamente r ecuaciones independientes. Luego, es suficiente resolver las r primeras ecuaciones del sistema; sus soluciones satisfacen automáticamente a las restantes $m-r$ ecuaciones.

Ahora pueden darse dos casos.

1. $r = n$. El sistema formado por las r primeras ecuaciones del sistema (20):

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2r}x_r = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rr}x_r = b_r, \end{cases}$$

se puede resolver empleando, por ejemplo, las fórmulas de Cramer. En este caso el sistema posee una solución *única*. Luego, es *compatible* y *determinado*.

2. $r < n$. Tomemos las r primeras ecuaciones del sistema y dejando en el primer miembro las r primeras incógnitas, pasemos las restantes al segundo miembro:

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1r}x_r=b_1-a_{1,r+1}x_{r+1}-\ldots-a_{1n}x_n, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2r}x_r=b_2-a_{2,r+1}x_{r+1}-\ldots-a_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1+a_{r2}x_2+\ldots+a_{rr}x_r=b_r-a_{r,r+1}x_{r+1}-\ldots-a_{rn}x_n. \end{array}\right. \qquad (21)$$

Podemos dar valores *cualesquiera* a las «incógnitas libres» $x_{r+1}, x_{r+2}, \ldots, x_n$ obteniendo del sistema (21) valores correspondientes para las incógnitas $x_1, x_2, \ldots, x_r$. Este es el caso de un sistema *compatible*, pero *indeterminado*. Las fórmulas generales para la solución se pueden obtener resolviendo el sistema (21) respecto a $x_1, x_2, \ldots, x_r$, por medio, por ejemplo, de las fórmulas de Cramer.

Ahora estamos en condiciones de responder a la pregunta aun pendiente: ¿qué puede decirse acerca de un sistema de n ecuaciones con n incógnitas cuyo determinante es *igual* a *cero*? Para un sistema de este tipo el rango de la matriz de los coeficientes $r < n$, ya que el único menor de orden n de esta matriz es, por hipótesis, igual a cero. Si el rango de la matriz ampliada B de este sistema es también igual a r, el sistema es *compatible*, pero *indeterminado* ya que $r < n$; si el rango de la matriz B es mayor que r, el sistema es *incompatible*.

Ejemplo 7. *Resuélvanse los siguientes sistemas de ecuaciones:*

1. $\left\{\begin{array}{r} x_1+2x_2+3x_3=2, \\ x_1-x_2+x_3=0, \\ x_1+3x_2-x_3=-2, \\ 3x_1+4x_2+3x_3=0. \end{array}\right.$ 2. $\left\{\begin{array}{r} x_1+2x_2+3x_3-x_4=0, \\ x_1-x_2+x_3+2x_4=4, \\ x_1+5x_2+5x_3-4x_4=-4, \\ x_1+8x_2+7x_3-7x_4=-8. \end{array}\right.$

3. $\left\{\begin{array}{r} x_1+2x_2+3x_3-x_4=0, \\ x_1-x_2+x_3+2x_4=4, \\ x_1+5x_2+5x_3-4x_4=-4, \\ x_1+8x_2+7x_3-7x_4=6. \end{array}\right.$

Solución.

1. Tenemos $r(A)=3$ y $r(B)=3$; el sistema es *compatible* y *determidado*. Puesto que

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 1 & 3 & -1 \end{vmatrix} = 14 \neq 0,$$

encontramos de las tres ecuaciones primeras del sistema, empleando, por ejemplo, las fórmulas de Cramer,

$x_1 = -1, x_2 = 0$ y $x_3 = 1.$

2. Tenemos $r(A) = 2$ y $r(B) = 2$; el sistema es *compatible* pero *indeterminado*. Puesto que el determinante

$$\begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} = -3 \neq 0,$$

de las dos primeras ecuaciones del sistema

$$\begin{cases} x_1+2x_2 = -3x_3+\ x_4, \\ x_1-\ x_2 = 4-x_3-2x_4 \end{cases}$$

obtenemos

$$x_1 = \frac{8}{3} - \frac{5}{3}x_3 - x_4 \text{ y } x_2 = -\frac{4}{3} - \frac{2}{3}x_3 + x_4,$$

donde se pueden asignar valores *cualesquiera* a las incógnitas x_3 y x_4.

3. Tenemos $r(A) = 2$ y $r(B) = 3$ y el sistema es *incompatible*.

§ 10. Sistemas homogéneos

Las *ecuaciones lineales homogéneas* son las ecuaciones en las que el segundo miembro es igual a cero:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \tag{22}$$

Este sistema es siempre *compatible* ya que posee, por ejemplo, la solución **nula**:

$$x_1 = 0, x_2 = 0, \ldots, x_n = 0$$

(es decir, la solución en la que los valores de **todas** las incógnitas son iguales a cero).

Es importante saber bajo qué condiciones el sistema homogéneo (22) tiene soluciones no nulas. La respuesta a este problema viene dada por el teorema siguiente.

Teorema 9. *Para que el sistema* (22) *tenga soluciones no nulas es necesario y suficiente que el rango r de la matriz de los coeficientes sea menor que n.*

Efectivamente, siendo $r = n$, el sistema (22) tiene, como se deduce de la demostración del teorema 8, una solución **única** que es, por consiguiente, la **nula**

$$x_1 = 0, \quad x_2 = 0, \ \ldots, \ x_n = 0.$$

En cambio, siendo $r < n$, el sistema (22) resulta **indeterminado** (porque no puede ser incompatible) y, por consiguiente, tiene un **conjuto infinito** de soluciones y, en particular, un conjunto infinito de soluciones no nulas.

Del teorema demostrado se deduce directamente este otro.

Teorema 10. *Para que un sistema homogéneo de n ecuaciones lineales con n incógnitas posee soluciones no nulas es necesario y suficiente que su determinante D sea igual a cero.*

Demostración. La condición

$$D = 0$$

es aquí *necesaria*, ya que siendo $D \neq 0$ el sistema tendrá una solución única, es decir, la solución nula. Esta condición es también *suficiente*, ya que siendo $D = 0$ el rango de la matriz de los coeficientes del sistema es $r < n$ y el sistema tiene un conjunto infinito de soluciones (no nulas).

Ejemplo 8. *Hállese la condición de pertenencia de dos rectas del espacio*

$$\frac{x-a_1}{l_1} = \frac{y-b_1}{m_1} = \frac{z-c_1}{n_1}$$

y

$$\frac{x-a_2}{l_2} = \frac{y-b_2}{m_2} = \frac{z-c_2}{n_2}$$

a un mismo plano (en el sistema rectangular de coordenadas).

Solución. Supongamos que estas rectas pertenecen a un plano

$$Ax+By+Cz+D = 0. \tag{23}$$

El vector (A, B, C) es entonces ortogonal a este plano y, por consiguiente, es ortogonal al vector $(a_2-a_1, \ b_2-b_1, \ c_2-c_1)$ que pertenece al plano, es decir,

$$A(a_2-a_1)+B(b_2-b_1)+C(c_2-c_1) = 0. \tag{24a}$$

Además, ambas rectas son ortogonales al vector (A, B, C), es decir,

$$\begin{aligned} Al_1+Bm_1+Cn_1 &= 0, \\ Al_2+Bm_2+Cn_2 &= 0. \end{aligned} \tag{24b}$$

Hemos obtenido el sistema (24a) y (24b) de tres ecuaciones lineales con tres incógnitas A, B y C. Para que tenga soluciones no nulas, es decir, para que exista el plano (23) que contenga las dos rectas dadas, es necesario y suficiente que el determinante de este sistema sea igual a cero:

$$\begin{vmatrix} a_2-a_1 & b_2-b_1 & c_2-c_1 \\ l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \end{vmatrix} = 0.$$

Esta es precisamente la condición que queríamos obtener.

Sea

$$x_1 = \alpha_1,\ x_2 = \alpha_2,\ \ldots,\ x_n = \alpha_n$$

una solución no nula del sistema homogéneo (22). Podemos considerar esta solución como una fila $e_1 = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ formada por n elementos. En este caso la fila

$$ce_1 = (c\alpha_1, c\alpha_2, \ldots, c\alpha_n)$$

será, evidentemente, una solución del sistema (22). Además, siendo

$$e_2 = (\beta_1, \beta_2, \ldots, \beta_n)$$

otra solución cualquiera del sistema (22), la combinación lineal de estas soluciones

$$c_1e_1+c_2e_2 = (c_1\alpha_1+c_2\beta_1,\ c_1\alpha_2+c_2\beta_2,\ \ldots,\ c_1\alpha_n+c_2\beta_n)$$

también será una solución del sistema cualesquiera que sean los valores de c_1 y c_2, ya que de

$$a_{i1}\alpha_1+a_{i2}\alpha_2+\ldots+a_{in}\alpha_n = 0,$$
$$a_{i1}\beta_1+a_{i2}\beta_2+\ldots+a_{in}\beta_n = 0$$

resulta

$$a_{i1}(c_1\alpha_1+c_2\beta_1)+a_{i2}(c_1\alpha_2+c_2\beta_2)+\ldots$$
$$\ldots+a_{in}(c_1\alpha_n+c_2\beta_n) = 0.$$

Luego, *cualquier combinación lineal de unas soluciones del sistema homogéneo* (22) *será también una solución del mismo*. Es interesante por esto hallar unas soluciones linealmente independientes del sistema (22) en términos de las cuales se expresen linealmente todas las demás soluciones de éste.

Un sistema linealmente independiente $e_1, e_2, \ldots, e_k$ de soluciones de las ecuaciones (22) se llama *fundamental* si toda solución del sistema (22) es una combinación lineal de las soluciones $e_1, e_2, \ldots$ $\ldots, e_k$.

Teorema 11 (sobre la existencia de sistemas fundamentales de soluciones). *Si el rango r de la matriz de los coeficientes del sistema de ecuaciones* (22) *es menor que n, este sistema posee sistemas fundamentales de soluciones.*

Demostración. Supongamos que el rango r de la matriz de los coeficientes del sistema (22) es menor que n y aceptemos, para concretar, que el menor D que figura *en el ángulo superior de la izquierda* de la matriz A es diferente de cero:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

Pasando al segundo miembro de las r primeras ecuaciones del sistema (22) las incógnitas libres $x_{r+1}, \dots, x_n$, obtenemos el sistema

$$\left\{\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r &= -a_{1,\,r+1}x_{r+1} - \dots - a_{1n}x_n, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2r}x_r &= -a_{2,\,r+1}x_{r+1} - \dots - a_{2n}x_n, \\ \dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rr}x_r &= -a_{r,\,r+1}x_{r+1} - \dots - a_{rn}x_n. \end{aligned}\right. \tag{25}$$

Asignando a las incógnitas libres los valores

$$x_{r+1} = 1,\ x_{r+2} = 0,\ \dots,\ x_n = 0,$$

obtenemos los valores correspondientes $x_1 = \alpha_1$, $x_2 = \alpha_2$, $\dots$, $x_r = \alpha_r$ de las r primeras incógnitas. Esto nos ofrece una fila solución

$$(\alpha_1, \alpha_2, \dots, \alpha_r, 1, 0, \dots, 0).$$

Análogamente, asignando a las incógnitas libres los valores

$$x_{r+1} = 0,\ x_{r+2} = 1,\ \dots,\ x_n = 0,$$

y calculando los respectivos valores $x_1 = \beta_1, x_2 = \beta_2, \dots, x_r = \beta_r$ de las restantes incógnitas, obtenemos la fila

$$(\beta_1, \beta_2, \dots, \beta_r, 0, 1, \dots, 0),$$

etc. De esta forma obtendremos un total de $k = n - r$ soluciones del sistema (25):

$$\begin{aligned} e_1 &= (\alpha_1, \alpha_2, \dots, \alpha_r, 1, 0, \dots, 0), \\ e_2 &= (\beta_1, \beta_2, \dots, \beta_r, 0, 1, \dots, 0), \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ e_k &= (\xi_1, \xi_2, \dots, \xi_r, 0, 0, \dots, 1). \end{aligned} \tag{26}$$

Estas k filas son linealmente independientes, ya que forman una matriz

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_r & 1 & 0 & \dots & 0 \\ \beta_1 & \beta_2 & \dots & \beta_r & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \xi_1 & \xi_2 & \dots & \xi_r & 0 & 0 & \dots & 1 \end{bmatrix}$$

de rango igual exactamente a k. (Esta matriz contiene un menor de orden k —por ejemplo, el formado por las últimas k columnas— diferente de cero.)

Demostremos ahora que las soluciones $e_1, e_2, \dots, e_k$ de (26) constituyen efectivamente un sistema *fundamental*. Es suficiente demostrar para ello que *toda solución del sistema* (22) *se expresa linealmente en términos de* $e_1, e_2, \dots, e_k$. Sea, pues,

$$e = (\vartheta_1, \vartheta_2, \dots, \vartheta_r, \vartheta_{r+1}, \dots, \vartheta_n)$$

una solución cualquiera del sistema (22). Consideremos la fila

$$e_0 = e - \vartheta_{r+1}e_1 - \vartheta_{r+2}e_2 - \dots - \vartheta_n e_k.$$

Es fácil ver que todos los elementos que figuran en las k últimas posiciones de esta fila son iguales a cero, es decir,

$$e_0 = (\varrho_1, \varrho_2, \dots, \varrho_r, 0, 0, \dots, 0).$$

La fila e_0 es una solución del sistema (22), porque es una combinación lineal de soluciones. Pero como los valores de todas las incógnitas libres en e_0 son iguales a cero, del sistema (25) —cuyo determinante es diferente de cero y que, además, es en este caso homogéneo— obtenemos que los valores de las restantes incógnitas en e_0 son también iguales a cero, es decir, que e_0 es la fila nula:

$$e_0 = e - \vartheta_{r+1}e_1 - \vartheta_{r+2}e_2 - \dots - \vartheta_n e_k = (0, 0, \dots, 0),$$

de donde resulta que

$$e = \vartheta_{r+1}e_1 + \vartheta_{r+2}e_2 + \dots + \vartheta_n e_k$$

que es lo que se quería demostrar.

Observemos que para obtener un sistema fundamental de soluciones podemos asignar a las incógnitas libres otros valores cualesquiera siempre que el determinante correspondiente de orden k sea diferente de cero. De esta forma podemos obtener tantos sistemas fundamentales de soluciones (cada uno compuesto de $k = n-r$ filas) como se quiera. Se puede demostrar que todo sistema fundamental de soluciones de las ecuaciones (22) está formado por exactamente $n-r$ elementos (esto se deducirá de los resultados del capítulo siguiente).

Luego, podemos decir que la *solución general* del sistema (22) de ecuaciones lineales homogéneas es de la forma

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_k e_k,$$

donde $e_1, e_2, \ldots, e_k$ es un sistema fundamental cualquiera de soluciones y $\alpha_1, \alpha_2, \ldots, \alpha_k$ son unos números arbitrarios.

Hagamos una observación más de importancia para lo sucesivo.

Observación. Consideremos un sistema de ecuaciones

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m \end{array}\right. \tag{27}$$

y su correspondiente sistema de ecuaciones homogéneas

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=0. \end{array}\right. \tag{28}$$

Sea $e_1 = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ una solución determinada del sistema (27) y sea $e_2 = (\beta_1, \beta_2, \ldots, \beta_n)$ otra solución cualquiera del mismo. La diferencia

$$e_1-e_2 = (\alpha_1-\beta_1, \alpha_2-\beta_2, \ldots, \alpha_n-\beta_n)$$

es entonces una solución del sistema (28): de

$$a_{i1}\alpha_1+a_{i2}\alpha_2+\ldots+a_{in}\alpha_n = b_i$$

y de

$$a_{i1}\beta_1+a_{i2}\beta_2+\ldots+a_{in}\beta_n = b_i$$

resulta

$$a_{i1}(\alpha_1-\beta_1)+a_{i2}(\alpha_2-\beta_2)+\ldots+a_{in}(\alpha_n-\beta_n) = b_i-b_i = 0.$$

Recíprocamente, siendo $e_3 = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ una solución cualquiera del sistema homogéneo (28), la fila $e_1+e_3 = (\alpha_1+\gamma_1, \alpha_2+\gamma_2, \ldots, \alpha_n+\gamma_n)$ satisface al sistema (27): de

$$a_{i1}\alpha_1+a_{i2}\alpha_2+\ldots+a_{in}\alpha_n = b_i$$

y de

$$a_{i1}\gamma_1+a_{i2}\gamma_2+\ldots+a_{in}\gamma_n = 0$$

resulta

$$a_{i1}(\alpha_1+\gamma_1)+a_{i2}(\alpha_2+\gamma_2)+\ldots+a_{in}(\alpha_n+\gamma_n) = b_i+0 = b_i.$$

De aquí se deduce que todas las soluciones del sistema (27) se pueden obtener agregando a una solución cualquiera del mismo todas las soluciones posibles del sistema homogéneo (28). En otras palabras, *la solución general del sistema* (27) *es igual a la suma de la solución general del sistema* (28) *y de una solución cualquiera, pero fija, del sistema* (27).

§ 11. Método de Gauss

Las fórmulas de Cramer, de gran interés teórico, no tienen, sin embargo, importancia seria en la práctica ya que su aplicación está sujeta a cálculos demasiado voluminosos. Para la solución práctica de los sistemas de ecuaciones lineales se emplea con mayor frecuencia el método de Gauss que consiste en la eliminación sucesiva de las incógnitas según el esquema siguiente. Para resolver el sistema de ecuaciones

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n = b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n = b_m \end{array}\right. \tag{29}$$

se escribe la matriz ampliada de este sistema

$$\bar{B} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right],$$

donde la raya separa la columna de los términos independientes, y después se realizan transformaciones elementales con las *filas* de la matriz $\bar{B}$: se permite cambiar el orden de las filas (lo que equivale a cambiar el orden de las ecuaciones), multiplicar las filas por cualesquiera números diferentes de cero (lo que corresponde a la multiplicación de las respectivas ecuaciones por estos números) y agregar a cualquier fila de la matriz B cualquier otra fila multiplicada por un número arbitrario (lo que corresponde a la adición de una ecuación multiplicada por este número a otra ecuación del sistema). Después de estas transformaciones se obtiene toda vez una matriz ampliada de un sistema nuevo, *equivalente al inicial*. Se trata de reducir la matriz B a la forma lo más sencilla posible que permita indicar directamente la solución del sistema.

Consideremos más detalladamente el método de Gauss en el caso de los sistemas 1, 2 *y* 3 *del ejemplo* 7.

1.

$$\left\{\begin{aligned} x_1+2x_2+3x_3 &= 2,\\ x_1-x_2+x_3 &= 0,\\ x_1+3x_2-x_3 &= -2,\\ 3x_1+4x_2+3x_3 &= 0. \end{aligned}\right. \tag{30}$$

La matriz ampliada de este sistema es

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2\\ 1 & -1 & 1 & 0\\ 1 & 3 & -1 & -2\\ 3 & 4 & 3 & 0 \end{array}\right].$$

Restando la primera fila de la segunda y de la tercera y restando la primera fila triplicada de la cuarta, obtenemos la matriz

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2\\ 0 & -3 & -2 & -2\\ 0 & 1 & -4 & -4\\ 0 & -2 & -6 & -6 \end{array}\right].$$

Esta matriz es la matriz ampliada del sistema

$$\left\{\begin{aligned} x_1+2x_2+3x_3 &= 2,\\ -3x_2-2x_3 &= -2,\\ x_2-4x_3 &= -4,\\ -2x_2-6x_3 &= -6, \end{aligned}\right. \tag{31}$$

que se obtiene del sistema dado (30) al restar la primera ecuación de la segunda y de la tercera y al restar la primera ecuación triplicada de la cuarta. Luego, el sistema (31) es un corolario del sistema (30) y toda solución del sistema (30) satisface también al sistema (31). Pero, recíprocamente, el sistema (30) se puede obtener del sistema (31) mediante transformaciones análogas: agregando la primera ecuación a la segunda y a la tercera y sumando la primera ecuación triplicada a la cuarta. Luego, el sistema (30) es, a su vez, un corolario del sistema (31), es decir, ambos sistemas son *equivalentes* por tener las mismas soluciones.

Agregando ahora la tercera fila triplicada a la segunda y la tercera fila duplicada a la cuarta, obtenemos

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & 0 & -14 & -14 \\ 0 & 1 & -4 & -4 \\ 0 & 0 & -14 & -14 \end{array}\right].$$

Restando la segunda fila de la cuarta y dividiéndola por -14, tendremos

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & -4 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Pero ésta es la matriz ampliada del sistema

$$\left\{\begin{aligned} x_1+2x_2+3x_3 &= 2, \\ x_3 &= 1, \\ x_2-4x_3 &= -4, \end{aligned}\right.$$

que es equivalente al sistema dado (30); luego, la solución del sistema (30) es

$$x_3 = 1, \quad x_2 = -4+4x_3 = 0, \quad x_1 = 2-2x_2-3x_3 = 2-3 = \\ = -1.$$

En este caso el rango de la matriz ampliada es igual al rango de la matriz de los coeficientes y es igual, evidentemente, a tres.

2.

$$\begin{cases} x_1+2x_2+3x_3-\ x_4 = \ \ 0, \\ x_1-\ x_2+\ x_3+2x_4 = \ \ 4, \\ x_1+5x_2+5x_3-4x_4 = -4, \\ x_1+8x_2+7x_3-7x_4 = -8. \end{cases}$$

Escribiendo la matriz ampliada de este sistema, obtenemos, después de las transformaciones evidentes:

$$\left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 1 & -1 & 1 & 2 & 4 \\ 1 & 5 & 5 & -4 & -4 \\ 1 & 8 & 7 & -7 & -8 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 3 & 2 & -3 & -4 \\ 0 & 6 & 4 & -6 & -8 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

de donde se deduce que nuestro sistema es equivalente al sistema

$$\begin{cases} x_1+2x_2+3x_3-\ x_4 = 0, \\ \quad -3x_2-2x_3+3x_4 = 4, \end{cases}$$

y que, por consiguiente,

$$x_2 = -\frac{4}{3}-\frac{2}{3}x_3+x_4,$$

$$x_1 = -2x_2-3x_3+x_4 = \frac{8}{3}-\frac{5}{3}x_3-x_4.$$

Aquí el rango de la matriz ampliada es igual al rango de la matriz de los coeficientes y es igual, evidentemente, a dos.

3.

$$\begin{cases} x_1+2x_2+3x_3-\ x_4 = \ \ 0, \\ x_1-\ x_2+\ x_3+2x_4 = \ \ 4, \\ x_1+5x_2+5x_3-4x_4 = -4, \\ x_1+8x_2+7x_3-7x_4 = \ \ 6. \end{cases}$$

Tenemos, obviamente,

$$\left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 1 & -1 & 1 & 2 & 4 \\ 1 & 5 & 5 & -4 & -4 \\ 1 & 8 & 7 & -7 & 6 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 3 & 2 & -3 & -4 \\ 0 & 6 & 4 & -6 & 6 \end{array}\right] \sim$$

$$\sim \left[\begin{array}{rrrr|r} 1 & 2 & 3 & -1 & 0 \\ 0 & -3 & -2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 14 \end{array}\right],$$

es decir, el sistema es incompatible, ya que el sistema que le equivale contiene la ecuación

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 14$$

(correspondiente a la última fila). Es fácil ver que el rango de la matriz de los coeficientes es igual aquí a dos, mientras que el rango de la matriz ampliada es igual a tres.

Ejemplo 9. *Resuélvase, empleando el método de Gauss, el sistema homogéneo de ecuaciones*

$$\begin{cases} x_1+2x_2+3x_3+4x_4+5x_5=0, \\ 2x_1+3x_2+4x_3+5x_4+x_5=0, \\ 3x_1+4x_2+5x_3+x_4+2x_5=0, \\ x_1+3x_2+5x_3+12x_4+9x_5=0, \\ 4x_1+5x_2+6x_3-3x_4+3x_5=0 \end{cases}$$

y hállese su sistema fundamental de soluciones.

Solución. Consideremos la matriz ampliada del sistema (la columna nula se puede, claro está, omitir). Después de unas transformaciones comprensibles, encontramos

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 1 & 3 & 5 & 12 & 9 \\ 4 & 5 & 6 & -3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -2 & -3 & -9 \\ 0 & -2 & -4 & -11 & -13 \\ 0 & 1 & 2 & 8 & 4 \\ 0 & -3 & -6 & -19 & -17 \end{bmatrix} \sim$$

$$\sim \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & -2 & -3 & -9 \\ 0 & 0 & 0 & -5 & 5 \\ 0 & 0 & 0 & 5 & -5 \\ 0 & 0 & 0 & -10 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 9 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

es decir, el sistema dado equivale al siguiente:

$$\begin{cases} x_1+2x_2+3x_3+4x_4+5x_5=0, \\ \qquad x_2+2x_3+3x_4+9x_5=0, \\ \qquad\qquad\quad -x_4+x_5=0. \end{cases}$$

Aquí $r=3$ y tres de las incógnitas se pueden expresar mediante las restantes, por ejemplo, así:

$$x_4=x_5,$$
$$x_2=-2x_3-3x_4-9x_5=-2x_3-12x_5,$$
$$x_1=-2x_2-3x_3-4x_4-5x_5=x_3+15x_5.$$

Un sistema fundamental se puede obtener dando a las incógnitas libres x_3 y x_5 los valores $x_3=1$, $x_5=0$ (en este caso $x_1=1$, $x_2=-2$, $x_4=0$) y los valores $x_3=0$, $x_5=1$ (en este caso $x_1=15$, $x_2=-12$, $x_4=1$). Así obtenemos el sistema fundamental de soluciones:

$$e_1=(1,\ -2,\ 1,\ 0,\ 0),$$
$$e_2=(15,\ -12,\ 0,\ 1,\ 1).$$

La solución general del sistema es de la forma

$$e=c_1e_1+c_2e_2=(c_1+15c_2,\ -2c_1-12c_2,\ c_1,\ c_2,\ c_2),$$

donde c_1 y c_2 son unos números arbitrarios.

CAPÍTULO II

ESPACIO DE n DIMENSIONES

§ 1. Definición de un espacio vectorial

Comenzaremos con un ejemplo bien conocido por el lector. El concepto de *vector*, o de segmento orientado, desempeña un papel importante en la Geometría. Los vectores pueden ser sumados entre sí y multiplicados por números. La suma $\overline{OC}$ de los vectores $\overline{OA}$ y $\overline{OB}$ es, por definición, la diagonal del paralelogramo $OACB$ (*fig. 1, a;* esta definición se puede extender también al caso en el que las rectas OA y OB coinciden) y el producto $\overline{OD}$ del vector $\overline{OA}$ por un número α se define por las condiciones: $OD = |\alpha| \cdot OA$ y los vectores $\overline{OD}$ y $\overline{OA}$ están orientados en una misma dirección, si $\alpha > 0$, y en direcciones opuestas si $\alpha < 0$ (fig. 1, *b*).

Pero el conjunto de todos los vectores planos o el conjunto de todos los vectores espaciales representan solamente unos ejemplos (aunque muy importantes) de espacios vectoriales.

En el capítulo I hemos visto que siendo

$$e_1 = (\alpha_1, \alpha_2, \ldots, \alpha_n) \quad \text{y} \quad e_2 = (\beta_1, \beta_2, \ldots, \beta_n)$$

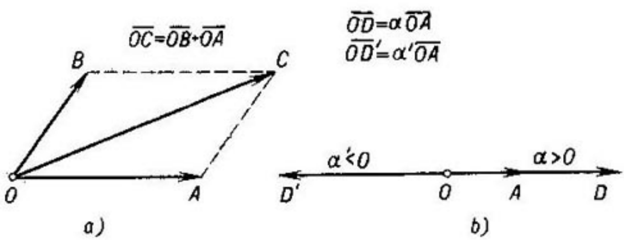


Fig. 1

dos soluciones de un sistema de ecuaciones lineales homogéneas, la suma de estas soluciones

$$e_1+e_2 = (\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_n+\beta_n)$$

y el producto de cualquiera de ellas, digamos de e_1, por un número arbitrario c

$$ce_1 = (c\alpha_1, c\alpha_2, \ldots, c\alpha_n)$$

también serán soluciones de este mismo sistema. Frecuentemente nos encontramos en las Matemáticas con situaciones análogas cuando se tiene un conjunto de elementos que pueden *ser sumados entre sí* y *multiplicados por números* obteniéndose de esta forma elementos del mismo conjunto. Por ejemplo, los *polinomios* en t de coeficientes reales pueden ser sumados entre sí y multiplicados por números reales obteniéndose polinomios del mismo tipo. Si el grado de los polinomios que se suman y se multiplican por números *no sobrepasa un número dado n*, los polinomios que se obtienen serán también en este caso de grado no mayor que n. Las funciones arbitrarias de t se pueden sumar entre sí y multiplicar por números resultando de nuevo funciones de t. Si las funciones a las que se aplican estas operaciones son *continuas* en un segmento $[a, b]$ (o en toda la recta numérica), las funciones que resultan tendrán la misma propiedad.

Finalmente, los propios *números* se pueden, por supuesto, sumar entre sí y multiplicar por números; es más, en lugar de un número se puede considerar pares, ternas y, en general, *colecciones ordenadas (filas)* formadas por n números:

$$(x_1, x_2, \ldots, x_n)$$

(si antes estas filas aparecían como soluciones de un sistema dado de ecuaciones lineales, ahora nada se exige de ellas). Las filas se pueden sumar entre sí:

$$(x_1, x_2, \ldots, x_n)+(y_1, y_2, \ldots, y_n) = \\ = (x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$$

y multiplicar por números:

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$$

obteniéndose toda vez una fila del mismo tipo.

Todas estas situaciones son distintos ejemplos de *espacios vectoriales* (siendo el último ejemplo de importancia primordial para lo sucesivo). Con el fin de abarcar estos y todos los demás casos posibles, introduciremos la definición siguiente.

Definición 1. *Un conjunto R de elementos x, y, z, ... se llama* ***espacio vectorial*** *o* ***espacio lineal,*** *si para cualesquiera dos elementos x e y del mismo está definida la suma* $x+y \in R$[1] *y si para todo elemento* $x \in R$ *y todo número real* α *está definido el producto* $\alpha x \in R$ *de modo que se cumplen las siguientes condiciones:*

1. $x+y = y+x$ *para todos los* $x, y \in R$.

2. $(x+y)+z = x+(y+z)$ *para todos los* $x, y, z \in R$.

3. *Existe un elemento* $0 \in R$ *(elemento nulo) tal que* $x+0 = x$ *para todos los elementos* $x \in R$.

4. *Para todo elemento* $x \in R$ *existe un elemento* $-x$ (llamado *opuesto de x*) *tal que* $x+(-x) = 0$.

5. $1 \cdot x = x$.

6. $\alpha(\beta x) = (\alpha\beta)x$[2].

7. $(\alpha+\beta)x = \alpha x+\beta x$.

8. $\alpha(x+y) = \alpha x+\alpha y$.

Los elementos de un espacio vectorial se llaman ***vectores.***

Ejemplos. Podemos hablar del espacio vectorial P_n de los polinomios de grado no mayor que n con coeficientes reales, o del espacio vectorial C de las funciones continuas en un segmento dado $[a, b]$, o del espacio vectorial de las soluciones de un sistema dado de ecuaciones lineales homogéneas, o, finalmente, del espacio vectorial de filas formadas por n elementos.

El espacio R suele llamarse espacio *real* (lineal o vectorial), ya que en R está definida la multiplicación de sus elementos por los números reales. Si los elementos de R se pueden multiplicar por los números complejos, se obtiene un espacio vectorial *complejo*. Como ejemplo de este espacio puede servir el conjunto de todas las filas formadas por n números complejos o el conjunto de todos los

---

[1] Los símbolos $\in$, $\subseteq$, $\subset$ se denominan símbolos de inclusión. La notación $a \in A$ significa que a es un *elemento* del conjunto A. La notación $A \subseteq B$ significa que el conjunto A es una *parte* del conjunto B (es decir, que todo elemento a de A pertenece, también a B); la notación $A \subset B$ significa que el conjunto A es una *parte regular* del conjunto B, es decir, que A está contenido en B pero no coincide con él.

[2] Para no confundir los vectores con los números convendremos en indicar —en todos los casos que puedan ofrecer dudas— los números por letras griegas y los vectores por letras latinas.

polinomios de coeficientes complejos (con más detalle se trata este tema en el § 5 del capítulo V).

Gracias a la condición 2 podemos hablar de la suma de $x+y+z = (x+y)+z$ (ó $x+(y+z)$ que viene a ser lo mismo) de tres o de un número mayor de elementos de R.

Se llama *diferencia* $x-y$ de los vectores x e y un vector z tal que $x = y+z$. Es fácil ver que $x-y = x+(-y)$.

Efectivamente,

$$y+[x+(-y)] = (y+x)+(-y) = (-y)+(y+x) = \\ = [(-y)+y]+x = 0+x = x.$$

Además, de la definición 1 se deducen directamente las proposiciones siguientes.

1. Unicidad del elemento nulo. Supongamos que en el espacio R existen dos elementos nulos 0_1 y 0_2. Puesto que para cualquier x de R se tiene $x+0_1 = x$ y $x+0_2 = x$, resulta, en particular, que $0_2+0_1 = 0_2$ y que $0_1+0_2 = 0_1$, de donde, debido a la igualdad $0_1+0_2 = 0_2+0_1$, se obtiene

$$0_1 = 0_2.$$

2. Unicidad del elemento opuesto. Supongamos que un elemento x tiene dos elementos opuestos y y z; tenemos entonces $x+y = 0$ y $x+z = 0$. Por consiguiente,

$$y+x+z = y+(x+z) = y+0 = y$$

e

$$y+x+z = (y+x)+z = 0+z = z,$$

de donde resulta

$$y = z.$$

3. Para todo elemento $x \in R$ se tiene $0x = 0$[1]. En efecto, tenemos $0x = (0+0)x = 0x+0x$ cualquiera que sea x. Agregando $-0x$ al primer y al segundo miembro de la última igualdad, obtenemos

$$0 = 0x.$$

[1] El mismo símbolo 0 se emplea aquí para indicar un número (en el primer miembro) y un vector (en el segundo miembro). Tanto aquí como en lo sucesivo, siempre estará claro del propio contexto que es lo que representa el símbolo 0, el número cero o el vector nulo.

4. Para cualquier número real α y para $0 \in R$ se tiene $\alpha 0 = 0$. Efectivamente, $\alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0$. Agregando $-\alpha 0$ al primer y al segundo miembro de esta igualdad, obtenemos

$$0 = \alpha 0.$$

5. Si se tiene $\alpha x = 0$, resulta que o bien $\alpha = 0$ o bien $x = 0$. En efecto, sea $\alpha \neq 0$; entonces,

$$x = 1 \cdot x = \left(\frac{1}{\alpha}\alpha\right)x = \frac{1}{\alpha}(\alpha x) = \frac{1}{\alpha}0 = 0.$$

6. Para todo x el elemento $(-1)x$ es el opuesto de x. Efectivamente,

$$x+(-1)x = 1 \cdot x + (-1)x = [1+(-1)]x = 0 \cdot x = 0$$

y, por consiguiente,

$$(-1)x = -x.$$

§ 2. Dimensión y base

Definición 2. *Los vectores $a_1, a_2, \ldots, a_k$ de un espacio lineal R se llaman* ***linealmente dependientes*** *si existen unos números $\alpha_1, \alpha_2, \ldots, \alpha_k$ no todos iguales simultáneamente a cero tales que*

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0.$$

Los vectores que no son linealmente dependientes se llaman ***linealmente independientes.***

Siendo los vectores $a_1, a_2, \ldots, a_k$ linealmente dependientes

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k = 0$$

y siendo, por ejemplo, $\alpha_k \neq 0$, se tiene

$$a_k = -\frac{\alpha_1}{\alpha_k}a_1 - \frac{\alpha_2}{\alpha_k}a_2 - \ldots - \frac{\alpha_{k-1}}{\alpha_k}a_{k-1},$$

es decir,

$$a_k = \xi_1 a_1 + \xi_2 a_2 + \ldots + \xi_{k-1}a_{k-1}, \tag{1}$$

donde $\xi_i = -\frac{\alpha_i}{\alpha_k}$. Si tiene lugar la igualdad (1) se dice que el vector a_k es una *combinación lineal* de los vectores $a_1, a_2, \ldots, a_{k-1}$ y también que el vector a_k *se expresa linealmente* en términos de $a_1, a_2, \ldots, a_{k-1}$. Luego, siendo los vectores $a_1, a_2, \ldots, a_k$ linealmente

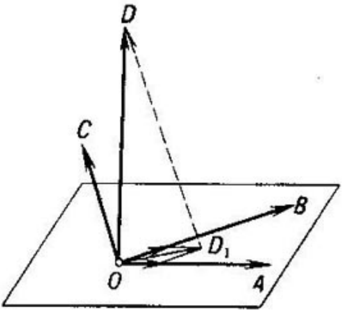


Fig. 2

dependientes, al menos uno de ellos se expresa linealmente en términos de los restantes. Está claro que la recíproca es también válida, es decir, si uno de los vectores se expresa linealmente en términos de los demás, estos vectores son linealmente dependientes en su conjunto.

Ejemplos. En el plano existen tantos pares de vectores linealmente independientes como se quiera: cualesquiera dos vectores no colineales, es decir, no paralelos a una misma recta, son linealmente independientes. Pero cualesquiera tres vectores del plano son linealmente dependientes.

En el espacio cualesquiera tres vectores no coplanares (es decir, no paralelos a un mismo plano) a, b y c son linealmente independientes (ya que siendo $\alpha a+\beta b+\gamma c = 0$ y siendo, por ejemplo, $\gamma \neq 0$, resulta $c = -\frac{\alpha}{\gamma} a - \frac{\beta}{\gamma} b$, es decir, el vector c es coplanar con los vectores a y b). Sin embargo, cualesquiera cuatro vectores a, b, c y d del espacio serán linealmente dependientes. En efecto, sean $a = \overline{OA}$, $b = \overline{OB}$ y $c = \overline{OC}$ unos vectores no coplanares y sea $d = \overline{OD}$ otro vector cualquiera; trazando la recta $DD_1 \parallel OC$ (fig. 2; el punto D_1 pertenece al plano AOB) y descomponiendo el vector $\overline{OD_1}$, que es coplanar con a y b, respecto a estos vectores, obtenemos evidentemente:

$$\overline{OD_1} = \alpha a+\beta b,$$
$$\overline{OD} = \overline{OD_1}+\overline{D_1D}$$

y

$$\overline{D_1D} = \gamma c,$$

es decir,

$$d = \alpha a + \beta b + \gamma c,$$

de donde resulta

$$d - \alpha a - \beta b - \gamma c = 0.$$

(Si los vectores a, b y c son coplanares, existen unos números α, β y γ no todos iguales simultáneamente a cero tales que

$$\alpha a + \beta b + \gamma c = 0;$$

pero se tiene entonces

$$\alpha a + \beta b + \gamma c + 0 \cdot d = 0$$

y los cuatro vectores son también linealmente dependientes.)

Definición 3. *Un espacio lineal R se llama n-dimensional si en él se pueden encontrar n vectores linealmente independientes pero es imposible encontrar más de n vectores linealmente independientes.*

Es decir, la **dimensión de un espacio** es el número máximo de vectores linealmente independientes que éste contiene. Convendremos en indicar por $d(R)$ la dimensión del espacio R.

La dimensión del conjunto de todos los vectores planos es igual a dos y la dimensión del conjunto de los vectores espaciales es igual a tres.

Los espacios que tienen dimensión finita se llaman espacios de *dimensión finita*. Un espacio en el que se puedan encontrar tantos vectores linealmente independientes como se quiera se llama espacio de *dimensión infinita*. Es un ejemplo de un espacio de dimensión infinita el conjunto P de todos los polinomios en t de coeficientes reales o el conjunto C de todas las funciones de t continuas en un segmento dado $[a, b]$ (o continuas en toda la recta numérica).

Definición 4. *Toda colección de n vectores linealmente independientes de un espacio n-dimensional R se llama **base** de este espacio.*

Teorema 1. *Todo vector x de un espacio lineal n-dimensional R se puede representar como una combinación lineal de los vectores de la base y, además, esta representación es única.*

Demostración. Sea $e_1, e_2, \ldots, e_n$ una base cualquiera de un espacio n-dimensional R y sea $x \in R$. Puesto que cualesquiera

$n+1$ vectores del espacio R (¡de n dimensiones!) son linealmente dependientes, resultan también dependientes, en particular, los vectores $e_1, e_2, \ldots, e_n, x$, es decir, existen unos números $\alpha_1, \alpha_2, \ldots$ $\ldots, \alpha_n, \alpha$ no todos iguales simultáneamente a cero tales que

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n + \alpha x = 0.$$

Además, es $\alpha \neq 0$, ya que de lo contrario sería diferente de cero uno de los números $\alpha_1, \alpha_2, \ldots, \alpha_n$ y, por consiguiente, los vectores e_1, $e_2, \ldots, e_n$ resultarían linealmente dependientes. Luego,

$$x = -\frac{\alpha_1}{\alpha} e_1 - \frac{\alpha_2}{\alpha} e_2 - \ldots - \frac{\alpha_n}{\alpha} e_n.$$

Tomando $-\frac{\alpha_i}{\alpha} = x_i$, tendremos

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Esta representación de x mediante $e_1, e_2, \ldots, e_n$ es *única*, ya que siendo $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$ y $x = y_1 e_1 + y_2 e_2 + \ldots + y_n e_n$, se tiene

$$(y_1 - x_1) e_1 + (y_2 - x_2) e_2 + \ldots + (y_n - x_n) e_n = 0$$

y, debido a la independencia lineal de los vectores $e_1, e_2, \ldots, e_n$, resulta

$$y_1 = x_1, \ y_2 = x_2, \ \ldots, \ y_n = x_n.$$

Los números $x_1, x_2, \ldots, x_n$ se llaman *coordenadas* del vector x en la base $e_1, e_2, \ldots, e_n$. Por lo tanto, el teorema 1 afirma que dada una base de un espacio vectorial n-dimensional R, todo vector de R tiene coordenadas en esta base (determinadas unívocamente). Además, queda claro que si coinciden las coordenadas de dos vectores x e y, estos vectores son iguales, ya que en este caso

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n = y.$$

Por esto, para determinar un vector basta indicar sus coordenadas $x_1, x_2, \ldots, x_n$ y se dice: vector $x = (x_1, x_2, \ldots, x_n)$.

Supongamos que se tienen dos vectores dados mediante sus coordenadas en una base. Entonces, *al sumar estos vectores se suman sus coordenadas respectivas:* si

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n \quad \text{e} \quad y = y_1 e_1 + y_2 e_2 + \ldots + y_n e_n,$$

se tiene

$$x + y = (x_1 + y_1) e_1 + (x_2 + y_2) e_2 + \ldots + (x_n + y_n) e_n.$$

Al multiplicar un vector por un número todas sus coordenadas se multiplican por este número: si

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n,$$

se tiene

$$\alpha x = (\alpha x_1)\, e_1 + (\alpha x_2)\, e_2 + \ldots + (\alpha x_n)\, e_n.$$

Todas las coordenadas del vector nulo son iguales a cero, ya que, debido a la independencia lineal de los vectores $e_1, e_2, \ldots, e_n$, de la igualdad $\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0$ resulta que todos los $\alpha_i = 0$. El vector opuesto del vector $x = (x_1, x_2, \ldots, x_n)$ es igual a $(-x_1, -x_2, \ldots, -x_n)$, pues

$$(x_1, x_2, \ldots, x_n) + (-x_1, -x_2, \ldots, -x_n) = (0, 0, \ldots, 0) = 0.$$

Teorema 2. *Si $e_1, e_2, \ldots, e_n$ son unos vectores linealmente independientes del espacio R y si todo vector $x \in R$ se expresa linealmente en términos de $e_1, e_2, \ldots, e_n$, estos vectores constituyen una base de R.*

Demostración. Los vectores $e_1, e_2, \ldots, e_n$ son, por hipótesis, linealmente independientes. Resta demostrar que en el espacio R no hay más de n vectores linealmente independientes. Tomemos cualesquiera $m > n$ vectores $a_1, a_2, \ldots, a_m$ de R. Cada uno de ellos se puede, por hipótesis, expresar linealmente en términos de $e_1, e_2, \ldots, e_n$:

$$a_1 = \alpha_{11}e_1 + \alpha_{21}e_2 + \ldots + \alpha_{n1}e_n,$$
$$a_2 = \alpha_{12}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{n2}e_n,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_m = \alpha_{1m}e_1 + \alpha_{2m}e_2 + \ldots + \alpha_{nm}e_n.$$

Consideremos la matriz

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nm} \end{bmatrix}.$$

Como quiera que el número de filas de esta matriz es igual a n, su rango es no mayor que n y, por consiguiente, entre sus columnas no hay más de n linealmente independientes. Pero por ser $m > n$, las m columnas de esta matriz son linealmente dependientes. Esto significa que los vectores $a_1, a_2, \ldots, a_m$ son también linealmente dependientes. Hemos demostrado que el espacio R es n-dimensional y que $e_1, e_2, \ldots, e_n$ es una de sus bases.

Del teorema 2 se desprende que el espacio R^n de las filas ordenadas formadas por n números es n-dimensional. Efectivamente, las n filas

$$e_1 = (1, 0, \ldots, 0),$$
$$e_2 = (0, 1, \ldots, 0),$$
$$\cdots\cdots\cdots\cdots$$
$$e_n = (0, 0, \ldots, 1)$$

son linealmente independientes, ya que de la igualdad

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = (\alpha_1, \alpha_2, \ldots, \alpha_n) = (0, 0, \ldots, 0)$$

resulta $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Por otra parte, toda fila $e = (\xi_1, \xi_2, \ldots, \xi_n)$ se expresa linealmente en términos de $e_1, e_2, \ldots, e_n$:

$$e = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n.$$

Las filas $e_1, e_2, \ldots, e_n$ forman, por consiguiente, una base del espacio R^n.

El espacio P_n de los polinomios de grado no mayor que n es de dimensión $n+1$. En efecto, los polinomios

$$1, t, t^2, \ldots, t^n$$

son linealmente independientes y todo polinomio en t de grado no mayor que n se expresa en términos de éstos de un modo evidente.

Teorema 3. *En un espacio lineal de dimensión finita todo conjunto de vectores linealmente independientes puede ser incluido en una base.*

Demostración. Sean $e_1, e_2, \ldots, e_k$ unos vectores linealmente independientes de un espacio R. Si todos los restantes vectores de R se expresan linealmente en términos de los vectores $e_1, e_2, \ldots, e_k$, estos últimos ya constituyen una base en virtud del teorema 2. Si existe un vector e_{k+1} que no se expresa linealmente en términos de $e_1, e_2, \ldots, e_k$, los $k+1$ vectores $e_1, e_2, \ldots, e_k, e_{k+1}$ serán linealmente independientes. Efectivamente, si tiene lugar la igualdad

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k + \alpha e_{k+1} = 0$$

debe ser $\alpha \neq 0$ —debido a la independencia lineal de los vectores $e_1, e_2, \ldots, e_k$— y, por consiguiente, el vector e_{k+1} se expresa linealmente en términos de $e_1, e_2, \ldots, e_k$.

Agreguemos el vector e_{k+1} a los vectores $e_1, e_2, \ldots, e_k$. Si todos los vectores del espacio R se expresan linealmente en términos de $e_1, e_2, \ldots, e_k, e_{k+1}$, estos últimos ya constituyen una base. Si existe un vector e_{k+2} que no se expresa linealmente en términos de $e_1, e_2, \ldots, e_k, e_{k+1}$, lo agregaremos a éstos; el sistema nuevo de vectores $e_1, e_2, \ldots, e_k, e_{k+1}, e_{k+2}$ será linealmente independiente, etc.

Este proceso no puede prolongarse indeterminadamente, ya que el espacio R es, por hipótesis, de dimensión finita y, por consiguiente, en él no puede haber un conjunto infinito $e_1, e_2, e_3, \ldots$ de vectores linealmente independientes. Por esto obtendremos, al fin y al cabo, un sistema linealmente independiente de vectores $e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_n$ en términos del cual se expresarán linealmente todos los demás vectores de R. Este sistema será, en virtud del teorema 2, una base del espacio R que contiene los vectores dados $e_1, e_2, \ldots, e_k$.

§ 3. Isomorfismo de espacios lineales

Sea R un espacio lineal n-dimensional y sea $e_1, e_2, \ldots, e_n$ una base del mismo. Para todo vector $x \in R$ existe, según el teorema 1, una representación única

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Si ponemos en correspondencia al vector x la fila $(x_1, x_2, \ldots, x_n)$, resulta, por lo visto en el § 2, que al sumar unos vectores se suman las filas que les corresponden y que al multiplicar un vector por un número se multiplica por este mismo número la fila que le corresponde.

Es decir, partiendo de la definición general de un espacio vectorial n-dimensional, hemos llegado a la conclusión de que este espacio tiene la misma estructura, en cierto sentido, que el espacio formado por todas las filas de n números. Por consiguiente, todos los espacios vectoriales n-dimensionales tienen la misma estructura; son, como suele decirse, *isomorfos* entre sí. El sentido exacto de este término viene dado en la definición siguiente.

Definición 5. *Dos espacios vectoriales R y R' se llaman **isomorfos** si se puede establecer entre sus elementos una correspondencia biyectiva tal que siendo $x \leftrightarrow x'$* (siendo x correspondiente a x') *e $y \leftrightarrow y'$, donde $x, y \in R$ y $x', y' \in R'$, se tenga*

$$x+y \leftrightarrow x'+y'$$

y

$$\alpha x \leftrightarrow \alpha x'$$

cualquiera que sea α (o, abreviando, $(x+y)' = x'+y'$ y $(\alpha x)' = \alpha x'$).

Tiene lugar el teorema siguiente.

Teorema 4. *Para que dos espacios vectoriales sean isomorfos es necesario y suficiente que tengan la misma dimensión.*

Demostración de la suficiencia. Sean dados dos espacios lineales n-dimensionales R y R'. Escojamos en cada uno de ellos una base

$$e_1, e_2, \ldots, e_n \text{ en } R \text{ y } e_1', e_2', \ldots, e_n' \text{ en } R'.$$

Pongamos en correspondencia al vector x, cuyas coordenadas en la base $e_1, e_2, \ldots, e_n$ son $x_1, x_2, \ldots, x_n$, el vector x' de R' que tiene estas *mismas coordenadas* en la base $e_1', e_2', \ldots, e_n'$. Al sumar unos vectores sus coordenadas correspondientes se suman y al multiplicar un vector por un número sus coordenadas se multiplican por este mismo número; por lo tanto, si

$$x \leftrightarrow x' \text{ e } y \leftrightarrow y',$$

se tiene

$$x+y \leftrightarrow x'+y'$$

y

$$\alpha x \leftrightarrow \alpha x'$$

cualquiera que sea α.

Luego, R' es isomorfo a R.

Demostración de la necesidad. Para demostrar que dos espacios lineales R y R' de diferentes dimensiones no son isomorfos entre sí, observemos, ante todo, que en una correspondencia «isomorfa» entre dos espacios al vector nulo de un espacio le corresponde el vector nulo del otro espacio. Efectivamente, sea 0 el vector nulo de R y sea $0'$ el vector de R' que le corresponde; sea x' un vector cualquiera de R' y sea $x \leftrightarrow x'$, donde $x \in R$. En estas condiciones la suma

$$0+x \leftrightarrow 0'+x'.$$

Pero $0+x = x$; además, $x \leftrightarrow x'$ y como la correspondencia entre R y R' es biyectiva, resulta que

$$0'+x' = x',$$

es decir, que $0'$ es el vector nulo de R'.

Siendo los espacios R y R' isomorfos y siendo correspondientes los vectores $a_1, a_1, \ldots, a_k$ de R y los vectores $a_1', a_2', \ldots, a_k'$ de R', la dependencia lineal de los vectores $a_1, a_2, \ldots, a_k$ implica la dependencia lineal de los vectores $a_1', a_2', \ldots, a_k'$ y viceversa. Efectivamente, supongamos, por ejemplo, que $\alpha_1 a_1+\alpha_2 a_2+\ldots+\alpha_k a_k = 0$. Entonces al vector $\alpha_1 a_1+\alpha_2 a_2+\ldots+\alpha_k a_k$ del espacio R que es igual al vector $0 \in R$ le corresponde en el espacio R' el vector $\alpha_1 a'_1 + \alpha_2 a'_2 + \ldots + \alpha_k a'_k$ y, por consiguiente,

$$\alpha_1 a'_1 + \alpha_2 a'_2 + \ldots + \alpha_k a'_k = 0'.$$

Por lo tanto, el número máximo de vectores linealmente independientes que tienen los espacios isomorfos es el mismo y, por consiguiente, las dimensiones de estos espacios son iguales.

En virtud del teorema 4, *la dimensión es la única característica de un espacio vectorial de dimensión finita*. Todos los espacios vectoriales n-dimensionales son totalmente idénticos en cuanto a su estructura algebraica. Podemos decir, por consiguiente, que un espacio vectorial n-dimensional es el espacio de todas las filas de n números. Por esto un espacio vectorial n-dimensional puede ser indicado simplemente por R^n: es legal emplear una misma notación para todos los espacios vectoriales n-dimensionales porque todos los espacios vectoriales n-dimensionales son iguales (isomorfos).

§ 4. Cambio de base

Supongamos que en el espacio R^n se tienen dos bases

$$e_1, e_2, \ldots, e_n \quad \text{y} \quad e'_1, e'_2, \ldots, e'_n.$$

Convengamos en denominar base *antigua* a la primera y base *nueva* a la segunda. Todo elemento de la base nueva se puede, de acuerdo al teorema 1, expresar linealmente en términos de la base antigua

$$\begin{aligned} e'_1 &= a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n, \\ e'_2 &= a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ e'_n &= a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n. \end{aligned} \tag{2}$$

Podemos decir que los vectores básicos nuevos se obtienen de los antiguos mediante la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

(con la particularidad de que los coeficientes de las descomposiciones de estos vectores respecto a los vectores de la base antigua forman las **columnas** de esta matriz). La matriz A se llama *matriz del cambio* de la base $e_1, e_2, \ldots, e_n$ por la base $e'_1, e'_2, \ldots, e'_n$.

El determinante de la matriz A es diferente de cero, ya que de lo contrario sus columnas y, por consiguiente, los vectores $e'_1, e'_2, \ldots$ $\ldots, e'_n$ resultarían linealmente dependientes.

Recíprocamente, si el determinante de la matriz A es diferente de cero, sus columnas son linealmente independientes y, por consiguiente, los vectores $e'_1, e'_2, \ldots, e'_n$ —que se obtienen de los vectores básicos $e_1, e_2, \ldots, e_n$ mediante la matriz A— son linealmente independientes, es decir, forman una base. Luego, es una matriz del cambio *cualquier* matriz cuadrada de orden n con el determinante distinto de cero.

Veamos ahora cómo se relacionan entre sí las coordenadas de un mismo vector en las bases antigua y nueva. Sea

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

y sea al mismo tiempo

$$x = x'_1e'_1 + x'_2e'_2 + \ldots + x'_ne'_n.$$

Tomando en lugar de $e'_1, e'_2, \ldots, e'_n$ sus expresiones (2) en términos de $e_1, e_2, \ldots, e_n$, obtenemos

$$\begin{aligned} x = x'_1(a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n) + x'_2(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n) + \\ + \ldots + x'_n(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n) = (a_{11}x'_1 + a_{12}x'_2 + \ldots + \\ + a_{1n}x'_n)\, e_1 + (a_{21}x'_1 + a_{22}x'_2 + \ldots + a_{2n}x'_n)\, e_2 + \ldots + (a_{n1}x'_1 + \\ + a_{n2}x'_2 + \ldots + a_{nn}x'_n)\, e_n, \end{aligned}$$

y, debido a que la descomposición del vector x según la base $e_1, e_2, \ldots, e_n$ es única, de aquí se deduce que

$$\begin{aligned} x_1 &= a_{11}x'_1 + a_{12}x'_2 + \ldots + a_{1n}x'_n, \\ x_2 &= a_{21}x'_1 + a_{22}x'_2 + \ldots + a_{2n}x'_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= a_{n1}x'_1 + a_{n2}x'_2 + \ldots + a_{nn}x'_n. \end{aligned}$$

Luego, las coordenadas antiguas del vector x se obtienen de sus coordenadas nuevas mediante la misma matriz A, pero los coeficientes de las descomposiciones respectivas forman ahora las *filas*

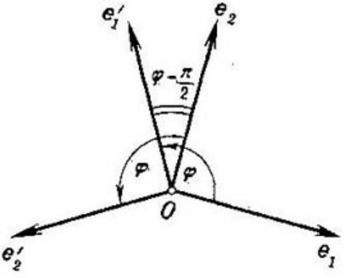


Fig. 3

de esta matriz. Es fácil ver así mismo que, recíprocamente, las coordenadas nuevas $x_1', x_2', \ldots, x_n'$ del vector x se expresan linealmente en términos de sus coordenadas antiguas $x_1, x_2, \ldots, x_n$.

Ejemplo. Sean e_1, e_2 los vectores unitarios en la dirección de los ejes de un sistema cartesiano rectangular de coordenadas. Hagamos girar los ejes de coordenadas en un ángulo φ en el sentido contrario al del movimiento de las manecillas de un reloj y sean e_1' y e_2' los nuevos vectores básicos. Los ángulos que el vector e_1' forma con los vectores e_1 y e_2 son iguales a φ y a $\varphi-\frac{\pi}{2}$, respectivamente (fig. 3). Por ello las coordenadas de este vector en la base e_1, e_2 son iguales a $\cos\varphi$ y a $\cos\left(\varphi-\frac{\pi}{2}\right)=\operatorname{sen}\varphi$, es decir, $e_1'=\cos\varphi\cdot e_1+$ $+\operatorname{sen}\varphi\cdot e_2$. Análogamente, los ángulos que el vector e_2' forma con los vectores e_1 y e_2 son iguales a $\frac{\pi}{2}+\varphi$ y a φ, respectivamente; sus coordenadas en la base e_1, e_2 son iguales a $\cos\left(\frac{\pi}{2}+\varphi\right)=-\operatorname{sen}\varphi$ y a $\cos\varphi$ y, por consiguiente, $e_2'=-\operatorname{sen}\varphi\cdot e_1+\cos\varphi\cdot e_2$.

Luego, la matriz del cambio será en este caso la siguiente:

$$A=\begin{bmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{bmatrix}$$

y las expresiones de las coordenadas antiguas a través de las coordenadas nuevas serán

$$x_1=\cos\varphi\cdot x_1'-\operatorname{sen}\varphi\cdot x_2',$$
$$x_2=\operatorname{sen}\varphi\cdot x_1'+\cos\varphi\cdot x_2'.$$

§ 5. Subespacios de un espacio lineal

Definición 6. *Se llama **subespacio** de un espacio vectorial R a todo conjunto R_1 de elementos de éste que forma él mismo un espacio vectorial respecto a las operaciones de adición y de multiplicación por número introducidas en R.*

Para comprobar que un conjunto R_1 de elementos de un espacio lineal R es un subespacio de este último es necesario probar que cualesquiera que sean dos vectores x e y de R_1 la suma $x+y$ también pertenece a R_1 y que cualesquiera que sean el vector x de R_1 y el número real α el producto αx pertenece también a R_1. Demostremos que estas condiciones son al mismo tiempo suficientes. Efectivamente, los axiomas 1 y 2 y los axiomas de 5 a 8 de un espacio vectorial, que son válidos en R, se cumplirán, en particular, para los elementos de R_1. Resta probar que el vector nulo pertenece a R_1 y que para todo x de R_1 el vector $-x$ también pertenece a R_1. Pero, siendo $x \in R_1$, tenemos que los productos $0 \cdot x = 0$ y $(-1)\,x = -x$ también pertenecen a R_1.

La dimensión de cualquier subespacio de un espacio vectorial no sobrepasa la dimensión del propio espacio, ya que los vectores linealmente independientes del subespacio R_1 serán también independientes en todo el espacio y, por consiguiente, el número máximo de vectores linealmente independientes de un subespacio no sobrepasa la dimensión de todo el espacio.

Ejemplos. En el espacio ordinario de tres dimensiones (comprendido como el conjunto de los vectores que le pertenecen), todos los planos y todas las rectas que pasan por el origen de coordenadas serán subespacios. El propio espacio R y el conjunto formado solamente por el vector nulo son subespacios de cualquier espacio R. En el espacio P_n de los polinomios de grado no mayor que n, son subespacios, por ejemplo, todos los polinomios P_k con $k < n$ (ya que al sumar y al multiplicar por números los polinomios de grado no mayor que k obtenemos polinomios de este mismo tipo). (Los polinomios cuyo grado es exactamente igual a n **no forman** un subespacio, ya que al sumar estos polinomios se puede obtener un polinomio de grado inferior debido a que los coeficientes principales de estos polinomios pueden simplificarse.) Por otra parte, cualquiera de los espacios P_n es un subespacio del espacio P de *todos* los polinomios de coeficientes reales y este último es, a su vez, un subespacio del espacio C de las *funciones continuas* (ya que los polinomios son funciones continuas y al sumarlos y multiplicarlos por números resultan de nuevo polinomios).

Consideremos un sistema de ecuaciones lineales homogéneas con la matriz de los coeficientes de rango r:

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{array}\right. \tag{3}$$

y supongamos que en el espacio lineal R^n se ha fijado una base cualquiera. Considerando toda solución $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ del sistema (3) como un vector del espacio R^n, obtenemos de los resultados del § 10 del capítulo I que *el conjunto de todas las soluciones del sistema* (3) *constituye un subespacio lineal k-dimensional* (*donde* $k = n-r$) de R^n *y que cualquier sistema fundamental de soluciones es una base de este subespacio.*

Demostremos que, recíprocamente, *todo subespacio de un espacio vectorial se determina por un sistema de ecuaciones lineales homogéneas.* Sea R_1 un subespacio k-dimensional de R^n y sea $e'_1, e'_2, \ldots, e'_k$ una de sus bases. Complementemos este sistema linealmente independiente de vectores hasta obtener una base $e'_1, e'_2, \ldots$ $\ldots, e'_k, e'_{k+1}, \ldots, e'_n$ de todo el espacio R^n. Es fácil ver que, siendo $x'_1, x'_2, \ldots, x'_n$ las coordenadas de un vector x respecto a esta base nueva, el subespacio R_1 se determina por las ecuaciones

$$x'_{k+1} = 0,\ x'_{k+2} = 0,\ \ldots,\ x'_n = 0.$$

Además, si $x_1, x_2, \ldots, x_n$ son las coordenadas del vector x respecto a la base antigua $e_1, e_2, \ldots, e_n$, tienen lugar, por lo visto en el § 4, las igualdades

$$\begin{array}{l} x'_{k+1} = \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n, \\ x'_{k+2} = \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x'_n = \alpha_{n-k,1}\, x_1 + \alpha_{n-k,2}\, x_2 + \ldots + \alpha_{n-k,n}\, x_n, \end{array}$$

donde α_{ij} son unos números; luego, el subespacio R_1 se determina en la base $e_1, e_2, \ldots, e_n$ por el siguiente sistema de ecuaciones lineales homogéneas

$$\begin{cases} \alpha_{11}x_1+\alpha_{12}x_2+\ldots+\alpha_{1n}x_n=0, \\ \alpha_{21}x_1+\alpha_{22}x_2+\ldots+\alpha_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{n-k,1}x_1+\alpha_{n-k,2}x_2+\ldots+\alpha_{n-k,n}x_n=0. \end{cases}$$

§ 6. Intersección y suma de subespacios

Definición 7. Sean R_1 y R_2 dos subespacios de un espacio vectorial R. *Se llama intersección* $R_3 = R_1 \cap R_2$ *de estos subespacios al conjunto de todos los vectores de* R *que pertenecen a la vez a* R_1 *y a* R_2.

Es fácil ver que *la intersección de dos subespacios* R_1 *y* R_2 *es un subespacio* (que está contenido en R_1 y en R_2).

Definición 8. Si R_1 y R_2 son unos subespacios de un espacio lineal R, se llama *suma* $R_4 = R_1 + R_2$ *de estos subespacios al conjunto de todos los vectores de tipo* $u+v$, *donde* $u \in R_1$ *y* $v \in R_2$.

La suma de dos subespacios es un subespacio (que posiblemente coincide con R). Efectivamente, siendo x, $y \in R_4$ y α un número cualquiera, tenemos

$$x = u_1 + v_1 \quad \text{e} \quad y = u_2 + v_2,$$

donde $u_1, u_2 \in R_1$ y $v_1, v_2 \in R_2$. Pero en este caso

$$x+y = (u_1+u_2)+(v_1+v_2),$$

donde $u_1+u_2 \in R_1$ y $v_1+v_2 \in R_2$, es decir, $x+y \in R_4$. Además, $\alpha x = \alpha u_1 + \alpha v_1$, donde $\alpha u_1 \in R_1$ y $\alpha v_1 \in R_2$, y, por consiguiente, $\alpha x \in R_4$.

El subespacio R_1 (así como el subespacio R_2) *está contenido en* R_4, ya que todo elemento $x \in R_1$ puede ser representado en la forma $x+0$, donde $x \in R_1$ y $0 \in R_2$.

Teorema 5. *Siendo* R_1 *y* R_2 *unos subespacios de un espacio lineal* R *y siendo* $R_3 = R_1 \cap R_2$ y $R_4 = R_1 + R_2$, *se tiene*

$$d(R_1)+d(R_2) = d(R_3)+d(R_4). \tag{4}$$

Demostración. Tomemos una base cualquiera

$$e_1, e_2, \ldots, e_k \tag{5}$$

del subespacio R_3. Complementemos el conjunto (5) de vectores que pertenecen a la vez a R_1 y a R_2 hasta obtener una base

$$e_1,\ e_2,\ \ldots,\ e_k,\ f_{k+1},\ \ldots,\ f_p \tag{6}$$

de R_1 y una base

$$e_1,\ e_2,\ \ldots,\ e_k,\ g_{k+1},\ \ldots,\ g_q \tag{7}$$

de R_2 (basándonos en el teorema 3). Probemos que los vectores

$$e_1,\ e_2,\ \ldots,\ e_k,\ f_{k+1},\ \ldots,\ f_p,\ g_{k+1},\ \ldots,\ g_q \tag{8}$$

son linealmente independientes. Con esto quedará demostrado, según el teorema 2, que constituyen una base de R_4, ya que siendo $z \in R_4$ se tiene $z = x+y$, donde $x \in R_1$ e $y \in R_2$ y, por consiguiente, x se expresa linealmente en términos de los vectores (6) e y se expresa linealmente en términos de los vectores (7), es decir, el vector z se expresa linealmente en términos de los vectores (8).

Supongamos que los vectores (8) son linealmente dependientes:

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_k e_k+\beta_{k+1} f_{k+1}+\ldots+\beta_p f_p+\gamma_{k+1} g_{k+1}+\ldots$$
$$\ldots+\gamma_q g_q = 0. \tag{9}$$

El vector

$$a = \alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_k e_k+\beta_{k+1} f_{k+1}+\ldots+\beta_p f_p,$$

que es igual a $-(\gamma_{k+1} g_{k+1}+\ldots+\gamma_q g_q)$, debe pertenecer en este caso tanto a R_1 como a R_2 y, por consiguiente, pertenece a la intersección R_3. Luego, debe expresarse linealmente en términos de los vectores básicos (5) del subespacio R_3; sea

$$a = \sigma_1 e_1+\sigma_2 e_2+\ldots+\sigma_k e_k.$$

De aquí resulta, debido a la unicidad de la descomposición del vector a respecto a la base del subespacio R_1:

$$\sigma_i = \alpha_i,\ i = 1,\ 2,\ \ldots,\ k \text{ y } \beta_{k+1} = \beta_{k+2} = \ldots = \beta_p = 0;$$

de la igualdad (9) se deduce entonces que

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_k e_k+\gamma_{k+1} g_{k+1}+\ldots+\gamma_q g_q = 0$$

y, como los vectores (7) son linealmente independientes, resulta

$$\alpha_i = 0,\ i = 1, 2,\ \ldots,\ k \text{ y } \gamma_{k+1} = \gamma_{k+2} = \ldots = \gamma_q = 0.$$

Luego, los vectores (8) constituyen una base del subespacio R_4 y, por consiguiente, la dimensión de éste es igual al número de estos vectores:

$$k+(p-k)+(q-k) = p+q-k.$$

Pero $d(R_1) = p$, $d(R_2) = q$ y $d(R_3) = k$. Hemos demostrado que *la suma de las dimensiones de dos subespacios es igual a la dimensión de la suma de estos subespacios más la dimensión de la intersección de los mismos.*

Así, dos subespacios *bidimensionales* R_1 y R_2 del espacio R^4 de cuatro dimensiones pueden tener como intersección el vector nulo; en este caso su suma coincide con todo el espacio y la igualdad (4) se convierte en $2+2 = 0+4$. Pueden tener como intersección una recta (un subespacio unidimensional); la suma es entonces de tres dimensiones y esto corresponde a la igualdad $2+2 = 1+3$. Finalmente, R_1 y R_2 pueden coincidir; en este caso la intersección y la suma de ellos son también bidimensionales y la igualdad (4) da $2+2 = 2+2$.

Dos subespacios *tridimensionales* de R^4 o bien tienen como intersección un plano (un subespacio bidimensional) y entonces $3+3 = 2+4$, o bien coinciden: $3+3 = 3+3$. (No puede haber más casos, ya que la suma de estos subespacios es a lo sumo de cuatro dimensiones.)

Si R_1 es un subespacio *bidimensional* y R_2 es un subespacio *tridimensional* de R^4, estos subespacios tienen como intersección una recta $(2+3 = 1+4)$ o bien R_1 está contenido en R_2 $(2+3 = 2+3)$.

Definición 9. *Si un espacio R es la suma de dos subespacios suyos R_1 y R_2 y la intersección R_3 de éstos está formada sólo por el vector nulo, se dice que R_4 es la* **suma directa** *de los subespacios R_1 y R_2* y se escribe

$$R = R_1 \oplus R_2.$$

Siendo $R = R_1 \oplus R_2$, es evidente que

$$d(R) = d(R_1)+d(R_2)$$

Teorema 6. *Si $R = R_1 \oplus R_2$, todo vector de R se representa de un modo único en la forma $u+v$, donde $u \in R_1$ y $v \in R_2$.*

Demostración. Según la definición de la suma de subespacios, todo vector de R se representa en la forma $u+v$, donde $u \in R_1$ y $v \in R_2$. Supongamos que un vector x de R ha sido descompuesto en dos sumas de este tipo

$$x = u_1+v_1 = u_2+v_2.$$

El vector $u_1-u_2 = v_2-v_1$ pertenece entonces simultáneamente a R_1 y a R_2, es decir, pertenece a R_3 y, por consiguiente, es igual a cero, de donde resulta que $u_1 = u_2$ y $v_1 = v_2$.

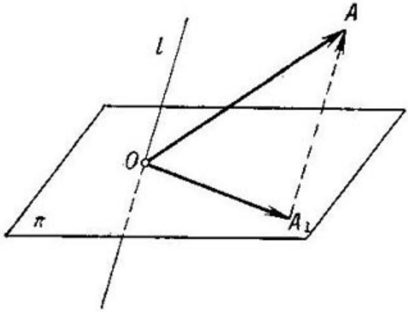


Fig. 4

Así, el espacio corriente de tres dimensiones R^3 es la suma directa de cualquier plano π (que pasa por el origen de coordenadas) y de cualquier recta l que no pertenece a este plano (y que pasa por el origen de coordenadas), ya que todo vector $\overline{OA}$ de R^3 se puede representar como la suma de un vector colineal con l y de un vector coplanar con el plano π (véase la fig. 4, donde $AA_1 \parallel l$), estando, además, la intersección $\pi \cap l$ compuesta solamente del vector nulo. El espacio R^3 se descompone también en la suma de dos cualesquiera de sus planos que no coinciden (y que pasan por el origen de coordenadas); pero esta suma no será directa.

Sea R un espacio vectorial cualquiera y sea $a_1, a_2, \ldots, a_k \in R$. El conjunto de todas las combinaciones lineales

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k$$

de estos vectores es, obviamente, un subespacio de R. Diremos que este subespacio *está generado por los vectores* $a_1, a_2, \ldots, a_k$. Este subespacio suele llamarse también *cápsula lineal* de los vectores $a_1, a_2, \ldots, a_k$.

§ 7. Definición de un espacio afín

Más de una vez hemos interpretado en las secciones anteriores el concepto general de espacio vectorial tomando como ejemplo el plano (corriente) o el espacio (corriente de tres dimensiones). Hablando en rigor, estas interpretaciones no son, sin embargo, totalmente exactas: el concepto principal de la Geometría elemental es el *punto* y todas las imágenes geométricas pueden ser comprendidas como *conjuntos de puntos;* por otra parte, en la definición de un espacio vectorial los puntos no figuran.

En el curso elemental de la geometría el concepto del *vector* aparece después del concepto del punto: se llama vector a un par ordenado de puntos $\overline{AB}$ (a un segmento orientado) y después se definen las condiciones de igualdad de los vectores y las reglas de adición y de multiplicación por número.

Ahora tendremos que proceder de un modo diferente. Disponiendo ya de la definición de espacio vectorial, la completaremos introduciendo en la consideración los *puntos*. El conjunto que así resultará (formado por vectores y puntos) —que suele llamarse espacio *vectorial puntual* o espacio *afín*— tendrá ya más semejanza con el espacio que es objeto de estudio en la Geometría elemental, aunque no coincidirá todavía por completo con éste. En efecto, el propio concepto de espacio «afín» presupone que este espacio esta desprovisto de *métrica*, es decir, de un método de medir las longitudes y los ángulos. Este concepto resultará completamente idéntico (por lo menos en los casos de dos y de tres dimensiones) al espacio corriente sólo después de introducir en él la métrica correspondiente (véase el capítulo IV).

Definición 10. *Supongamos que se tiene un espacio vectorial R* (continuaremos indicando sus elementos por letras latinas minúsculas) *y, además, un conjunto de elementos que serán llamados* ***puntos*** y que se indicarán por letras latinas mayúsculas, y supongamos que *a todo par ordenado M, N de puntos se pone en correspondencia uno y sólo un vector x de R* (aun cuando a diferentes pares de puntos puede ponerse en correspondencia un mismo vector); en este caso escribiremos $\overline{MN} = x$. Supongamos que *esta correspondencia entre los puntos y los vectores posee las siguientes propiedades:*

1. *Para todo punto M y para todo vector x existe un punto N, y sólo uno, tal que $\overline{MN} = x$.*

2. *Para cualesquiera tres puntos M, N y P se tiene*

$$\overline{MN}+\overline{NP} = \overline{MP}.$$

Todos los puntos y todos los vectores en conjunto forman un ***espacio afín.***

Se dice que un espacio afín es *n-dimensional* si el espacio vectorial R que le corresponde es n-dimensional.

Es decir, un espacio afín A es un conjunto de elementos de dos géneros: de puntos y de vectores y la relación entre ellos se da mediante la operación de *construcción de vectores*. Un vector arbitrario x se puede construir a partir de cualquier punto M obteniendo así un punto determinado N y entonces $\overline{MN} = x$. El punto M se llama *origen* y el punto N se llama *extremo* del vector $\overline{MN}$.

Son casi evidentes las proposiciones siguientes.

1. *Si* $\overline{MN} = \overline{QP}$, *se tiene* $\overline{MQ} = \overline{NP}$. Esto resulta de la igualdad

$$\overline{MN}+\overline{NP} = \overline{MP} = \overline{MQ}+\overline{QP}.$$

En particular, puesto que

$$\overline{MN}+\overline{NN} = \overline{MN} = \overline{MM}+\overline{MN},$$

se tiene $\overline{MM} = \overline{NN}$, es decir, *son iguales todos los vectores cuyo origen coincide con el extremo*.

2. *El vector cuyo origen coincide con el extremo es el vector nulo*, porque de la igualdad $\overline{MN}+\overline{NN} = \overline{MN}$ resulta que $\overline{NN} = 0$.

3. *El vector* $\overline{NM}$ es el opuesto de $\overline{MN}$, ya que $\overline{MN}+\overline{NM} = \overline{MM} = 0$ y, por consiguiente, $\overline{NM} = -\overline{MN}$.

§ 8. Introducción de coordenadas en un espacio afín

En un espacio afín n-dimensional A las *coordenadas de los puntos* se pueden introducir del modo siguiente. Tomemos un punto cualquiera O a título del *origen de coordenadas*. Para todo vector x existe entonces, en virtud de la condición 1 de la definición 10, un punto X, y sólo uno, tal que $\overline{OX} = x$. Así quedará establecida una correspondencia biyectiva entre todos los puntos y todos los vectores de A: al punto X le corresponderá el vector $x = \overline{OX}$ que tiene por extremo este punto (en el espacio corriente tridimensional esto corresponde a la construcción de todos los vectores a partir del origen de coordenadas).

Tomemos después una *base* cualquiera $e_1, e_2, \ldots, e_n$ en el espacio vectorial R correspondiente a A. Todo vector x de A quedará determinado entonces por la fila de sus coordenadas $x = (x_1, x_2, \ldots, x_n)$. Asignaremos estas mismas coordenadas al punto X que corresponde al vector x, escribiendo $X(x_1, x_2, \ldots, x_n)$.

Luego, si escogemos en un espacio afín n-dimensional A un *sistema de coordenadas* (es decir, un punto O como el origen de coordenadas y una base $e_1, e_2, \ldots, e_n$ del espacio vectorial R correspondiente a A), resulta que a todo punto de A corresponderá unívocamente una fila formada por n números que son las coordenadas de este punto. Todas las coordenadas del punto O son iguales a cero, ya que a este punto le corresponde, obviamente, el vector nulo 0.

Siendo $X(x_1, x_2, \ldots, x_n)$ e $Y(y_1, y_2, \ldots, y_n)$ dos puntos de un espacio afín A, de la igualdad

$$\overline{OX}+\overline{XY}=\overline{OY}$$

se tiene

$$\overline{XY}=\overline{OY}-\overline{OX},$$

es decir, *las coordenadas del vector $\overline{XY}$ son iguales a las diferencias entre las coordenadas de su extremo y de su origen.*

Se puede demostrar que, igual que en el caso de los espacios vectoriales, todos los espacios afines de una misma dimensión tienen, en cierto sentido, una estructura idéntica.

Sean A y A' dos espacios afines de una misma dimensión n. Tomemos en cada uno de ellos un sistema de coordenadas y pongamos en correspondencia a todo punto de A el punto de A' con las mismas coordenadas y a todo vector de A el vector de A' con las mismas coordenadas. Siendo $X \leftrightarrow X'$ e $Y \leftrightarrow Y'$, donde $X(x_1, x_2, \ldots, x_n)$, $Y(y_1, y_2, \ldots, y_n)$ y $X, Y \in A$ y $X', Y' \in A'$, tendremos entonces que el vector $\overline{XY}=(y_1-x_1, y_2-x_2, \ldots, y_n-x_n)$ corresponderá al vector $\overline{X'Y'} \in A'$ (puesto que tienen las mismas coordenadas) con la particularidad de que esta correspondencia entre los espacios vectoriales R y R' (correspondientes a A y a A') será, evidentemente, isomorfa.

Definición 11. *Dos espacios afines A y A' se llaman isomorfos, si entre sus elementos se puede establecer una correspondencia biyectiva, que a los puntos del primer espacio asigna los puntos del segundo y a los vectores del primer espacio asigna los vectores del segundo, tal que:*

1. *Siendo $X \leftrightarrow X'$ e $Y \leftrightarrow Y'$, se tenga $\overline{XY} \leftrightarrow \overline{X'Y'}$.*

2. *La correspondencia que con esto se establece entre los espacios vectoriales R y R'* (correspondientes a A y a A') *sea isomorfa* (en el sentido de la definición 5).

Teorema 7. *Para que dos espacios afines A y A' sean isomorfos es necesario y suficiente que coincidan sus dimensiones.*

Hemos demostrado ya que dos espacios afines de una misma dimensión son isomorfos. La recíproca es obvia, ya que, siendo isomorfos los espacios vectoriales R y R', éstos deben ser, en virtud del teorema 4, de una misma dimensión n. Esta misma dimensión tienen, por definición, los espacios afines correspondientes A y A'.

Luego, todos los espacios afines de una misma dimensión son idénticos y la dimensión de un espacio afín es la *única* característica del mismo. Por esto podemos indicar en lo sucesivo por A^n todo espacio afín n-dimensional.

§ 9. Cambio de sistema de coordenadas

Veamos como se transforman las coordenadas de un punto de un espacio afín A^n al pasar a un sistema nuevo de coordenadas.

Supongamos primero que varía sólo el origen de coordenadas. Supongamos que el origen nuevo ha sido colocado en el punto O' cuyas coordenadas en el sistema antiguo son $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Para todo punto X de A^n tenemos

$$\overline{OO'}+\overline{O'X} = \overline{OX}. \tag{10}$$

Las coordenadas del vector $\overline{OX} = (x_1, x_2, \ldots, x_n)$ son las coordenadas del punto X en el sistema antiguo de coordenadas; las coordenadas del vector $\overline{O'X}=(x_1', x_2', \ldots, x_n')$ son las coordenadas del punto X en el sistema nuevo; las coordenadas del vector $\overline{OO'}$ son las coordenadas del punto O' en el sistema antiguo, es decir, son $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. De la igualdad (10) obtenemos

$$\overline{O'X} = \overline{OX}-\overline{OO'}$$

o, pasando a las coordenadas,

$$(x_1', x_2', \ldots, x_n') = (x_1, x_2, \ldots, x_n)-(\alpha_1, \alpha_2, \ldots, \alpha_n),$$

de donde

$$x_i' = x_i-\alpha_i, \quad i = 1, 2, \ldots, n,$$

es decir, *las nuevas coordenadas de un punto se obtienen restando de sus coordenadas antiguas las coordenadas del origen nuevo en el sistema antiguo de coordenadas.*

Supongamos ahora que el origen de coordenadas no varía, pero que en el espacio vectorial R, correspondiente a A, se escoge una base nueva con la matriz del cambio

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}, \tag{11}$$

es decir, la base antigua formada por los vectores $e_1, e_2, \ldots, e_n$ se sustituye por una nueva formada por los vectores $e'_1, e'_2, \ldots, e'_n$, donde

$$e'_i = a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n, \quad i = 1, 2, \ldots, n.$$

Puesto que las coordenadas del punto X vienen a ser, por definición, las coordenadas del vector $x = \overline{OX}$, resulta, según hemos visto en el § 4, que las coordenadas antiguas del punto se expresan a través de sus coordenadas nuevas por las fórmulas

$$x_i = a_{i1}x'_1 + a_{i2}x'_2 + \ldots + a_{in}x'_n, \text{ donde } i = 1, 2, \ldots, n.$$

En el caso general, en el que el origen de coordenadas O se traslada al punto $O'(\alpha_1, \alpha_2, \ldots, \alpha_n)$ y la base $e_1, e_2, \ldots, e_n$ se sustituye mediante la matriz del cambio (11) por una base nueva, las coordenadas antiguas $x_1, x_2, \ldots, x_n$ de un punto arbitrario X y sus coordenadas nuevas $x'_1, x'_2, \ldots, x'_n$ están ligadas por las relaciones

$$\begin{aligned} x_1 &= a_{11}x'_1 + a_{12}x'_2 + \ldots + a_{1n}x'_n + \alpha_1, \\ x_2 &= a_{21}x'_1 + a_{22}x'_2 + \ldots + a_{2n}x'_n + \alpha_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= a_{n1}x'_1 + a_{n2}x'_2 + \ldots + a_{nn}x'_n + \alpha_n. \end{aligned}$$

§ 10. Variedades lineales

Sea dado un espacio vectorial R^n en el que se ha escogido una base.

Consideremos un sistema (compatible) de ecuaciones lineales, en el caso general no homogéneas:

$$\left\{\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}\right. \tag{12}$$

tal que el rango de la matriz de los coeficientes sea igual a r y sea $k = n-r$.

Definición 12. *El conjunto de los vectores de un espacio R^n cuyas coordenadas satisfacen el sistema de ecuaciones lineales* (12) *se llama* ***variedad lineal.***

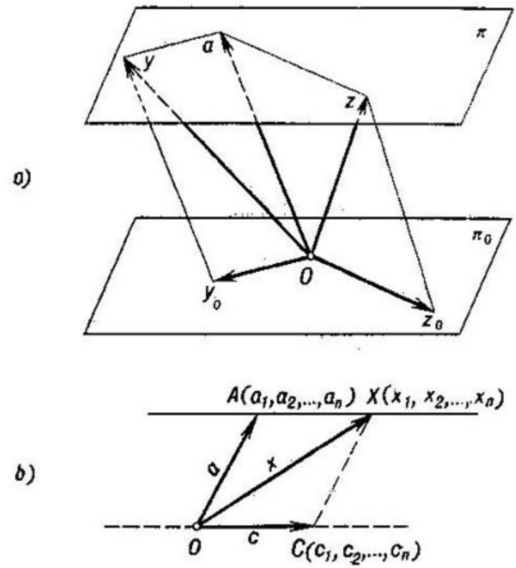


Fig. 5

Según la observación hecha al final del § 10 del capítulo I, la solución general x del sistema (12) es igual a la suma de la solución general x_0 del sistema homogéneo correspondiente (es decir, con los mismos coeficientes de las incógnitas)

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n\ =0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n\ =0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=0 \end{array}\right. \tag{13}$$

y de una solución $a=(a_1, a_2, \ldots, a_n)$ cualquiera, pero fija, del sistema (12). Luego, *la variedad lineal formada por las soluciones del sistema* (12) *se obtiene agregando un mismo vector a a todo vector del subespacio formado por las soluciones del sistema homogéneo correspondiente* (13) (véase la fig. 5*a*, donde los extremos de los vectores que forman la variedad lineal pertenecen a un plano π que se obtiene trasladando paralelamente el subespacio π_0 en el vector a).

Demostremos que, recíprocamente, al agregar un mismo vector a a todo vector de un subespacio R_1 se obtiene una variedad lineal. Supongamos que el subespacio R_1 se determina por el sistema (13) de ecuaciones lineales homogéneas y sea $a = (a_1, a_2, \ldots, a_n)$. Tomemos

$$\begin{aligned} &a_{11}a_1 + a_{12}a_2 + \ldots + a_{1n}a_n = b_1, \\ &a_{21}a_1 + a_{22}a_2 + \ldots + a_{2n}a_n = b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{m1}a_1 + a_{m2}a_2 + \ldots + a_{mn}a_n = b_m \end{aligned}$$

y consideremos el sistema de ecuaciones

$$\left\{\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{aligned}\right. \tag{14}$$

El vector a es una solución de este sistema (en general, no homogéneo). Por consiguiente, la variedad lineal que determina el sistema (14) coincide con el conjunto $R_1 + a$ de los vectores dados.

La variedad lineal (12) se llama k-dimensional si es k-dimensional el subespacio (13) que le corresponde.

§ 11. Planos k-dimensionales en un espacio afín

Supongamos que en un espacio afín A^n se ha tomado un sistema de coordenadas. Consideremos de nuevo el sistema de ecuaciones (12).

Definición 13. *El conjunto de todos los puntos de A^n cuyas coordenadas satisfacen al sistema de ecuaciones* (12) *se llama plano k-dimensional; los planos unidimensionales suelen llamarse también rectas y los planos* $(n-1)$*-dimensionales,* ***hiperplanos.***

Está claro que todo hiperplano (para éstos se tiene $r = 1$) se puede determinar por una sola ecuación lineal

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b.$$

En el espacio de tres dimensiones los hiperplanos son los planos corrientes y en el plano corriente, son simplemente las rectas.

Se puede demostrar que al pasar a un sistema nuevo de coordenadas en A^n, los puntos que satisfacen al sistema de ecuaciones (12) satisfarán a un sistema nuevo de ecuaciones lineales cuya matriz de los coeficientes es también de rango igual a r.

Sea π un plano k-dimensional determinado por el sistema de ecuaciones (12). El sistema correspondiente (13) de ecuaciones lineales homogéneas también determina un plano k-dimensional π_0 que «pasa por el origen de coordenadas». Si construimos todos los vectores a partir del origen de coordenadas, los vectores cuyos extremos pertenecen a π_0 forman un subespacio y los vectores cuyos extremos pertenecen a π forman una variedad lineal k-dimensional. Esta variedad se obtiene agregando un mismo vector a a todos los vectores del subespacio π_0. Podemos decir, por esto, que *el plano k-dimensional π se obtiene de π_0* ***mediante una translación paralela de vector*** *a*. Esto nos sirve de base para introducir la definición siguiente.

Definición 14. *Dos planos k-dimensionales se dicen* ***paralelos*** *si son equivalentes* (es decir, tienen las mismas soluciones) *los sistemas homogéneos correspondientes a los sistemas que determinan estos planos. Se dice que un plano k-dimensional π_1 y un plano l-dimensional π_2 son paralelos* (siendo $l > k$) *si π_1 es paralelo a un plano k-dimensional contenido en π_2* (en este caso los sistemas que determinan a π_1 y a π_2 son tales que el sistema homogéneo correspondiente a π_2 es un **corolario** del sistema homogéneo correspondiente a π_1).

Sea de nuevo π un plano k-dimensional determinado por el sistema de ecuaciones (12). En términos vectoriales la *solución general* del sistema (12) toma la forma

$$x = \alpha_1 c_1 + \alpha_2 c_2 + \ldots + \alpha_k c_k + a, \tag{15}$$

donde

$$x_0 = \alpha_1 c_1 + \alpha_2 c_2 + \ldots + \alpha_k c_k$$

es la solución general del sistema homogéneo correspondiente (13) y a es un vector fijo (una de las soluciones del sistema (12)). Siendo $c_i = (c_{i1}, c_{i2}, \ldots, c_{in})$ para $i = 1, 2, \ldots, k$ y $a = (a_1, a_2, \ldots, a_n)$ y expresando la igualdad (15) en coordenadas, obtenemos *las ecuaciones paramétricas de un plano k-dimensional:*

$$x_j = \alpha_1 c_{1j} + \alpha_2 c_{2j} + \ldots + \alpha_n c_{nj} + a_j, \quad j = 1, 2, \ldots, n.$$

Si el rango de la matriz de los coeficientes del sistema (12) es igual a $n-1$, hemos convenido en llamar *recta* al plano (unidimensional) correspondiente. En este caso la solución general del sistema (12) es, en forma vectorial,

$$x = \alpha c + a, \tag{16}$$

donde $x_0 = \alpha c$ es la solución general del sistema homogéneo correspondiente y $a = (a_1, a_2, \ldots, a_n)$ es un vector fijo (fig. 5, *b*). Siendo $c = (c_1, c_2, \ldots, c_n)$ y $x = (x_1, x_2, \ldots, x_n)$ y expresando la igualdad (16) en coordenadas, obtenemos *las ecuaciones paramétricas de una recta:*

$$x_1 = \alpha c_1 + a_1,$$
$$x_2 = \alpha c_2 + a_2,$$
$$\cdots\cdots\cdots\cdots$$
$$x_n = \alpha c_n + a_n,$$

que excluyendo el parámetro α pueden ser reducidas a la forma

$$\frac{x_1 - a_1}{c_1} = \frac{x_2 - a_2}{c_2} = \ldots = \frac{x_n - a_n}{c_n}.$$

(Éstas son *las ecuaciones canónicas de una recta.* Tienen sentido incluso en el caso en que se anulan algunos de los denominadores: en este caso también son iguales a cero los numeradores correspondientes.)

Si $A(a_1, a_2, \ldots, a_n)$ y $B(b_1, b_2, \ldots, b_n)$ son dos puntos del espacio A^n, *la recta AB* que pasa por estos puntos se determina, obviamente, por las ecuaciones

$$\frac{x_1 - a_1}{b_1 - a_1} = \frac{x_2 - a_2}{b_2 - a_2} = \ldots = \frac{x_n - a_n}{b_n - a_n}. \tag{17}$$

Indicando por β cada una de las razones iguales de (17), obtenemos *las ecuaciones paramétricas de la recta AB* en la forma

$$x_i = a_i + \beta(b_i - a_i)$$

es decir,

$$x_i = (1 - \beta)\, a_i + \beta b_i, \quad i = 1, 2, \ldots, n.$$

Tomando $1 - \beta = \alpha$, tendremos

$$x_i = \alpha a_i + \beta b_i, \text{ donde } \alpha + \beta = 1. \tag{18}$$

Si $0 \leqslant \alpha, \beta \leqslant 1$, se dice que el punto correspondiente X pertenece *al segmento AB.*

Si el rango de la matriz de los coeficientes del sistema (12) es igual a $n-2$, el plano que este sistema determina es *bidimensional* y la solución general de este sistema en forma vectorial es

$$x = \alpha_1 c_1 + \alpha_2 c_2 + a,$$

donde $x_0 = \alpha_1 c_1 + \alpha_2 c_2$ es la solución general del sistema homogéneo correspondiente y $a = (a_1, a_2, \ldots, a_n)$ es un vector fijo. Siendo $c_i = (c_{i1}, c_{i2}, \ldots, c_{in})$, $i = 1, 2$, las coordenadas de los puntos $X = (x_1, x_2, \ldots, x_n)$ de este plano se determinan por las fórmulas

$$x_k = \alpha_1 c_{1k} + \alpha_2 c_{2k} + a_k, \quad k = 1, 2, \ldots, n$$

(que representan *las ecuaciones paramétricas de un plano bidimensional*).

Supongamos que en el espacio afín A^n se han tomado dos planos: el plano k-dimensional π determinado por el sistema de ecuaciones (12) y el plano l-dimensional π' determinado por el sistema de ecuaciones

$$\left\{\begin{array}{l} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n = d_1, \\ c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n = d_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_{s1}x_1 + c_{s2}x_2 + \ldots + c_{sn}x_n = d_s. \end{array}\right. \tag{19}$$

La *intersección* de estos planos (es decir, el conjunto de los puntos que pertenecen a la vez a π y a π') quedará determinada entonces por el sistema formado de todas las ecuaciones del sistema (12) y de todas las ecuaciones del sistema (19) y, por consiguiente, también será, en general, un plano (que, en particular, puede estar formado por un punto sólo e incluso puede no contener ningún punto, si el sistema unido resulta incompatible). Es fácil ver que *todo plano k-dimensional es la intersección de $r = n-k$ hiperplanos.*

§ 12. Conjuntos convexos en un espacio afín

Definición 15. *Un conjunto de puntos de un espacio afín se llama* ***convexo*** *si con dos cualesquiera puntos suyos A y B contiene también todos los puntos del segmento AB.*

Es fácil ver que *la intersección de cualquier número de conjuntos convexos es un conjunto convexo.*

Definición 16. *Un conjunto de puntos de un espacio afín se llama* ***acotado*** *si son acotadas las coordenadas de todos sus puntos en un sistema de coordenadas* (es fácil ver que éstas serán entonces acotadas en todos los sistemas de coordenadas).

Sea dado en un espacio afín A^n un hiperplano

$$a_1x_1+a_2x_2+\ldots+a_nx_n=b. \tag{20}$$

Este hiperplano divide todos los puntos de A^n en dos *semiespacios:* A_1 que es el conjunto de todos los puntos para los cuales $a_1x_1+a_2x_2+\ldots+a_nx_n\geqslant b$ y A_2 que es el conjunto de los puntos para los cuales $a_1x_1+a_2x_2+\ldots+a_nx_n\leqslant b$. La intersección de los semiespacios A_1 y A_2 es el propio hiperplano (20).

Teorema 8. *Todo semiespacio de un espacio afín A^n es un conjunto convexo.*

Demostración. Sean $P(p_1, p_2, \ldots, p_n)$ y $Q(q_1, q_2, \ldots, q_n)$ dos puntos de A^n pertenecientes, por ejemplo, a A_1; en este caso

$$a_1p_1+a_2p_2+\ldots+a_np_n\geqslant b$$

y

$$a_1q_1+a_2q_2+\ldots+a_nq_n\geqslant b.$$

Si $X(x_1, x_2, \ldots, x_n)$ es un punto cualquiera del segmento PQ, tenemos según las fórmulas (18)

$$x_i=\alpha p_i+\beta q_i, \quad i=1, 2, \ldots, n,$$

donde $\alpha, \beta\geqslant 0$ y $\alpha+\beta=1$. Para este punto X tenemos

$$\begin{aligned} a_1x_1+a_2x_2+\ldots+a_nx_n &= a_1(\alpha p_1+\beta q_1)+\\ +a_2(\alpha p_2+\beta q_2)+\ldots+a_n(\alpha p_n+\beta q_n) &= \alpha(a_1p_1+a_2p_2+\ldots+a_np_n)+\\ &+\beta(a_1q_1+a_2q_2+\ldots+a_nq_n)\geqslant \alpha b+\beta b=(\alpha+\beta)\,b=b, \end{aligned}$$

es decir, un punto arbitrario X del segmento PQ pertenece a A_1.

El hiperplano (20) *es un conjunto convexo* por ser la intersección de dos conjuntos convexos A_1 y A_2. *Todo plano k-dimensional de A^n es convexo* porque es la intersección de varios hiperplanos.

Supongamos que en A^n se tienen m semiespacios definidos por las desigualdades

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n\geqslant b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n\geqslant b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n\geqslant b_m. \end{array}\right. \tag{21}$$

(Todos los signos de desigualdad son aquí del mismo sentido, ya que esta situación puede siempre alcanzarse multiplicando, siempre que esto sea necesario, ambos miembros de la igualdad por -1.) La

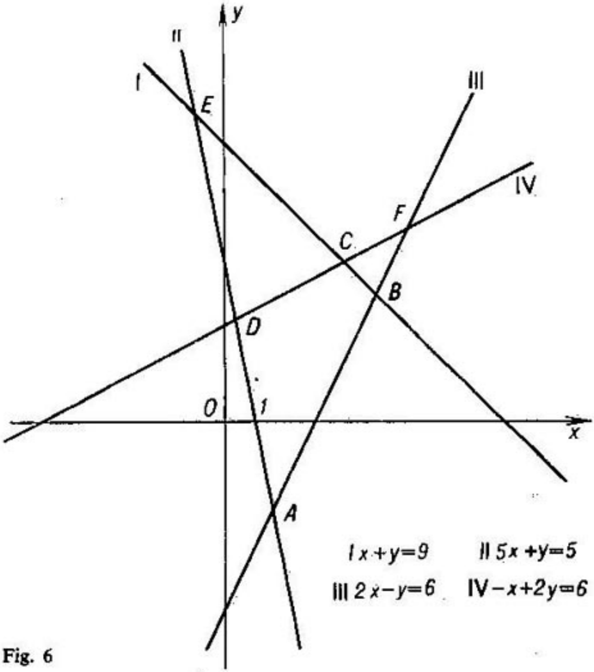


Fig. 6

intersección de estos semiespacios —que es denominada *región poliedra convexa*— determina *el conjunto de las soluciones del sistema de desigualdades lineales* (21). Si esta intersección es acotada se llama *poliedro (convexo) del espacio n-dimensional* A^n.

Así, sobre el plano (bidimensional) el sistema de desigualdades

$$\begin{cases} x+y \leqslant 9, \\ 5x+y \geqslant 5, \\ 2x-y \leqslant 6, \\ 2y-x \leqslant 6 \end{cases} \tag{22}$$

determina un cuadrilatero (convexo) $ABCD$ (fig. 6). El sistema de desigualdades

$$\begin{cases} x+y \geqslant 9, \\ 5x+y \leqslant 5, \\ 2x-y \geqslant 6, \\ 2y-x \geqslant 6 \end{cases}$$

(todas las desigualdades del sistema (22) han cambiado de signo) determina un conjunto vacío de puntos. El sistema de desigualdades

$$\begin{cases} x+y \leqslant 9, \\ 5x+y \geqslant 5, \\ 2x-y \leqslant 6 \end{cases}$$

determina la región triangular ABE y el sistema de desigualdades

$$\begin{cases} x+y \leqslant 9, \\ 5x+y \geqslant 5, \\ 2y-x \leqslant 6 \end{cases}$$

determina una región triangular no acotada.

CAPÍTULO III

APLICACIONES LINEALES

§ 1. Definición y ejemplos

Definición 1. Se dice que *en un espacio lineal R está definida una aplicación $\mathcal{A}$ si a todo vector $x \in R$ corresponde un vector determinado* $\mathcal{A}(x)$ que generalmente se indicará por $\mathcal{A}x$. *Una aplicación $\mathcal{A}$ se llama **lineal** si*

1. $\mathcal{A}(x+y) = \mathcal{A}x+\mathcal{A}y$,
2. $\mathcal{A}(\alpha x) = \alpha\mathcal{A}x$.

para cualesquiera dos vectores x e y de R y para un número arbitrario real α. El vector $\mathcal{A}x$ se llama imagen del vector x.

Tomemos en el espacio R una base $e_1, e_2, \ldots, e_n$. Siendo $x = x_1e_1+x_2e_2+\ldots+x_ne_n$ y siendo $\mathcal{A}$ una aplicación lineal, tenemos

$$\mathcal{A}x = x_1\mathcal{A}e_1+x_2\mathcal{A}e_2+\ldots+x_n\mathcal{A}e_n.$$

Pero $\mathcal{A}e_i$ (donde $i = 1, 2, \ldots, n$) es también un vector de R, es decir, se puede descomponer $\mathcal{A}e_i$ respecto a la base $e_1, e_2, \ldots, e_n$; sea

$$\mathcal{A}e_i = a_{1i}e_1+a_{2i}e_2+\ldots+a_{ni}e_n.$$

Entonces

$$\mathcal{A}x = x_1(a_{11}e_1+a_{21}e_2+\ldots+a_{n1}e_n)+x_2(a_{12}e_1+a_{22}e_2+\ldots$$
$$\ldots+a_{n2}e_n)+\ldots+x_n(a_{1n}e_1+a_{2n}e_2+\ldots+a_{nn}e_n)=(a_{11}x_1+a_{12}x_2+$$
$$+\ldots+a_{1n}x_n)e_1+(a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n)e_2+\ldots+(a_{n1}x_1+$$
$$+a_{n2}x_2+\ldots+a_{nn}x_n)\,e_n.$$

Si $x_1', x_2', \ldots, x_n'$ son las coordenadas del vector (transformado) $\mathcal{A}x$ en esta misma base $e_1, e_2, \ldots, e_n$, es decir, si

$$\mathcal{A}x = x_1'e_1 + x_2'e_2 + \ldots + x_n'e_n,$$

tenemos debido a la unicidad de la descomposición de un vector según una base

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n. \end{aligned} \tag{1}$$

En una base dada $e_1, e_2, \ldots, e_n$ a toda aplicación lineal $\mathcal{A}$ corresponde de esta forma una matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}.$$

La i-ésima columna de la matriz A está formada por los coeficientes de la descomposición del vector $\mathcal{A}e_i$ según la base $e_1, e_2, \ldots, e_n$ y los coeficientes de las descomposiciones (1) de las coordenadas del vector $\mathcal{A}x$ respecto a las coordenadas del vector x forman las filas de la matriz A.

Dada una base de un espacio R de n dimensiones, además de corresponder toda aplicación lineal $\mathcal{A}$ a una determinada matriz A, resulta que, recíprocamente, *toda matriz cuadrada A de orden n puede ser considerada como la matriz de cierta aplicación lineal.* Efectivamente, sea $e_1, e_2, \ldots, e_n$ una base del espacio R y sea dada una matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

(escribiremos abreviadamente $A = [a_{ik}]$). Indiquemos por $\mathcal{A}$ la aplicación que transforma un vector cualquiera

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

en el vector

$$\mathcal{A}x = x'_1e_1 + x'_2e_2 + \ldots + x'_ne_n,$$

donde

$$x'_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \text{ para } i = 1, 2, \ldots, n.$$

Demostremos que esta aplicación es *lineal*. En efecto, esta aplicación transforma otro vector cualquiera

$$y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$$

en

$$\mathcal{A}y = y'_1e_1 + y'_2e_2 + \ldots + y'_ne_n,$$

donde

$$y'_i = a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{in}y_n,$$

y por ello el vector

$$x+y = (x_1+y_1)e_1 + (x_2+y_2)e_2 + \ldots + (x_n+y_n)\,e_n$$

se transforma en el vector

$$\mathcal{A}(x+y) = z_1e_1 + z_2e_2 + \ldots + z_ne_n,$$

donde

$$z_i = a_{i1}(x_1+y_1) + a_{i2}(x_2+y_2) + \ldots + a_{in}(x_n+y_n) = x'_i + y'_i.$$

Luego,

$$\mathcal{A}(x+y) = \mathcal{A}x + \mathcal{A}y.$$

Además, cualquiera que sea α se tiene

$$\alpha x = (\alpha x_1)\,e_1 + (\alpha x_2)\,e_2 + \ldots + (\alpha x_n)\,e_n$$

y

$$\mathcal{A}(\alpha x) = t_1e_1 + t_2e_2 + \ldots + t_ne_n,$$

donde

$$t_i = a_{i1}(\alpha x_1) + a_{i2}(\alpha x_2) + \ldots + a_{in}(\alpha x_n) = \alpha x'_i\,.$$

Por consiguiente,

$$\mathcal{A}(\alpha x) = \alpha\mathcal{A}x,$$

es decir, $\mathcal{A}$ es una aplicación lineal.

Por esto, *dada una base en un espacio vectorial R^n, a toda aplicación lineal le corresponde una determinada matriz cuadrada de orden n y, recíprocamente, a toda matriz de este tipo le corresponde una determinada aplicación lineal.* Debido a ello indicaremos con una misma letra la aplicación lineal y la matriz que le corresponde (en la base dada); así, $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ son unas aplicaciones lineales y $A, B, C, \ldots$ son las matrices que les corresponden. La matriz A se denomina *matriz de la aplicación lineal* $\mathcal{A}$.

Es fácil ver que

$$\mathcal{A}0 = 0$$

cualquiera que sea la aplicación lineal $\mathcal{A}$. Siendo la aplicación $\mathcal{A}$ tal que $\mathcal{A}x = 0$ *sólo* para $x = 0$, se dice que esta aplicación es *no degenerada;* en el caso contrario la aplicación $\mathcal{A}$ es *degenerada.*

Sea $A = [a_{ik}]$ la matriz de una aplicación lineal $\mathcal{A}$. Consideremos el sistema siguiente de ecuaciones lineales homogéneas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0. \end{cases}$$

Para que exista una solución no nula de este sistema (y, por consiguiente, para que exista un vector no nulo $x = x_1e_1 + x_2e_2 + \ldots$ $\ldots + x_ne_n$ tal que $\mathcal{A}x = 0$) es necesario y suficiente, según hemos visto en el teorema 10 del capítulo 1, que el determinante de la matriz A (que indicaremos por $|A|$) sea igual a cero. Luego, *para que una aplicación $\mathcal{A}$ sea no degenerada es necesario y suficiente que el determinante de la matriz A de esta aplicación sea diferente de cero* (¡como quiera que se escoja la base!). Una matriz cuyo determinante es diferente de cero se llama *matriz no degenerada.*

Consideremos algunos ejemplos.

1. Sea $\mathcal{A}$ una rotación de todos los vectores $\overline{OX}$ del plano xOy (o simplemente una rotación del plano xOy) alrededor del origen de coordenadas en un ángulo φ en dirección opuesta a la de las saetas de un reloj. Esta aplicación es lineal porque da lo mismo sumar primero los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ y girar después la suma de éstos en el ángulo φ o girar primero los vectores y sumarlos después (fig. 7, *a*); igualmente da lo mismo multiplicar primero el vector $\boldsymbol{a}$ por un número α y girarlo después en el ángulo φ o realizar estas operaciones en el orden inverso (fig. 7, *b*).

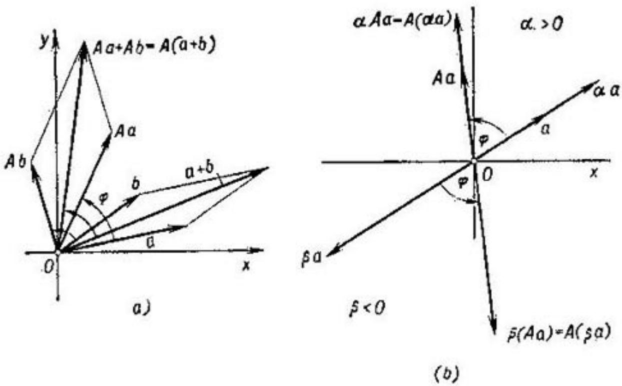


Fig. 7

Supongamos que los vectores básicos son unitarios y recíprocamente ortogonales. El vector $\mathcal{A}e_1$ es un vector unitario que forma el ángulo φ con e_1 y el ángulo $\varphi-\frac{\pi}{2}$ con e_2. Por consiguiente,

$$\mathcal{A}e_1=\cos\varphi\cdot e_1+\operatorname{sen}\varphi\cdot e_2 .$$

El vector unitario $\mathcal{A}e_2$ forma el ángulo $\frac{\pi}{2}+\varphi$ con e_1 y el ángulo φ con e_2. Por consiguiente,

$$\mathcal{A}e_2=-\operatorname{sen}\varphi\cdot e_1+\cos\varphi\cdot e_2 .$$

Es decir,

$$A=\begin{bmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{bmatrix}.$$

2. Sea $\mathcal{A}$ una *rotación del espacio* corriente *de tres dimensiones* en un ángulo φ alrededor del eje Oz. Siendo e_1, e_2 y e_3 los vectores unitarios del sistema cartesiano rectangular de coordenadas, tendremos

$$\mathcal{A}e_1=\cos\varphi\cdot e_1+\operatorname{sen}\varphi\cdot e_2 ,$$
$$\mathcal{A}e_2=-\operatorname{sen}\varphi\cdot e_1+\cos\varphi\cdot e_2 ,$$
$$\mathcal{A}e_3=e_3 ,$$

es decir, la matriz de esta aplicación será

$$A = \begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. Sea $\mathcal{A}a$ la *proyección ortogonal* del vector *a sobre el plano* xOy en el espacio corriente de tres dimensiones. Esta aplicación es *lineal* debido a que la proyección de una suma de vectores es igual a la suma de la proyecciones de los sumandos y a que la proyección del producto de un vector por un número es igual al producto de la proyección del vector por este número. Es evidente que escogiendo la base igual que en el ejemplo 2 tendremos

$$\mathcal{A}e_1 = e_1, \quad \mathcal{A}e_2 = e_2, \quad \mathcal{A}e_3 = 0$$

y, por consiguiente,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

4. Sea $\mathcal{A}a$ el vector *simétrico del vector a respecto al plano* xOy en el espacio corriente de tres dimensiones. Esta aplicación es, obviamente, lineal. Además,

$$\mathcal{A}e_1 = e_1, \quad \mathcal{A}e_2 = e_2, \quad \mathcal{A}e_3 = -e_3$$

y la matriz de esta aplicación es

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

5. Tomemos en el espacio P_n de los polinomios en t de grado no mayor que n

$$\mathcal{A}(x(t)) = x'(t).$$

Esta «aplicación de derivación» es lineal según se deduce de las reglas principales del cálculo diferencial. Para hallar la matriz de esta aplicación tomemos como base, por ejemplo, los vectores

$$e_0 = 1,\ e_1 = t,\ e_2 = \frac{t^2}{2!},\ \ldots,\ e_n = \frac{t^n}{n!}.$$

Entonces

$$\mathcal{A}e_0 = 0,\ \mathcal{A}e_1 = e_0,\ \mathcal{A}e_2 = e_1,\ \ldots,\ \mathcal{A}e_n = e_{n-1}$$

y

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

6. Indiquemos por $\mathcal{E}$ la así llamada *aplicación idéntica* del espacio lineal R definida por la igualdad $\mathcal{E}x = x$ *cualquiera* que sea $x \in R$. Luego, $\mathcal{E}e_i = e_i$ para todo $i = 1, 2, \ldots, n$ y, por consiguiente, la matriz de la aplicación $\mathcal{E}$ en *cualquier* base es de la forma

$$E = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}.$$

7. Indiquemos por O la así llamada *aplicación nula* de un espacio lineal R definida por la igualdad $Ox = 0$ para todo $x \in R$. La matriz de esta aplicación está formada solamente por ceros

$$O = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Está claro que las aplicaciones de los ejemplos 1, 2, 4 y 6 son no degeneradas y las aplicaciones de los ejemplos 3, 5 y 7 son degeneradas.

Teorema 1. *Una aplicación lineal de un espacio vectorial transforma todo subespacio en un subespacio.*

Demostración. Sea R_1 un subespacio de un espacio vectorial R^n. Indiquemos por $\mathcal{A}R_1$ el conjunto de todos los vectores que son imágenes de los vectores de R_1 en la aplicación lineal $\mathcal{A}$. Debemos demostrar que $\mathcal{A}R_1$ es un subespacio. Supongamos que los vectores x e y pertenecen a $\mathcal{A}R_1$. Esto significa que $x = \mathcal{A}x'$ e $y = \mathcal{A}y'$, donde $x' \in R_1$ e $y' \in R_1$. Pero en estas condiciones resulta que $x+y = \mathcal{A}x'+\mathcal{A}y' = \mathcal{A}(x'+y') \in \mathcal{A}R_1$, (ya que $x'+y' \in R_1$) y que $\alpha x = \alpha\mathcal{A}x' = \mathcal{A}(\alpha x') \in \mathcal{A}R_1$ (ya que $\alpha x' \in R_1$) para todo α. Por consiguiente, $\mathcal{A}R_1$ es un subespacio.

(Es fácil ver que la dimensión de $\mathcal{A}R_1$ no sobrepasa la dimensión de R_1).

Teorema 2. *Una aplicación lineal de un espacio vectorial transforma toda variedad lineal en una variedad lineal.*

Demostración. Sea M una variedad lineal de R^n. Existen entonces un subespacio R_1 y un vector a tales que $M = R_1+a$ (véase la pág. 88). Si $\mathcal{A}$ es una aplicación lineal, se tiene $\mathcal{A}M = \mathcal{A}R_1+\mathcal{A}a$. El conjunto $\mathcal{A}R_1$ es, según el teorema 1, un subespacio lineal y, por consiguiente, el conjunto $\mathcal{A}M$ es una variedad lineal (véase la pág. 89).

Sea A^n un espacio afín de n dimensiones, sea R^n su correspondiente espacio vectorial y sea $\mathcal{A}$ una aplicación lineal dada en éste. Esta aplicación se puede extender, con el razonamiento que sigue, también a los *puntos* de A^n. Supongamos que en A^n se ha escogido un sistema de coordenadas. Si el vector $x = x_1e_1+x_2e_2+\ldots+x_ne_n$ se transforma en la aplicación $\mathcal{A}$ en el vector $\mathcal{A}x = x'_1e_1+x'_2e_2+\ldots+x'_ne_n$, se toma por definición que el punto $X(x_1, x_2, \ldots, x_n)$ (que es el extremo del vector $\overline{OX} = x$) se transforma en el punto $X'(x'_1, x'_2, \ldots, x'_n)$ (que es el extremo del vector $\overline{OX'} = \mathcal{A}x$).

Del teorema 2 se deduce directamente que *una aplicación lineal de un espacio afín transforma los planos k-dimensionales en planos* (de dimensión no mayor). En particular, las rectas se transforman en rectas o en puntos.

Se llama *aplicación afín* una aplicación biyectiva[1] del conjunto de puntos de un espacio afín n-dimensional sobre sí mismo que transforma toda recta en una recta. Del teorema 2 resulta que *toda aplicación lineal no degenerada de un espacio afín es una aplicación afín*. Se puede demostrar que toda aplicación afín se reduce a una aplicación lineal (no degenerada) seguida, posiblemente, de una traslación paralela (véase la pág. 264).

---

[1] Se dice que se tiene una *aplicación biyectiva* de un conjunto M sobre un conjunto N (en particular, de M sobre M) si a todo elemento $a \in M$ se pone en correspondencia un determinado elemento $b \in N$ correspondiendo, además, todo elemento $b \in N$ a un determinado elemento $a \in M$. Se dice también en este caso que entre los conjuntos A y B se ha establecido una *correspondencia biyectiva* (véase las pág. 72 y 85).

§ 2. Operaciones con aplicaciones lineales

A. Adición de aplicaciones lineales

*Si $\mathcal{A}$ y $\mathcal{B}$ son dos aplicaciones lineales de un espacio lineal R, se llama **suma** $\mathcal{A}+\mathcal{B}$ a una aplicación $\mathcal{C}$ tal que*

$$\mathcal{C}x = \mathcal{A}x+\mathcal{B}x$$

para todo $x \in R$. Es fácil ver que la suma de unas aplicaciones lineales es también una aplicación *lineal*. Si las matrices de las aplicaciones lineales $\mathcal{A}$ y $\mathcal{B}$ (en una base determinada) son $A = [a_{ik}]$ y $B = [b_{ik}]$, respectivamente, la matriz de la aplicación $\mathcal{C} = \mathcal{A}+\mathcal{B}$ es $C = [c_{ik}]$, donde

$$c_{ik} = a_{ik}+b_{ik}.$$

La matriz C se llama *suma de las matrices* A y B. Es decir, se toma por definición que

$$[a_{ik}]+[b_{ik}] = [a_{ik}+b_{ik}].$$

(Se pueden sumar, naturalmente, sólo matrices de un mismo orden.)

Es evidente que la operación de adición de aplicaciones lineales (y de adición de matrices) posee las siguientes propiedades:

1. $\mathcal{A}+\mathcal{B} = \mathcal{B}+\mathcal{A}$.
2. $(\mathcal{A}+\mathcal{B})+\mathcal{C} = \mathcal{A}+(\mathcal{B}+\mathcal{C})$.
3. $\mathcal{A}+\mathcal{O} = \mathcal{A}$ para toda aplicación $\mathcal{A}$.
4. Indicando por $-\mathcal{A}$ la aplicación definida por la condición $(-\mathcal{A})x = -\mathcal{A}x$ para todo $x \in R$, resulta que $-\mathcal{A}$ es una aplicación lineal y que

$$(-\mathcal{A})+\mathcal{A} = \mathcal{O}.$$

Indicaremos por $-A$ la matriz de la aplicación $-\mathcal{A}$; está claro que siendo $A = [a_{ik}]$ se tiene $-A = [-a_{ik}]$.

B. Multiplicación de una aplicación lineal por un número

*Si $\mathcal{A}$ es una aplicación lineal de un espacio R y α es un número, se llama **producto de** $\mathcal{A}$ por α a la aplicación $\alpha\mathcal{A}$ tal que*

$$(\alpha\mathcal{A})\,x = \alpha\,\mathcal{A}x$$

para todo vector $x \in R$.

Está claro que $\alpha\mathcal{A}$ es también una aplicación lineal y que su matriz αA se obtiene multiplicando por α todos los elementos de la matriz A de la aplicación $\mathcal{A}$:

$$\alpha[a_{ik}] = [\alpha a_{ik}].$$

La matriz αA se denomina *producto de la matriz A por el número α.*

Es evidente que para la multiplicación de una aplicación lineal por un número son válidas las identidades siguientes:

1. $1\cdot\mathcal{A} = \mathcal{A}$.
2. $\alpha(\beta\mathcal{A}) = (\alpha\beta)\,\mathcal{A}$.
3. $(\alpha+\beta)\,\mathcal{A} = \alpha\mathcal{A}+\beta\mathcal{A}$.
4. $\alpha(\mathcal{A}+\mathcal{B}) = \alpha\mathcal{A}+\alpha\mathcal{B}$.

Para la multiplicación de una matriz por un número son válidas unas identidades análogas.

C. Multiplicación de aplicaciones lineales

Se llama ***producto*** *$\mathcal{A}\mathcal{B}$ de unas aplicaciones $\mathcal{A}$ y $\mathcal{B}$ a la aplicación $\mathcal{C}$ tal que*

$$\mathcal{C}x = \mathcal{A}\,(\mathcal{B}x)$$

para todo vector $x \in R$.

Por consiguiente, la multiplicación de aplicaciones se reduce *a la realización sucesiva de las mismas*, con la particularidad de que primero se realiza la aplicación $\mathcal{B}$ y después el vector obtenido $\mathcal{B}x$ se somete a la aplicación $\mathcal{A}$.

Por ejemplo, siendo $\mathcal{A}$ una rotación del plano en ángulo φ en el sentido opuesto al del movimiento de las saetas de un reloj y siendo $\mathcal{B}$ una rotación en un ángulo ψ (en el mismo sentido), resulta que $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$ es una rotación de ángulo $\varphi+\psi$. Si $\mathcal{A}$ es la reflexión simétrica del plano xOy respecto al eje Ox y $\mathcal{B}$ es la reflexión simétrica respecto al eje Oy, resulta que $\mathcal{A}^2 = \mathcal{B}^2 = \mathcal{E}$ es la aplicación idéntica y que $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A} = \mathcal{S}$ es la simetría respecto al origen de coordenadas. Si $\mathcal{A}$ es una proyección ortogonal de todo el espacio sobre un plano o sobre una recta, tenemos que $\mathcal{A}^2 = \mathcal{A}$. Si $\mathcal{A}$ es la derivación en el espacio de los polinomios, la aplicación $\mathcal{A}^2$ consiste en tomar la segunda derivada.

El producto de unas aplicaciones lineales es también una aplicación lineal.

Efectivamente,

1. $(\mathcal{A}\mathcal{B})(x+y) = \mathcal{A}(\mathcal{B}(x+y)) = \mathcal{A}(\mathcal{B}x+\mathcal{B}y) =$
$$= \mathcal{A}(\mathcal{B}x)+\mathcal{A}(\mathcal{B}y) = (\mathcal{A}\mathcal{B})\,x+(\mathcal{A}\mathcal{B})\,y.$$
2. $(\mathcal{A}\mathcal{B})(\alpha x) = \mathcal{A}(\mathcal{B}(\alpha x)) = \mathcal{A}(\alpha\mathcal{B}x) =$
$$= \alpha\mathcal{A}(\mathcal{B}x) = \alpha(\mathcal{A}\mathcal{B})\,x.$$

Veamos cómo se expresa la matriz C de la aplicación lineal $\mathcal{C} = \mathcal{A}\mathcal{B}$ mediante las matrices $A = [a_{ik}]$ y $B = [b_{ik}]$. Tenemos
$$\begin{aligned}\mathcal{C}e_k = \mathcal{A}(\mathcal{B}e_k) = \mathcal{A}(b_{1k}e_1+b_{2k}e_2+\ldots+b_{nk}e_n) &= \\ = b_{1k}\mathcal{A}e_1+b_{2k}\mathcal{A}e_2+\ldots+b_{nk}\mathcal{A}e_n &= \\ = b_{1k}(a_{11}e_1+a_{21}e_2+\ldots+a_{n1}e_n)&+ \\ +b_{2k}(a_{12}e_1+a_{22}e_2+\ldots+a_{n2}e_n)&+\ldots \\ \ldots+b_{nk}(a_{1n}e_1+a_{2n}e_2+\ldots+a_{nn}e_n) &= \\ = (a_{11}b_{1k}+a_{12}b_{2k}+\ldots+a_{1n}b_{nk})\,e_1&+ \\ +(a_{21}b_{1k}+a_{22}b_{2k}+\ldots+a_{2n}b_{nk})\,e_2&+\ldots \\ \ldots+(a_{n1}b_{1k}+a_{n2}b_{2k}+\ldots+a_{nn}b_{nk})\,e_n&;\end{aligned}$$
es decir, si $\mathcal{C}e_k = c_{1k}e_1+c_{2k}e_2+\ldots+c_{nk}e_n$, se tiene
$$c_{ik} = a_{i1}b_{1k}+a_{i2}b_{2k}+\ldots+a_{in}b_{nk}.$$
Vemos, pues, que *para obtener el elemento de la matriz C que aparece en el cruce de su i-ésima fila y de su k-ésima columna es preciso multiplicar cada elemento de la i-ésima fila de la matriz A por el elemento correspondiente de la k-ésima columna de la matriz B y sumar todos los productos obtenidos.* La matriz C se denomina *producto de las matrices A* y *B*.

Ejemplos.

1. $$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & 2 \\ 3 & 2 & -1 \end{bmatrix}\cdot\begin{bmatrix} 4 & 1 & 2 \\ 0 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 5 & 7 \\ -2 & 2 & 1 \\ 11 & -1 & 7 \end{bmatrix}$$
$$\left(=\begin{bmatrix} 1\cdot4+2\cdot0+3\cdot1 & 1\cdot1+2(-1)+3\cdot2 & 1\cdot2+2\cdot1+3\cdot1 \\ (-1)\cdot4+1\cdot0+2\cdot1 & (-1)\cdot1+1(-1)+2\cdot2 & (-1)\cdot2+1\cdot1+2\cdot1 \\ 3\cdot4+2\cdot0+(-1)\cdot1 & 3\cdot1+2(-1)+(-1)\cdot2 & 3\cdot2+2\cdot1+(-1)\cdot1 \end{bmatrix}\right).$$

El producto de estas mismas matrices tomadas en el orden contrario es igual a

$$\begin{bmatrix} 4 & 1 & 2 \\ 0 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & 2 \\ 3 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 9 & 13 & 12 \\ 4 & 1 & -3 \\ 2 & 6 & 6 \end{bmatrix}.$$

Vemos que *la operación de multiplicación de matrices es* (en general) *no conmutativa.*

2. $$\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} \cos\psi & -\operatorname{sen}\psi \\ \operatorname{sen}\psi & \cos\psi \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\varphi\cos\psi - \operatorname{sen}\varphi\operatorname{sen}\psi & -\cos\varphi\operatorname{sen}\psi - \operatorname{sen}\varphi\cos\psi \\ \operatorname{sen}\varphi\cos\psi + \cos\varphi\operatorname{sen}\psi & -\operatorname{sen}\varphi\operatorname{sen}\psi + \cos\varphi\cos\psi \end{bmatrix} =$$

$$= \begin{bmatrix} \cos(\varphi+\psi) & -\operatorname{sen}(\varphi+\psi) \\ \operatorname{sen}(\varphi+\psi) & \cos(\varphi+\psi) \end{bmatrix}.$$

Este segundo ejemplo tiene la siguiente interpretación geométrica: *la realización sucesiva de dos rotaciones de ángulos φ y ψ de un plano equivale a una rotación de ángulo $\varphi+\psi$.*

Consideremos las propiedades de las operaciones de multiplicación de aplicaciones lineales y de multiplicación de matrices.

1. *Si $\mathcal{A}$, $\mathcal{B}$ y $\mathcal{C}$ son unas aplicaciones lineales de un espacio R, se tiene*

$$(\mathcal{A}\mathcal{B})\mathcal{C} = \mathcal{A}(\mathcal{B}\mathcal{C}).$$

Efectivamente, tenemos para todo vector $x \in R$

$$[(\mathcal{A}\mathcal{B})\mathcal{C}]\,x = (\mathcal{A}\mathcal{B})(\mathcal{C}x) = \mathcal{A}(\mathcal{B}(\mathcal{C}x))$$

y

$$[\mathcal{A}(\mathcal{B}\mathcal{C})]\,x = \mathcal{A}[(\mathcal{B}\mathcal{C})\,x] = \mathcal{A}(\mathcal{B}(\mathcal{C}x));$$

es decir, *la operación de multiplicación de aplicaciones lineales* (y, por consiguiente, de multiplicación de matrices) *cumple la ley asociativa.*

2. *Para todo aplicación lineal $\mathcal{A}$ se tiene*

$$\mathcal{A}\mathcal{E} = \mathcal{E}\mathcal{A} = \mathcal{A}.$$

La matriz E de la aplicación idéntica $\mathcal{E}$ (véase la pág. 102) se llama *matriz unidad.* Para cualquier matriz A (del mismo orden que E) se tiene

$$AE = EA = A.$$

3. *Las operaciones de multiplicación y de adición de aplicaciones lineales están ligadas por las leyes distributivas:*

$$(\mathcal{A}+\mathcal{B})\,\mathcal{C} = \mathcal{A}\mathcal{C}+\mathcal{B}\mathcal{C}$$

y

$$\mathcal{C}(\mathcal{A}+\mathcal{B}) = \mathcal{C}\mathcal{A}+\mathcal{C}\mathcal{B},$$

ya que para todo vector $x \in R$ se tiene

$$((\mathcal{A}+\mathcal{B})\,\mathcal{C})\,x = (\mathcal{A}+\mathcal{B})\,(\mathcal{C}x) = \\ = \mathcal{A}(\mathcal{C}x)+\mathcal{B}(\mathcal{C}x) = (\mathcal{A}\mathcal{C})\,x+ \\ +(\mathcal{B}\mathcal{C})\,x = (\mathcal{A}\mathcal{C}+\mathcal{B}\mathcal{C})\,x$$

y

$$(\mathcal{C}(\mathcal{A}+\mathcal{B}))\,x = \mathcal{C}((\mathcal{A}+\mathcal{B})\,x) = \\ =\mathcal{C}(\mathcal{A}x+\mathcal{B}x) = \mathcal{C}(\mathcal{A}x)+\mathcal{C}(\mathcal{B}x) = \\ = (\mathcal{C}\mathcal{A})\,x+(\mathcal{C}\mathcal{B})\,x = (\mathcal{C}\mathcal{A}+\mathcal{C}\mathcal{B})\,x.$$

En el caso de las matrices tienen lugar unas igualdades análogas.

Recordemos las leyes fundamentales de adición y de multiplicación de los números.

Leyes de adición

1. La adición es *conmutativa:*

$$a+b = b+a$$

cualesquiera que sean los dos números a y b.

2. La adición es *asociativa:*

$$a+(b+c) = (a+b)+c$$

cualesquiera que sean los tres números a, b y c.

3. Existe un número 0 tal que

$$a+0 = a$$

para cualquier a.

4. Para todo número a existe un número $-a$ tal que

$$a+(-a) = 0.$$

Leyes de multiplicación

1. La multiplicación es *conmutativa:*

$$ab = ba$$

cualesquiera que sean los dos números a y b.

2. La multiplicación es *asociativa:*

$$a(bc) = (ab)c$$

cualesquiera que sean los tres números a, b y c.

3. Existe un número 1 tal que

$$a \cdot 1 = a$$

cualquiera que sea a.

4. Para todo número $a \neq 0$ existe un número $1/a = a^{-1}$ tal que $a \cdot a^{-1} = 1$.

Además tiene lugar la ley *distributiva* que relaciona la adición y la multiplicación:

$$a(b+c) = ab + ac$$

cualesquiera que sean los tres números a, b y c.

Hemos demostrado que en el caso de la adición y de la multiplicación de aplicaciones lineales y de matrices son válidas todas estas leyes excepto la primera y la cuarta leyes de la multiplicación. Resulta que la multiplicación de matrices (véase el ejemplo 1) y, por consiguiente, también la multiplicación de aplicaciones lineales son, en general, operaciones *no conmutativas.*

En cuanto a la existencia de la aplicación lineal inversa (de la matriz inversa) es válida la proposición siguiente. *Para toda aplicación lineal $\mathcal{A}$ **no degenerada** existe la aplicación lineal $\mathcal{A}^{-1}$ **(inversa de** $\mathcal{A}$) tal que*

$$\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} = \mathcal{E}$$

(y, respectivamente, *para toda matriz A de determinante diferente de cero existe la matriz A^{-1} **inversa** de A tal que $AA^{-1} = A^{-1}A = E$*).

Demostremos esta proposición. Sea dada una aplicación lineal no degenerada $\mathcal{A}$ con la matriz A en una base determinada e_1, $e_2, \ldots, e_n$. Demostremos primero que existe la *matriz inversa* de A, es decir, la matriz A^{-1} tal que

$$AA^{-1} = A^{-1}A = E;$$

entonces la aplicación lineal $\mathcal{A}^{-1}$ de matriz A^{-1} en esta misma base $e_1, e_2, \ldots, e_n$ será la inversa de $\mathcal{A}$, ya que la realización sucesiva de las aplicaciones $\mathcal{A}^{-1}$ y $\mathcal{A}$ representará la aplicación lineal de matriz unidad, es decir, la aplicación idéntica.

Sea, pues, A una matriz de determinante diferente de cero. Consideremos la matriz formada por los complementos algebraicos de los elementos correspondientes de la matriz A:

$$\begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \ldots & A_{nn} \end{bmatrix}.$$

Trasponiéndola obtenemos la matriz A^* que se denomina la *adjunta* de la matriz A:

$$A^* = \begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}.$$

Multiplicando las matrices A y A^*, obtenemos

$$AA^* = A^*A = \begin{bmatrix} |A| & 0 & \ldots & 0 \\ 0 & |A| & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & |A| \end{bmatrix}$$

(según los teoremas 3 y 4 del capítulo I). Luego, la matriz

$$A^{-1} = \frac{A^*}{|A|}$$

será la *inversa* de A.

Ejemplo. *Hállese la matriz inversa de*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix},$$

Solución. El determinante de la matriz A es igual a 4. Los complementos algebraicos de sus elementos son: $A_{11} = 2$, $A_{12} = -2$, $A_{13} = 2$, $A_{21} = -2$,

$A_{22} = 4$, $A_{23} = -2$, $A_{31} = -8$, $A_{32} = 10$, $A_{33} = -4$ y, por consiguiente,

$$A^{-1} = \frac{1}{4}\begin{bmatrix} 2 & -2 & -8 \\ -2 & 4 & 10 \\ 2 & -2 & -4 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 & -4 \\ -1 & 2 & 5 \\ 1 & -1 & -2 \end{bmatrix}.$$

Observemos que *siendo las aplicaciones $\mathcal{A}$ y $\mathcal{B}$ no degeneradas también será no degenerado el producto de las mismas* (ya que la igualdad $(\mathcal{A}\mathcal{B})\,x = \mathcal{A}(\mathcal{B}x) = 0$ implica $\mathcal{B}x = 0$ y, por consiguiente, $x = 0$) y que además

$$(\mathcal{A}\mathcal{B})^{-1} = \mathcal{B}^{-1}\mathcal{A}^{-1}$$

(y en el caso de las matrices $(AB)^{-1} = B^{-1}A^{-1}$), ya que

$$(\mathcal{A}\mathcal{B})(\mathcal{B}^{-1}\mathcal{A}^{-1}) = \mathcal{A}[\mathcal{B}(\mathcal{B}^{-1}\mathcal{A}^{-1})] =$$
$$= \mathcal{A}[(\mathcal{B}\mathcal{B}^{-1})\,\mathcal{A}^{-1}] = \mathcal{A}(\mathcal{E}\mathcal{A}^{-1}) = \mathcal{A}\mathcal{A}^{-1} = \mathcal{E}.$$

Teorema 3. *El determinante del producto de dos matrices es igual al producto de los determinantes de los factores:* si $AB = C$, se tiene

$$|C| = |A|\,|B|.$$

Demostración. Sea

$$AB = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} =$$
$$= \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}.$$

En estas condiciones, como sabemos,

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}.$$

El determinante de la matriz C es igual a

$$|C| = \begin{vmatrix} \sum\limits_{j_1=1}^{n} a_{1j_1}b_{j_1 1} & \sum\limits_{j_2=1}^{n} a_{1j_2}b_{j_2 2} & \dots & \sum\limits_{j_n=1}^{n} a_{1j_n}b_{j_n n} \\ \sum\limits_{j_1=1}^{n} a_{2j_1}b_{j_1 1} & \sum\limits_{j_2=1}^{n} a_{2j_2}b_{j_2 2} & \dots & \sum\limits_{j_n=1}^{n} a_{2j_n}b_{j_n n} \\ \dots & \dots & \dots & \dots \\ \sum\limits_{j_1=1}^{n} a_{nj_1}b_{j_1 1} & \sum\limits_{j_2=1}^{n} a_{nj_2}b_{j_2 2} & \dots & \sum\limits_{j_n=1}^{n} a_{nj_n}b_{j_n n} \end{vmatrix}.$$

y puede ser representado, según la cuarta propiedad de los determinantes, en forma de la suma

$$\sum_{j_1, j_2, \dots, j_n} \begin{vmatrix} a_{1j_1}b_{j_1 1} & a_{1j_2}b_{j_2 2} & \dots & a_{1j_n}b_{j_n n} \\ a_{2j_1}b_{j_1 1} & a_{2j_2}b_{j_2 2} & \dots & a_{2j_n}b_{j_n n} \\ \dots & \dots & \dots & \dots \\ a_{nj_1}b_{j_1 1} & a_{nj_2}b_{j_2 2} & \dots & a_{nj_n}b_{j_n n} \end{vmatrix},$$

donde los índices $j_1, j_2, \dots, j_n$ recorren independientemente todos los valores $1, 2, 3, \dots, n$ (esta suma contiene un total de n^n sumandos). Sin embargo, se puede aceptar que en el determinante que aparece bajo el signo de la suma todos los índices $j_1, j_2, \dots, j_n$ son *diferentes*, ya que los determinantes que tienen índices iguales j_k son iguales a cero por tener columnas proporcionales. Por consiguiente, en esta suma quedan sólo $n!$ sumandos correspondientes a *diferentes* colecciones de $j_1, j_2, \dots, j_n$. Sacando ahora fuera del signo del determinante el factor común de los elementos de cada una de las columnas, obtenemos

$$|C| = \sum_{j_1, j_2, \dots, j_n} b_{j_1 1}b_{j_2 2} \dots b_{j_n n} \begin{vmatrix} a_{1j_1} & a_{1j_2} & \dots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \dots & a_{2j_n} \\ \dots & \dots & \dots & \dots \\ a_{nj_1} & a_{nj_2} & \dots & a_{nj_n} \end{vmatrix}, \qquad (2)$$

donde la suma se efectúa por todas las permutaciones posibles $j_1, j_2, \dots, j_n$ de los números $1, 2, \dots, n$.

Cambiemos entre sí las columnas del determinante que figura bajo el signo de la última suma de manera que los segundos índices de sus elementos aparezcan en el orden de crecimiento. Esto se puede realizar mediante un número determinado de trasposiciones de columnas. Para pasar de una permutación a otra de la misma paridad se requiere un número par de trasposiciones y para pasar a una permutación de paridad opuesta se necesita un número impar de trasposiciones; como, además, la permutación 1, 2, ..., n es par, resulta que el determinante del segundo miembro de la igualdad (2) es igual a $(-1)^{[j_1, j_2, \ldots, j_n]} |A|$. Tenemos, por consiguiente,

$$|C| = \sum_{j_1, j_2, \ldots, j_n} (-1)^{[j_1, j_2, \ldots, j_n]} b_{j_1 1} b_{j_2 2} \ldots b_{j_n n} |A| = |A|\,|B|.$$

Ejemplo. En el ejemplo 1 de la pág. 106 hemos multiplicado las matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & 2 \\ 3 & 2 & -1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 4 & 1 & 2 \\ 0 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix}.$$

Tenemos en este caso $|A| = -10$, $|B| = -9$ y $|AB| = |BA| = 90 = |A| \cdot |B|$.

§ 3. Cambio de base

Supongamos que una aplicación lineal $\mathcal{A}$ tiene en una base $e_1, e_2, \ldots, e_n$ la matriz $A = [a_{ik}]$ y que en la base $e_1', e_2', \ldots, e_n'$ tiene otra matriz $A' = [a_{ik}']$. Veamos cómo están relacionadas entre sí las matrices A y A'.

Indiquemos por $C = [c_{ik}]$ la matriz del cambio de la base $e_1, e_2, \ldots, e_n$ por la base $e_1', e_2', \ldots, e_n'$. Entonces,

$$e_i' = c_{1i}e_1 + c_{2i}e_2 + \ldots + c_{ni}e_n.$$

Consideremos la matriz C como la matriz de una aplicación lineal $\mathcal{C}$ en la base $e_1, e_2, \ldots, e_n$. Es obvio entonces que

$$\mathcal{C}e_i = c_{1i}e_1 + c_{2i}e_2 + \ldots + c_{ni}e_n = e_i',$$

es decir, la aplicación lineal $\mathcal{C}$ transforma los vectores $e_1, e_2, \ldots, e_n$ en los vectores $e_1', e_2', \ldots, e_n'$ respectivamente.

El determinante de la matriz de la aplicación $\mathcal{C}$ es diferente de cero (§4 del capítulo II) y, por consiguiente, para $\mathcal{C}$ existe la aplicación inversa $\mathcal{C}^{-1}$ tal que $\mathcal{C}^{-1}e_1' = e_1, \mathcal{C}^{-1}e_2' = e_2, \ldots, \mathcal{C}^{-1}e_n' = e_n$. Por hipótesis,

$$\mathcal{A}e_1' = a_{1i}'e_1' + a_{2i}'e_2' + \ldots + a_{ni}'e_n'.$$

Aplicando a ambos miembros de esta igualdad la aplicación $\mathcal{C}^{-1}$, obtenemos

$$\mathcal{C}^{-1}\mathcal{A}e_i' = a_{1i}'e_1 + a_{2i}'e_2 + \ldots + a_{ni}'e_n.$$

Introduciendo en el primer miembro de la última igualdad las relaciones $e'_i = \mathcal{C}e_i$ encontramos

$$\mathcal{C}^{-1}\mathcal{A}\mathcal{C}e_i = a'_{1i}e_1 + a'_{2i}e_2 + \ldots + a'_{ni}e_n,$$

es decir, la matriz A' es la matriz de la aplicación $\mathcal{C}^{-1}\mathcal{A}\mathcal{C}$ en la base $e_1, e_2, \ldots, e_n$. Pero, por otro lado, la matriz de esta aplicación es igual al producto de las matrices de las aplicaciones $\mathcal{C}^{-1}$, $\mathcal{A}$ y $\mathcal{C}$ en la base $e_1, e_2, \ldots, e_n$, es decir,

$$A' = C^{-1}AC.$$

De aquí se deduce, en particular, que *el determinante de la matriz de una aplicación lineal no depende de la base:*

$$|A'| = |C^{-1}AC| = |C^{-1}|\,|A|\,|C| = |C^{-1}|\,|C|\,|A| = $$
$$= |C^{-1}C|\,|A| = |E|\,|A| = |A|.$$

Ejemplo. Una aplicación $\mathcal{A}$ tiene en una base e_1, e_2 la matriz $A = \begin{bmatrix} 6 & -2 \\ 6 & -1 \end{bmatrix}$. Determínese la matriz de esta aplicación en la base $e'_1 = e_1 + 2e_2$, $e'_2 = 2e_2 + 3e_3$.

Solución. La matriz del cambio es en este caso $C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ y su matriz inversa es $C^{-1} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}$. Luego,

$$A' = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 6 & -2 \\ 6 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 6 \\ 4 & 9 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

§ 4. Matrices rectangulares

Una matriz formada de m filas y de n columnas se denomina *$[m\times n]$-matriz*. Se puede definir la **operación de adición de $[m\times n]$-matrices** tomando

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ b_{m1} & b_{m2} & \ldots & b_{mn} \end{bmatrix} =$$
$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \ldots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \ldots & a_{2n}+b_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \ldots & a_{mn}+b_{mn} \end{bmatrix},$$

y la **operación de multiplicación de una $[m\times n]$-matriz por un número** α mediante la igualdad

$$\alpha\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{bmatrix}.$$

Las matrices rectangulares pueden ser consideradas como las matrices de las *aplicaciones lineales* de un espacio vectorial *en otro.* Supongamos concretamente que se tienen dos espacios lineales R^n y R^m, en general, de distintas dimensiones y que a todo vector $x \in R^n$ se pone en correspondencia un vector $\mathcal{A}x \in R^m$ de manera que se cumplan las condiciones siguientes:

1. $\mathcal{A}(x+y) = \mathcal{A}x+\mathcal{A}y.$
2. $\mathcal{A}(\alpha x) = \alpha\mathcal{A}x.$

Tomemos una base $e_1, e_2, \dots, e_n$ en el espacio R^n y una base $f_1, f_2, \dots, f_m$ en el espacio R^m. El vector $\mathcal{A}e_i$ pertenece a R^m y, por consiguiente, puede ser descompuesto por la base $f_1, f_2, \dots, f_m$; sea

$$\begin{aligned} \mathcal{A}e_1 &= a_{11}f_1+a_{21}f_2+\dots+a_{m1}f_m, \\ \mathcal{A}e_2 &= a_{12}f_1+a_{22}f_2+\dots+a_{m2}f_m, \\ &\dots\dots\dots\dots\dots\dots \\ \mathcal{A}e_n &= a_{1n}f_1+a_{2n}f_2+\dots+a_{mn}f_m. \end{aligned}$$

Es decir, a la aplicación lineal $\mathcal{A}$ del espacio R^n en el espacio R^m le corresponde la *matriz rectangular*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Es fácil ver que las $[m\times n]$-matrices (en particular, las matrices cuadradas de orden n) *forman un espacio lineal* R. Indiquemos por e_{ik} la $[m\times n]$-matriz en la que $a_{ik} = 1$ mientras que todos los demás elementos son iguales a cero. Queda claro, entonces, que todas estas matrices son linealmente independientes y que cualquier $[m\times n]$-matriz es una combinación lineal de aquéllas. Por consiguiente, *la dimensión del espacio R es igual a mn*. En particular, el espacio formado por todas las matrices cuadradas de orden n es de dimensión n^2.

Por analogía con la multiplicación de matrices cuadradas se puede definir también la **multiplicación de matrices rectangulares.** Esta multiplicación puede ser realizada sólo en el caso en el que la longitud de la fila del primer factor sea igual a la longitud de la columna del segundo, es decir, cuando el número de columnas del primer factor sea igual al número de filas del segundo. El producto de una $[m\times n]$-matriz por una $[n\times p]$-matriz será, evidentemente, una $[m\times p]$-matriz. En particular, el producto de una $[m\times n]$-matriz por una $[n\times 1]$-matriz, es decir, por una columna, será una $[m\times 1]$-matriz, es decir, una columna, y el producto de una $[1\times m]$-matriz, es decir, de una fila, por una $[m\times n]$-matriz será una $[1\times n]$-matriz, es decir, una fila.

Ejemplos.

$$1.\ \begin{bmatrix} -1 & 3 & 2 & 0 \\ 1 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & -3 & 3 \end{bmatrix} = \begin{bmatrix} 10 & 0 & 7 \\ 2 & -5 & 4 \end{bmatrix}.$$

$$2.\ \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

$$3.\ [1 \quad 2 \quad 3] \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} = [14 \quad 20].$$

Es fácil comprender el sentido de la operación de multiplicación de matrices rectangulares. Supongamos que se tienen tres espacios lineales R^n, R^m y R^p, en general, de distintas dimensiones y dos aplicaciones lineales: la aplicación $\mathcal{A}$ que transforma R^m en R^p y la aplicación $\mathcal{B}$ que transforma R^n en R^m. La aplicación $\mathcal{A}\cdot\mathcal{B}$ que pone en correspondencia a todo vector $x \in R^n$ un vector $\mathcal{A}(\mathcal{B}x)$ del espacio R^p se denomina *producto* de las aplicaciones $\mathcal{A}$ y $\mathcal{B}$. Es fácil ver que $\mathcal{A}\mathcal{B}$ es una aplicación lineal de R^n en R^p y si a la aplicación $\mathcal{A}$ le corresponde la $[p\times m]$-matriz A y a la aplicación

$\mathcal{B}$ le corresponde la $[m \times n]$-matriz B, la matriz de la aplicación $\mathcal{AB}$ será la $[p \times n]$-matriz AB.

Es fácil demostrar que para las operaciones de adición y de multiplicación de matrices rectangulares se cumplen las relaciones siguientes:

1. $A+B = B+A$.
2. $(A+B)+C = A+(B+C)$.
3. $A+0 = A$, donde 0 es la matriz rectangular con todos los elementos iguales a cero (la matriz nula).
4. $A+(-A) = 0$, donde $-A$ es la matriz en la que todos los elementos son los opuestos de los elementos correspondientes de la matriz A.
5. $A(BC) = (AB)C$.
6. $(A+B)C = AC+BC$.
7. $C(A+B) = CA+CB$,

siempre que las matrices sean tales que todas estas operaciones puedan ser realizadas. Además, siendo A una $[m \times n]$-matriz cualquiera, se tiene

$$AE_n = A \quad \text{y} \quad E_m A = A,$$

donde E_n es la matriz unidad de orden n y E_m es la matriz unidad de orden m.

Si $\mathcal{A}$ es una aplicación lineal de un espacio R, a la igualdad $\mathcal{A}x = y$, donde $x, y \in R$, le corresponde en toda base la igualdad matricial

$$AX = Y,$$

donde A es la matriz de la aplicación $\mathcal{A}$, X es la columna formada por las coordenadas del vector x e Y es la columna formada por las coordenadas del vector y.

Consideremos un sistema de ecuaciones lineales

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n = b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n = b_m. \end{array}\right. \tag{3}$$

Indiquemos por A la matriz de los coeficientes de las incógnitas de este sistema, por X la columna formada por las incógnitas y por B la columna formada por los segundos miembros. El sistema (3) puede ser representado entonces en forma de una ecuación matricial

$$AX = B.$$

Si la matriz A es cuadrada y su determinante es diferente de cero, existe la matriz inversa A^{-1}. Multiplicando ambos miembros de la última igualdad por A^{-1} *a la izquierda*, obtenemos

$$A^{-1}(AX) = A^{-1}B, \text{ de donde } X = A^{-1}B.$$

Desarrollando esta fórmula llegamos a las *fórmulas de Cramer* (compárese con la pág. 39).

§ 5. Rango y defecto de una aplicación lineal

Definición 2. *Sea $\mathcal{A}$ una aplicación lineal de un espacio R. El conjunto D de todos los vectores de forma $\mathcal{A}x$, donde $x \in R$, se llama* ***dominio de valores*** *de la aplicación $\mathcal{A}$ y el conjunto M de todos los vectores x tales que $\mathcal{A}x = 0$ se llama* ***núcleo*** *de la aplicación.*

Demostremos que *el dominio de valores y el núcleo de una aplicación lineal son subespacios de R.*

Efectivamente, si $x, y \in D$, se tiene $x = \mathcal{A}x_1$ e $y = \mathcal{A}y_1$, donde $x_1, y_1 \in R$, y por esto

$$x+y = \mathcal{A}x_1 + \mathcal{A}y_1 = \mathcal{A}(x_1+y_1),$$

donde $x_1+y_1 \in R$; luego, $x+y \in D$; además,

$$\alpha x = \alpha \mathcal{A}x_1 = \mathcal{A}(\alpha x_1),$$

donde $\alpha x_1 \in R$, y por consiguiente, $\alpha x \in D$.

Por otro lado, si $x, y \in M$, es decir, si $\mathcal{A}x = 0$ y $\mathcal{A}y = 0$, se tiene

$$\mathcal{A}(x+y) = \mathcal{A}x + \mathcal{A}y = 0$$

y

$$\mathcal{A}(\alpha x) = \alpha \mathcal{A}x = 0,$$

es decir, $x+y \in M$ y $\alpha x \in M$.

La dimensión del dominio de valores de una aplicación $\mathcal{A}$ coincide con el rango de la matriz A (y se denomina ***rango de la aplicación*** *$\mathcal{A}$).* Efectivamente, el subespacio D es generado por los vectores

$$\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_n, \tag{4}$$

donde $e_1, e_2, \ldots, e_n$ es cualquier base del espacio R, y, por consiguiente, la dimensión de D es igual al número máximo de vectores linealmente independientes que contiene el sistema (4), es decir, es igual al número máximo de columnas linealmente independientes de la matriz A (véase la pág. 44).

La dimensión del núcleo M se denomina **defecto** *de la aplicación lineal* $\mathcal{A}$.

Teorema 4. *La suma del rango y del defecto de una aplicación lineal* $\mathcal{A}$ *es igual a la dimensión n del espacio.*

Demostración. Sea r el rango de la aplicación lineal $\mathcal{A}$. Entre los vectores $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_n$ existen entonces r vectores linealmente independientes y todos los restantes se expresan linealmente en términos de éstos. Sean $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_r$ estos vectores.

Indiquemos por L el subespacio que generan en R los vectores $e_1, e_2, \ldots, e_r$ y demostremos que *la intersección del subespacio L* (de dimensión r) *con el núcleo M consta sólo del vector nulo*. Efectivamente, si $x \in L \cap M$, se tiene $x \in L$, es decir,

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r,$$

y $x \in M$, es decir,

$$\mathcal{A}x = \alpha_1 \mathcal{A}e_1 + \alpha_2 \mathcal{A}e_2 + \ldots + \alpha_r \mathcal{A}e_r = 0.$$

Pero los vectores $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_r$ son linealmente independientes; luego, $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$ y $x = 0$.

Demostremos ahora que *los subespacios L y M generan todo el espacio R* (es decir, la suma de estos subespacios coincide con R). Sea x un vector cualquiera de R. Entonces $\mathcal{A}x \in D$ y, por consiguiente,

$$\mathcal{A}x = \beta_1 \mathcal{A}e_1 + \beta_2 \mathcal{A}e_2 + \ldots + \beta_r \mathcal{A}e_r.$$

El vector $y = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_r e_r$ pertenece, obviamente, a L y la diferencia $z = x - y \in M$, ya que

$$\mathcal{A}z = \mathcal{A}(x-y) = \mathcal{A}x - \mathcal{A}y = 0.$$

Hemos encontrado que $x = y + z$, donde $y \in L$ y $z \in M$.

Es decir, el espacio R es igual a la *suma directa* de los subespacios L y M y, por consiguiente, su dimensión n es igual a la suma de las dimensiones de estos subespacios (véase la pág. 81).

§ 6. Aplicación lineal no degenerada

Una aplicación lineal $\mathcal{A}$ se llama *no degenerada* si de la igualdad $\mathcal{A}x = 0$ se deduce que $x = 0$ (§1). Además, hemos demostrado (§ 2) que el determinante de la matriz de una aplicación lineal no degenerada es diferente de cero cualquiera que sea la base y que *para toda aplicación lineal no degenerada* $\mathcal{A}$ *existe la aplicación* *lineal inversa* $\mathcal{A}^{-1}$. Recíprocamente, *si para la aplicación lineal* $\mathcal{A}$ *existe la aplicación inversa* $\mathcal{A}^{-1}$, *esta aplicación es no degenerada,* ya que aplicando a ambos miembros de la igualdad $\mathcal{A}x = 0$ la aplicación $\mathcal{A}^{-1}$, obtenemos

$$\mathcal{A}^{-1}(\mathcal{A}x) = \mathcal{A}^{-1}0 = 0,$$

pero

$$\mathcal{A}^{-1}(\mathcal{A}x) = (\mathcal{A}^{-1}\mathcal{A})x = x$$

y, por consiguiente $x = 0$.

El rango de una aplicación lineal no degenerada de un espacio $R^n = R$ *es igual a* n, ya que el determinante de su matriz es diferente de cero; *el defecto de una aplicación lineal no degenerada es igual a cero*. Recíprocamente, *toda aplicación lineal de rango* n *del espacio* R *será,* evidentemente, *no degenerada*. El dominio de valores de una aplicación lineal no degenerada es n-dimensional y, por consiguiente, coincide con todo el R: una aplicación lineal no degenerada transforma R *sobre todo el* R. El núcleo de una aplicación lineal no degenerada consta sólo del vector nulo. Una aplicación lineal no degenerada es *biyectiva*, ya que de la igualdad $\mathcal{A}x = \mathcal{A}y$ resulta $\mathcal{A}(x-y) = 0$ y, por consiguiente, $x-y = 0$, es decir, $x = y$.

Una aplicación lineal no degenerada transforma vectores linealmente independientes en vectores linealmente independientes. Efectivamente, si los vectores $e_1, e_1, \ldots, e_k$ son linealmente independientes y

$$\alpha_1 \mathcal{A}e_1+\alpha_2\mathcal{A}e_2+\ldots+\alpha_k\mathcal{A}e_k =$$
$$= \mathcal{A}(\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_k e_k) = 0.$$

se tiene

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_k e_k = 0 \text{ y } \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0.$$

Demostremos que siendo $\mathcal{A}$ una aplicación lineal de rango r y $\mathcal{B}$ una aplicación lineal no degenerada, ambas aplicaciones $\mathcal{A}\mathcal{B}$ y $\mathcal{B}\mathcal{A}$ serán de rango r (y, por consiguiente, el producto de una matriz de rango r por una matriz de determinante diferente de cero será una matriz de rango r).

Efectivamente, el dominio de valores de la aplicación lineal (no degenerada) $\mathcal{B}$ coincide con todo el espacio y, por consiguiente, el dominio de valores de la aplicación $\mathcal{AB}$ es r-dimensional, es decir, el rango de la aplicación $\mathcal{AB}$ es igual a r.

Por otra parte, el dominio de valores de la aplicación $\mathcal{A}$ es r-dimensional y como la aplicación $\mathcal{B}$, por ser no degenerada, transforma vectores linealmente independientes en vectores linealmente independientes, resulta que el dominio de valores de la aplicación $\mathcal{BA}$ es también r-dimensional, es decir, el rango de la aplicación $\mathcal{BA}$ es igual a r.

§ 7. Subespacios invariantes. Vectores propios y valores propios de una aplicación lineal

Sea R_1 un subespacio de un espacio lineal R y sea $\mathcal{A}$ una aplicación lineal de R. La imagen $\mathcal{A}x$ de un vector x de R_1 puede no pertenecer, en general, a R_1. Son de interés especial los subespacios cuyos vectores no salen en la aplicación $\mathcal{A}$ de estos subespacios.

Definición 3. *Un subespacio R_1 de un espacio R se llama* ***invariante*** *respecto a una aplicación lineal $\mathcal{A}$ si la imagen $\mathcal{A}x$ de todo vector x de R_1 pertenece a R_1.*

Ejemplos. 1. Sea $\mathcal{A}$ una rotación del espacio corriente de tres dimensiones alrededor del eje Oz. En este caso el plano xOy y el eje Oz, por ejemplo, serán subespacios invariantes.

2. Si $\mathcal{A}$ es la proyección ortogonal del mismo espacio R^3 sobre el plano xOy, los subespacios invariantes son: el plano xOy, todos los planos que pasan por el eje Oz, el propio eje Oz y todas las rectas del plano xOy que pasan por el origen de coordenadas.

3. En el espacio P_n de los polinomios de grado no mayor que n los subespacios P_k son invariantes respecto a la derivación para todo k, $0 \leqslant k \leqslant n$.

4. En cualquier espacio todo subespacio es invariante respecto a las aplicaciones idéntica y nula.

4. En cualquier espacio y para cualquier aplicación lineal el propio espacio R y el subespacio formado sólo por el vector nulo son invariantes.

Demostremos que *la intersección y la suma de subespacios invariantes respecto a una aplicación lineal $\mathcal{A}$ son invariantes respecto a $\mathcal{A}$.*

Efectivamente, si los subespacios R_1 y R_2 son invariantes respecto a $\mathcal{A}$ y si $x \in R_1 \cap R_2$, resulta que $x \in R_1$ y $x \in R_2$ y, por consiguiente $\mathcal{A}x \in R_1$ y $\mathcal{A}x \in R_2$, es decir, $\mathcal{A}x \in R_1 \cap R_2$.

Por otra parte, si $x \in R_1+R_2$, se tiene $x = u+v$, donde $u \in R_1$ y $v \in R_2$. Pero entonces $\mathcal{A}u \in R_1$ y $\mathcal{A}v \in R_2$ y $\mathcal{A}x = \mathcal{A}u+ +\mathcal{A}v \in R_1+R_2$.

Tienen un interés particular los subespacios invariantes de *una dimensión*. Si R_1 es un subespacio de este tipo y $x \in R_1$, también $\mathcal{A}x \in R_1$ y, por consiguiente, $\mathcal{A}x = \lambda x$. Siendo y otro vector cualquiera de R_1, se tiene $y = \alpha x$ y

$$\mathcal{A}y = \mathcal{A}(\alpha x) = \alpha\mathcal{A}x = \alpha(\lambda x) = \lambda(\alpha x) = \lambda y.$$

Definición 4. *Un vector $x \neq 0$ se llama* ***vector propio*** *de una aplicación lineal $\mathcal{A}$, si existe un número λ tal que $\mathcal{A}x = \lambda x$; este número λ se denomina* ***valor propio*** *de la aplicación $\mathcal{A}$ (de la matriz A) correspondiente al vector x.*

Según acabamos de ver, siendo R_1 un subespacio invariante unidimensional de R, todo vector de R_1 es un vector propio de la aplicación $\mathcal{A}$, además, con el mismo valor propio. Recíprocamente, si x es un vector propio de una aplicación $\mathcal{A}$, el subespacio unidimensional R_1 que este vector genera (formado por todos los vectores de tipo αx) será, obviamente, invariante respecto a $\mathcal{A}$.

Supongamos que una aplicación lineal $\mathcal{A}$ tiene n vectores propios linealmente independientes $e_1, e_2, \ldots, e_n$ con los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectivamente. Tomemos los vectores $e_1, e_2, \ldots, e_n$ por básicos; debido a las igualdades

$$\begin{aligned} \mathcal{A}e_1 &= \lambda_1 e_1, \\ \mathcal{A}e_2 &= \lambda_2 e_2, \\ &\ldots\ldots\ldots \\ \mathcal{A}e_n &= \lambda_n e_n, \end{aligned}$$

la matriz de la aplicación $\mathcal{A}$ será de la forma

$$A = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

(toda matriz de este tipo se llama *diagonal*). Está claro que tiene lugar también la afirmación recíproca: *si la matriz A de una aplicación lineal $\mathcal{A}$ es diagonal en una base determinada, todos los vectores de esta base son vectores propios de la aplicación $\mathcal{A}$.* Sin embargo, no toda aplicación lineal ni mucho menos de un espacio n-dimensional posee n vectores propios linealmente independientes.

Teorema 5. *Los vectores propios de una aplicación $\mathcal{A}$ correspondientes a valores propios, distintos dos a dos, son linealmente independientes.*

Realizaremos la demostración por inducción según el número de vectores propios considerados. Para un vector x esto queda claro, ya que, por definición del vector propio, éste es diferente de cero (y, por consiguiente, de la igualdad $\alpha x = 0$ resulta $\alpha = 0$).

Supongamos que nuestra afirmación es válida para $k-1$ vectores $x_1, x_2, \ldots, x_{k-1}$ y aceptemos que k vectores propios $x_1, x_2, \ldots, x_k$, correspondientes a los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_k$ distintos dos a dos, son linealmente dependientes

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0. \tag{5}$$

Aplicando a ambos miembros de esta igualdad la aplicación $\mathcal{A}$, obtenemos

$$\alpha_1 \mathcal{A} x_1 + \alpha_2 \mathcal{A} x_2 + \ldots + \alpha_k \mathcal{A} x_k = 0,$$

es decir,

$$\alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \ldots + \alpha_k \lambda_k x_k = 0. \tag{6}$$

Por otra parte, multiplicando la igualdad (5) por λ_k, tendremos

$$\alpha_1 \lambda_k x_1 + \alpha_2 \lambda_k x_2 + \ldots + \alpha_k \lambda_k x_k = 0. \tag{7}$$

Restando la igualdad (7) de la igualdad (6), encontramos

$$\alpha_1(\lambda_1 - \lambda_k)\, x_1 + \alpha_2(\lambda_2 - \lambda_k)\, x_2 + \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k) x_{k-1} = 0;$$

puesto que, por suposición, todos los λ_i son distintos y, por la hipótesis de inducción, los vectores $x_1, x_2, \ldots, x_{k-1}$ son linealmente independientes, resulta

$$\alpha_1 = \alpha_2 = \ldots = \alpha_{k-1} = 0$$

y entonces de la igualdad (5) tenemos

$$\alpha_k x_k = 0 \quad \text{y} \quad \alpha_k = 0.$$

¿Cómo determinar los valores propios y los vectores propios de una aplicación lineal? Supongamos que x es un vector propio de una aplicación lineal $\mathcal{A}$ y que λ es el valor propio que le corresponde. En este caso, $\mathcal{A}x = \lambda x$. Tomemos en el espacio R una base cualquiera $e_1, e_2, \ldots, e_n$ y sea

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

y $A = [a_{ik}]$ la matriz de la aplicación $\mathcal{A}$ en esta base. En estas condiciones (véase el § 1 del capítulo III),

$$\mathcal{A}x = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)\, e_1 + (a_{21}x_1 + a_{22}x_2 + \ldots$$
$$\ldots + a_{2n}x_n)\, e_2 + \ldots + (a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n)\, e_n =$$
$$= \lambda(x_1e_1 + x_2e_2 + \ldots + x_ne_n),$$

de donde, debido a la independencia lineal de los vectores e_1, $e_2, \ldots, e_n$, resulta

$$\begin{cases} (a_{11}-\lambda)\, x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + (a_{22}-\lambda)\, x_2 + \ldots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn}-\lambda)\, x_n = 0. \end{cases} \tag{8}$$

Para que exista una solución no nula de este sistema (homogéneo) es necesario y suficiente que su determinante sea igual a cero:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix} = 0$$

(teorema 10 del capítulo I). El primer miembro de la última igualdad representa en sí un polinomio de grado n en λ que se denomina *polinomio característico* de la aplicación $\mathcal{A}$ en la base $e_1, e_2, \ldots, e_n$. Este polinomio es el determinante de la matriz $A-\lambda E$. Será demostrado (teorema 6) que, en realidad, este polinomio *no depende de cómo se elija la base* y, por esto, puede ser denominado *polinomio característico de la aplicación* $\mathcal{A}$.

Hemos demostrado que *todo valor propio de una aplicación $\mathcal{A}$ es raíz de su polinomio característico*. Recíprocamente, *toda raíz del polinomio característico de una aplicación $\mathcal{A}$ es un valor propio de la misma:* los vectores propios correspondientes se determinan del sistema (8) que tiene en este caso necesariamente soluciones no nulas, ya que su determinante es igual a cero.

Teorema 6. *El polinomio característico de una aplicación lineal no depende de la elección de la base.*

Demostración. Sea $\varphi_e(\lambda) = |A-\lambda E|$ el polinomio característico de la aplicación $\mathcal{A}$ en la base $e_1, e_2, \ldots, e_n$; supongamos que la base nueva $e'_1, e'_2, \ldots, e'_n$ se obtiene de la antigua mediante la matriz

C. En este caso, el polinomio característico de la aplicación $\mathcal{A}$ en la base $e_1', e_2', \ldots, e_n'$ es

$$\varphi_{e'}(\lambda) = |A' - \lambda E| = |C^{-1}AC - \lambda E| = |C^{-1}AC - C^{-1}\lambda EC| =$$
$$= |C^{-1}(A - \lambda E)C| = |C^{-1}|\,|A - \lambda E|\,|C| = |A - \lambda E| = \varphi_e(\lambda).$$

Aquí nos hemos basado en que $|C^{-1}|\,|C| = |C^{-1}C| = |E| = 1$.

Sea

$$\varphi(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}\alpha_1\lambda^{n-1} + \ldots + \alpha_n$$

el polinomio característico de la aplicación $\mathcal{A}$. Es fácil ver que α_1 es igual a la suma $a_{11} + a_{22} + \ldots + a_{nn}$ de los elementos diagonales de la matriz A (esta suma se llama *traza* de la matriz A). Por otra parte, $\alpha_n = \varphi(0)$ es el *determinante* de la matriz A; luego, *para que una aplicación $\mathcal{A}$ sea no degenerada es necesario y suficiente que $\varphi(0)$ sea diferente de cero, es decir, que $\mathcal{A}$ no tenga valores propios nulos.*

Ejemplo. Hállense los valores propios y los vectores propios de la aplicación lineal $\mathcal{A}$ de matriz

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}.$$

Solución. El polinomio característico de la aplicación $\mathcal{A}$ es

$$\varphi(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 5 & 4-\lambda \end{vmatrix} = \lambda^2 - 5\lambda - 6.$$

Sus raíces son $\lambda_1 = 6$ y $\lambda_2 = -1$. Los vectores propios se determinan de dos sistemas de ecuaciones

$$\begin{cases} (1-\lambda_i)\,x_1 + 2x_2 = 0, \\ 5x_1 + (4-\lambda_i)\,x_2 = 0 \end{cases}$$

$(i = 1, 2)$; puesto que el determinante de estos sistemas es igual a cero, ambos se reducen a una ecuación.

Para $\lambda = 6$ ésta es la ecuación $5x_1 - 2x_2 = 0$ y de ella encontramos $x_1 : x_2 = 2 : 5$ y a título de vector propio correspondiente a $\lambda = 6$ podemos tomar $a_1 = (2, 5)$ (o cualquier otro vector multiplo de a_1).

Para $\lambda = -1$ tenemos la ecuación $x_1 + x_2 = 0$, de donde resulta $x_1 : x_2 = -1$ y el vector propio correspondiente es $a_2 = (1, -1)$ (o cualquier vector multiplo de éste).

Para la aplicación *idéntica* **todos** los vectores no nulos del espacio son, obviamente, vectores propios con el valor propio igual a la *unidad*.

Para la aplicación *nula* **todos** los vectores no nulos del espacio son vectores propios con el valor propio igual a *cero*.

Hallemos también los valores propios y los vectores propios de las aplicaciones de 1 a 5 del § 1.

1. *Rotación del plano de ángulo* φ.

El polinomio característico es

$$\varphi(\lambda)=\begin{vmatrix}\cos\varphi-\lambda & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi-\lambda\end{vmatrix}=\lambda^2-2\cos\varphi\cdot\lambda+1.$$

Sus raíces son complejas: $\lambda_{1,\,2}=\cos\varphi\pm i\operatorname{sen}\varphi$. Luego, esta aplicación no tiene, en general (no siendo φ múltiplo de π), valores propios (reales).

Si $\varphi=2\pi k$, la aplicación es la *idéntica* y **todo** vector es un vector propio (con $\lambda=1$).

Si $\varphi=(2k+1)\pi$, la aplicación es la *simetría central* y **cualquier** vector del plano será, obviamente, un vector propio con el valor propio igual a -1.

2. *Rotación de ángulo* φ *del espacio de tres dimensiones alrededor del eje* Oz. El polinomio característico es

$$\varphi(\lambda)=\begin{vmatrix}\cos\varphi-\lambda & -\operatorname{sen}\varphi & 0\\ \operatorname{sen}\varphi & \cos\varphi-\lambda & 0\\ 0 & 0 & 1-\lambda\end{vmatrix}=(1-\lambda)(\lambda^2-2\cos\varphi\cdot\lambda+1).$$

Sus raíces son $\lambda_1=1$, $\lambda_{2,\,3}=\cos\varphi\pm i\operatorname{sen}\varphi$.

Supongamos que φ no es múltiplo de 2π. Los vectores propios correspondientes al valor propio $\lambda=1$ se determinan del sistema de ecuaciones

$$\begin{cases}(\cos\varphi-1)\,x_1-\operatorname{sen}\varphi\cdot x_2=0,\\ \operatorname{sen}\varphi\cdot x_1+(\cos\varphi-1)\,x_2=0.\end{cases}$$

Puesto que el determinante de este sistema (homogéneo)

$$\begin{vmatrix}\cos\varphi-1 & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi-1\end{vmatrix}=1-2\cos\varphi+1=4\operatorname{sen}^2\frac{\varphi}{2}$$

es diferente de cero, su solución es $x_1=x_2=0$; x_3 es en este caso arbitrario y un vector propio será, por ejemplo, el vector $e_3=(0,0,1)$ (y todos los vectores multiplos de éste).

Si $\varphi=2\pi k$, la aplicación es la *idéntica*.

Si $\varphi=(2k+1)\pi$, la aplicación es *la simetría respecto al eje* Oz. Al valor propio $\lambda=1$ le corresponde, igual que antes, el vector propio $e_3=(0,0,1)$. Los vectores propios correspondientes al valor propio $\lambda_{2,3}=-1$ se determinan del sistema de ecuaciones

$$\begin{cases}(\cos\pi+1)\,x_1-\operatorname{sen}\pi\cdot x_2=0,\\ \operatorname{sen}\pi\cdot x_1+(\cos\pi+1)\,x_2=0,\\ \qquad\qquad\qquad 2x_3=0,\end{cases}$$

que se reduce a una ecuación: $x_3 = 0$; por consiguiente, para $\lambda = -1$ los vectores propios serán todos los vectores de tipo $\alpha e_1 + \beta e_2$.

3. *Proyección del espacio de tres dimensiones sobre el plano xOy*. El polinomio característico es

$$\varphi(\lambda) = -\lambda(1-\lambda)^2.$$

Sus raíces son $\lambda_1 = 0$ y $\lambda_{2,3} = 1$. Para $\lambda = 0$ los vectores propios se determinan del sistema $x_1 = 0$ y $x_2 = 0$. Son los vectores múltiplos de e_3. Para $\lambda = 1$ el sistema se reduce a una ecuación $x_3 = 0$ y los vectores propios son los vectores de tipo $\alpha e_1 + \beta e_2$.

4. *Simetría del espacio de tres dimensiones respecto al plano xOy*. El polinomio característico es

$$\varphi(\lambda) = -(1+\lambda)(1-\lambda)^2.$$

Sus raíces son $\lambda_1 = -1$, y $\lambda_{2,3} = 1$. Para $\lambda = -1$ se obtiene el sistema $x_1 = 0$ y $x_2 = 0$ y los vectores propios son los vectores múltiplos de e_3. Para $\lambda = 1$ el sistema se reduce a una ecuación $x_3 = 0$ y los vectores propios son todos los vectores del plano xOy.

5. *Derivación en el espacio de los polinomios P_{n-1}*. El polinomio característico es $\varphi(\lambda) = (-1)^n \lambda^n$. Sus raíces son $\lambda_{1,2,\ldots,n} = 0$. Los vectores propios correspondientes se determinan del sistema

$$x_2 = 0,\ x_3 = 0,\ \ldots,\ x_n = 0$$

y, por consiguiente, todos los vectores propios son multiplos de e_0 (es decir, los polinomios de grado cero).

Volvamos a considerar el ejemplo (pág. 125) de la aplicación lineal de matriz $A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$. Según hemos visto, su polinomio característico es $\varphi(\lambda) = \lambda^2 - 5\lambda - 6$. El cuadrado de la matriz A es igual a

$$A^2 = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} = \begin{bmatrix} 11 & 10 \\ 25 & 26 \end{bmatrix}.$$

«Introduciendo» en $\varphi(\lambda)$ la matriz A en lugar de λ, obtenemos

$$\varphi(A) = A^2 - 5A - 6E = \begin{bmatrix} 11 & 10 \\ 25 & 26 \end{bmatrix} - 5\begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} - 6\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

es decir, la matriz nula. Vemos, pues, que la matriz A es una «raíz» de su polinomio característico. Esto es válido para cualquier matriz cuadrada.

Teorema 7 (teorema de Hamilton — Cayley). *Toda matriz cuadrada es una raíz de su polinomio característico.*

Demostración. Consideremos el polinomio $\psi(\lambda) = |\lambda E - A|$, que sólo difiere del polinomio característico $\varphi(\lambda)$ de la matriz A sólo en el factor $(-1)^n$, y demostremos que $\psi(A) = 0$.

Indiquemos por B la matriz adjunta de $\lambda E - A$ (§ 2 del capítulo III). Tenemos, entonces, la siguiente identidad respecto a λ

$$(\lambda E - A)\,B = |\lambda E - A|\,E.$$

Los elementos de la matriz B son polinomios en λ de grado no mayor que $n-1$. Sea

$$B = \begin{bmatrix} b_{11}^0+b_{11}^1\lambda+\ldots+b_{11}^{n-1}\lambda^{n-1} & b_{12}^0+b_{12}^1\lambda+\ldots+b_{12}^{n-1}\lambda^{n-1} & \ldots & b_{1n}^0+b_{1n}^1\lambda+\ldots+b_{1n}^{n-1}\lambda^{n-1} \\ b_{21}^0+b_{21}^1\lambda+\ldots+b_{21}^{n-1}\lambda^{n-1} & b_{22}^0+b_{22}^1\lambda+\ldots+b_{22}^{n-1}\lambda^{n-1} & \ldots & b_{2n}^0+b_{2n}^1\lambda+\ldots+b_{2n}^{n-1}\lambda^{n-1} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ b_{n1}^0+b_{n1}^1\lambda+\ldots+b_{n1}^{n-1}\lambda^{n-1} & b_{n2}^0+b_{n2}^1\lambda+\ldots+b_{n2}^{n-1}\lambda^{n-1} & \ldots & b_{nn}^0+b_{nn}^1\lambda+\ldots+b_{nn}^{n-1}\lambda^{n-1} \end{bmatrix} = B_0+\lambda B_1+\lambda^2 B_2+\ldots+\lambda^{n-1}B_{n-1},$$

donde

$$B_k = \begin{bmatrix} b_{11}^k & b_{12}^k & \ldots & b_{1n}^k \\ b_{21}^k & b_{22}^k & \ldots & b_{2n}^k \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1}^k & b_{n2}^k & \ldots & b_{nn}^k \end{bmatrix}.$$

Si $\psi(\lambda) = \lambda^n+\alpha_1\lambda^{n-1}+\ldots+\alpha_n$, tenemos

$$(\lambda E-A)(B_0+\lambda B_1+\lambda^2 B_2+\ldots+\lambda^{n-1}B_{n-1}) = \psi(\lambda)E = \\ = (\lambda^n+\alpha_1\lambda^{n-1}+\ldots+\alpha_n)E.$$

Como esta relación es una identidad respecto a λ, deben coincidir los coeficientes de las potencias iguales de λ en ambos miembros:

$$\begin{aligned} B_{n-1} &= E, \\ B_{n-2}-AB_{n-1} &= \alpha_1 E, \\ B_{n-3}-AB_{n-2} &= \alpha_2 E, \\ &\ldots\ldots \\ B_1-AB_2 &= \alpha_{n-2}E, \\ B_0-AB_1 &= \alpha_{n-1}E, \\ -AB_0 &= \alpha_n E. \end{aligned}$$

Multiplicando estas igualdades a la izquierda por $A^n, A^{n-1},\ldots, A, E$, respectivamente, y sumándolas, obtendremos en el primer miembro la matriz nula y en el segundo miembro la expresión $A^n+\alpha_1 A^{n-1}+\ldots+\alpha_n E = \psi(A)$, de donde resulta precisamente que $\psi(A) = 0$.

Según hemos visto (ejemplo 1 de la pág. 126), no toda aplicación lineal posee al menos un vector propio y, por consiguiente, al menos un subespacio invariante de dimensión uno. Al mismo tiempo tiene lugar el teorema siguiente.

Teorema 8. *Toda aplicación lineal posee un subespacio invariante de una o de dos dimensiones.*

Demostración. Si el polinomio característico de la aplicación $\mathcal{A}$ tiene al menos una raíz real, esta aplicación posee un vector propio y, por consiguiente, un subespacio invariante de dimensión uno. Supongamos ahora que todas las raíces del polinomio característico de la aplicación $\mathcal{A}$ son complejas y sea $\lambda = \alpha + i\beta$ una de ellas. Resolviendo para este valor de λ el sistema de ecuaciones (8), obtendremos las soluciones (complejas) respectivas $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, $\ldots$, $z_n = x_n + iy_n$ y, por consiguiente, tendrán lugar las igualdades

$$\begin{aligned}
&a_{11}(x_1+iy_1)+a_{12}(x_2+iy_2)+\ldots+a_{1n}(x_n+iy_n) = \\
&\qquad\qquad\qquad\qquad = (\alpha+i\beta)(x_1+iy_1), \\
&a_{21}(x_1+iy_1)+a_{22}(x_2+iy_2)+\ldots+a_{2n}(x_n+iy_n) = \\
&\qquad\qquad\qquad\qquad = (\alpha+i\beta)(x_2+iy_2), \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{n1}(x_1+iy_1)+a_{n2}(x_2+iy_2)+\ldots+a_{nn}(x_n+iy_n) = \\
&\qquad\qquad\qquad\qquad = (\alpha+i\beta)(x_n+iy_n).
\end{aligned}$$

Igualando las partes reales e imaginarias, obtenemos dos sistemas de igualdades:

$$\begin{aligned}
&a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n = \alpha x_1-\beta y_1, \\
&a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n = \alpha x_2-\beta y_2, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n = \alpha x_n-\beta y_n,
\end{aligned} \tag{9}$$

y

$$\begin{aligned}
&a_{11}y_1+a_{12}y_2+\ldots+a_{1n}y_n = \alpha y_1+\beta x_1, \\
&a_{21}y_1+a_{22}y_2+\ldots+a_{2n}y_n = \alpha y_2+\beta x_2, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{n1}y_1+a_{n2}y_2+\ldots+a_{nn}y_n = \alpha y_n+\beta x_n.
\end{aligned} \tag{10}$$

Consideremos dos vectores (reales)

$$u = x_1e_1+x_2e_2+\ldots+x_ne_n$$

y

$$v = y_1e_1+y_2e_2+\ldots+y_ne_n.$$

Las igualdades (9) muestran que $\mathcal{A}u = \alpha u-\beta v$ y las igualdades (10) muestran que $\mathcal{A}v = \alpha v+\beta u$; pero en este caso el subespacio

R_1 generado por los vectores u y v es invariante respecto a $\mathcal{A}$ ya que siendo $x \in R_1$, es decir, siendo $x = \xi u + \eta v$, también el vector

$$\mathcal{A}x = \xi\mathcal{A}u + \eta\mathcal{A}v = \xi(\alpha u - \beta v) + \eta(\alpha v + \beta u) = (\alpha\xi + \beta\eta)u + \\ + (\alpha\eta - \beta\xi)v$$

pertenece a R_1. Este subespacio es bidimensional porque si los vectores u y v fuesen linealmente dependientes, tendríamos $v = \gamma u$ y

$$\mathcal{A}u = \alpha u - \beta v = (\alpha - \beta\gamma)u,$$

es decir, el vector u sería un vector propio de la aplicación $\mathcal{A}$ correspondiente al valor propio real $\alpha - \beta\gamma$.

Teorema 9. *Si $\mathcal{A}$ es una aplicación lineal no degenerada y R_1 es un subespacio invariante respecto a $\mathcal{A}$, R_1 es también invariante respecto a $\mathcal{A}^{-1}$.*

Demostración. Sea $e_1, e_2, \ldots, e_r$ una base del subespacio R_1. En este caso los vectores $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_r$ también pertenecen a R_1 (por ser R_1 invariante), son linealmente independientes (véase el § 6) y, por consiguiente, también constituyen una base de R_1, es decir, un vector cualquiera $x \in R_1$ puede ser representado en la forma

$$x = \alpha_1\mathcal{A}e_1 + \alpha_2\mathcal{A}e_2 + \ldots + \alpha_r\mathcal{A}e_r.$$

Pero entonces también

$$\mathcal{A}^{-1}x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_r e_r \in R_1.$$

Fijémonos, finalmente, en un detalle que será útil en lo sucesivo.

Observación. Sea $\mathcal{A}$ una aplicación lineal cualquiera de un espacio R de n dimensiones que se descompone en la *suma directa* $R = R_1 \oplus R_2$ de subespacios suyos R_1 y R_2 *invariantes ambos respecto* a $\mathcal{A}$; sea $e_1, e_2, \ldots, e_r$ una base de R_1 y sea $e_{r+1}, e_{r+2}, \ldots, e_n$ una base de R_2. Puesto que los subespacios R_1 y R_2 son invariantes tienen lugar las igualdades

$$\begin{aligned}
\mathcal{A}e_1 &= a_{11}e_1 + a_{21}e_2 + \ldots + a_{r1}e_r,\\
\mathcal{A}e_2 &= a_{12}e_1 + a_{22}e_2 + \ldots + a_{r2}e_r,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\mathcal{A}e_r &= a_{1r}e_1 + a_{2r}e_2 + \ldots + a_{rr}e_r,\\
\mathcal{A}e_{r+1} &= a_{r+1,\,r+1}\, e_{r+1} + a_{r+2,\,r+1}\, e_{r+2} + \ldots + a_{n,\,r+1}\, e_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\mathcal{A}e_n &= a_{r+1,\,n}\, e_{r+1} + a_{r+2,\,n}\, e_{r+2} + \ldots + a_{nn}e_n
\end{aligned}$$

(ya que $\mathcal{A}e_k \in R_1$ para $k = 1, 2, \ldots, r$ y $\mathcal{A}e_k \in R_2$ para $k = r+1, r+2, \ldots, n$). Por ello la matriz de la aplicación $\mathcal{A}$ en la base $e_1, e_2, \ldots, e_n$ de todo el espacio adquiere la forma

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1r} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2r} & 0 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{r1} & a_{r2} & \ldots & a_{rr} & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & a_{r+1,\,r+1} & a_{r+1,\,r+2} & \ldots & a_{r+1,\,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 0 & a_{n,\,r+1} & a_{n,\,r+2} & \ldots & a_{nn} \end{bmatrix}.$$

Se puede decir que la matriz A se descompone en «células»

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

donde A_1 es la matriz de la aplicación $\mathcal{A}$ en el subespacio R_1, A_2 es la matriz de la aplicación $\mathcal{A}$ en el subespacio R_2 y las matrices rectangulares que figuran en el ángulo inferior izquierdo y en el ángulo superior derecho están formadas por ceros.

Por consiguiente, *conociendo las matrices A_1 y A_2 de la aplicación $\mathcal{A}$ en los subespacios R_1 y R_2 podemos formar con ellas la matriz de la aplicación $\mathcal{A}$ en todo el espacio R.*

CAPÍTULO IV

ESPACIO EUCLÍDEO

§ 1. Producto escalar

Hemos definido los espacios lineales, cuyos vectores pueden ser sumados y multiplicados por números, hemos introducido los conceptos de dimensión, de base y de aplicación lineal y ahora introduciremos la *métrica* de estos espacios, es decir, *el método que permite medir las longitudes y los ángulos*. Existen diferentes formas de definir la métrica de un espacio lineal. Optaremos por la forma que se basa en el concepto de *producto escalar*. En el espacio corriente de tres dimensiones se llama producto escalar de dos vectores al producto de sus longitudes por el coseno del ángulo que forman estos vectores. Por consiguiente, el *producto escalar* (x, y) de dos vectores x e y es una función de dos variables (vectores) que, como se sabe, posee las propiedades siguientes:

1. *Para cualesquiera dos vectores* x *e* y *se tiene*

$$(x, y) = (y, x)$$

(**propiedad conmutativa** del producto escalar).

2. *Para todo vector* x *y cualquier número real* α *se tiene*

$$(\alpha x, y) = \alpha(x, y)$$

(**propiedad asociativa** del producto escalar respecto a la multiplicación de los vectores por los números).

3. *Para cualesquiera tres vectores* x, y y z *se tiene*

$$(x+y, z) = (x, z)+(y, z)$$

(**propiedad distributiva** del producto escalar respecto a la adición de los vectores).

4. *Para el cuadrado escalar de cualquier vector* x *se tiene* $(x, x) \geqslant 0$ *y de la igualdad* $(x, x) = 0$ *se deduce* que $x = 0$ (**propiedad positiva** del producto escalar).

En el caso de un espacio lineal general no disponemos de los conceptos de longitud y de ángulo y, por ello, introduciremos el producto escalar axiomáticamente.

Definición 1. *Diremos que en un espacio lineal R está definido el* ***producto escalar*** *si a todo par de vectores x e y de R corresponde un número (x, y) que cumple las condiciones* 1, 2 *y* 3. (Estas condiciones pueden ser llamadas ***axiomas*** del producto escalar.)

Un espacio vectorial provisto de un producto escalar que, además de las condiciones 1, 2 *y* 3, *cumple también la condición* 4 *se denomina espacio vectorial* ***euclídeo****.*

De las igualdades 1, 2 y 3 se deducen fácilmente las relaciones siguientes:

$$2'.\ (x, \alpha y) = (\alpha y, x) = \alpha(y, x) = \alpha(x, y).$$

$$3'.\ (z, x+y) = (x+y, z) = (x, z)+(y, z) = (z, x)+(z, y).$$

Ejemplos.

1. Supongamos que en un espacio vectorial de n dimensiones se ha fijado una base determinada; en este caso el producto escalar de los vectores

$$x = (x_1, x_2, \ldots, x_n) \text{ e } y = (y_1, y_2, \ldots, y_n)$$

se puede definir mediante la igualdad

$$(x, y) = x_1y_1+x_2y_2+\ldots+x_ny_n \tag{1}$$

(las condiciones 1, 2, 3 y 4 se comprueban directamente).

2. En el espacio P de los polinomios en t y en el espacio C de las funciones continuas sobre el segmento $[a, b]$ el producto escalar se puede definir mediante la igualdad

$$(x, y) = \int_a^b x(t)y(t)\,dt \tag{2}$$

(es evidente que se cumplen las condiciones 1, 2 y 3; también se cumple la condición 4, ya que una función continua no negativa de integral igual a cero es idénticamente igual a cero).

Definición 2. *En un espacio euclídeo se llama* ***longitud*** *o* ***módulo*** *de un vector x a la raíz cuadrada de su cuadrado escalar:*

$$|x| = \sqrt{(x, x)}.$$

El ***ángulo*** *φ entre los vectores x e y se define mediante la igualdad*

$$\cos\varphi = \frac{(x, y)}{|x|\,|y|}. \tag{3}$$

Demostremos que en cualquier espacio euclídeo el ángulo puede ser definido de esta forma, es decir, que *el coseno del ángulo entre vectores, introducido mediante la fórmula* (3), *no sobrepasa por su valor absoluto a la unidad.*

Debemos comprobar que

$$-1 \leqslant \frac{(x, y)}{|x|\,|y|} \leqslant 1.$$

Es suficiente demostrar para ello la desigualdad de **Cauchy — Buniakovski:** *en un espacio euclídeo R se tiene para todos los x e y*

$$(x, y)^2 \leqslant |x|^2\,|y|^2.$$

Demostración. Sea α un número real arbitrario. Debido a la condición 4, para el vector $x-\alpha y$ se tiene

$$(x-\alpha y,\ x-\alpha y) \geqslant 0,$$

de donde, basándose en las condiciones 1, 2 y 3, obtenemos

$$(x, x)-2\alpha(x, y)+\alpha^2(y, y) \geqslant 0.$$

Aquí figura un trinomio de segundo grado en α. Puesto que él es no negativo *cualquiera que sea* el valor de α, no puede tener dos raíces reales distintas y, por consiguiente, su discriminante es no positivo:

$$(x, y)^2-|x|^2\,|y|^2 \leqslant 0$$

que es lo que se quería demostrar.

Es fácil ver que la igualdad $(x, y)^2 = |x|^2\,|y|^2$ se alcanza si, y sólo si, existe un número α tal que $x-\alpha y = 0$, es decir, si los vectores x e y son proporcionales:

$$x = \alpha y.$$

La desigualdad de Cauchy – Buniakovski, aplicada al espacio provisto del producto escalar (1), lleva a la «desigualdad de Cauchy»:

$$(a_1b_1+a_2b_2+\ldots+a_nb_n)^2 \leqslant$$
$$\leqslant (a_1^2+a_2^2+\ldots+a_n^2)(b_1^2+b_2^2+\ldots+b_n^2),$$

válida para cualesquiera números a_i y b_i, y aplicada al espacio C con el producto escalar (2), lleva a la «desigualdad de Buniakovski»:

$$\left[\int_a^b x(t)\,y(t)\,dt\right]^2 \leqslant \int_a^b [x(t)]^2\,dt\cdot\int_a^b [y(t)]^2\,dt,$$

válida para cualesquiera dos funciones continuas $x(t)$ e $y(t)$.

En un espacio euclídeo se cumple también la *desigualdad triangular:*

$$|x+y| \leqslant |x|+|y|.$$

Demostración. Empleando la desigualdad de Cauchy—Buniakovski, encontramos

$$|x+y|^2 = (x+y, x+y) = (x, x)+2(x, y)+(y, y) =$$
$$= |x|^2+2(x, y)+|y|^2 \leqslant |x|^2+2|x|\,|y|+|y|^2 = (|x|+|y|)^2,$$

de donde resulta

$$|x+y| \leqslant |x|+|y|.$$

Dos vectores x e y se llaman ***ortogonales*** *si el producto escalar (x, y) de estos vectores es igual a cero.*

En cualquier espacio euclídeo es válido el **«teorema de Pitágoras»**: *si los vectores x e y son ortogonales, se tiene*

$$|x+y|^2 = |x|^2+|y|^2.$$

Efectivamente, siendo $(x, y) = 0$ se tiene

$$|x+y|^2 = |x|^2+2(x, y)+|y|^2 = |x|^2+|y|^2.$$

§ 2. Base ortonormal

Definición 3. *Una base $e_1, e_2, \ldots, e_n$ de un espacio euclídeo se llama* ***ortogonal,*** *si*

$$(e_i, e_k) = 0 \text{ para } i \neq k.$$

Si se tiene, además,

$$|e_i| = 1 \text{ para } i = 1, 2, \ldots, n,$$

se dice que la base es ortonormal.

Lema. *Los vectores no nulos ortogonales dos a dos son linealmente independientes.*

Demostración. Supongamos que los vectores no nulos $x_1, x_2, \ldots, x_m$ son ortogonales dos a dos: $(x_i, x_k) = 0$ para $i \neq k$. Supongamos que

$$\alpha_1 x_1+\alpha_2 x_2+\ldots+\alpha_m x_m = 0.$$

Multiplicando ambos miembros de esta igualdad escalarmente por x_i con $i = 1, 2, \ldots, m$, tendremos

$$\alpha_1(x_1, x_i)+\alpha_2(x_2, x_i)+\ldots+\alpha_m(x_m, x_i) = 0;$$

puesto que $(x_i, x_k) = 0$ para $i \neq k$ y $(x_i, x_k) \neq 0$ para todo $i = 1, 2, \ldots, m$, de aquí resulta que $\alpha_i = 0$ para todo $i = 1, 2, \ldots, m$.

Teorema 1. En todo espacio euclídeo R existen bases ortonormales.

Demostración. Sea $g_1, g_2, \ldots, g_n$ una base cualquiera del espacio R. Pongamos $f_1 = g_1$ y $f_2 = g_2 + \alpha f_1$ y escojamos α de modo que los vectores f_1 y f_2 sean ortogonales

$$(g_2+\alpha f_1, f_1) = (g_2, f_1)+\alpha(f_1, f_1) = 0;$$

luego,

$$\alpha = -\frac{(g_2, f_1)}{(f_1, f_1)}.$$

Pueto que $f_1 \neq 0$, el denominador (f_1, f_1) de esta fracción es diferente de cero. El vector f_2 que hemos obtenido es no nulo, ya que los vectores g_1 y g_2 son linealmente independientes.

Supongamos ahora que hemos encontrado ya unos vectores no nulos $f_1, f_2, \ldots, f_{k-1}$ ortogonales dos a dos. Tomemos

$$f_k = g_k + \lambda_1 f_1 + \lambda_2 f_2 + \ldots + \lambda_{k-1} f_{k-1}$$

y escojamos los números $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ de modo que el vector f_k sea ortogonal a $f_1, f_2, \ldots, f_{k-1}$. Para ello es necesario que se cumplan las igualdades

$$(f_k, f_i) = (g_k, f_i) + \lambda_i(f_i, f_i) = 0$$

para $i = 1, 2, \ldots, k-1$; de aquí resulta

$$\lambda_i = -\frac{(g_k, f_i)}{(f_i, f_i)}.$$

El denominador (f_i, f_i) de esta fracción es distinto de cero, ya que todos los vectores f_i con $i = 1, 2, \ldots, k-1$ son no nulos por hipótesis. Puesto que los vectores $g_1, g_2, \ldots, g_k$ son linealmente independientes, el vector obtenido f_k también es no nulo.

Continuemos esta construcción hasta encontrar el último vector (no nulo)

$$f_n = g_n + \xi_1 f_1 + \xi_2 f_2 + \ldots + \xi_{n-1} f_{n-1},$$

ortogonal a todos los vectores anteriores $f_1, f_2, \ldots, f_{n-1}$. Los vectores $f_1, f_2, \ldots, f_n$ son, en virtud del último lema, linealmente independientes y, por consiguiente, constituyen una base (ortogonal). Dividiendo ahora todo vector f_i por su módulo, obtendremos una base ortonormal formada por los vectores

$$e_1 = \frac{f_1}{|f_1|}, \quad e_2 = \frac{f_2}{|f_2|}, \ \ldots, e_n = \frac{f_n}{|f_n|}.$$

Es fácil ver que *siendo la base* $g_1, g_2, \ldots, g_n$ *ortogonal resulta* $f_1 = g_1$, $f_2 = g_2$, $\ldots, f_n = g_n$ *y que siendo esta base ortonormal resulta* $e_1 = g_1, e_2 = g_2, \ldots, e_n = g_n$.

El método que hemos aplicado para obtener un sistema ortonormal de vectores a partir de un sistema linealmente independiente dado lleva el nombre de *proceso de ortogonalización.*

Observación. *Si R_1 es un subespacio de R y $e_1, e_2, \ldots, e_k$ es una base ortonormal de R_1, los vectores $e_1, e_2, \ldots, e_k$ pueden ser incluidos en una base ortonormal de todo el espacio.* Para demostrar esto es suficiente completar el sistema $e_1, e_2, \ldots, e_k$ hasta obtener una base del espacio R y realizar la ortogonalización del conjunto obtenido de vectores comenzando por $e_1, e_2, \ldots, e_k$.

Ejemplo. *Hállese una base ortogonal en el espacio de los polinomios de grado no mayor que 4 definidos sobre el segmento* $[-1, 1]$.

Solución. Tomemos por base inicial

$$g_0 = 1, g_1 = t, g_2 = t^2, g_3 = t^3, g_4 = t^4.$$

Pongamos

$$f_0 = g_0 = 1$$

y $f_1 = g_1 + \alpha f_0$. Puesto que $(g_1, f_0) = \int_{-1}^{1} t\,dt = 0$, resulta $\alpha = 0$ y

$$f_1 = g_1 = t.$$

Pongamos después $f_2 = g_2 + \beta f_0 + \gamma f_1$. Tenemos $(g_2, f_0) = \int_{-1}^{1} t^2 dt = \frac{2}{3}$, $(f_0, f_0) = \int_{-1}^{1} dt = 2$, de donde $\beta = -\frac{1}{3}$, y $(g_2, f_1) = \int_{-1}^{1} t^3\,dt = 0$, es decir, $\gamma = 0$. Por consiguiente,

$$f_2 = t^2 - \frac{1}{3}.$$

Sea $f_3 = g_3 + \lambda f_0 + \mu f_1 + \nu f_2$. Tenemos $(g_3, f_0) = \int_{-1}^{1} t^3\,dt = 0$, de donde

$\lambda = 0$; $(g_3, f_1) = \int\limits_{-1}^{1} t^4\, dt = \frac{2}{5}$ y $(f_1, f_1) = \int\limits_{-1}^{1} t^2\, dt = \frac{2}{3}$, de donde $\mu = -\frac{3}{5}$,

y $(g_3, f_2) = \int\limits_{-1}^{1} \left(t^5 - \frac{1}{3} t^3\right) dt = 0$, es decir, $\nu = 0$. Por consiguiente,

$$f_3 = t^3 - \frac{3}{5} t.$$

Pongamos, finalmente, $f_4 = g_4 + \xi f_0 + \eta f_1 + \zeta f_2 + \varrho f_3$. Puesto que $(g_4, f_0) = \int\limits_{-1}^{1} t^4\, dt = \frac{2}{5}$ y $(f_0, f_0) = 2$, resulta $\xi = -\frac{1}{5}$; tenemos, además, $(g_4, f_1) = \int\limits_{-1}^{1} t^5\, dt = 0$, es decir, $\eta = 0$. Tenemos después

$$(g_4, f_2) = \int\limits_{-1}^{1} t^4 \left(t^2 - \frac{1}{3}\right) dt = \frac{2}{7} - \frac{2}{15} = \frac{16}{105}$$

y

$$(f_2, f_2) = \int\limits_{-1}^{1} \left(t^4 - \frac{2}{3} t^2 + \frac{1}{9}\right) dt = \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \frac{8}{45},$$

de donde $\zeta = -\frac{6}{7}$; finalmente, $(g_4, f_3) = \int\limits_{-1}^{1} t^4 \left(t^3 - \frac{3}{5} t\right) dt = 0$, es decir, $\varrho = 0$. Por consiguiente,

$$f_4 = t^4 - \frac{1}{5} - \frac{6}{7}\left(t^2 - \frac{1}{3}\right) = t^4 - \frac{6}{7} t^2 + \frac{3}{35}.$$

Los polinomios obtenidos

$$f_0, f_1, f_2, f_3 \text{ y } f_4$$

coinciden (salvo un factor) con los cinco primeros *polinomios de Legendre* que desempeñan un papel importante en diferentes secciones de la Física matemática.

Hallemos *la expresión en coordenadas del producto escalar*. Sea $e_1, e_2, \ldots, e_n$ una base cualquiera de un espacio R provisto de producto escalar y sea

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n \quad \text{e} \quad y = y_1 e_1 + y_2 e_2 + \ldots + y_n e_n.$$

Tenemos

$$(x, y) = \sum_{i,k} x_i y_k (e_i, e_k) = \sum_{i,k} g_{ik} x_i y_k,$$

donde $g_{ik} = (e_i, e_k)$. Si el espacio R es euclídeo y $e_1, e_2, \ldots, e_n$ es una base ortonormal, se tiene $(e_i, e_k) = 0$ para $i \neq k$ y $(e_i, e_i) = 1$ para todo $i = 1, 2, \ldots, n$ y, por consiguiente,

$$(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Recíprocamente, si en una base $e_1, e_2, \ldots, e_n$ el producto escalar de los vectores

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n \quad \text{e} \quad y = y_1 e_1 + y_2 e_2 + \ldots + y_n e_n$$

es igual a

$$x_1 y_1 + x_2 y_2 + \ldots + x_n y_n,$$

es fácil ver que esta base es **ortonormal**, ya que en este caso $(e_i, e_i) = 1$ y $(e_i, e_k) = 0$ para $i \neq k$. Es más, *si en una base determinada el cuadrado escalar de un vector cualquiera* $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$ *es igual a* $(x, x) = x_1^2 + x_2^2 + \ldots + x_n^2$, *esta base es ortonormal*, ya que $(e_i, e_i) = 1$ para todo i y para $i \neq k$ tenemos

$$(e_i + e_k, e_i + e_k) = \\ = 0^2 + 0^2 + \ldots + 1^2 + 0^2 + \ldots + 1^2 + 0^2 + \ldots + 0^2 = 2$$

(las unidades figuran en las posiciones i-ésima y k-ésima) y

$$(e_i + e_k, e_i + e_k) = (e_i, e_i) + 2(e_i, e_k) + (e_k, e_k) = 2 + 2(e_i, e_k),$$

de donde resulta $(e_i, e_k) = 0$.

Sea $e_1, e_2, \ldots, e_n$ una base ortonormal y sea $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$. Multiplicando escalarmente ambos miembros de la última igualdad por e_i, obtenemos $(x, e_i) = x_i$, es decir, la i-ésima coordenada del vector x en una base ortonormal es igual al producto escalar de x por el vector básico e_i. (Este producto escalar puede ser llamado *proyección* del vector x sobre el vector e_i. Por consiguiente, las coordenadas de un vector en una base ortonormal son sus proyecciones sobre los vectores básicos.)

Definición 4. *Dos espacios R y R' provistos de producto escalar se llaman isomorfos, si entre sus elementos se puede establecer una correspondencia biyectiva tal que siendo $x \leftrightarrow x'$ e $y \leftrightarrow y'$, donde x,*

$y \in R$ y $x', y' \in R'$, no sólo $x+y \leftrightarrow x'+y'$ y $\alpha x \leftrightarrow \alpha x'$ sino también $(x, y) = (x', y')$.

Teorema 2. *Todos los espacios (vectoriales) euclídeos de n dimensiones son isomorfos.*

Demostración. Sean R y R' dos espacios euclídeos de n dimensiones. Tomemos en cada uno de ellos una base ortonormal ($e_1, e_2, \ldots, e_n$ en R y $e'_1, e'_2, \ldots, e'_n$ en R') y pongamos en correspondencia a todo elemento de R el elemento de R' *con las mismas coordenadas;* entonces, como se sabe (véase el § 3 del capítulo II), a la suma de elementos de R le corresponde la suma de los elementos respectivos de R' y al producto de un elemento de R por un número le corresponde el producto del elemento respectivo de R' por el mismo número. Además, si

$x = x_1e_1+x_2e_2+\ldots+x_ne_n$ e $y = y_1e_1+y_2e_2+\ldots+y_ne_n$ (y, por consiguiente, $x' = x_1e'_1+x_2e'_2+\ldots+x_ne'_n$ e $y' = y_1e'_1+y_2e'_2+\ldots+y_ne'_n$), tenemos para los productos escalares

$$(x, y) = x_1y_1+x_2y_2+\ldots+x_ny_n = (x', y').$$

Luego, los espacios R y R' tienen la misma estructura: los vectores correspondientes tienen longitudes iguales $\left(|x| = \sqrt{(x, x)} = \sqrt{(x', x')} = |x'|\right)$ y los ángulos entre los pares de vectores correspondientes son iguales

$$\left(\cos\left(\widehat{x, y}\right) = \frac{(x, y)}{|x|\,|y|} = \frac{(x', y')}{|x'|\,|y'|} = \cos\left(\widehat{x', y'}\right)\right).$$

§ 3. Complemento ortogonal

Definición 5. *Dos subespacios R_1 y R_2 de un espacio euclídeo R se llaman **recíprocamente ortogonales** si todo vector de R_1 es ortogonal a todo vector de R_2* (en este caso escribiremos $R_1 \perp R_2$).

Así, en el espacio corriente de tres dimensiones son ortogonales un plano π que pasa por el origen de coordenadas (el plano se comprende como el conjunto de todos los vectores que pertenecen a π) y la recta l perpendicular al plano (y que también pasa por el origen de coordenadas) (fig. 8, *a*). Al contrario, dos planos π_1 y π_2 perpendiculares en el sentido de la Geometría elemental (fig. 8, *b*) no son subespacios ortogonales en el sentido de la definición dada, pues, de $a_1 \in \pi_1$ y $a_2 \in \pi_2$ no resulta aun que $a_1 \perp a_2$.

Para que dos subespacios R_1 y R_2 sean recíprocamente ortogonales es necesario y suficiente que todos los vectores básicos de uno sean ortogonales a todos los vectores básicos del otro. La necesidad se deduce de la definición 5; para demostrar la suficiencia, supongamos que $e_1, e_2, \ldots, e_k$ es una base de R_1 y que $f_1, f_2, \ldots, f_m$ es una base de R_2 con la particularidad de que $(e_i, f_j) = 0$ para todo $i = 1,$

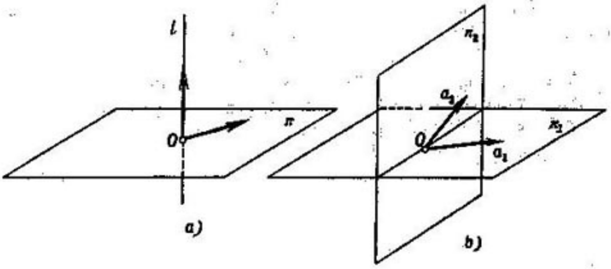


Fig. 8

$2, \ldots, k$ y $j = 1, 2, \ldots, m$; en este caso para todo $x = x_1e_1 + x_2e_2 + \ldots + x_ke_k$ y para todo $y = y_1f_1 + y_2f_2 + \ldots + y_mf_m$ se tiene

$$(x, y) = \sum_{i,j} x_iy_j(e_i, f_j) = 0$$

y, por consiguiente, estos vectores son ortogonales.

Demostremos que *la intersección de dos subespacios recíprocamente ortogonales es el vector nulo.*

Efectivamente, sean R_1 y R_2 dos subespacios recíprocamente ortogonales de R. Si el vector $x \in R_1 \cap R_2$, se tiene $x \in R_1$ y $x \in R_2$; pero entonces $(x, x) = 0$ y, por consiguiente, $x = 0$.

Sea R_1 un subespacio arbitrario de un espacio euclídeo R. Tomemos en R_1 una base ortonormal $e_1, e_2, \ldots, e_r$ y complementémosla hasta obtener una base ortonormal $e_1, e_2, \ldots, e_r, e_{r+1}, \ldots, e_n$ de todo el espacio. Los vectores $e_{r+1}, \ldots, e_n$ generan un subespacio R_2 de $n-r$ dimensiones que es, obviamente, ortogonal a R_1.

Demostremos que *todo vector x de R que es ortogonal a R_1 pertenece a R_2.* Efectivamente, si el vector

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

es ortogonal a R_1, se tiene

$$(x, e_i) = x_i = 0 \quad \text{para} \quad i = 1, 2, \ldots, r$$

y, por consiguiente,

$$x = x_{r+1}e_{r+1} + \ldots + x_ne_n \in R_2.$$

Definición 6. *El subespacio R_2 formado por todos los vectores de R ortogonales a todos los vectores de R_1 se llama* **complemento ortogonal** *de R_1*; indicaremos este subespacio R_2 por $R_1^{\perp}$.

Es fácil ver que el complemento ortogonal de un subespacio r-dimensional es de $n-r$ dimensiones y que el complemento ortogonal de $R_1^{\perp}$ coincide con R_1, es decir,

$$(R_1^{\perp})^{\perp} = R_1.$$

Los subespacios R_1 y $R_1^{\perp}$ generan todo el espacio R y la intersección de ellos es el vector nulo. Por consiguiente, *el espacio euclídeo R es la suma directa de cualquier subespacio suyo y del complemento ortogonal de éste:*

$$R = R_1 \oplus R_1^{\perp}.$$

Luego, todo vector x de R puede ser representado unívocamente en la forma de una suma $x = y+z$, donde $y \in R_1$ y $z \in R_1^{\perp}$ (teorema 6 del capítulo II). El vector y puede ser denominado *proyección ortogonal del vector x sobre el subespacio R_1*. Se acepta que el **ángulo entre el vector x y el subespacio R_1** es igual al ángulo entre el vector x y su proyección y sobre R_1; por consiguiente, el coseno de este ángulo es igual a

$$\frac{(x, y)}{|x|\,|y|} = \frac{(y+z, y)}{|x|\,|y|} = \frac{(y, y)}{|x|\,|y|} = \frac{|y|^2}{|x|\,|y|} = \frac{|y|}{|x|}.$$

Consideremos de nuevo un sistema de ecuaciones lineales homogéneas

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n = 0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n = 0. \end{array}\right. \tag{4}$$

Se puede dar la siguiente interpretación geométrica a este sistema. *En un espacio euclídeo R^n se han fijado* (en una base ortonormal) *m vectores $a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, m$.* El problema consiste en *hallar todos los vectores $x = (x_1, x_2, \ldots, x_n)$ ortogonales a cada uno de los vectores $a_1, a_2, \ldots, a_m$.*

Sea r el rango de la matriz $A = [a_{ik}]$. Si el vector x es ortogonal a todos los vectores a_i, también es ortogonal al espacio r-dimensional R_1 que éstos generan. Por consiguiente, los vectores soluciones x forman el *complemento ortogonal* $R_1^{\perp}$ del subespacio R_1. La dimensión de $R_1^{\perp}$ (es decir, el número máximo de soluciones linealmente independientes del sistema (4)) es igual, como hemos visto, a $n-r$. Todo sistema fundamental de soluciones de las ecuaciones (4) es una base del subespacio $R_1^{\perp}$.

§ 4. Espacio euclídeo (puntual vectorial)

Sea A^n un espacio afín n-dimensional y sea R^n su espacio vectorial correspondiente provisto de una métrica euclídea (es decir, de un producto escalar que satisface a las condiciones 1, 2, 3 y 4 del § 1). En el espacio A^n se puede definir la *distancia* entre dos cualesquiera de sus puntos M y N tomándola igual al módulo del vector $\overline{MN}$ y el *ángulo* MPN tomándolo igual al ángulo entre los vectores $\overline{PM}$ y $\overline{PN}$.

El espacio A^n con la métrica introducida de esta forma se denomina simplemente ***espacio euclídeo*** (a diferencia del espacio euclídeo vectorial introducido anteriormente).

Por consiguiente, un espacio euclídeo n-dimensional se puede definir mediante los cinco siguientes grupos de axiomas:

I. **Axiomas de adición de vectores** (axiomas 1, 2, 3 y 4 de la pág. 64).

II. **Axiomas de multiplicación de un vector por un número** (axiomas 5, 6, 7 y 8 de la pág. 64).

III. **Axioma de dimensión:** *existen n vectores linealmente independientes, pero no existe más de n vectores linealmente independientes* (pág. 66).

IV. **Axiomas que relacionan los vectores y los puntos** (axiomas 1 y 2 de la pág. 84).

V. **Axiomas del producto escalar** (axiomas 1, 2, 3 y 4 de la pág. 132).

Se puede demostrar que todos los espacios euclídeos n-dimensionales «tienen la misma estructura» (son isomorfos). En particular, para $n = 2$ tenemos el plano corriente y para $n = 3$, el espacio corriente de tres dimensiones.

Supongamos que en un espacio A^n se tiene un plano k-dimensional R_1 que *pasa por el origen de coordenadas*

$$\left\{\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n = 0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n = 0 \end{array}\right. \qquad (5)$$

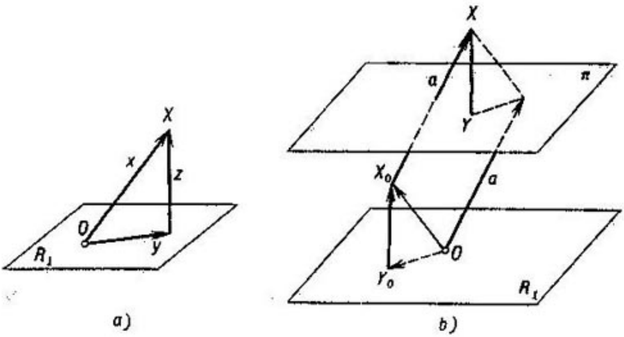


Fig. 9

(es decir, un subespacio k-dimensional) y un punto $X(\xi_1, \xi_2, \ldots, \xi_n)$. El vector $x = \overline{OX}$ puede ser representado entonces en la forma $x = y + z$, donde $y \in R_1$ y $z \in R_1^{\perp}$ (fig. 9, a). La longitud del vector z se denomina *distancia entre el punto X y el subespacio R_1*.

Supongamos ahora que en A^n se tiene un plano k-dimensional π *cualquiera*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

y un punto $X(\xi_1, \xi_2, \ldots, \xi_n)$. El plano π se obtiene trasladando paralelamente en un vector a el subespacio R_1 correspondiente a este plano definido por el sistema de ecuaciones (5). El punto X se obtiene entonces trasladando en el mismo vector a un punto determinado X_0 (y, por consiguiente, $\overline{OX} = \overline{OX_0} + a$; véase la fig. 9, b). *La distancia entre el punto X y el plano k-dimensional π* se toma igual a la distancia entre el punto X_0 y el subespacio R_1. Un plano k-dimensional π_1 y un plano l-dimensional π_2 son *ortogonales* si son ortogonales los subespacios π_1^0 y π_2^0 que les corresponden. En este caso todo vector $\overline{MN}$, donde $M, N \in \pi_1$, es ortogonal a todo vector $\overline{PQ}$, donde $P, Q \in \pi_2$ (fig. 10).

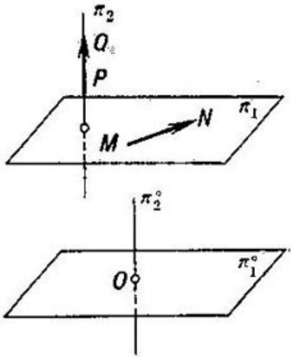


Fig. 10

En un espacio euclídeo A^n se llama ***hiperesfera*** al conjunto de todos los puntos que se encuentran a una misma distancia r (*radio* de la hiperesfera) de un punto fijo Q *(centro)*. La ecuación de una hiperesfera de radio r y de centro en el punto $Q(\alpha_1, \alpha_2, \ldots, \alpha_n)$ en un sistema de coordenadas de base ortonormal es, evidentemente,

$$(x_1-\alpha_1)^2+(x_2-\alpha_2)^2+\ldots+(x_n-\alpha_n)^2 = r^2.$$

Queda claro así que la hiperesfera es un caso particular de una superficie de segundo grado (véase el capítulo VII).

CAPÍTULO V

APLICACIONES LINEALES EN UN ESPACIO EUCLÍDEO

§ 1. Aplicación conjugada de una dada

Definición 1. *Sea $\mathcal{A}$ una aplicación lineal de un espacio euclídeo R. La aplicación lineal $\mathcal{A}^*$ tal que*

$$(\mathcal{A}x, y) = (x, \mathcal{A}^*y)$$

para todos los $x, y \in R$ se llama ***conjugada de*** $\mathcal{A}$.

Para toda aplicación lineal $\mathcal{A}$ existe no más de una aplicación conjugada $\mathcal{A}^*$. Esto se deduce del lema siguiente.

Lema. *Si $\mathcal{B}$ y $\mathcal{C}$ son unas aplicaciones lineales de un espacio R y si*

$$(x, \mathcal{B}y) = (x, \mathcal{C}y)$$

para todos los x e y de R, se tiene $\mathcal{B} = \mathcal{C}$.

Demostración. De la igualdad $(x, \mathcal{B}y) = (x, \mathcal{C}y)$ se deduce que

$$(x, \mathcal{B}y)-(x, \mathcal{C}y) = (x, \mathcal{B}y-\mathcal{C}y) = 0,$$

es decir, que $(x, (\mathcal{B}-\mathcal{C})y) = 0$ para todos los x e y.

Tomando, en particular, $x = (\mathcal{B}-\mathcal{C})y$, obtenemos

$$((\mathcal{B}-\mathcal{C})y, (\mathcal{B}-\mathcal{C})y) = 0$$

y, por consiguiente, en virtud de la propiedad 4 del producto escalar se tiene $(\mathcal{B}-\mathcal{C})y = 0$, es decir, $\mathcal{B}y = \mathcal{C}y$ para todo $y \in R$. Pero esto significa precisamente que $\mathcal{B} = \mathcal{C}$.

Demostremos ahora que *para toda aplicación lineal $\mathcal{A}$* ***existe*** *la aplicación conjugada $\mathcal{A}^*$.* Sea $A = [a_{ik}]$ la matriz de la aplicación lineal $\mathcal{A}$ en una base *ortonormal* $e_1, e_2, \ldots, e_n$, sea A' la matriz traspuesta de A y sea $\mathcal{A}^*$ la aplicación lineal de matriz A' en esta misma base. Es evidente entonces que

$$(\mathcal{A}\ e_i,\ e_k) = (a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n,\ e_k) = a_{ki}$$

y

$$(e_i,\ \mathcal{A}^*\ e_k) = (e_i,\ a_{k1}e_1 + a_{k2}e_2 + \ldots + a_{kn}e_n) = a_{ki},$$

es decir, $(\mathcal{A}e_i,\ e_k) = (e_i,\ \mathcal{A}^*e_k)$ para todos los i y k. Pero en este caso, siendo $x = \sum_{i=1}^{n} x_ie_i$ e $y = \sum_{k=1}^{n} y_ke_k$, se tiene

$$(\mathcal{A}x,\ y) = \left(\mathcal{A}\sum_{i=1}^{n} x_ie_i,\ \sum_{k=1}^{n} y_ke_k\right) = \sum_{i,\,k} x_iy_k(\mathcal{A}e_i,\ e_k)$$

y

$$(x,\ \mathcal{A}^*y) = \left(\sum_{i=1}^{n} x_ie_i,\ \mathcal{A}^*\sum_{k=1}^{n} y_ke_k\right) = \sum_{i,\,k} x_iy_k(e_i,\ \mathcal{A}^*e_k) =$$
$$= \sum_{i,\,k} x_iy_k(\mathcal{A}e_i,\ e_k) = (\mathcal{A}x,\ y),$$

es decir, la aplicación $\mathcal{A}^*$ es la *conjugada de* $\mathcal{A}$.

Hemos demostrado que *para toda aplicación lineal* $\mathcal{A}$ *de un espacio euclídeo existe una aplicación única conjugada* $\mathcal{A}^*$, *cuya matriz en cualquier base ortonormal es la traspuesta de la matriz de la aplicación* $\mathcal{A}$.

Sean $\mathcal{A}$ y $\mathcal{B}$ unas aplicaciones lineales cualesquiera de un espacio euclídeo R y sea $\mathcal{E}$ la aplicación idéntica. Entonces:

1. $\mathcal{E}^* = \mathcal{E}$, ya que

$$(x,\ \mathcal{E}^*\ y) = (\mathcal{E}\ x,\ y) = (x,\ y) = (x,\ \mathcal{E}\ y)$$

(en otras palabras la matriz de la aplicación $\mathcal{E}$ en cualquier base coincide con su traspuesta: $E' = E$).

2. $(\mathcal{A}^*)^* = \mathcal{A}$, ya que indicando $(\mathcal{A}^*)^*$ simplemente por $\mathcal{A}^{**}$ tenemos

$$(\mathcal{A}x,\ y) = (x,\ \mathcal{A}^*\ y) = (\mathcal{A}^*\ y,\ x) = (y,\ \mathcal{A}^{**}x) = (\mathcal{A}^{**}x,\ y)$$

(en otras palabras, la doble trasposición no altera la matriz: $(A')' = A$).

3. $(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^*+\mathcal{B}^*$, ya que

$$(x, (\mathcal{A}+\mathcal{B})^*y) = ((\mathcal{A}+\mathcal{B})x, y) = (\mathcal{A}x+\mathcal{B}x, y) =$$
$$= (\mathcal{A}x, y)+(\mathcal{B}x, y) = (x, \mathcal{A}^* y)+(x, \mathcal{B}^* y) =$$
$$= (x, \mathcal{A}^* y+\mathcal{B}^* y) = (x, (\mathcal{A}^*+\mathcal{B}^*)y)$$

(en otras palabras, la matriz traspuesta de una suma es igual a la suma de las matrices traspuestas de los sumandos: $(A+B)' = A'+B'$).

4. $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$, ya que

$$(x, (\mathcal{A}\mathcal{B})^*y) = ((\mathcal{A}\mathcal{B})x, y) = (\mathcal{A}(\mathcal{B}x), y) = (\mathcal{B}x, \mathcal{A}^*y) =$$
$$= (x, \mathcal{B}^*(\mathcal{A}^*y)) = (x, (\mathcal{B}^*\mathcal{A}^*)y)$$

(y, por consiguiente, para las matrices tiene lugar la igualdad $(AB)' = B'A'$).

5. Si existe $\mathcal{A}^{-1}$, se tiene $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}$, ya que de la igualdad $\mathcal{A}\mathcal{A}^{-1} = \mathcal{E}$ y de los puntos 4 y 1 resulta $(\mathcal{A}\mathcal{A}^{-1})^* = (\mathcal{A}^{-1})^*\mathcal{A}^* = \mathcal{E}^* = \mathcal{E}$, es decir, $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}$ (y, por consiguiente, para las matrices se tiene $(A^{-1})' = (A')^{-1}$).

§ 2. Aplicación autoconjugada

Definición 2. *Toda aplicación que coincide con su conjugada se llama* ***autoconjugada*** *o simétrica.*

Si $\mathcal{A}$ es una aplicación autoconjugada, para todos los x e y de R se cumple la igualdad

$$(\mathcal{A}x, y) = (x, \mathcal{A} y).$$

Sea $A = [a_{ik}]$ la matriz de una aplicación autoconjugada $\mathcal{A}$ en una base *ortonormal*; entonces, $A' = A$, es decir, $a_{ik} = a_{ki}$. Toda matriz de este tipo se llama *simétrica*.

La aplicación idéntica es autoconjugada, ya que $\mathcal{E}^* = \mathcal{E}$.

La suma de unas aplicaciones autoconjugadas es una aplicación autoconjugada, ya que, siendo $\mathcal{A}^* = \mathcal{A}$ y $\mathcal{B}^* = \mathcal{B}$, se tiene

$$(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^*+\mathcal{B}^* = \mathcal{A}+\mathcal{B}.$$

La aplicación inversa de una aplicación autoconjugada no degenerada es una aplicación autoconjugada ya que, siendo $\mathcal{A}^* = \mathcal{A}$, se tiene

$$(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1} = \mathcal{A}^{-1}.$$

Para que el producto de unas aplicaciones autoconjugadas sea autoconjugada es necesario y suficiente que estas aplicaciones *conmuten*. Efectivamente, siendo $\mathcal{A}^* = \mathcal{A}$ y $\mathcal{B}^* = \mathcal{B}$, se tiene $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{B}\mathcal{A}$ lo que es igual a $\mathcal{A}\mathcal{B}$ si, y sólo si, las aplicaciones $\mathcal{A}$ y $\mathcal{B}$ conmutan, es decir, si $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$.

Teorema 1. *Si el subespacio R_1 es invariante respecto a una aplicación lineal $\mathcal{A}$, su complemento ortogonal $R_1^\perp$ es invariante respecto a la aplicación $\mathcal{A}^*$ conjugada de $\mathcal{A}$.*

Demostración. Sea x un vector cualquiera de $R_1^\perp$ y sea y un vector cualquiera de R_1. Puesto que $\mathcal{A}y \in R_1$ y $x \in R_1^\perp$, es decir, $x \perp \mathcal{A}y$, tenemos

$$(\mathcal{A}^*x, y) = (x, \mathcal{A}\, y) = 0.$$

Por consiguiente, $\mathcal{A}^*\, x \in R_1^\perp$ y $R_1^\perp$ es invariante respecto de $\mathcal{A}^*$.

Corolario. *Si $\mathcal{A}$ es una aplicación autoconjugada y R_1 es un subespacio invariante respecto de $\mathcal{A}$, también $R_1^\perp$ es invariante respecto de $\mathcal{A}$.*

Efectivamente, en virtud del teorema 1, $R_1^\perp$ es invariante respecto de $\mathcal{A}^*$; pero $\mathcal{A}^* = \mathcal{A}$; luego, $R_1^\perp$ es invariante respecto de $\mathcal{A}$.

Teorema 2. *Todas las raíces del polinomio característico de una aplicación autoconjugada $\mathcal{A}$ son reales.*

Demostración. Sea $\alpha+i\beta$ una raíz compleja del polinomio característico de una aplicación autoconjugada $\mathcal{A}$. Como puede verse de la demostración del teorema 8 del capítulo III, en el espacio R existe entonces un subespacio bidimensional generado por unos vectores u y v tal que

$$\begin{aligned} \mathcal{A}u &= \alpha u - \beta v, \\ \mathcal{A}v &= \beta u + \alpha v, \end{aligned} \tag{1}$$

donde $\beta \neq 0$ y los vectores u y v no son nulos simultáneamente. (Si el propio espacio R es bidimensional y no existen en él vectores propios, este subespacio coincide con todo el R.) Multiplicando escalarmente la primera de las igualdades (1) por v y la segunda por u, obtenemos

$$\begin{aligned} (\mathcal{A}u, v) &= \alpha(u, v) - \beta(v, v) \\ (u, \mathcal{A}v) &= \beta(u, u) + \alpha(u, v). \end{aligned}$$

Puesto que $(\mathcal{A}u, v) = (u, \mathcal{A}v)$, resulta que $\beta(|u|^2+|v|^2) = 0$, es decir, $\beta = 0$ lo que contradice a lo supuesto.

Teorema 3. *La matriz de una aplicación autoconjugada puede ser reducida en una base ortonormal determinada a la forma diagonal.*

Demostración. Sea λ_1 uno de los valores propios de la aplicación autoconjugada $\mathcal{A}$. Por lo visto en el teorema 2, λ_1 es real. Indiquemos por e_1 el vector propio correspondiente a λ_1; entonces $\mathcal{A}e_1 = \lambda_1 e_1$. Podemos aceptar que el vector e_1 es de longitud unitaria, ya que de lo contrario podríamos sustituirlo por el vector $\frac{e_1}{|e_1|}$, es decir, por el vector propio unitario correspondiente a este mismo valor propio λ_1.

Indiquemos por R_1 el subespacio de dimensión uno generado por el vector e_1. Su complemento ortogonal $R_1^{\perp}$ será invariante respecto de $\mathcal{A}$ y, además, la aplicación $\mathcal{A}$ continuará siendo autoconjugada en este complemento. Sea λ_2 un valor propio (real) de la aplicación $\mathcal{A}$ en el subespacio $R_1^{\perp}$; indiquemos por e_2 el vector propio (unitario) correspondiente; entonces

$$\mathcal{A}e_2 = \lambda_2 e_2.$$

Sea R_2 el subespacio (invariante) generado por los vectores e_1 y e_2; el subespacio $R_2^{\perp}$ también será invariante respecto de $\mathcal{A}$. Continuando este razonamiento, encontraremos n vectores unitarios propios de la aplicación $\mathcal{A}$ ortogonales dos a dos (y, por consiguiente, linealmente independientes). En la base formada por estos vectores la matriz de $\mathcal{A}$ se reduce a la forma diagonal

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

El último teorema ofrece la interpretación geométrica de las aplicaciones autoconjugadas: si

$$x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

es un vector arbitrario de R, se tiene

$$\mathcal{A}x = x_1 \lambda_1 e_1 + x_2 \lambda_2 e_2 + \dots + x_n \lambda_n e_n.$$

Es decir, en la transformación de los puntos que corresponde a la aplicación $\mathcal{A}$ el punto $X(x_1, x_2, \dots, x_n)$ se transforma en el punto $X'(\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n)$ y, por consiguiente, esta transformación se reduce, en la base formada por los vectores propios de la aplica-

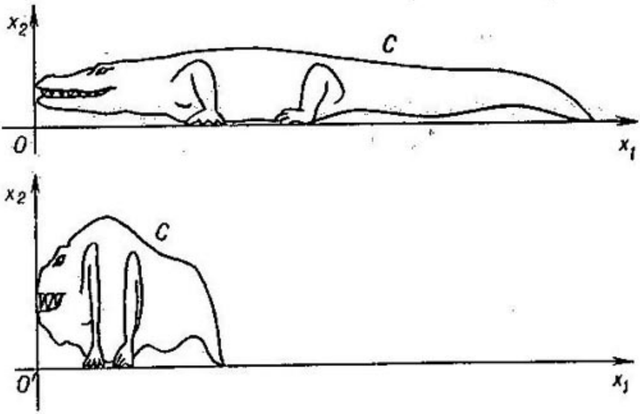


Fig. 11

ción $\mathcal{A}$, a n dilataciones a lo largo de los ejes de coordenadas con coeficientes iguales, respectivamente, a $\lambda_1, \lambda_2, \ldots, \lambda_n$ (véase la fig. 11 en la que se representa la acción de una aplicación autoconjugada de valores propios $\lambda_1 = \frac{1}{3}$ y $\lambda_2 = 2$ sobre la figura C del plano euclídeo).

§ 3. Aplicación ortogonal

Definición 3. *Una aplicación lineal $\mathcal{A}$ de un espacio euclídeo R se llama **ortogonal** si*

$$(\mathcal{A}x, \mathcal{A}y) = (x, y)$$

para todos los x e y de R.

Es decir, toda aplicación ortogonal conserva el producto escalar y, por consiguiente, *conserva las longitudes de los vectores y los ángulos entre éstos.*

Basta exigir, sin embargo, que la aplicación lineal *conserve las longitudes* de los vectores para que resulte ortogonal. Efectivamente, supongamos que

$$|\mathcal{A}x| = |x|$$

para *todo* x. En este caso

$$|\mathcal{A}(x+y)| = |x+y|.$$

Pero

$$|\mathcal{A}(x+y)|^2 = (\mathcal{A}(x+y), \mathcal{A}(x+y)) = (\mathcal{A}x+\mathcal{A}y, \mathcal{A}x+\mathcal{A}y) =$$
$$= (\mathcal{A}x, \mathcal{A}x)+2(\mathcal{A}x, \mathcal{A}y)+(\mathcal{A}y, \mathcal{A}y) = |\mathcal{A}x|^2+2(\mathcal{A}x, \mathcal{A}y)+$$
$$+|\mathcal{A}y|^2$$

y

$$|x+y|^2 = (x+y, x+y) = (x, x)+2(x, y)+(y, y) = |x|^2+$$
$$+2(x, y)+|y|^2,$$

de donde tomando en consideración las igualdades $|\mathcal{A}x| = |x|$, $|\mathcal{A}y| = |y|$ y $|\mathcal{A}(x+y)| = |x+y|$ obtenemos

$$(\mathcal{A}x, \mathcal{A}y) = (x, y).$$

Está claro que *toda aplicación ortogonal transforma cualquier base ortonormal en una base ortonormal*. Demostremos que es válida también la afirmación recíproca: *toda aplicación lineal $\mathcal{A}$ que transforma al menos una base ortonormal en una base ortonormal es ortogonal*. Efectivamente, supongamos que la aplicación $\mathcal{A}$ transforma una base ortonormal $e_1, e_2, \ldots, e_n$ en una base ortonormal $e'_1, e'_2, \ldots, e'_n$, es decir, $\mathcal{A}e_i = e'_i$ para $i = 1, 2, \ldots, n$. En este caso para $x = x_1e_1+x_2e_2+\ldots+x_ne_n$ e $y = y_1e_1+y_2e_2+\ldots$ $\ldots+y_ne_n$
tenemos

$$\mathcal{A}x = x_1e'_1+x_2e'_2+\ldots+x_ne'_n,$$
$$\mathcal{A}y = y_1e'_1+y_2e'_2+\ldots+y_ne'_n$$

y

$$(\mathcal{A}x, \mathcal{A}y) = x_1y_1+x_2y_2+\ldots+x_ny_n = (x, y).$$

Si $\mathcal{A}$ es una aplicación ortogonal y $\mathcal{A}^*$ es su aplicación conjugada, se tiene

$$(x, y) = (\mathcal{A}x, \mathcal{A}y) = (x, \mathcal{A}^*(\mathcal{A}y)) = (x, (\mathcal{A}^*\mathcal{A})y)$$

para todos los x e y de R. Luego, $\mathcal{A}^*\mathcal{A} = \mathcal{E}$ (lema del § 1), es decir,

$$\mathcal{A}^* = \mathcal{A}^{-1}.$$

Además, la igualdad $\mathcal{A}^*\mathcal{A} = \mathcal{E}$ o $\mathcal{A}^* = \mathcal{A}^{-1}$ es una condición **necesaria** y **suficiente** para que la aplicación lineal $\mathcal{A}$ sea ortogonal. De aquí resulta, en particular, que *toda aplicación ortogonal es no degenerada.*

La aplicación inversa de una aplicación ortogonal es también ortogonal, ya que siendo $\mathcal{A}^* = \mathcal{A}^{-1}$, se tiene

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

(de otra forma: una aplicación ortogonal es una aplicación biyectiva del espacio R sobre sí mismo que conserva las longitudes de los vectores; pero en este caso la aplicación inversa posee la misma propiedad).

La suma de unas aplicaciones ortogonales no es, en general, una aplicación ortogonal. En cambio, *el producto de unas aplicaciones ortogonales* conserva las longitudes de los vectores y, por consiguiente, *es también ortogonal*. De otra forma:

$$(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{B}^{-1}\mathcal{A}^{-1} = (\mathcal{A}\mathcal{B})^{-1}.$$

Sea $A = [a_{ik}]$ la matriz de una aplicación ortogonal $\mathcal{A}$ en una base *ortonormal* $e_1, e_2, \ldots, e_n$; puesto que $\mathcal{A}^*\mathcal{A} = \mathcal{E}$, tenemos

$$A'A = E,$$

es decir,

$$\begin{bmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

De aquí resulta que

$$a_{1i}a_{1k} + a_{2i}a_{2k} + \ldots + a_{ni}a_{nk} = 0$$

para $i \neq k$ y que

$$a_{1i}^2 + a_{2i}^2 + \ldots + a_{ni}^2 = 1$$

para todo $i = 1, 2, \ldots, n$, es decir, *las columnas de la matriz A —consideradas como vectores— forman ellas mismas un sistema* *ortonormal*. (Claro está, lo mismo resulta tmbién de que toda aplicación ortogonal transforma una base ortonormal en una base ortonormal: las imágenes $\mathcal{A}e_1, \mathcal{A}e_2, \ldots, \mathcal{A}e_n$ de los vectores básicos $e_1, e_2, \ldots, e_n$ constituyen una base ortonormal; por consiguiente, $(\mathcal{A}e_i, \mathcal{A}e_k) = a_{1i}a_{1k}+a_{2i}a_{2k}+\ldots+a_{ni}a_{nk} = 0$ para $i \neq k$

y $(\mathcal{A}e_i, \mathcal{A}e_i) = a_{1i}^2+a_{2i}^2+\ldots+a_{ni}^2 = 1$.)

Si $\mathcal{A}$ es una aplicación ortogonal, la aplicación $\mathcal{A}^* = \mathcal{A}^{-1}$ es también ortogonal; luego, las columnas de la matriz A', es decir, *las filas de la matriz A, también constituyen un sistema ortonormal:*

$$a_{i1}a_{k1}+a_{i2}a_{k2}+\ldots+a_{in}a_{kn} = 0$$

para $i \neq k$ y

$$a_{i1}^2+a_{i2}^2+\ldots+a_{in}^2 = 1.$$

Toda matriz A tal que

$$A' = A^{-1}$$

se llama *matriz ortogonal*. Por consiguiente, la matriz de una aplicación ortogonal es ortogonal en cualquier base ortonormal; recíprocamente, si en alguna base ortonormal la matriz de una aplicación $\mathcal{A}$ es ortogonal ($A^{-1} = A'$), se tiene $\mathcal{A}^{-1} = \mathcal{A}^*$ y la aplicación $\mathcal{A}$ también es ortogonal.

El determinante de una matriz ortogonal es igual a ± 1. Efectivamente, de la igualdad $AA' = E$ resulta

$$|AA'| = |A|\,|A'| = |E| = 1.$$

y puesto que $|A'| = |A|$, se tiene

$$|A|^2 = 1 \quad \text{y} \quad |A| = \pm 1.$$

Teorema 4. *Los valores propios de una aplicación ortogonal son iguales a* ± 1.

Demostración. Sea x un vector propio de una aplicación ortogonal $\mathcal{A}$ y sea λ el valor propio que le corresponde. Entonces,

$$(x, x) = (\mathcal{A}x, \mathcal{A}x) = (\lambda x, \lambda x) = \lambda^2(x, x),$$

de donde obtenemos (ya que $(x, x) \neq 0$)

$$\lambda^2 = 1 \quad \text{y} \quad \lambda = \pm 1.$$

Teorema 5. *Si el subespacio R_1 es invariante respecto a una aplicación ortogonal $\mathcal{A}$, su complemento ortogonal $R_1^{\perp}$ también es invariante respecto a $\mathcal{A}$.*

Demostración. Puesto que $\mathcal{A}$ es una aplicación ortogonal, tenemos $\mathcal{A}^{-1} = \mathcal{A}^*$. Según el teorema 1, el subespacio $R_1^\perp$ es invariante respecto a la aplicación $\mathcal{A}^* = \mathcal{A}^{-1}$; pero en este caso también será invariante, en virtud del teorema 9 del capítulo III, respecto a $(\mathcal{A}^{-1})^{-1} = \mathcal{A}$.

Analicemos *qué es lo que representa una aplicación ortogonal arbitraria*. Sea primero $\mathcal{A}$ una aplicación ortogonal de la **recta** R^1 y sea $e \in R^1$. Entonces, se tiene $\mathcal{A}e \in R^1$ y, por consiguiente, $\mathcal{A}e = \lambda e$, donde $\lambda = \pm 1$, es decir, $\mathcal{A}e = \pm e$ y $\mathcal{A}$ es o bien *la aplicación idéntica* o bien *la simetría central*.

Sea ahora $\mathcal{A}$ una aplicación ortogonal del **plano** y sea

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

su matriz en una base ortonormal. Como sabemos, debe ser

$$\begin{aligned} a_{11}^2 + a_{21}^2 &= 1, \\ a_{12}^2 + a_{22}^2 &= 1, \\ a_{11}a_{12} + a_{21}a_{22} &= 0. \end{aligned}$$

De las dos primeras igualdades resulta que existen unos φ y ψ tales que

$$a_{11} = \cos\varphi, \quad a_{21} = \operatorname{sen}\varphi \quad \text{y} \quad a_{12} = \cos\psi, \quad a_{22} = \operatorname{sen}\psi.$$

La tercera igualdad implica entonces

$$\cos\varphi \cdot \cos\psi + \operatorname{sen}\varphi \cdot \operatorname{sen}\psi = \cos(\psi - \varphi) = 0,$$

de donde resulta que

$$\psi - \varphi = \frac{\pi}{2} \text{ ó } \frac{3\pi}{2}.$$

En el primer caso se tiene $a_{12} = \cos\psi = -\operatorname{sen}\varphi$, $a_{22} = \operatorname{sen}\psi = \cos\varphi$ y

$$A = \begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}, \tag{2}$$

es decir, la aplicación $\mathcal{A}$ es *la rotación de ángulo φ alrededor del origen de coordenadas*. (En particular, para $\varphi = 0$ se tiene *la aplicación idéntica* y para $\varphi = \pi$ se tiene *la simetría respecto al origen de coordenadas*.)

En el segundo caso tenemos $a_{12} = \operatorname{sen} \varphi$, $a_{22} = -\cos \varphi$ y

$$A = \begin{bmatrix} \cos \varphi & \operatorname{sen} \varphi \\ \operatorname{sen} \varphi & -\cos \varphi \end{bmatrix}.$$

Esta matriz es simétrica y, por consiguiente, la aplicación ortogonal $\mathcal{A}$ es *autoconjugada*, es decir, en una base ortonormal (en general, nueva) su matriz se reduce a la forma diagonal

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

donde $\lambda_i = \pm 1$. El determinante de esta matriz es igual a $\lambda_1 \lambda_2$ y, por otro lado, debe ser igual a

$$\begin{vmatrix} \cos \varphi & \operatorname{sen} \varphi \\ \operatorname{sen} \varphi & -\cos \varphi \end{vmatrix} = -1;$$

luego uno de los valores λ_i es igual a -1 y el otro a $+1$ y la matriz de la aplicación $\mathcal{A}$ es de la forma

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

Un vector cualquiera x, igual en la base nueva a $x_1 e_1 + x_2 e_2$, se transforma en $x' = x_1 e_1 - x_2 e_2$. Se trata, pues, de *la simetría respecto a la recta* determinada por el vector e_1, es decir, por el primer vector básico de la base nueva (fig. 12).

Por consiguiente, *una aplicación ortogonal del plano es o bien una* ***rotación*** *de ángulo φ alrededor del origen de coordenadas* (en particular, **la aplicación idéntica o la simetría central;** el determinante de esta aplicación es igual a $+1$) *o bien una simetría axial* (con el determinante igual a -1).

De lo demostrado se desprenden, en particular, dos teoremas de la geometría elemental (plana):

1. *El producto de dos simetrías axiales es una rotación alrededor del punto de intersección de los ejes de simetría* (ya que es una aplicación ortogonal de determinante igual a $+1$).

2. *El producto de una rotación y de una simetría respecto a un eje que pasa por el centro de rotación es una simetría respecto a un eje nuevo que pasa por el mismo punto* (ya que es una aplicación ortogonal de determinante igual a -1).

Pasemos ahora al caso general de una aplicación ortogonal de un espacio de n dimensiones.

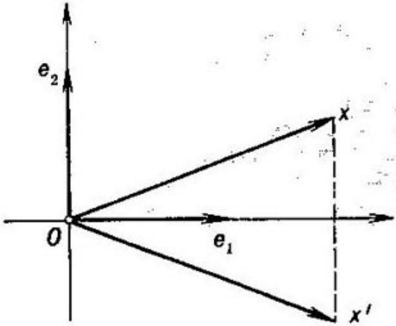


Fig. 12

Teorema 6. *La matriz de una aplicación ortogonal se reduce en una base ortonormal determinada a la forma*

$$\begin{bmatrix} 1 &&&&&&&&&&&&& \\ & 1 &&&&&&&&&&&& \\ && \ddots &&&&&&&&&&& \\ &&& 1 &&&&&&&&&& \\ &&&& -1 &&&&&&&&& \\ &&&&& \ddots &&&&&&&& \\ &&&&&& -1 &&&&&&& \\ &&&&&&& \cos\varphi_1 & -\operatorname{sen}\varphi_1 &&&&& \\ &&&&&&& \operatorname{sen}\varphi_1 & \cos\varphi_1 &&&&& \\ &&&&&&&&& \cos\varphi_2 & -\operatorname{sen}\varphi_2 &&& \\ &&&&&&&&& \operatorname{sen}\varphi_2 & \cos\varphi_2 &&& \\ &&&&&&&&&&& \ddots && \\ &&&&&&&&&&&& \cos\varphi_k & -\operatorname{sen}\varphi_k \\ &&&&&&&&&&&& \operatorname{sen}\varphi_k & \cos\varphi_k \end{bmatrix} \qquad (3)$$

(Todos los demás elementos de esta matriz son iguales a cero.)

Aplicaremos para la **demostración** el método de inducción matemática. Hemos visto ya que el teorema es válido para $n = 1$ y $n = 2$. Supongamos que el teorema es válido para todos los espacios de dimensión menor que n. Sea R un espacio euclídeo n-dimensional y sea $\mathcal{A}$ una aplicación ortogonal. Pueden darse dos casos.

1. *La aplicación $\mathcal{A}$ tiene un valor propio real $\lambda = \pm 1$* (necesariamente se da este caso si, por ejemplo, n es impar). Sea e_1 el correspondiente vector (unitario) propio (de modo que $\mathcal{A}\, e_1 = \pm e_1$)

y sea R_1 el subespacio de dimensión uno generado por el vector e_1. Según el teorema 5, el subespacio $R_1^{\perp}$ de $n-1$ dimensiones es invariante respecto de $\mathcal{A}$. Está claro que también en este subespacio $\mathcal{A}$ es una aplicación ortogonal. Por la hipótesis de inducción, en $R_1^{\perp}$ se puede encontrar una base ortonormal $e_2, e_3, \ldots, e_n$ en la que la matriz de la aplicación $\mathcal{A}$ se reduce a la forma (3). Teniendo en cuenta la observación hecha en el § 7 del capítulo III, obtenemos que (alterando posiblemente la numeración de los vectores básicos) la matriz de la aplicación $\mathcal{A}$ en todo el espacio R se reduce a la forma (3) en una base ortonormal determinada.

2. *La aplicación $\mathcal{A}$ no tiene valores propios reales*. Por lo visto en el teorema 8 del capítulo III, existe en R un subespacio invariante R_1 de dos dimensiones. Según hemos demostrado anteriormente en el plano R_1 se puede encontrar una base ortonormal e_1 y e_2 en la que la matriz de la aplicación $\mathcal{A}$ se reduce a la forma (2). (El segundo caso $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ no puede presentarse, ya que hemos aceptado que la aplicación $\mathcal{A}$ no tiene valores propios reales.)

El subespacio $R_1^{\perp}$ es invariante respecto a $\mathcal{A}$ y en él $\mathcal{A}$ es, por supuesto, ortogonal. Por la hipótesis de inducción, se puede encontrar en $R_1^{\perp}$ una base ortonormal $e_3, e_4, \ldots, e_n$ en la que la matriz de la aplicación $\mathcal{A}$ se reduce a la forma (3). (En este caso n es necesariamente par y la diagonal principal de esta matriz no contiene los números $+1$ y -1.) En virtud de la observación del § 7 del capítulo III, resulta que la matriz de la aplicación $\mathcal{A}$, considerada en todo el espacio R, se reduce a la forma (3) en la base ortonormal $e_1, e_2, \ldots, e_n$.

Del último teorema se puede extraer la interpretación geométrica de una aplicación ortogonal. Puesto que toda matriz de tipo (3) es el producto de varias matrices de tipo

$$\begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & & -1 & & & & \\ & & & & & 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & & 1 \end{bmatrix} \tag{4}$$

y de varias matrices de tipo

$$\begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & & \cos\varphi & -\operatorname{sen}\varphi & & & \\ & & & & \operatorname{sen}\varphi & \cos\varphi & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix}, \tag{5}$$

resulta que toda aplicación ortogonal $\mathcal{A}$ se puede efectuar realizando sucesivamente varias *simetrías respecto a los «hiperplanos coordenados»* (la matriz de cada una de estas simetrías es de la forma (4)) y varias *rotaciones alrededor de unos «ejes de $n-2$ dimensiones»*, es decir, unas rotaciones que representan giros idénticos que se realizan simultáneamente en todos los planos bidimensionales perpendiculares al «eje» de rotación de $n-2$ dimensiones (la matriz de cada una de estas rotaciones es de la forma (5)).

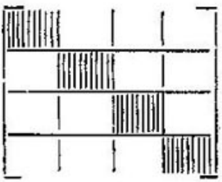
Uniendo dos $+1$ o dos -1 consecutivas de la matriz (3) en «células»

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos 0 & -\operatorname{sen} 0 \\ \operatorname{sen} 0 & \cos 0 \end{bmatrix} \text{ y } \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} =$$

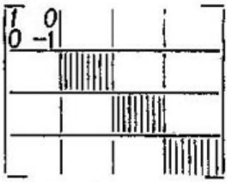
$$= \begin{bmatrix} \cos\pi & -\operatorname{sen}\pi \\ \operatorname{sen}\pi & \cos\pi \end{bmatrix},$$

obtendremos (después de cambiar, posiblemente, la numeración de los vectores básicos) cuatro tipos de matrices ortogonales (se han sombreado las «células» de tipo $\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$, donde φ puede ser igual, en particular, a cero o a π; todos los elementos de las células vacías son iguales a cero):

para n par

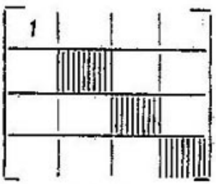


y



si $|A| = 1$ si $|A| = -1$.

para n impar



y

si $|A| = 1$ si $|A| = -1$.

§ 4. Aplicación lineal no degenerada arbitraria

Lema. *Si $\mathcal{A}$ es una aplicación lineal no degenerada arbitraria de un espacio euclídeo, la aplicación $\mathcal{A}^*\mathcal{A}$* (así como la aplicación *$\mathcal{A}\mathcal{A}^*$) es una aplicación autoconjugada y todos sus valores propios son positivos.*

Demostración. La aplicación $\mathcal{B} = \mathcal{A}^*\mathcal{A}$ es autoconjugada, ya que

$$\mathcal{B}^* = (\mathcal{A}^*\mathcal{A})^* = \mathcal{A}^*\mathcal{A}^{**} = \mathcal{A}^*\mathcal{A} = \mathcal{B}.$$

Sea ahora $e_1, e_2, \ldots, e_n$ una base ortonormal tal que la matriz de la aplicación autoconjugada $\mathcal{B}$ se reduce en ella a la forma diagonal

$$B = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \hdotsfor{4} \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

En esta base se tiene $\mathcal{B}e_i = \lambda_i e_i$. Por ello, si $x = x_1e_1+x_2e_2+ +\ldots+x_ne_n$ es un vector cualquiera de R, resulta

$$(\mathcal{B}x, x) = (x_1\mathcal{B}e_1+x_2\mathcal{B}e_2+\ldots+x_n\mathcal{B}e_n,\ x_1e_1+x_2e_2+\ldots+ +x_ne_n) =$$
$$= (x_1\lambda_1e_1+x_2\lambda_2e_2+\ldots+x_n\lambda_ne_n, x_1e_1+x_2e_2+\ldots+x_ne_n) = = \lambda_1x_1^2+\lambda_2x_2^2+\ldots+\lambda_nx_n^2.$$

Por otro lado,

$$(\mathcal{B}x, x) = (\mathcal{A}^*\mathcal{A}x, x) = (\mathcal{A}x, \mathcal{A}x) > 0,$$

ya que para $x \neq 0$ es $\mathcal{A}x \neq 0$ (pues la aplicación $\mathcal{A}$ es no degenerada); por consiguiente, $\lambda_1x_1^2+\lambda_2x_2^2+\ldots+\lambda_nx_n^2 > 0$ *para cualesquiera* $x_1, x_2, \ldots, x_n$. Pero entonces todo $\lambda_i > 0$ ya que, si al menos uno de estos coeficientes, digamos λ_k, fuese no positivo, para el vector $x = e_k$ resultaría $(\mathcal{B}x, x) = \lambda_k \leqslant 0$ lo cual es imposible.

Teorema 6. *Toda aplicación lineal no degenerada $\mathcal{A}$ de un espacio euclídeo puede ser representada como el producto*

$$\mathcal{A} = \mathcal{H}\mathcal{C}$$

de una aplicación autoconjugada $\mathcal{C}$ y una aplicación ortogonal $\mathcal{H}$.

Demostración. Tomemos como base del espacio R aquella en la que la matriz de la aplicación (autoconjugada) $\mathcal{B} = \mathcal{A}^*\mathcal{A}$ se reduce a la forma diagonal

$$B = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

donde todos los $\lambda_i > 0$, e indiquemos por $\mathcal{C}$ la aplicación cuya matriz en esta misma base es de la forma

$$C = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \ldots & 0 \\ 0 & \sqrt{\lambda_2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sqrt{\lambda_n} \end{bmatrix}.$$

Queda claro que $\mathcal{C}$ es una aplicación autoconjugada no degenerada y que $\mathcal{C}^2 = \mathcal{B}$. Si ahora $\mathcal{A} = \mathcal{H}\mathcal{C}$, tenemos $\mathcal{H} = \mathcal{A}\mathcal{C}^{-1}$

y resta probar solamente que la aplicación $\mathcal{H}$ es ortogonal. Pero esto resulta de la igualdad

$$\mathcal{H}^*\mathcal{H} = (\mathcal{A}\mathcal{C}^{-1})^* \mathcal{A}\mathcal{C}^{-1} = (\mathcal{C}^{-1})^* \mathcal{A}^*\mathcal{A}\mathcal{C}^{-1} =$$
$$= \mathcal{C}^{-1}\mathcal{B}\mathcal{C}^{-1} = \mathcal{C}^{-1}\mathcal{C}^2\mathcal{C}^{-1} = \mathcal{E}$$

y con esto queda demostrado el teorema.

Podríamos demostrar de la misma forma que la aplicación $\mathcal{A}$ puede ser representada en la forma $\mathcal{A} = \mathcal{C}_1\mathcal{H}_1$, donde $\mathcal{C}_1$ es una aplicación autoconjugada y $\mathcal{H}_1$ es una aplicación ortogonal. Por consiguiente, toda aplicación lineal no degenerada se reduce a varias simetrías respecto a hiperplanos, a varias rotaciones alrededor de «ejes» de $n-2$ dimensiones y a varias dilataciones a lo largo de rectas recíprocamente ortogonales.

§ 5. Espacio lineal complejo

Todo cuanto hemos expuesto hasta el momento se refería al espacio lineal real. *Un espacio lineal complejo* es un conjunto de elementos para los cuales están definidas dos operaciones: —la de adición y la de multiplicación por números complejos— y se cumplen las condiciones de la 1 a la 8 del § 1 del capítulo II. *Un espacio euclídeo complejo* es un espacio lineal complejo en el cual para todo par de vectores x e y está definido el producto escalar (en general, complejo) (x, y) con la particularidad de que se cumplen las condiciones siguientes[1]:

1. $(x, y) = \overline{(y, x)}$[2];
2. $(\alpha x, y) = \alpha (x, y)$;
3. $(x+y, z) = (x, z)+(y, z)$;
4. (x, x) es un número no negativo y de la igualdad $(x, x) = 0$ se deduce que $x = 0$[3];

De las condiciones 1, 2 y 3 obtenemos:

2'. $(x, \alpha y) = \overline{(\alpha y, x)} = \overline{\alpha (y, x)} = \bar{\alpha}\overline{(y, x)} = \bar{\alpha}(x, y)$,

3'. $(z, x+y) = \overline{(x+y, z)} = \overline{(x, z)+(y, z)} = \overline{(x, z)}+\overline{(y, z)} = (z, x)+(z, y)$.

Igual que antes la *longitud* del vector x se define por $\sqrt{(x, x)}$. Dos vectores x e y se llaman *ortogonales*, si $(x, y) = 0$. El concepto de *base ortonormal* se introduce igual que antes y de la misma forma se puede demostrar que toda base puede ser convertida en una ortonormal. Igual que en el caso del espacio

---

[1] En la literatura científica estos espacios se llaman *unitarios* y el término «espacio euclídeo complejo» se emplea en otro sentido.

[2] Por $\bar{\alpha}$ se designa el número conjugado de α.

[3] Observemos que de la condición 1 aplicada a dos vectores iguales $(x, x) = \overline{(x, x)}$ se deduce que el cuadrado escalar (x, x) de cualquier vector debe ser *real*.

real se introducen los conceptos de *aplicación lineal*, de su *matriz* en una base dada y de la *multiplicación* de aplicaciones lineales y de matrices. Tampoco varía la definición de los conceptos de *subespacio invariante*, de *vector propio* y de *valor propio* de una aplicación lineal. El primer teorema nuevo afirma:

En un espacio lineal complejo toda aplicación lineal tiene al menos un vector propio.

Esto se deduce del así llamado «teorema principal del Algebra», según el cual *toda ecuación de coeficientes complejos tiene al menos una raíz* (en general, compleja).

Una *aplicación autoconjugada* de un espacio euclídeo complejo se define por la condición: $(\mathcal{A}x, y) = (x, \mathcal{A}y)$ para todos los x e y. La matriz de una aplicación autoconjugada en cualquier base ortonormal satisface a la condición $A = \overline{A}'$, donde $\overline{A}$ es la matriz cuyos elementos son los conjugados de los elementos correspondientes de la matriz A.

Demostremos de nuevo que *todos los valores propios de una aplicación autoconjugada son reales.* Sea x un vector propio de una aplicación autoconjugada $\mathcal{A}$ y sea λ su correspondiente valor propio; en este caso,

$$(\mathcal{A}x, x) = (x, \mathcal{A}x)$$

o

$$(\lambda x, x) = (x, \lambda x),$$

de donde resulta

$$\lambda(x, x) = \overline{\lambda}(x, x)$$

y, puesto que $(x, x) \neq 0$, se tiene $\lambda = \overline{\lambda}$, es decir, λ es real.

Una aplicación lineal de un espacio euclídeo complejo se llama *unitaria* si $(\mathcal{A}x, \mathcal{A}y) = (x, y)$ para todos los $x, y \in R$. Si $A = [a_{ik}]$ es la matriz de una aplicación unitaria en una base ortonormal, se tiene

$$A\overline{A}' = \overline{A}'A = E,$$

es decir, $a_{i1}\bar{a}_{i1}+a_{i2}\bar{a}_{i2}+\ldots+a_{in}\bar{a}_{in} = 1$ y $a_{i1}\bar{a}_{k1}+a_{i2}\bar{a}_{k2}+\ldots+a_{in}\bar{a}_{kn} = 0$ para $i \neq k$. Una matriz de este tipo se llama *unitaria*.

Demostremos que *en un espacio euclídeo complejo el valor absoluto de todos los valores propios de una aplicación unitaria es igual a la unidad.* Sea $\mathcal{A}$ una aplicación unitaria, sea x un vector propio de la misma y sea λ el valor propio correspondiente. En estas condiciones

$$(x, x) = (\mathcal{A}x, \mathcal{A}x) = (\lambda x, \lambda x) = \lambda\overline{\lambda}(x, x).$$

Puesto que $(x, x) \neq 0$, se tiene $\lambda\overline{\lambda} = 1$, es decir, $|\lambda| = 1$.

Se puede demostrar que *en un espacio euclídeo complejo la matriz de una aplicación unitaria se reduce a la forma diagonal*

$$\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

donde todos los valores propios λ_i son de módulo igual a la unidad.

Los valores propios de una aplicación autoconjugada son reales; los valores propios de una aplicación unitaria son de módulo igual a la unidad. Entre todas las aplicaciones lineales de un espacio euclídeo complejo las aplicaciones autoconjugadas desempeñan, en cierto sentido, el mismo papel que desempeñan los números reales en el conjunto de todos los números complejos con la particularidad de que las aplicaciones autoconjugadas con todos los valores propios positivos (que se denominan aplicaciones *definidas positivas*) desempeñan el papel de los números reales y las aplicaciones unitarias desempeñan el papel de los números (complejos) de módulo igual a la unidad. Por ejemplo, es fácil demostrar que una aplicación lineal arbitraria $\mathcal{A}$ de un espacio euclídeo complejo puede ser representada en la forma $\mathcal{A} = \mathcal{B} + \mathcal{C} i$, donde $\mathcal{B}$ y $\mathcal{C}$ son aplicaciones autoconjugadas e $i = \sqrt{-1}$. Mayor contenido tiene otro teorema, análogo al teorema sobre la representación de cualquier número complejo en la forma del producto de un número real positivo por un número de módulo igual a la unidad («forma trigonométrica» de un número complejo):

Toda aplicación lineal no degenerada $\mathcal{A}$ de un espacio euclídeo complejo puede ser representada en la forma $\mathcal{A} = \mathcal{H}\mathcal{C}$, donde $\mathcal{C}$ es una aplicación definida positiva y $\mathcal{H}$ es una aplicación unitaria.

Para el caso del espacio euclídeo real el teorema correspondiente ha sido demostrado anteriormente. En el caso del espacio complejo la demostración es análoga.

CAPÍTULO VI

FORMAS BILINEALES Y CUADRÁTICAS

Los resultados de los cinco primeros parágrafos de este capítulo se refieren a un espacio lineal *arbitrario*. La métrica (euclídea) sólo se emplea en el último, es decir, en el sexto parágrafo.

§ 1. Función lineal y forma lineal

Definición 1. *Se dice que en un espacio vectorial R está dada una* ***función lineal*** *$f(x)$ si a todo vector $x \in R$ le corresponde un número $f(x)$ con la particularidad de que se cumplen las condiciones siguientes:*

1. $f(x+y) = f(x)+f(y)$,
2. $f(\alpha x) = \alpha f(x)$,

donde x e y son unos vectores arbitrarios de R y α es un número real cualquiera.

Para encontrar *la expresión de una función lineal en coordenadas* tomemos en el espacio R una base $e_1, e_2, \ldots, e_n$. Si $x = x_1e_1+x_2e_2+ +\ldots+x_ne_n$ es un vector cualquiera de R, se tiene

$$f(x) = f(x_1e_1+x_2e_2+\ldots+x_ne_n) = x_1f(e_1)+x_2f(e_2)+\ldots+x_nf(e_n).$$

Por consiguiente, siendo fija la base, toda función lineal $f(x)$ se representa mediante una ***forma lineal***

$$f(x) = a_1x_1+a_2x_2+\ldots+a_nx_n,$$

donde $x_1, x_2, \ldots, x_n$ son las coordenadas del vector x y $a_i = f(e_i)$ son unos coeficientes que no dependen de x.

Ejemplo. Sea $a = (a_1, a_2, \ldots, a_n)$ un vector cualquiera de un espacio euclídeo R y sea $f(x) = (a, x)$ para todo x. En este caso $f(x)$ es una función lineal, pues,

$$f(x+y) = (a, x+y) = (a, x)+(a, y) = f(x)+f(y)$$

y

$$f(\alpha x) = (a, \alpha x) = \alpha(a, x) = \alpha f(x).$$

(En el caso de los espacios euclídeos este ejemplo tiene un carácter general: si en R se toma una base ortonormal $e_1, e_2, \ldots, e_n$, *para toda función lineal $f(x)$ se puede indicar un vector a tal que $f(x) =$ $= (a, x)$ cualquiera que sea x.*)

§ 2. Función bilineal. Formas bilineal y cuadrática

Definición 2. *Una función de dos variables $A(x, y)$ definida en un espacio lineal R se llama **bilineal** si es lineal respecto a y para x fijo y es lineal respecto a x para y fijo.*

Es decir, si $A(x, y)$ es una función bilineal, se tiene

$$A(x+y, z) = A(x, z)+A(y, z),$$
$$A(\alpha x, y) = \alpha A(x, y),$$
$$A(z, x+y) = A(z, x)+A(z, y),$$
$$A(x, \alpha y) = \alpha A(x, y),$$

para todos los $x, y, z \in R$ y para cualquier número real α.

El producto escalar (x, y) es un ejemplo de una función bilineal.

Hallemos *la expresión de una función bilineal en coordenadas*. Sea $e_1, e_2, \ldots, e_n$ una base del espacio R y sea

$$x = x_1e_1+x_2e_2+\ldots+x_ne_n \text{ e } y = y_1e_1+y_2e_2+\ldots+y_ne_n.$$

En estas condiciones,

$$\begin{aligned} A(x, y) &= \\ &= A(x_1e_1+x_2e_2+\ldots+x_ne_n, y_1e_1+y_2e_2+\ldots+y_ne_n) = \\ &\qquad = \sum_{i,k} x_iy_kA(e_i, e_k) = \sum_{i,k} a_{ik}x_iy_k, \end{aligned}$$

donde los coeficientes $a_{ik} = A(e_i, e_k)$ dependen sólo de la base y no dependen de x e y. Por consiguiente, en una base fija una función bilineal se representa mediante una *forma bilineal*, es decir, mediante una expresión de tipo $\sum_{i,k} a_{ik}x_iy_k$. La matriz $A = [a_{ik}]$ se llama *matriz de esta forma bilineal*. En particular, el producto escalar (x, y) se representa mediante la forma bilineal

$$\sum_{i,k=1}^{n} g_{ik}x_iy_k, \quad \text{donde} \quad g_{ik} = (e_i, e_k).$$

Veamos *cómo varía la matriz de una forma bilineal al cambiar la base*. Supongamos que en la base $e_1, e_2, \ldots, e_n$

$$A(x, y) = \sum_{i,k} a_{ik}x_iy_k, \quad \text{donde} \quad a_{ik} = A(e_i, e_k),$$

y sea $e_1', e_2', \ldots, e_n'$ una base nueva en la que

$$A(x, y) = \sum_{p,q} b_{pq}x_p'y_q', \quad \text{donde} \quad b_{pq} = A(e_p', e_q').$$

Tomemos $A = [a_{ik}]$ y $B = [b_{ik}]$ y sea $C = [c_{ik}]$ la matriz del cambio de la base antigua por la nueva; en estas condiciones, se tiene

$$\begin{aligned}
&e_p' = c_{1p}e_1 + c_{2p}e_2 + \ldots + c_{np}e_n,\\
&e_q' = c_{1q}e_1 + c_{2q}e_2 + \ldots + c_{nq}e_n\\
&\text{y } b_{pq} = A(e_p', e_q') =\\
&\qquad = A(c_{1p}e_1 + c_{2p}e_2 + \ldots + c_{np}e_n, c_{1q}e_1 + c_{2q}e_2 + \ldots + c_{nq}e_n) =\\
&\qquad\qquad = \sum_{i,k} c_{ip}c_{kq}A(e_i, e_k) = \sum_{i,k} c_{ip}c_{kq}a_{ik} = \sum_{i,k} c_{ip}a_{ik}c_{kq}.
\end{aligned}$$

Indicando c_{ip} por d_{pi}, obtenemos

$$b_{qp} = \sum_{i,k} d_{pi}a_{ik}c_{kq}.$$

La matriz $C' = [d_{pi}]$ es la traspuesta de la matriz $C = [c_{ip}]$. Además, puesto que $\sum_k a_{ik}c_{kq}$ es el elemento que figura en la i-ésima fila y en la q-ésima columna de la matriz AC, resulta que

$$\sum_{i,k} d_{pi}a_{ik}c_{kq} = \sum_i d_{pi}\left(\sum_k a_{ik}c_{kq}\right)$$

es el elemento que figura en la p-ésima fila y en la q-ésima columna de la matriz $C'(AC) = C'AC$. Por consiguiente,

$$B = C'AC.$$

Observemos que la matriz del cambio C (y también la matriz C') es no degenerada (es decir, su rango es n); luego el rango de la matriz B es igual al rango de la matriz A (véase el § 6 del capítulo III). Por consiguiente, *el rango de la matriz de una forma bilineal no depende de la selección de la base* y, por ello, puede ser denominado *rango de la forma bilineal.*

Una función bilineal $A(x, y)$ se llama *simétrica* si

$$A(x, y) = A(y, x)$$

para todos los x e y de R. En este caso se tiene $a_{ik} = a_{ki}$, es decir, la matriz $[a_{ik}]$ de la forma bilineal correspondiente es *simétrica* (en cualquier base). El producto escalar es un ejemplo de una función bilineal simétrica. Este ejemplo es de carácter general ya que, recíprocamente, toda función bilineal simétrica $A(x, y)$ satisface, obviamente, a las condiciones 1, 2 y 3 del § 1 del capítulo IV y, por consiguiente, puede ser considerada como un producto escalar.

Si en una forma bilineal simétrica $A(x, y)$ se toma $y = x$, se obtiene una *forma cuadrática* $A(x, x)$. La forma cuadrática determina unívocamente a la forma bilineal simétrica que la genera. Efectivamente, sea $A(y, x) = A(x, y)$ para todos los x e y; entonces,

$$A(x+y, x+y) = A(x, x)+2A(x, y)+A(y, y),$$

de donde

$$A(x, y) = \tfrac{1}{2}\left[A(x+y, x+y)-A(x, x)-A(y, y)\right].$$

Una función bilineal $A(x, y)$ se llama *antisimétrica si*

$$A(x, y) = -A(y, x)$$

para todos los $x, y \in R$. En una base fija, una función antisimétrica se representa por una *forma antisimétrica*

$$A(x, y) = \sum_{i,k=1}^{n} a_{ik}x_iy_k,$$

donde, como es fácil ver, $a_{ik} = -a_{ki}$ para todos los i y k y, en particular, $a_{ii} = 0$ para todos los i. Así en el espacio de tres dimensiones la expresión de una forma antisimétrica es

$$\alpha(x_1y_2-x_2y_1)+\beta(x_1y_3-x_3y_1)+\gamma(x_2y_3-x_3y_2).$$

Sea $A(x, y)$ una función bilineal cualquiera. Es obvio que $B(x, y) = A(x, y)+A(y, x)$ es entonces una función simétrica mientras que $C(x, y) = A(x, y) - A(y, x)$ es una función antisimétrica. Pero

$$A(x, y) = \tfrac{1}{2}[B(x, y) + C(x, y)];$$

por consiguiente, *toda función bilineal puede ser representada como la suma de una función simétrica y una función antisimétrica.*

§ 3. Reducción de una forma cuadrática a la suma de cuadrados

Teorema 1. *Sea $A(x, x)$ una función cuadrática cualquiera en un espacio vectorial de n dimensiones. Existe entonces una base en la que esta forma se reduce a la suma de cuadrados* (es decir, una base en la que todos los coeficientes de los productos de las coordenadas distintas del vector x son iguales a cero).

Emplearemos para la *demostración* la inducción según el número de variables que contiene la forma. Si en $A(x, x)$ figura una coordenada solamente, se tiene

$$A(x, x) = a_{11}x_1^2$$

y nuestra afirmación es evidente. Supongámosla válida para todas las formas cuadráticas que dependen de $m-1$ coordenadas y consideremos una forma cuadrática

$$A(x, x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + \ldots + a_{mm}x_m^2.$$

que depende de m coordenadas.

Si existe aquí al menos un cuadrado con coeficiente diferente de cero, por ejemplo, si $a_{mm} \neq 0$, agrupamos todos los términos que contienen x_m

$$2a_{1m}\, x_1x_m + 2a_{2m}x_2x_m + \ldots + 2a_{m-1,\,m}\, x_{m-1}x_m + a_{mm}x_m^2$$

y «despejamos el cuadrado perfecto»:

$$\begin{aligned} 2a_{1m}x_1x_m + 2a_{2m}x_2x_m + \ldots + 2a_{m-1,\,m}\, x_{m-1}x_m + a_{mm}x_m^2 = \\ = \frac{1}{a_{mm}}(a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{m-1,\,m}\, x_{m-1} + a_{mm}x_m)^2 - \\ - \frac{1}{a_{mm}}(a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{m-1,\,m}\, x_{m-1})^2. \end{aligned}$$

Tendremos entonces

$$A(x, x) = \frac{1}{a_{mm}}(a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{mm}x_m)^2 + B(x, x),$$

donde la forma cuadrática $B(x, x)$ contiene solamente $m-1$ coordenadas $x_1, x_2, \ldots, x_{m-1}$. Pongamos

$$\begin{aligned} y_1 &= x_1,\, y_2 = x_2,\, \ldots,\, y_{m-1} = x_{m-1}, \\ y_m &= a_{1m}x_1 + a_{2m}x_2 + \ldots + a_{mm}x_m, \\ y_{m+1} &= x_{m+1},\, \ldots,\, y_n = x_n. \end{aligned}$$

Puesto que el determinante

$$\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{1m} & a_{2m} & a_{3m} & \ldots & a_{mm} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & \ldots & 1 \end{bmatrix} = a_{mm} \neq 0, \qquad (1)$$

la introducción de estas nuevas coordenadas corresponde a un cambio de base (con la matriz del cambio igual a la inversa de la matriz del determinante (1)).

Según la hipótesis de inducción, la forma $B(x, x)$ que depende de $m-1$ variables $x_1, x_2, \ldots, x_{m-1}$ puede ser reducida, mediante un cambio de base, a la suma de cuadrados. Con esto quedará también reducida a la suma de cuadrados la forma $A(x, x)$.

Hemos supuesto que al menos uno de los cuadrados de la forma $A(x, x)$ tiene el coeficiente no nulo. Si esto no tiene lugar, es decir, si todos los $a_{ii} = 0$, aceptemos, por ejemplo, que $a_{12} \neq 0$ y pongamos

$$\begin{aligned} x_1 &= y_1 + y_2, \\ x_2 &= y_1 - y_2, \\ x_3 &= y_3, \\ &\ldots\ldots \\ x_n &= y_n; \end{aligned}$$

esto corresponde al cambio de base

$$\begin{aligned} e_1' &= e_1 + e_2, \\ e_2' &= e_1 - e_2, \\ e_3' &= e_3, \\ &\ldots\ldots \\ e_n' &= e_n \end{aligned}$$

con la matriz del cambio

$$\begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & -1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

(el determinante de esta matriz es igual a $-2 \neq 0$). El producto x_1x_2 se convertirá entonces en $y_1^2 - y_2^2$ y llegaremos al primer caso.

Hemos demostrado que dada en un espacio vectorial R de n dimensiones una forma cuadrática cualquiera, se puede encontrar una base de R en la que la forma se reduce a la suma de cuadrados:

$$A(x, x) = a_1 x_1'^2 + a_2 x_2'^2 + \dots + a_n x_n'^2, \tag{2}$$

donde $x_1', x_2', \dots, x_n'$ son las coordenadas del vector x en la base nueva. Los coeficientes a_i pueden ser tanto positivos como negativos; algunos de estos coeficientes pueden ser iguales a cero. Si realizamos, además, la sustitución $\sqrt{|a_i|}\, x_i' = z_i$, obtendremos para la forma cuadrática $A(x, x)$ la expresión

$$A(x, x) = \pm z_1^2 \pm z_2^2 \pm \dots \pm z_m^2,$$

(donde el coeficiente de cada una de las variables nuevas z_1, $z_2, \dots, z_m$ es igual o bien a $+1$ o bien a -1) o, alterando la numeración de los vectores básicos, la expresión

$$A(x, x) = z_1^2 + z_2^2 + \dots + z_p^2 - z_{p+1}^2 - \dots - z_{p+q}^2.$$

Ejemplo 1. *Redúzcase a la suma de cuadrados la forma cuadrática* $x_1^2 + 2x_1x_2 + 3x_2^2 + 4x_1x_3 + 6x_2x_3 + 5x_3^2$.

Solución.

$$A(x, x) = (x_1 + x_2 + 2x_3)^2 + 2x_2^2 + 2x_2x_3 + x_3^2 =$$
$$= (x_1 + x_2 + 2x_3)^2 + (x_2 + x_3)^2 + x_2^2.$$

§ 4. Ley de inercia de las formas cuadráticas

Reduciendo de diferentes modos una forma cuadrática $A(x, x)$ a la suma de cuadrados, obtendremos en la fórmula (2) diferentes coeficientes. Sin embargo, tiene lugar el siguiente hecho importante:

Teorema 2 (ley de inercia de las formas cuadráticas). *Si una forma cuadrática se reduce a la suma de cuadrados en dos bases* *diferentes, el número de los cuadrados positivos, así como el número de los cuadrados negativos, es el mismo en ambos casos.*

Demostración (por reducción al absurdo). Supongamos que la forma cuadrática $A(x, x)$ tiene en una base $e_1, e_2, \ldots, e_n$ la expresión

$$A(x, x) = x_1^2 + x_2^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2, \tag{3}$$

donde x_i son las coordenadas del vector x en esta base, y que en otra base $e'_1, e'_2, \ldots, e'_n$

$$A(x, x) = x_1'^2 + x_2'^2 + \ldots + x_k'^2 - x_{k+1}'^2 - \ldots - x_{k+m}'^2, \tag{4}$$

donde x'_i son las coordenadas del vector x en la base nueva. Sea, por ejemplo, $p > k$. Consideremos en el espacio R el subespacio R_1 generado por los vectores $e_1, e_2, \ldots, e_p$ y el subespacio R_2 generado por los vectores $e'_{k+1}, e'_{k+2}, \ldots, e'_n$. Puesto que la suma de sus dimensiones, igual a $p+(n-k)$, es mayor que n, resulta que la intersección de estos subespacios es de dimensión no nula (teorema 5 del § 6 del capítulo II), es decir, existe un vector $x \neq 0$ que pertenece a $R_1 \cap R_2$. Este vector puede ser representado tanto en la forma

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_p e_p$$

como en la forma

$$x = \beta_{k+1} e'_{k+1} + \beta_{k+2} e'_{k+2} + \ldots + \beta_n e'_n.$$

Para el vector x tenemos según la fórmula (3)

$$A(x, x) = \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_p^2 > 0$$

ya que al menos uno $\alpha_i \neq 0$; al mismo tiempo según la fórmula (4)

$$A(x, x) = -\beta_{k+1}^2 - \beta_{k+2}^2 - \ldots - \beta_{k+m}^2 \leqslant 0$$

(aquí la desigualdad no es estricta porque $k+m$ puede ser menor que n).

Hemos llegado a una contradicción y ello demuestra que debe ser $p \leqslant k$. Análogamente se demuestra que no puede tener lugar la desigualdad $p < k$. Luego, $p = k$. De la misma forma se demuestra que $q = m$.

Es fácil ver que *la suma $p+q$ es igual al rango r de la forma cuadrática $A(x, x)$.*

§ 5. Formas definidas

Definición 3. *Una forma cuadrática $A(x, x)$ se llama* ***definida positiva (negativa)*** *si $A(x, x) > 0$ ($A(x, x) < 0$) para todo $x \neq 0$ y se llama* ***semidefinida positiva (negativa)*** *si $A(x, x) \geq 0$ ($A(x, x) \leq 0$) para todo x.*

Por ejemplo, si $A(x, y) = (x, y)$ es el producto escalar en un espacio euclídeo, la forma cuadrática correspondiente $A(x, x) = (x, x)$ (el cuadrado escalar del vector x) es *definida positiva.*

Está claro que toda forma cuadrática definida positiva se reduce a la suma de cuadrados con coeficientes positivos y que toda forma semidefinida positiva se reduce a la suma de cuadrados con coeficientes no negativos (algunos de estos coeficientes pueden ser iguales a cero). Es importante el siguiente criterio de definición positiva de una forma.

Teorema 3 (criterio de Sylvester). *Para que una forma cuadrática $A(x, x)$ sea definida positiva es necesario y suficiente que sean positivos todos los menores «angulares» de la matriz $A = [a_{ik}]$, es decir, los menores $\Delta_1 = a_{11}$,*

$$\Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}, \ \ldots, \Delta_n = |A|.$$

deben ser positivos.

Para la **demostración** emplearemos la inducción según el número de variables que contiene la forma.

Para las formas cuadráticas que dependen de una variable se tiene $A(x, x) = a_{11}x_1^2$ y nuestra afirmación es obvia. Supongamos que ésta es válida para todas las formas cuadráticas que dependen de $m-1$ variables y consideremos una forma cuadrática $A(x, x) = \sum_{i, k=1}^{m} a_{ik}x_ix_k$ en m variables $x_1, x_2, \ldots, x_m$.

1. **Demostración de la necesidad.** Si representamos la forma definida positiva $A(x, x)$ como

$$A(x, x) = \sum_{i, k=1}^{m-1} a_{ik}x_ix_k + 2\sum_{i=1}^{m-1} a_{im}x_ix_m + a_{mm}x_m^2,$$

resulta que la forma cuadrática $B(x', x') = \sum_{i, k=1}^{m-1} a_{ik}x_ix_k$, que depende de $m-1$ variables $x_1, x_2, \ldots, x_{m-1}$, es definida positiva, ya que siendo $B(x', x') \leq 0$ para $x' = (x_1, x_2, \ldots, x_{m-1})$ tendríamos $A(x, x) \leq 0$ para $x = (x_1, x_2, \ldots, x_{m-1}, 0)$.

Según la hipótesis de inducción todos los menores angulares de la matriz de la forma cuadrática $B(x', x')$ son positivos, es decir,

$$\Delta_1 = a_{11} > 0,\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0,\ \ldots$$

$$\ldots, \Delta_{m-1} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,m-1} \\ a_{12} & a_{22} & \cdots & a_{2,m-1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1,m-1} & a_{2,m-1} & \cdots & a_{m-1,m-1} \end{vmatrix} > 0.$$

Resta demostrar que también $\Delta_m = |A| > 0$.

Como sabemos, la forma definida positiva $A(x, x)$ se reduce en una base $e'_1, e'_2, \ldots, e'_m$ a la suma de cuadrados:

$$A(x, x) = x_1'^2 + x_2'^2 + \ldots + x_m'^2.$$

El determinante de su matriz es igual en esta base nueva a 1 y, por consiguiente, es mayor que cero. Pero en el cambio de base la matriz de una forma bilineal se transforma de acuerdo a la fórmula (pág. 167)

$$B = C'AC,$$

donde A es su matriz en la base antigua, B es su matriz en la base nueva y C es la matriz del cambio de la base antigua por la nueva. Luego,

$$|B| = |C'|\,|A|\,|C| = |A|\,|C|^2.$$

Es decir, el signo del determinante de la matriz de una forma bilineal no depende de la base; por consiguiente, también en la base antigua tenemos

$$|A| = \Delta_m > 0.$$

2. Demostración de la suficiencia. Supongamos que todos los menores angulares de la matriz de una forma cuadrática $A(x, x)$ son positivos:

$$\Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_{m-1} > 0, \Delta_m = |A| > 0$$

y demostremos que la forma cuadrática $A(x, x)$ es definida positiva. De la hipótesis de inducción resulta, ante todo, que es definida positiva la forma cuadrática $B(x', x') = \sum_{i,k=1}^{m-1} a_{ik}x_ix_k$ de m−1 variables. Por consiguiente, $B(x', x')$ se reduce en una base nueva a la suma de cuadrados

$$B(x', x') = x_1'^2 + x_2'^2 + \ldots + x_{m-1}'^2.$$

Realizando el cambio correspondiente de las variables $x_1, x_2, \ldots x_{m-1}$ y tomando, además, $x_m = x'_m$, obtenemos

$$A(x, x) = x_1'^2 + x_2'^2 + \ldots + x_{m-1}'^2 + 2(b_{1m}x'_1x'_m + b_{2m}x'_2x'_m + \\ + \ldots + b_{m-1, m}x'_{m-1}x'_m) + a_{mm}x_m'^2,$$

donde b_{im} son unos coeficientes nuevos. Tenemos entonces

$$A(x, x) = (x'_1 + b_{1m}x'_m)^2 + (x'_2 + b_{2m}x'_m)^2 + \ldots + (x'_{m-1} + b_{m-1, m}x'_m)^2 + b\, x_m'^2,$$

donde, obviamente, $b = a_{mm} - b_{1m}^2 - b_{2m}^2 - \ldots - b_{m-1, m}^2$ y tomando

$$x'_i + b_{im}x'_m = y_i, \; x'_m = y_m$$

(lo que corresponde a un cambio de base, siendo el determinante de la matriz del cambio igual a la unidad), obtenemos

$$A(x, x) = y_1^2 + y_2^2 + \ldots + y_{m-1}^2 + by_m^2.$$

El determinante de la matriz de esta forma cuadrática es igual a b y como su signo, por lo demostrado en el punto 1, coincide con el signo de Δ_m, resulta que $b > 0$ y, por consiguiente la forma cuadrática $A(x, x)$ es definida positiva. Hemos demostrado el teorema.

Es fácil indicar ahora *las condiciones de definición negativa de una forma cuadrática* $A(x, x) = \sum_{i, k=1}^{n} a_{ik}x_ix_k$. Para que una forma cuadrática $A(x, x)$ sea definida negativa es necesario y suficiente que la forma cuadrática

$$-A(x, x) = \sum_{i, k=1}^{n} (-a_{ik})x_ix_k$$

sea definida positiva; luego, todos los menores angulares de la matriz

$$\begin{bmatrix} -a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{12} & -a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{1n} & -a_{2n} & \ldots & -a_{nn} \end{bmatrix},$$

es decir,

$$-a_{11} \begin{vmatrix} -a_{11} & -a_{12} \\ -a_{12} & -a_{22} \end{vmatrix}, \ldots, \begin{vmatrix} -a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{12} & -a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{1n} & -a_{2n} & \ldots & -a_{nn} \end{vmatrix},$$

deben ser positivos. Pero esto significa que

$$\Delta_1 = a_{11} < 0, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0,$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} < 0, \ldots$$

es decir, que *se alteran los signos de los menores angulares de la matriz A, siendo de signo menos el primero.*

Ejemplo 2. Al extremar la función

$$F(x, y, z) = 2x^2+y^2+11z^2-2xy+4xz-6yz-2y+8z, \tag{5}$$

obtenemos que sus derivadas parciales se anulan para

$x = 1, y = 2$ y $z = 0.$

La segunda diferencial toma la forma

$$d^2F = 2(2dx^2-2dx\,dy+dy^2+4\,dx\,dz-6\,dy\,dz+11\,dz^2).$$

En los paréntesis figura una forma cuadrática respecto a las diferenciales dx, dy y dz de las variables independientes. Los menores angulares de su matriz

$$\Delta_1 = 2,\ \Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = 1,\ \Delta_3 = \begin{vmatrix} 2 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 11 \end{vmatrix} = 1$$

son positivos. Por consiguiente, esta forma cuadrálica es definida positiva y la función (5) tiene *mínimo* en el punto (1, 2, 0),

§ 6. Formas bilineales y cuadráticas en un espacio euclídeo

Lema. *Sea R un espacio euclídeo y sea $C = [c_{ik}]$ la matriz del cambio de una base ortonormal $e_1, e_2, \ldots, e_n$ por otra base, también ortonormal, $e'_1, e'_2, \ldots, e'_n$. Entonces C es una matriz ortogonal.*

Demostración. Por hipótesis,

$$e'_i = c_{1i}e_1+c_{2i}e_2+\ldots+c_{ni}e_n, \quad i = 1, 2, \ldots, n.$$

Consideremos la aplicación lineal $\mathcal{C}$ de matriz C en la base $e_1, e_2, \ldots, e_n$. Tenemos

$$\mathcal{C}e_i = c_{1i}e_1+c_{2i}e_2+\ldots+c_{ni}e_n = e'_i, \quad i = 1, 2, \ldots, n.$$

Pero toda aplicación $\mathcal{C}$ que transforma al menos una base ortonormal en una base ortonormal es ortogonal (véase el § 3 del capítulo V). Por consiguiente, C es una matriz ortogonal.

Sea ahora $e_1, e_2, \ldots, e_n$ una base ortonormal de un espacio euclídeo R y sea $A(x, y)$ una función bilineal que se representa en esta base por la forma bilineal

$$A(x, y) = \sum_{i,\, k=1}^{n} a_{ik} x_i y_k,$$

donde $x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$ e $y = y_1e_1 + y_2e_2 + \ldots + y_ne_n$. Consideremos la aplicación lineal $\mathcal{A}$ de esta misma matriz A en esta misma base $e_1, e_2, \ldots, e_n$. Si C es la matriz del cambio de esta base por una base nueva $e_1', e_2', \ldots, e_n'$, la matriz A de la forma bilineal se convierte en

$$B = C'AC$$

y la matriz de la aplicación lineal $\mathcal{A}$ se convierte en

$$C^{-1}AC,$$

es decir, estas matrices se transforman, en general, de modo distinto. Sin embargo, si la base nueva $e_1', e_2', \ldots, e_n'$ también es ortonormal, la matriz del cambio C es ortogonal y se tiene $C' = C^{-1}$. En este caso la matriz de la forma bilineal $A(x, y)$ y la matriz de la aplicación lineal $\mathcal{A}$ se transforman del mismo modo. Por consiguiente, en un espacio euclídeo a toda función bilineal le corresponde una aplicación lineal completamente definida (que tiene la misma matriz en cualquier base ortonormal).

Si $A(x, y)$ es una función simétrica bilineal, la aplicación lineal correspondiente $\mathcal{A}$ será autoconjugada. Pero la matriz de toda aplicación autoconjugada tiene, en una base ortonormal determinada, la forma diagonal

$$\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

La forma bilineal $A(x, y)$ se reduce en esta misma base a la forma

$$\lambda_1 x_1' y_1' + \lambda_2 x_2' y_2' + \ldots + \lambda_n x_n' y_n'$$

(donde los coeficientes λ_i son los valores propios de la aplicación lineal $\mathcal{A}$) y la forma cuadrática correspondiente $A(x, x)$ se reduce a la suma de cuadrados

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \ldots + \lambda_n x_n'^2.$$

Ejemplo 3. *Redúzcase la forma cuadrática*

$$A(x, x) = 66x^2 - 24xy + 59y^2$$

en el espacio euclídeo R^2 a la suma de cuadrados mediante una aplicación ortogonal.

Solución. El polinomio característico de la matriz de esta forma es

$$\varphi(\lambda) = \begin{vmatrix} 66-\lambda & -12 \\ -12 & 59-\lambda \end{vmatrix} = \lambda^2 - 125\lambda + 3750.$$

Sus raíces son $\lambda_1 = 75$ y $\lambda_2 = 50$.

En la base nueva (formada por los vectores propios correspondientes a los valores propios λ_1 y λ_2) se tiene

$$A(x, x) = 75x_1'^2 + 50y'^2.$$

CAPÍTULO VII

ESTUDIO DE CURVAS Y DE SUPERFICIES DE SEGUNDO GRADO

En los tres primeros parágrafos de este capítulo se considera el espacio euclídeo de dos dimensiones (el plano) con la métrica corriente (euclídea). En el último parágrafo se considera el espacio euclídeo de tres dimensiones.

§ 1. Reducción de la ecuación general de una curva de segundo grado a la forma canónica

Tomemos en el plano un sistema cartesiano rectangular de coordenadas y consideremos la ecuación general de *segundo grado*

$$f(x, y) = a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_1x+2a_2y+a = 0. \qquad (1)$$

Como se sabe, para algunos valores determinados de los coeficientes esta ecuación representa una *elipse* $\left(\frac{x^2}{a^2}+\frac{y^2}{b^2} = 1\right)$, una *hipérbola* $\left(\frac{x^2}{a^2}-\frac{y^2}{b^2} = 1\right)$ o una *parábola* $(y^2 = 2px)$. Demostraremos que *la ecuación* (1) *es siempre la ecuación* de una de estas curvas: de una *elipse*, de una *hipérbola* o de una *parábola* (sin contar los casos degenerados, es decir, los casos de un *par de rectas* —que se da si el primer miembro de la ecuación se descompone en el producto de dos factores lineales—, de un *punto* o de un *«conjunto vacío»* que no contiene punto alguno).

Indiquemos por e_1 y e_2 los vectores unitarios de los ejes del sistema (rectangular) de coordenadas escogido. El grupo de los términos principales

$$a_{11}x^2+2a_{12}xy+a_{22}y^2 \qquad (2)$$

de la ecuación (1) puede ser considerado como una forma cuadrática de las coordenadas x e y del vector (x, y). Según hemos demostrado en el § 6 del capítulo VI, esta forma cuadrática se reduce en una base e_1', e_2' (también ortonormal) a la suma de cuadrados

$$\lambda_1 x'^2 + \lambda_2 y'^2, \tag{3}$$

donde λ_1 y λ_2 son los valores propios de la matriz

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \tag{4}$$

y e_1' y e_2' son los vectores propios que les corresponden.

Supongamos que el vector e_1' se obtiene del vector e_1 mediante un giro de ángulo φ en el sentido opuesto al del movimiento de las manecillas de un reloj. Puesto que el vector e_2 es ortogonal a e_1 y el vector e_2' es ortogonal a e_1', resulta que el vector e_2' se obtiene del vector e_2 o bien mediante un giro de ángulo φ o bien mediante un giro de ángulo φ seguido de la simetría respecto al origen de coordenadas. En el segundo caso sustituiremos e_2 por el vector $e_2'' = -e_2'$ que también será un vector propio de la matriz (4) correspondiente al mismo valor propio λ_2 ya que siendo $\mathcal{A}e_2' = \lambda_2 e_2'$ se tiene

$$\mathcal{A}e_2'' = \mathcal{A}(-e_2') = -\mathcal{A}e_2' = -\lambda_2 e_2' = \lambda_2 e_2''.$$

Por consiguiente, se puede aceptar que la base nueva e_1', e_2' se obtiene de la antigua mediante un giro de ángulo φ en el sentido opuesto al del movimiento de las manecillas de un reloj, es decir, que

$$e_1' = \cos\varphi \cdot e_1 + \operatorname{sen}\varphi \cdot e_2,$$
$$e_2' = -\operatorname{sen}\varphi \cdot e_1 + \cos\varphi \cdot e_2.$$

Pero en estas condiciones las coordenadas antiguas x e y (de un vector y, por ello, también del punto correspondiente) y las coordenadas nuevas x' e y' estarán ligadas por las relaciones

$$\begin{aligned} x &= \cos\varphi \cdot x' - \operatorname{sen}\varphi \cdot y', \\ y &= \operatorname{sen}\varphi \cdot x' + \cos\varphi \cdot y'. \end{aligned} \tag{5}$$

Introduciendo las expresiones (5) en la ecuación (1), reduciremos esta ecuación a la forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2b_1 x' + 2b_2 y' + b = 0, \tag{6}$$

donde b_1, b_2 y b son unos coeficientes nuevos. Esta operación lleva el nombre de *referencia de la curva a los ejes principales:* en lo sucesivo veremos que, si la curva (1) representa una elipse o una hipérbola, los nuevos ejes de coordenadas son paralelos a los ejes principales de la curva.

Los coeficientes λ_1 y λ_2 son los valores propios de la matriz (4) y pueden ser determinados de la ecuación

$$\varphi(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{12} & a_{22}-\lambda \end{vmatrix} = 0. \tag{7}$$

Estos valores son reales, pues la matriz (4) es simétrica (teorema II del capítulo V). El producto $\lambda_1\lambda_2$ de los valores propios es igual al término independiente $\varphi(0)$ de la ecuación de segundo grado (7) es decir, coincide con el determinante

$$\delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}.$$

Consideremos ahora por separado dos casos $\delta \neq 0$ y $\delta = 0$.

I. $\delta = \lambda_1\lambda_2 \neq 0$.

Transformemos la ecuación (6) del modo siguiente

$$\lambda_1\left(x'+\frac{b_1}{\lambda_1}\right)^2+\lambda_2\left(y'+\frac{b_2}{\lambda_2}\right)^2+c=0,$$

donde $c = b-\frac{b_1^2}{\lambda_1}-\frac{b_2^2}{\lambda_2}$. Realicemos la sustitución

$$x'' = x'+\frac{b_1}{\lambda_1}, \quad y'' = y'+\frac{b_2}{\lambda_2}.$$

Esto corresponde al traslado del origen de coordenadas al punto $\left(-\frac{b_1}{\lambda_1}, -\frac{b_2}{\lambda_2}\right)$ con la particularidad de que se conservan las direcciones de los ejes. La ecuación (6) se reduce entonces a la forma

$$\lambda_1 x''^2+\lambda_2 y''^2+c=0. \tag{8}$$

Supongamos primero que $\lambda_1\lambda_2 > 0$ (es decir, que $\delta > 0$). En este caso el lugar geométrico de los puntos cuyas coordenadas satisfacen a la ecuación (8) es una *elipse* (fig. 13, *a*) si c y λ_1 son de

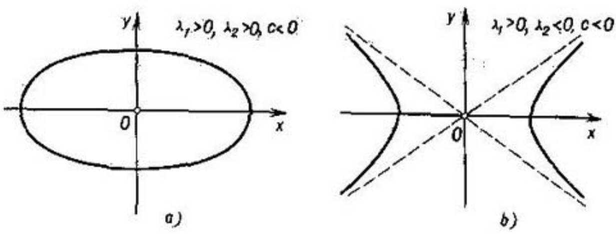


Fig. 13

signos opuestos, se reduce *a un punto* si $c = 0$ y *no contiene punto alguno* si c y λ_1 son del mismo signo[1].

Sea ahora $\lambda_1\lambda_2 < 0$ (es decir, $\delta < 0$); en este caso (8) es la ecuación de una *hipérbola* si $c \neq 0$ (fig. 13, *b*) y *de dos rectas que se cortan* si $c = 0$.

En el caso I la línea es *una curva de segundo grado con centro* (es fácil ver que el origen de coordenadas es para la curva (8) el **centro de simetría**).

II. $\delta = \lambda_1\lambda_2 = 0$ y sea, por ejemplo, $\lambda_2 \neq 0$. La ecuación (1) se reduce a

$$\lambda_2 y'^2 + 2b_1x' + 2b_2y' + b = 0. \tag{9}$$

Si $b_1 \neq 0$, tendremos formando el cuadrado perfecto

$$\lambda_2\left(y' + \frac{b_2}{\lambda_2}\right)^2 + 2b_1\left(x' + \frac{b}{2b_1} - \frac{b_2^2}{2\lambda_2 b_1}\right) = 0.$$

Trasladando el origen de coordenadas

$$x'' = x' + \frac{b}{2b_1} - \frac{b_2^2}{2\lambda_2 b_1}, \quad y'' = y' + \frac{b_2}{\lambda_2},$$

[1] En lugar del «punto» determinado por la ecuación $\lambda_1 x''^2 + \lambda_2 y''^2 = 0$, se habla también del par de «rectas imaginarias» $y'' = \pm i\sqrt{\frac{\lambda_1}{\lambda_2}}\, x''$ que se cortan en un punto real (es decir, en un punto corriente que existe realmente). El «conjunto vacío» de puntos $\lambda_1 x''^2 + \lambda_2 y''^2 + c = 0$, donde λ_1, λ_2 y c son del mismo signo, se llama también «elipse imaginaria».

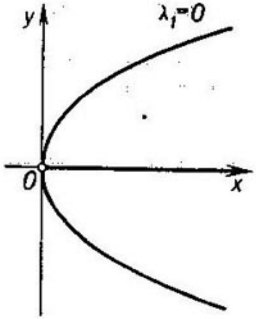


Fig. 14

reduciremos la ecuación (9) a la forma

$$\lambda_2 y''^2 + 2b_1 x'' = 0. \tag{10}$$

Esta es la ecuación canónica de una *parábola* (fig. 14).

En el caso en el que $b_1 = 0$, la ecuación (9) se reduce a la forma

$$\lambda_2\left(y' + \frac{b_2}{\lambda_2}\right)^2 + b - \frac{b_2^2}{\lambda_2} = 0$$

que mediante la sustitución

$$x'' = x', \quad y'' = y' + \frac{b_2}{\lambda_2}$$

se convierte en

$$\lambda_2 y''^2 + c = 0. \tag{11}$$

Esta ecuación representa dos rectas paralelas si $c\lambda_2 < 0$, *dos rectas confundidas* si $c = 0$ y un *«conjunto vacío de puntos»* (que no contiene punto alguno) si $c\lambda_2 > 0$[1].

§ 2. Invariantes de una curva de segundo grado

Se llama ***invariante*** *de una curva a toda expresión formada por los coeficientes de su ecuación que no varía al cambiar un sistema cartesiano rectangular de coordenadas por otro sistema del mismo*

[1] A veces se dice que la ecuación $\lambda_2 y''^2 + c = 0$, donde $\lambda_2 c > 0$, determina un «par de rectas paralelas imaginarias».

tipo, es decir, que no varía al realizar **rotaciones** y **traslaciones paralelas** de los ejes de coordenadas.

Teorema 1. *Son invariantes de una curva de segundo orden* (1) *la suma de los coeficientes de los cuadrados de las coordenadas*

$$s = a_{11}+a_{22},$$

el determinante formado por los coeficientes de los términos principales

$$\delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix},$$

y el determinante del tercer orden

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a \end{vmatrix}.$$

Demostración. Consideremos por separado los casos de traslación del origen de coordenadas y de rotación de los ejes coordenados. Supongamos primero que *el origen de coordenadas se traslada* a un punto de coordenadas (α, β) (conservándose las direcciones de los ejes). Tenemos entonces

$$x = x'+\alpha,$$
$$y = y'+\beta,$$

donde x' e y' son las coordenadas nuevas. Introduciendo estas expresiones de x y de y en la ecuación (1) obtenemos

$$a_{11}(x'+\alpha)^2+2a_{12}(x'+\alpha)(y'+\beta)+a_{22}(y'+\beta)^2+$$
$$+2a_1(x'+\alpha)+2a_2(y'+\beta)+a = 0,$$

es decir,

$$a_{11}x'^2+2a_{12}x'y'+a_{22}y'^2+2(a_{11}\alpha+a_{12}\beta+a_1)x'+2(a_{12}\alpha+a_{22}\beta+$$
$$+a_2)y'+(a_{11}\alpha^2+2a_{12}\alpha\beta+a_{22}\beta^2+2a_1\alpha+2a_2\beta+a) = 0. \qquad (12)$$

Vemos, pues, que el grupo de los términos principales no ha cambiado y, por ello, es evidente que s y δ son unos invariantes. Observemos que el coeficiente de x' es igual a $2(a_{11}\alpha+a_{12}\beta+a_1) = f'_x(\alpha, \beta)$, es decir, al valor de la derivada parcial del primer término de la ecuación (1) respecto a x calculado para $x = \alpha$ e $y = \beta$; el coeficiente de y' es igual a $f'_y(\alpha, \beta)$ y el término independiente es igual a $f(\alpha, \beta)$; por consiguiente, la ecuación transformada toma la forma definitiva

$$a_{11}x'^2+2a_{12}x'y'+a_{22}y'^2+f'_x(\alpha, \beta)x'+f'_y(\alpha, \beta)y'+f(\alpha, \beta)=0,$$

donde $f(x, y)$ es el polinomio que figura en el primer miembro de la ecuación de la curva.

Para la ecuación (12) el determinante Δ es igual a

$$\begin{vmatrix} a_{11} & a_{12} & a_{11}\alpha+a_{12}\beta+a_1 \\ a_{12} & a_{22} & a_{12}\alpha+a_{22}\beta+a_2 \\ a_{11}\alpha+a_{12}\beta+a_1 & a_{12}\alpha+a_{22}\beta+a_2 & a_{11}\alpha^2+2a_{12}\alpha\beta+a_{22}\beta^2+2a_1\alpha+2a_2\beta+a \end{vmatrix}.$$

Restando de la última fila de este determinante la primera multiplicada por α y la segunda multiplicada por β, obtenemos

$$\begin{vmatrix} a_{11} & a_{12} & a_{11}\alpha+a_{12}\beta+a_1 \\ a_{12} & a_{22} & a_{12}\alpha+a_{22}\beta+a_2 \\ a_1 & a_2 & a_1\alpha+a_2\beta+a \end{vmatrix}.$$

Realizando estas mismas operaciones con las columnas del determinante obtenido, encontraremos que es igual a

$$\begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a \end{vmatrix},$$

es decir, coincide con el determinante antiguo Δ. Por consiguiente, hemos demostrado también que Δ es un invariante respecto a las traslaciones del origen de coordenadas.

En el caso de *una rotación de los ejes* coordenados de ángulo φ la matriz del cambio de la base antigua por la nueva

$$\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$$

es ortogonal; por consiguiente, la matriz de la forma cuadrática

$$a_{11}x^2+2a_{12}xy+a_{22}y^2$$

se transforma igual que la matriz de la aplicación lineal correspondiente (§ 6 del capítulo VI). Pero en el caso de una aplicación lineal de matriz

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$$

los coeficientes de su polinomio característico

$$\varphi(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{12} & a_{22}-\lambda \end{vmatrix} =$$
$$= \lambda^2-(a_{11}+a_{22})\,\lambda+a_{11}a_{22}-a_{12}^2 = \lambda^2-s\lambda+\delta$$

no dependen de la selección de la base (teorema 6 del § 7 del capítulo III). Con esto queda demostrada la invariancia de s y de δ en las rotaciones de los ejes coordenados. Análogamente se puede demostrar que también el determinante Δ es un invariante.

En efecto, si tomamos una base nueva tal que

$$e_1' = \cos\varphi\cdot e_1+\operatorname{sen}\varphi\cdot e_2,$$
$$e_2' = -\operatorname{sen}\varphi\cdot e_1+\cos\varphi\cdot e_2,$$

las fórmulas de transformación de coordenadas serán

$$\begin{aligned} x &= \cos\varphi\cdot x'-\operatorname{sen}\varphi\cdot y', \\ y &= \operatorname{sen}\varphi\cdot x'+\cos\varphi\cdot y'. \end{aligned} \tag{13}$$

Consideremos una forma cuadrática de tres variables

$$F(x, y, z) = a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_1xz+2a_2yz+az^2$$

en una base ortonormal e_1, e_2, e_3 del espacio euclídeo de tres dimensiones R^3. Para $z = 1$ esta forma coincide con $f(x, y)$. Si realizamos un cambio ortogonal de matriz

$$\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

de la base e_1, e_2 y e_3, tendremos las siguientes fórmulas de transformación de las coordenadas de R^3

$$\begin{aligned} x &= \cos\varphi\cdot x'-\operatorname{sen}\varphi\cdot y', \\ y &= \operatorname{sen}\varphi\cdot x'+\cos\varphi\cdot y', \\ z &= z'. \end{aligned} \tag{14}$$

Si $f(x, y)$ se transforma mediante la sustitución (13) en

$$b_{11}x'^2+2b_{12}x'y'+b_{22}y'^2+2b_1x'+2b_2y'+a,$$

está claro que $F(x, y, z)$ se transforma mediante la sustitución (14) en

$$b_{11}x'^2+2b_{12}x'y'+b_{22}y'^2+2b_1x'z'+2b_2y'z'+az'^2.$$

El determinante de la matriz de una forma cuadrática no varía al realizar un cambio ortogonal de la base; por consiguiente, para la forma $F(x, y, z)$ tiene lugar la igualdad

$$\begin{vmatrix} b_{11} & b_{12} & b_1 \\ b_{12} & b_{22} & b_2 \\ b_1 & b_2 & a \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a \end{vmatrix}.$$

Pero su primer miembro es el determinante Δ para $f(x, y)$ en la base nueva e_1', e_2' y su segundo miembro es el mismo determinante en la base antigua. Por consiguiente, este determinante tampoco varía en las rotaciones de los ejes coordenados. Hemos demostrado completamente el teorema.

El valor de δ permite juzgar sobre el tipo de la curva: si $\delta > 0$, tenemos una curva de *tipo elíptico* (una elipse, un punto o «un conjunto vacío», es decir, una «elipse imaginaria»); si $\delta < 0$, tenemos una curva de *tipo hiperbólico* (una hipérbola o dos rectas reales que se cortan); si $\delta = 0$, tenemos una curva de *tipo parabólico* (una parábola o dos rectas paralelas que, posiblemente, se confunden o incluso no existen, es decir, son «imaginarias»).

La invariancia de las expresiones s, δ y Δ, establecida en el teorema 1, facilita la reducción de la ecuación de la curva a la forma canónica. Así, por ejemplo, en el caso de una curva central, es decir, para $\delta \neq 0$, una vez determinados los valores λ_1 y λ_2, la ecuación de la curva se reduce, como hemos visto, a la forma

$$\lambda_1 x''^2 + \lambda_2 y''^2 + c = 0.$$

Para esta ecuación tenemos

$$\Delta = \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & c \end{vmatrix} = \lambda_1 \lambda_2 c,$$

de donde resulta $\lambda_1 \lambda_2 c = \Delta$ o $\delta c = \Delta$, es decir, $c = \dfrac{\Delta}{\delta}$. Por consiguiente, la ecuación «canónica», es decir, simplificada, de una curva central de segundo grado es

$$\lambda_1 x^2 + \lambda_2 y^2 + \frac{\Delta}{\delta} = 0.$$

Si

$$\delta > 0 \quad \text{y} \quad \Delta \neq 0,$$

nuestra curva es una *elipse* o una «*elipse imaginaria*». Será una *elipse* (real) si λ_1 y $\dfrac{\Delta}{\delta}$ son de signos opuestos, es decir, si $\lambda_1 \dfrac{\Delta}{\delta} < 0$; pero como $\delta > 0$ y λ_1 es del mismo signo que s, esto ocurrirá si

$s\Delta < 0$. La curva será una «*elipse imaginaria*» en el caso en el que $s\Delta > 0$. Finalmente, si $\delta > 0$ y $\Delta = 0$, la curva es un *punto*.

Si $\delta < 0$, la curva representa una *hipérbola* para $\Delta \neq 0$ y se descompone en *dos rectas que se cortan* para $\Delta = 0$.

Para una *parábola* cuya ecuación ha sido reducida a la forma (10), se tiene

$$\Delta = \begin{vmatrix} 0 & 0 & b_1 \\ 0 & \lambda_2 & 0 \\ b_1 & 0 & 0 \end{vmatrix} = -b_1^2\lambda_2,$$

de donde

$$b_1 = \pm\sqrt{-\frac{\Delta}{\lambda_2}}.$$

Aquí $b_1 \neq 0$ y, por consiguiente, $\Delta \neq 0$.

En el caso de *dos rectas paralelas* (distintas, confundidas o «imaginarias») la ecuación de la curva se reduce a la forma

$$\lambda_2 y^2 + c = 0.$$

En este caso se tiene

$$\Delta = \begin{vmatrix} 0 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & c \end{vmatrix} = 0.$$

Resumamos los resultados de los dos últimos parágrafos en el cuadro siguiente:

$\delta > 0$ Curva de tipo elíptico	$\Delta \neq 0$	$s\Delta < 0$. Elipse
		$s\Delta > 0$. «Elipse imaginaria»
	$\Delta = 0$	Punto (o dos «rectas imaginarias» que se cortan en este punto)
$\delta < 0$ Curva de tipo hiperbólico	$\Delta \neq 0$	Hipérbola
	$\Delta = 0$	Dos rectas que se cortan
$\delta = 0$ Curva de tipo parabólico	$\Delta \neq 0$	Parábola
	$\Delta = 0$	Dos rectas paralelas (distintas, confundidas o «imaginarias»)

De este cuadro se ve, en particular, que *el determinante Δ es igual a cero si, y sólo si, la curva se descompone en dos rectas* (reales o «imaginarias»). Por consiguiente, el determinante Δ permite decir si la curva se descompone o no se descompone en dos rectas, sin que sea necesario reducir la ecuación de la curva a la forma canónica.

Ejemplos. *Determínese el tipo de las curvas siguientes y redúzcase la ecuación de éstas a la forma canónica.*

1. $3x^2-2xy+3y^2+2x-4y+1=0;$
2. $3x^2-2xy+3y^2+2x-4y+2=0;$
3. $x^2+y^2+2x+1=0;$
4. $x^2+2xy-y^2-6x+4y-3=0;$
5. $x^2+3xy+2y^2+2x+5y-3=0;$
6. $x^2-2xy+y^2+4x-6y+1=0;$
7. $x^2+4xy+4y^2-2x-4y-3=0;$
8. $x^2+4xy+4y^2-2x-4y+1=0;$
9. $x^2+4xy+4y^2+2x+4y+2=0.$

Solución.

1. $\delta=\begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix}=8>0.$

Tenemos una curva de tipo elíptico. Puesto que

$$\Delta=\begin{vmatrix} 3 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 1 \end{vmatrix}=-3\neq 0,$$

la curva no se descompone.

Tenemos, además,

$$\varphi(\lambda)=\begin{vmatrix} 3-\lambda & 1 \\ -1 & 3-\lambda \end{vmatrix}=\lambda^2-6\lambda+8=0,$$

$$\lambda=3\pm 1,\quad \lambda_1=4,\quad \lambda_2=2.$$

La ecuación canónica de la curva es

$$4x'^2+2y'^2-\frac{3}{8}=0 \quad ó \quad \frac{32}{3}x'^2+\frac{16}{3}y'^2=1.$$

La curva es una *elipse;* sus semiejes son

$$a=\frac{\sqrt{3}}{4\sqrt{2}}\approx 0{,}33 \quad y \quad b=\frac{\sqrt{3}}{4}\approx 0{,}4.$$

2. $\delta=\begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix}=8>0.$

La curva es de tipo elíptico.

$$\Delta = \begin{vmatrix} 3 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 2 \end{vmatrix} = 5 \neq 0.$$

La curva no se descompone.

Además, $s = 6$, $\delta = 8$ y $\varphi(\lambda) = \lambda^2 - 6\lambda + 8$; $\lambda_1 = 4$ y $\lambda_2 = 2$. La ecuación canónica de la curva es

$$4x'^2 + 2y'^2 + \frac{5}{8} = 0.$$

Tenemos una *«elipse imaginaria»* (un «conjunto vacío» de puntos).

3. $\delta = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 > 0.$

La curva es de tipo elíptico.

$$\Delta = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = 0.$$

Esta curva, cuya ecuación puede ser representada en la forma

$$(x+1)^2 + y^2 = 0,$$

representa un *punto*

$$x = -1 \quad \text{e} \quad y = 0$$

y puede ser interpretada como dos rectas «imaginarias» que se cortan en este punto:

$$x + iy + 1 = 0 \quad \text{y} \quad x - iy + 1 = 0.$$

4. $\delta = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 < 0.$

La curva es de tipo hiperbólico.

$$\Delta = \begin{vmatrix} 1 & 1 & -3 \\ 1 & -1 & 2 \\ -3 & 2 & -3 \end{vmatrix} = -1 \neq 0.$$

La curva no se descompone.

Además, $s = 0$, $\delta = -2$ y $\varphi(\lambda) = \lambda^2 - 2$, $\lambda_1 = \sqrt{2}$ y $\lambda_2 = -\sqrt{2}$. La ecuación canónica de la curva es

$$\sqrt{2}\,x'^2 - \sqrt{2}\,y'^2 + \frac{1}{2} = 0$$

o

$$2\sqrt{2}\,y'^2 - 2\sqrt{2}\,x'^2 = 1.$$

La curva es una *hipérbola;* sus semiejes son: $a = b \approx 0{,}6$.

$$5.\ \delta = \begin{vmatrix} 1 & \frac{3}{2} \\ \frac{3}{2} & 2 \end{vmatrix} = -\frac{1}{4} < 0.$$

La curva es de tipo hiperbólico.

$$\Delta = \begin{vmatrix} 1 & \frac{3}{2} & 1 \\ \frac{3}{2} & 2 & \frac{5}{2} \\ 1 & \frac{5}{2} & -3 \end{vmatrix} = 0.$$

La curva se descompone en *dos rectas que se cortan.* Por consiguiente, el primer miembro de la ecuación de la curva se descompone en dos factores lineales. Para determinar estos factores se puede proceder del modo siguiente. Resolvamos la ecuación

$$x^2+3xy+2y^2+2x+5y-3 = x^2+(3y+2)\,x+2y^2+5y-3 = 0$$

respecto a x (el primer miembro de la ecuación se descompone, como sabemos, en factores lineales; luego, x se expresa racionalmente mediante y):

$$x = -\left(\frac{3}{2}y+1\right) \pm \sqrt{\frac{9}{4}y^2+3y+1-2y^2-5y+3} = -\left(\frac{3}{2}y+1\right) \pm \left(\frac{1}{2}y-2\right);$$

$$x_1 = -y-3, \quad x_2 = -2y+1.$$

Por consiguiente, el primer miembro de la ecuación se descompone en los factores siguientes: $(x+y+3)\,(x+2y-1) = 0$ y la curva se descompone en dos rectas

$$x+y+3 = 0 \quad \text{y} \quad x+2y-1 = 0.$$

$$6.\ \delta = \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} = 0.$$

La curva es de tipo parabólico.

$$\Delta = \begin{vmatrix} 1 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 1 \end{vmatrix} = -1 \neq 0.$$

La curva no se descompone. Además, $s = 2$, $\delta = 0$ y $\varphi(\lambda) = \lambda^2-2\lambda$, $\lambda_1 = 0$, $\lambda_2 = 2$, $b_1 = \pm\sqrt{\frac{1}{2}}$. La ecuación canónica de la curva es

$$2y'^2 \pm \sqrt{2}\,x' = 0 \quad \text{ó} \quad y'^2 = \pm\frac{1}{\sqrt{2}}\,x'.$$

Tenemos una *parábola.*

$$7.\ \delta = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0.$$

La curva es de tipo parabólico.

$$\Delta = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & -3 \end{vmatrix} = 0.$$

La curva se descompone en *dos rectas paralelas*

$$x+2y+1 = 0 \quad \text{y} \quad x+2y-3 = 0.$$

$$8.\ \delta = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0.$$

La curva es de tipo parabólico.

$$\Delta = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{vmatrix} = 0$$

y la curva se descompone en *dos rectas confundidas*

$$f(x, y) = (x+2y-1)^2 = 0,\ x+2y-1 = 0.$$

$$9.\ \delta = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0.$$

La curva es de tipo parabólico.

$$\Delta = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 2 \end{vmatrix} = 0$$

y la curva representa un «conjunto vacío» de puntos, es decir, se descompone en *dos «rectas imaginarias» paralelas:*

$$f(x, y) = (x+2y+1)^2+1 = (x+2y+1+i)\,(x+2y+1-i) = 0.$$

§ 3. Determinación del centro y de los ejes principales de una curva con centro. Determinación del vértice y del eje de una parábola

En este parágrafo aceptamos que $\Delta \neq 0$, es decir, que la curva *no se descompone* en dos rectas.

Supongamos dada una ecuación general de segundo grado (1). Determinemos los valores propios λ_1 y λ_2 de la matriz (4) y los vectores propios e_1' y e_2' correspondientes. Sabemos que en la base compuesta de estos vectores la forma cuadrática $a_{11}x^2+2a_{12}xy+$ $+a_{22}y^2$ se reduce a la suma de cuadrados $\lambda_1 x'^2+\lambda_2 y'^2$ y que la ecuación (1) se reduce a la forma (6). Como sabemos, los vectores propios e_1' y e_2' de la matriz (4) se determinan de los sistemas de ecuaciones

$$\begin{cases} a_{11}x_1+a_{12}y_1=\lambda_1 x_1, \\ a_{12}x_1+a_{22}y_1=\lambda_1 y_1 \end{cases} \text{ y } \begin{cases} a_{11}x_2+a_{12}y_2=\lambda_2 x_2, \\ a_{12}x_2+a_{22}y_2=\lambda_2 y_2; \end{cases}$$

puesto que el determinante

$$\begin{vmatrix} a_{11}-\lambda_i & a_{12} \\ a_{12} & a_{22}-\lambda_i \end{vmatrix}, \ i=1, 2,$$

es igual a cero, ambos sistemas se reducen a una ecuación:

$(a_{11}-\lambda_1)x_1+a_{12}y_1=0$ en el caso del primer sistema,

$(a_{11}-\lambda_2)x_2+a_{12}y_2=0$ en el caso del segundo sistema.

Por consiguiente, para $e_1'=(x_1, y_1)$ tenemos

$$\frac{y_1}{x_1}=\frac{\lambda_1-a_{11}}{a_{12}}$$

y para $e_2'=(x_2, y_2)$ tenemos

$$\frac{y_2}{x_2}=\frac{\lambda_2-a_{11}}{a_{12}}.$$

Luego, los coeficientes angulares de los nuevos ejes de coordenadas en el sistema antiguo son iguales a

$$k_1=\frac{\lambda_1-a_{11}}{a_{12}} \text{ (para el nuevo eje } x \text{ que corresponde a } \lambda_1)$$

y

$$k_2=\frac{\lambda_2-a_{11}}{a_{12}} \text{ (para el nuevo eje } y \text{ que corresponde a } \lambda_2).$$

Para reducir la ecuación de la curva a la forma canónica basta, como hemos visto, realizar a continuación una traslación del origen de coordenadas; por consiguiente, los valores k_1 y k_2 determinan *las direcciones de los ejes principales* de la curva (1).

Supongamos que la curva considerada de segundo grado tiene centro, es decir, aceptemos que $\delta \neq 0$. Para determinar el centro de la curva, es decir, el origen del nuevo sistema de coordenadas, nos basaremos en el siguiente razonamiento elemental. Si trasladamos el origen de coordenadas al punto (α, β) sin alterar las direcciones de los ejes, esto es, si tomamos

$$x = x'+\alpha,$$
$$y = y'+\beta,$$

la ecuación (1), como hemos visto, se transforma en

$$a_{11}x'^2+2a_{12}x'y'+a_{22}y'^2+f'_x(\alpha, \beta)x'+$$
$$+f'_y(\alpha, \beta)y'+f(\alpha, \beta) = 0.$$

Consideremos el sistema de ecuaciones

$$\left\{\begin{aligned} \frac{1}{2} f'_x(x, y) &= a_{11}x+a_{12}y+a_1 = 0, \\ \frac{1}{2} f'_y(x, y) &= a_{12}x+a_{22}y+a_2 = 0. \end{aligned}\right. \tag{15}$$

Puesto que su determinante δ es, por hipótesis, diferente de cero, el sistema tiene una solución (única) α, β. Si trasladamos el origen de coordenadas al punto (α, β), en la ecuación desaparecerán los términos de primer grado en x' e y' y, por consiguiente, el origen nuevo de coordenadas será el *centro* de la curva. Luego, *el centro de una curva de segundo grado con centro (de una elipse o de una hipérbola) se determina del sistema de ecuaciones* (15).

Consideremos ahora una curva de segundo grado *sin centro* (es decir, sea $\delta = 0$). Esta es una *parábola*, pues hemos aceptado que $\Delta \neq 0$. Sean $\lambda_1 = 0$ y λ_2 los valores propios de la matriz (4); las direcciones de los ejes nuevos se determinan igual que antes

$$k_1 = \frac{\lambda_1 - a_{11}}{a_{12}} = -\frac{a_{11}}{a_{12}} \text{ (para el eje } Ox' \text{ que corresponde a } \lambda_1)$$

y

$$k_2 = \frac{\lambda_2 - a_{11}}{a_{12}} \text{ (para el eje } Oy' \text{ que corresponde a } \lambda_2).$$

El nuevo origen de coordenadas, es decir, el *vértice* (α, β) de la parábola, puede ser determinado con el razonamiento siguiente. Para la parábola dada mediante la ecuación canónica $y^2 = 2px$, el eje Oy es la tangente en el vértice. El nuevo eje Oy tiene en los ejes coordenados antiguos el coeficiente angular igual a $k_2 = \dfrac{\lambda_2 - a_{11}}{a_{12}}$.

Puesto que coincide con la tangente a la parábola en el vértice (α, β), el coeficiente k_2 debe ser igual a la derivada y'_x en este punto. Para determinar y'_x derivemos respecto a x la ecuación (1) considerando que y es una función de x; obtenemos

$$f'_x(x, y)+f'_y(x, y)y'_x = 0,$$

de donde

$$y'_x = -\frac{f'_x(x, y)}{f'_y(x, y)}.$$

Es decir, en el vértice (α, β) se tiene

$$k_2 = -\frac{f'_x(\alpha, \beta)}{f'_y(\alpha, \beta)};$$

de donde

$$f'_x(\alpha, \beta)+k_2 f'_y(\alpha, \beta) = 0.$$

Por consiguiente, *las coordenadas del vértice* (α, β) *de una parábola se pueden determinar resolviendo el sistema de ecuaciones formado por la ecuación*

$$f'_x(x, y)+k_2 f'_y(x, y) = 0 \tag{16}$$

y la ecuación (1).

Veamos el *sentido geométrico* de la ecuación (16) que puede ser representada detalladamente en la forma

$$(a_{11}x+a_{12}y+a_1)+k_2(a_{12}x+a_{22}y+a_2) = 0.$$

Esta es la ecuación de una recta que pertenece al haz determinado por las rectas

$$a_{11}x+a_{12}y+a_1 = 0 \quad \text{y} \quad a_{12}x+a_{22}y+a_2 = 0.$$

Los coeficientes angulares $-\frac{a_{11}}{a_{12}}$ y $-\frac{a_{12}}{a_{22}}$ de estas rectas coinciden por ser $\delta = 0$ y son iguales a k_1; luego, estas rectas son paralelas al nuevo eje Ox. Por consiguiente, también la recta (16), que pertenece al haz que ellas determinan, es paralela al nuevo eje Ox. Como, además, esta recta pasa por el vértice, es *el eje de simetría de la parábola*, es decir, *su diámetro principal*.

Ejemplos. *Determínense los nuevos sistemas de coordenadas y constrúyanse las gráficas de las curvas* 1, 4 *y* 6 *del parágrafo anterior.*

Solución.

1. $3x^2-2xy+3y^2+2x-4y+1 = 0$, $\lambda_1 = 4$, $\lambda_2 = 2$.

Las direcciones de los ejes nuevos son

$$k_1 = \frac{4-3}{-1} = -1 \qquad (\text{eje } Ox'),$$

$$k_2 = \frac{2-3}{-1} = 1 \qquad (\text{eje } Oy').$$

El nuevo origen de coordenadas (el centro de la curva) se determina del sistema de ecuaciones

$$\begin{cases} 3x-\ y+1 = 0, \\ -x+3y-2 = 0, \end{cases}$$

de donde $x = -\frac{1}{8}$, $y = \frac{5}{8}$. La ecuación del eje Ox' (en el sistema antiguo de coordenadas) es

$$y-\frac{5}{8} = -\left(x+\frac{1}{8}\right) \quad \text{ó} \quad x+y = \frac{1}{2};$$

y la ecuación del eje Oy' es

$$y-\frac{5}{8} = x+\frac{1}{8} \quad \text{ó} \quad y-x = \frac{3}{4}.$$

En la fig. 15 se representa la gráfica de la curva.

4. $x^2+2xy-y^2-6x+4y-3 = 0$, $\lambda_1 = \sqrt{2}$, $\lambda_2 = -\sqrt{2}$.

Las direcciones de los ejes nuevos son

$$k_1 = \frac{\sqrt{2}-1}{1} = \sqrt{2}-1 \qquad (\text{eje } Ox'),$$

$$k_2 = -(\sqrt{2}+1) \qquad (\text{eje } Oy').$$

El nuevo origen de coordenadas se determina del sistema de ecuaciones

$$\begin{cases} x+y-3 = 0, \\ x-y+2 = 0, \end{cases}$$

de donde $x = \frac{1}{2}$ e $y = \frac{5}{2}$. La ecuación del eje Ox' (en el sistema antiguo de coordenadas) es

$$y-\frac{5}{2} = (\sqrt{2}-1)\left(x-\frac{1}{2}\right)$$

y la ecuación del eje Oy' es

$$y-\frac{5}{2} = -(\sqrt{2}+1)\left(x-\frac{1}{2}\right).$$

En la fig. 16 se representa la gráfica de la curva.

6. $x^2-2xy+y^2+4x-6y+1 = 0$, $\lambda_1 = 0$, $\lambda_2 = 2$.

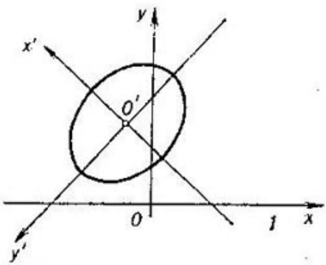


Fig. 15

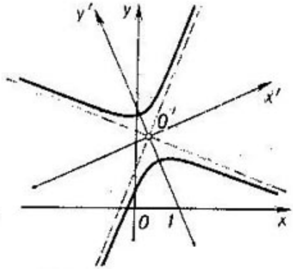


Fig. 16

Las direcciones de los nuevos ejes coordenados son

$$k_1 = \frac{-1}{-1} = 1 \qquad (\text{eje } Ox'),$$

$$k_2 = \frac{2-1}{-1} = -1 \qquad (\text{eje } Oy').$$

La ecuación del eje de la parábola es

$$(x-y+2)-(-x+y-3) = 0,$$

es decir,

$$x-y+\frac{5}{2} = 0.$$

Resolviendo esta ecuación y la ecuación de la curva, obtenemos las coordenadas del vértice de la parábola: $x = -\frac{31}{8}$ e $y = -\frac{11}{8}$. La ecuación del nuevo eje Oy es

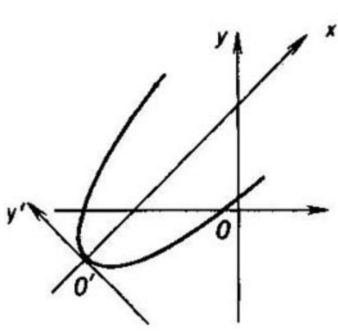


Fig. 17

$$y+\frac{11}{8} = -1\left(x+\frac{31}{8}\right) \quad \text{ó} \quad x+y+\frac{21}{4} = 0.$$

Hemos obtenido anteriormente la ecuación canónica de esta parábola en la forma $y'^2 = \pm\frac{1}{\sqrt{2}}x$. El signo en el segundo miembro se determina por la dirección de Ox'. De la ecuación de la curva obtenemos, expresando y mediante x,

$$y = x+3\pm\sqrt{2x+8}.$$

Por consiguiente, la curva se encuentra en la región $x \geqslant -4$, es decir, a la derecha del eje Oy'. Orientando el eje Ox' a la derecha, deberemos tomar en la ecuación canónica el signo más (y orientando el eje Ox' a la izquierda, el signo menos).

En la fig. 17 se representa la gráfica de la curva.

§ 4. Estudio de la ecuación general de una superficie de segundo grado

En este parágrafo sólo nos ocuparemos de la reducción de la ecuación general de una superficie de segundo grado a la forma canónica.

Supongamos que en un sistema de coordenadas cartesianas rectangulares del espacio se tiene una ecuación

$$f(x, y, z) = a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_{13}xz+2a_{23}yz+ \\ +a_{33}z^2+2a_1x+2a_2y+2a_3z+a = 0. \quad (17)$$

Consideremos la forma cuadrática de tres variables

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_{13}xz+2a_{23}yz+a_{33}z^2.$$

En una base determinada, también ortonormal, esta forma se reduce a la suma de cuadrados

$$\lambda_1x'^2+\lambda_2y'^2+\lambda_3z'^2.$$

La ecuación (17) se reduce, entonces, a la forma

$$\lambda_1x'^2+\lambda_2y'^2+\lambda_3z'^2+2b_1x'+2b_2y'+2b_3z'+b = 0.$$

Se pueden dar tres casos:

I. Todos los λ_i son diferentes de cero.
II. Uno de los valores de λ_i es igual a cero.
III. Dos de los valores de λ_i son iguales a cero.

Consideremos por separado cada uno de estos casos.

I. $\lambda_1\lambda_2\lambda_3 \neq 0$.

Igual que en el caso de una curva de segundo grado podemos deshacernos entonces de los términos de primer grado

$$\lambda_1\left(x'+\frac{b_1}{\lambda_1}\right)^2+\lambda_2\left(y'+\frac{b_2}{\lambda_2}\right)^2+\lambda_3\left(z'+\frac{b_3}{\lambda_3}\right)^2+c = 0.$$

Realizando la sustitución

$$x'' = x' + \frac{b_1}{\lambda_1},$$

$$y'' = y' + \frac{b_2}{\lambda_2},$$

$$z'' = z' + \frac{b_3}{\lambda_3},$$

es decir, realizando una traslación paralela de los ejes de coordenadas, obtenemos la ecuación

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 z''^2 + c = 0.$$

Esta es la ecuación de una *superficie* de segundo grado *con centro* (su *centro* es el nuevo origen de coordenadas).

Aceptemos que $c < 0$ (de lo contrario multiplicaremos la ecuación por -1). Siendo $c < 0$, se pueden dar cuatro casos:

1. $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$ *(un elipsoide)*.
2. $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 < 0$ *(un hiperboloide de una hoja)*.
3. $\lambda_1 > 0$, $\lambda_2 < 0$, $\lambda_3 < 0$ *(un hiperboloide de dos hojas)*.
4. $\lambda_1 < 0$, $\lambda_2 < 0$, $\lambda_3 < 0$ *(un «conjunto vacío» de puntos* que también se denomina «elipsoide imaginario»).

Si $c = 0$ y todos los λ_i son del mismo signo se tiene un *punto* (un «cono imaginario»); si $c = 0$ y los λ_i tienen signos diferentes, resulta un *cono*.

II. Uno de los coeficientes λ_i es igual a cero; supongamos, por ejemplo, que $\lambda_3 = 0$. Realizando la traslación correspondiente del origen de coordenadas, podemos reducir la ecuación de la superficie a la forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2b_3 z' + b = 0. \tag{18}$$

Aquí se pueden dar los casos $b_3 = 0$ y $b_3 \neq 0$.

Si $b_3 = 0$, la ecuación se reduce a la forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + b = 0.$$

Esta es la ecuación de una *superficie cilíndrica* cuya forma se determina por su directriz $\lambda_1 x'^2 + \lambda_2 y'^2 + b = 0$ en el plano $x'O'y'$ *(un cilindro elíptico, un cilindro hiperbólico, dos planos que se cortan, una recta* o *dos «planos imaginarios»* que se cortan a lo largo de esta recta real, *un «conjunto vacío» de puntos* o *un «cilindro elíptico imaginario»)*.

Si $b_3 \neq 0$, la ecuación (18) se reduce a la forma

$$\lambda_1 x''^2 + \lambda_2 y''^2 + 2b_3 z'' = 0.$$

Si es $\lambda_1\lambda_2 > 0$, se tiene *un paraboloide elíptico* y si $\lambda_1\lambda_2 < 0$, se tiene *un paraboloide hiperbólico*.

III. Dos de los números λ_i son iguales a cero; sea, por ejemplo, $\lambda_2 = 0$ y $\lambda_3 = 0$. La ecuación (17) se reduce a la forma

$$\lambda_1 x'^2 + 2b_2 y' + 2b_3 z' + b = 0. \tag{19}$$

Si $b_2 = 0$ y $b_3 = 0$, se tiene *dos planos paralelos*, distintos para $\lambda_1 b < 0$, confundidos para $b = 0$ e «imaginarios» para $\lambda_1 b > 0$.

Finalmente, si al menos uno de los coeficientes b_2, b_3 es diferente de cero, tomemos en la ecuación (19)

$$x' = x'',$$

$$y' = \frac{b_2 y'' + b_3 z''}{\sqrt{b_2^2 + b_3^2}},$$

$$z' = \frac{b_3 y'' - b_2 z''}{\sqrt{b_2^2 + b_3^2}};$$

es fácil ver que esto equivale a pasar a una base nueva (también ortonormal) siendo la matriz del cambio igual a

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \dfrac{b_2}{\sqrt{b_2^2+b_3^2}} & \dfrac{b_3}{\sqrt{b_2^2+b_3^2}} \\ 0 & \dfrac{b_3}{\sqrt{b_2^2+b_3^2}} & \dfrac{-b_2}{\sqrt{b_2^2+b_3^2}} \end{bmatrix}$$

La ecuación (19) se transforma en

$$\lambda_1 x''^2 + 2\sqrt{b_2^2+b_3^2}\, y'' + b = 0$$

y, puesto que $\sqrt{b_2^2+b_3^2} \neq 0$, esta última ecuación se transforma mediante una traslación del origen de coordenadas en

$$\lambda_1 x'''^2 + 2\sqrt{b_2^2+b_3^2}\, y''' = 0.$$

Tenemos *un cilindro parabólico*.

Observemos, sin demostrarlo, que (igual que en el caso de una curva de segundo grado) para la transformación de la ecuación de una superficie de segundo grado pueden ser empleados los *invariantes*. Estos serán ahora

$$s_1 = a_{11} + a_{22} + a_{33},$$

$$s_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{13} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{23} & a_{33} \end{vmatrix},$$

$$\delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$$

(estos números coinciden, salvo el signo, con los coeficientes del polinomio característico de la matriz

$$\left.\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}\right)$$

y el determinante

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_1 \\ a_{12} & a_{22} & a_{23} & a_2 \\ a_{13} & a_{23} & a_{33} & a_3 \\ a_1 & a_2 & a_3 & a \end{vmatrix}.$$

La ecuación de una superficie con centro se reduce a la forma

$$\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2 + \frac{\Delta}{\delta} = 0.$$

El determinante Δ se anula si, y sólo si, la superficie es cónica o cilíndrica (en particular, si se descompone en dos planos, distintos, confundidos o «imaginarios»).

CAPÍTULO VIII

CONCEPTOS PRINCIPALES DE LA TEORÍA ESPECIAL DE LA RELATIVIDAD

§ 1. Espacios bidimensionales provistos de producto escalar

Sea R un espacio lineal en el que está definido el *producto escalar*, es decir, a todo par de vectores x e y de R corresponde un número (real) (x, y) tal que

1) $(x, y) = (y, x)$,
2) $(\alpha x, y) = \alpha(x, y)$,
3) $(x+y, z) = (x, z)+(y, z)$

para todos los x, y, z de R y para todos los números reales α. Obsérvese que no se exige el cumplimiento de la condición 4 (pág. 132). Estos espacios se llaman *espacios provistos de una métrica cuadrática* ya que, dada una base, el cuadrado escalar (x, x) del vector x es una forma cuadrática de sus coordenadas.

Se llama *longitud* o *módulo* de un vector a la raíz cuadrada de su cuadrado escalar. Notemos que, hablando en términos generales, un vector no nulo puede ser de longitud nula o incluso imaginaria. (Si $(x, x) = -a^2 < 0$, se toma por definición que $|x| = ai$, donde $a > 0$ e $i = \sqrt{-1}$.)

Si en el espacio R se escoge una base, el producto escalar se representa por una forma bilineal simétrica

$$(x, y) = \sum_{i, k} g_{ik}x_i y_k$$

en las coordenadas x_i e y_k de los vectores x e y. La forma cuadrática correspondiente se reduce en una base (en general, nueva) a la suma de cuadrados

$$(x, x) = x_1^2+x_2^2+\ldots+x_p^2-x_{p+1}^2-\ldots-x_{p+q}^2$$

(§ 3 del capítulo VI). Además, el número p de los cuadrados positivos y el número q de los cuadrados negativos son unos invariantes del espacio R (ley de inercia de las formas cuadráticas, § 4 del capítulo VI) y determinan el tipo del espacio.

Así, para el espacio bidimensional R (el plano) p y q pueden tomar los valores siguientes:

1. $p = 2$ y $q = 0$;
1'. $p = 0$ y $q = 2$;
2. $p = 1$ y $q = 0$;
2'. $p = 0$ y $q = 1$;
3. $p = 1$ y $q = 1$.

En el caso 1 el cuadrado escalar de un vector cualquiera $x = x_1e_1 + x_2e_2$ en una base determinada (ortonormal) es igual a $x_1^2 + x_2^2$ y este espacio es *euclídeo*.

En el caso 1' se tiene $(x, x) = -x_1^2 - x_2^2$ y el espacio no difiere esencialmente del euclídeo.

En el caso 2 (o 2' que viene a ser casi lo mismo) la forma cuadrática (x, x) contiene sólo un cuadrado y en una base determinada se tiene $(x, x) = x_2^2$ (o $-x_2^2$, respectivamente). Este plano se llama *semieuclídeo*.

Finalmente, en el caso 3 la forma cuadrática (x, x) se reduce en una base determinada a la diferencia de cuadrados $x_1^2 - x_2^2$; este plano se llama *seudoeuclídeo*.

§ 2. Plano semieuclídeo

Sea R un *espacio vectorial* de dos dimensiones provisto de una métrica semieuclídea y sea e_1, e_2 una base de este espacio en la que el cuadrado escalar (x, x) de un vector cualquiera $x = x_1e_1 + x_2e_2$ es igual a x_2^2. En particular, se tiene

$$(e_1, e_1) = 0, \ (e_2, e_2) = 1$$

y

$$(e_1 + e_2, e_1 + e_2) = 1 = (e_1, e_1) + 2(e_1, e_2) + (e_2, e_2) = \\ = 2(e_1, e_2) + 1,$$

de donde resulta

$$(e_1, e_2) = 0.$$

Llamaremos *canónica* a toda base de esta índole.

Sean $x = x_1e_1 + x_2e_2$ e $y = y_1e_1 + y_2e_2$ dos vectores arbitrarios de R; el producto escalar (x, y) es igual a

$$x_1y_1(e_1, e_1) + (x_1y_2 + x_2y_1)(e_1, e_2) + x_2y_2(e_2, e_2) = x_2y_2$$

y el módulo del vector x es igual a

$$|x| = \sqrt{x_2^2} = |x_2|.$$

Sea e_1', e_2' otra base, *también canónica*, del espacio R y sea

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

la matriz del cambio de la primera base por la segunda, es decir,

$$e_1' = a_{11}e_1 + a_{21}e_2 \quad \text{y} \quad e_2' = a_{12}e_1 + a_{22}e_2.$$

Tenemos entonces

$$(e_1', e_1') = a_{21}^2 = (e_1, e_1) = 0,$$

de donde $a_{21} = 0$, y

$$(e_2', e_2') = a_{22}^2 = (e_2, e_2) = 1,$$

es decir, $a_{22} = \pm 1$. Por consiguiente, la matriz del cambio de una base canónica por otra es de la forma

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & \pm 1 \end{bmatrix}. \tag{1}$$

Fijemos ahora una base canónica cualquiera e_1, e_2 y definamos el *ángulo* entre los vectores $x = x_1e_1 + x_2e_2$ e $y = y_1e_1 + y_2e_2$ tomándolo igual a

$$\left| \frac{y_1}{y_2} - \frac{x_1}{x_2} \right|. \tag{2}$$

El ángulo así definido no resulta, en general, invariante respecto a un cambio por una base nueva (aun canónica). Veamos qué limitaciones debemos imponer a la matriz del cambio para que el ángulo (2) no dependa del sistema (canónico) de coordenadas. Al pasar a una base (canónica) nueva con la matriz del cambio (1) las coordenadas de los vectores x e y se transforman, respectivamente, en

$$x_1' = a_{11}x_1 + a_{12}x_2, \quad x_2' = \pm x_2$$

y

$$y_1' = a_{11}y_1 + a_{12}y_2, \quad y_2' = \pm y_2,$$

siendo iguales los signos de x_2' e y_2'. Según la definición (2), el ángulo entre los vectores x e y en la base nueva debe ser igual a

$$\left|\frac{y_1'}{y_2'}-\frac{x_1'}{x_2'}\right| = \left|\frac{a_{11}y_1+a_{12}y_2}{\pm y_2}-\frac{a_{11}x_1+a_{12}x_2}{\pm x_2}\right| = |a_{11}|\left|\frac{y_1}{y_2}-\frac{x_1}{x_2}\right|;$$

y tendrá el mismo valor que antes si, y sólo si, $a_{11}=\pm 1$. Por ello, una vez escogida una base canónica, consideraremos a continuación sólo aquellas bases e_1', e_2' para las cuales la matriz del cambio de la base e_1, e_2 es de la forma

$$\begin{bmatrix}\pm 1 & v\\ 0 & \pm 1\end{bmatrix}$$

(hemos tomado $a_{12}=v$).

Indiquemos por A_0 la matriz

$$A_0 = \begin{bmatrix}1 & v\\ 0 & 1\end{bmatrix}. \tag{3}$$

Es obvio entonces que

$$A_1 = \begin{bmatrix}-1 & v\\ 0 & 1\end{bmatrix} = A_0\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix};$$

$$A_2 = \begin{bmatrix}1 & v\\ 0 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}A_0;$$

$$A_3 = \begin{bmatrix}-1 & v\\ 0 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}A_0\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}.$$

Consideremos ahora un *espacio vectorial puntual* de dos dimensiones en el que la *distancia* entre los puntos $X(x_1, x_2)$ e $Y(y_1, y_2)$ se toma igual al módulo del vector $\overline{XY}=(y_1-x_1, y_2-x_2)$ en la métrica semieuclídea, es decir, se toma igual a

$$|y_2-x_2|$$

o, en otras palabras, a la longitud euclídea de la proyección del vector sobre el eje de las ordenadas. (En particular, la longitud de cualquier segmento paralelo a e_1 será igual a cero.) Se llama *circunferencia* de **radio** r y **de centro** en un punto dado $M(\alpha_1, \alpha_2)$ al conjunto de todos los puntos que se encuentran a una misma distancia semieuclídea r del punto M; este conjunto es el par de rectas paralelas al eje de las abscisas que se encuentran a una misma dis-

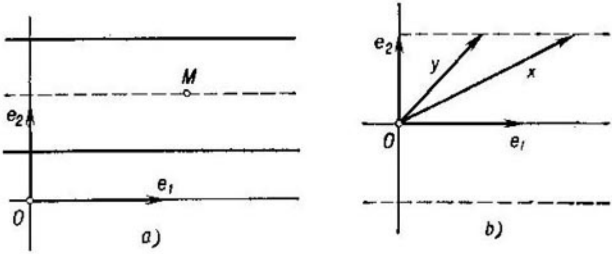


Fig. 18

tancia (euclídea) r del punto dado M (fig. 18, a). Cualquier punto de la recta que pasa por M y es paralela a estas rectas también será un centro de esta circunferencia. La ecuación de la circunferencia de radio r y de centro en el punto $M(\alpha_1, \alpha_2)$ es

$$(x_2-\alpha_2)^2 = r^2.$$

En particular, la ecuación de la «circunferencia unidad» (de la circunferencia de radio igual a la unidad) y de centro en el origen de coordenadas es

$$x_2^2 = 1.$$

Se llama ángulo entre dos rectas al ángulo entre los vectores paralelos a ellas. Si $x = (\xi, 1)$ e $y = (\eta, 1)$ son dos vectores de longitud igual a la unidad, el ángulo entre ellos es igual a

$$\left|\frac{\eta}{1}-\frac{\xi}{1}\right| = |\eta-\xi|$$

y se mide por el «arco» que estos vectores cortan en la «circunferencia unidad» (fig. 18, b). Observemos que en la métrica semieuclídea los ángulos suplementarios son iguales, ya que el ángulo entre los vectores $(\xi, 1)$ y $(-\eta, -1)$ es igual a $\left|\frac{\xi}{1}-\frac{-\eta}{-1}\right| = |\xi-\eta|$, es decir, es igual al ángulo entre los vectores $(\xi, 1)$ y $(\eta, 1)$.

Consideremos algunos teoremas de la «geometría semieuclídea elemental».

Teorema 1. *El lado mayor de un triángulo es igual a la suma de los dos otros lados.*

Efectivamente, puesto que $AB = A'B'$, $AC = A'C'$, $BC = B'C'$ (fig. 19, a) y $A'B' = A'C'+B'C'$, resulta que $AB = AC+BC$, es decir, $c = a+b$.

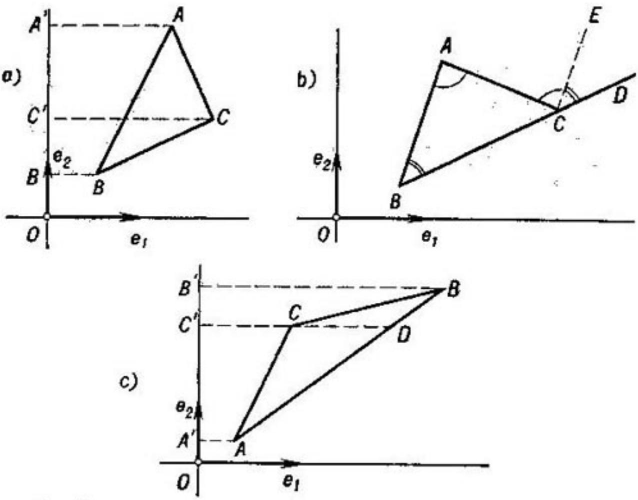


Fig. 19

Teorema 2. *El ángulo mayor de un triángulo es igual a la suma de los dos otros ángulos.*

Para demostrarlo tracemos la recta $CE \parallel BA$ (véase la fig. 19, *b*) Tenemos entonces $\angle ACE = A$ y $\angle ECD = B$ por ser ángulos de lados paralelos. Además, $\angle ACE + \angle ECD = \angle ACD = C$ y, por consiguiente,

$$C = A + B.$$

Teorema 3. *Los lados de un triángulo son proporcionales a los ángulos opuestos.*

Para demostrarlo tracemos la recta $CD \parallel e_1$ (fig. 19, *c*). Tenemos entonces $A = \frac{CD}{b}$ (donde CD es igual al valor absoluto de la diferencia de las abscisas de los puntos D y C, es decir, es igual a la longitud euclídea del segmento CD) y $B = \frac{CD}{a}$; luego $A \cdot b = B \cdot a$, de donde resulta

$$\frac{A}{a} = \frac{B}{b}.$$

$\left(\text{Análogamente se demuestra que } \frac{A}{a} = \frac{C}{c}.\right)$

Estas tres últimas afirmaciones permiten observar la cierta dualidad de los teoremas del plano semieuclídeo que se manifiesta en la paridad de los lados y de los ángulos de un triángulo. Si sustituimos en los enunciados de estos teoremas la palabra «lado» por la palabra «ángulo» y viceversa, el teorema 1 se convierte en el teorema 2 y el teorema 2 en el teorema 1; estos teoremas son duales uno del otro. El teorema 3 es el dual de sí mismo.

Esta dualidad no se da en el plano corriente (euclídeo), donde existen rectas paralelas (el ángulo entre las cuales es igual a cero) pero no existen «puntos paralelos» (la distancia entre los cuales es igual a cero). Esta «injusticia» queda salvada en la geometría semieuclídea, donde, además de rectas paralelas, existen también «puntos paralelos» (a saber, los puntos pertenecientes a una recta paralela a e_1).

§ 3. Plano seudoeuclídeo

Sea R un espacio vectorial de dos dimensiones provisto de una métrica seudoeuclídea y sea e_1, e_2 una base de este espacio en la que el cuadrado escalar de un vector cualquiera $x = x_1e_1+x_2e_2$ es igual a $x_1^2-x_2^2$. Tenemos, en particular,

$$(e_1, e_1) = 1, (e_2, e_2) = -1$$

y

$$(e_1+e_2,\ e_1+e_2) = 1-1 = 0 = (e_1,\ e_1)+2(e_1,\ e_2)+(e_2,\ e_2) = \\ = 2(e_1,\ e_2),$$

de donde resulta

$$(e_1,\ e_2) = 0.$$

Llamaremos *ortonormal* a toda base de este tipo. En una base ortonormal el producto escalar de los vectores

$$x = x_1e_1+x_2e_2 \quad \text{e} \quad y = y_1e_1+y_2e_2$$

es igual a

$$(x,\ y) = x_1y_1(e_1,\ e_1)+(x_1y_2+x_2y_1)(e_1,\ e_2)+x_2y_2(e_2,\ e_2) = \\ = x_1y_1-x_2y_2$$

y el módulo del vector x es igual a

$$|x| = \sqrt{x_1^2-x_2^2}\,.$$

Consideremos ahora un espacio puntual vectorial de dos dimensiones en el que la distancia entre los puntos $X(x_1, x_2)$ e $Y(y_1, y_2)$ se toma igual al módulo del vector $\overline{XY} = (y_1-x_1, y_2-x_2)$ en la métrica seudoeuclídea, es decir, se toma igual a

$$\sqrt{(y_1-x_1)^2-(y_2-x_2)^2}\,.$$

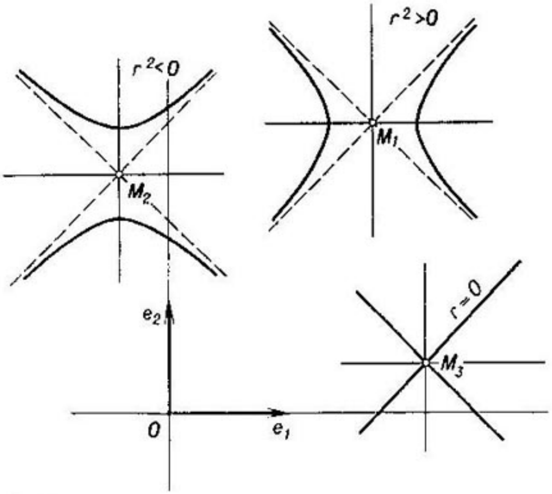


Fig. 20

La *circunferencia* de **radio** r y de **centro** en el punto $M(\alpha_1, \alpha_2)$ es el conjunto de todos los puntos que se encuentran a una misma distancia (seudoeuclídea) r del punto M. Por consiguiente, la ecuación de una circunferencia de radio r y de centro en el punto $M(\alpha_1, \alpha_2)$ es

$$(x_1-\alpha_1)^2-(x_2-\alpha_2)^2 = r^2.$$

(en la métrica euclídea esta ecuación representa una hipérbola si $r \neq 0$ y dos rectas que se cortan si $r = 0$; véase la fig. 20). *El radio de esta circunferencia puede ser positivo, nulo o incluso «imaginario puro».* Así, la ecuación de la circunferencia de radio *positivo* $r = a$ y de centro en el origen de coordenadas es

$$x_1^2-x_2^2 = a^2$$

(una hipérbola de eje real horizontal). La ecuación de la circunferencia de radio *imaginario* $r = ai$ (y del mismo centro) es

$$x_2^2-x_1^2 = a^2$$

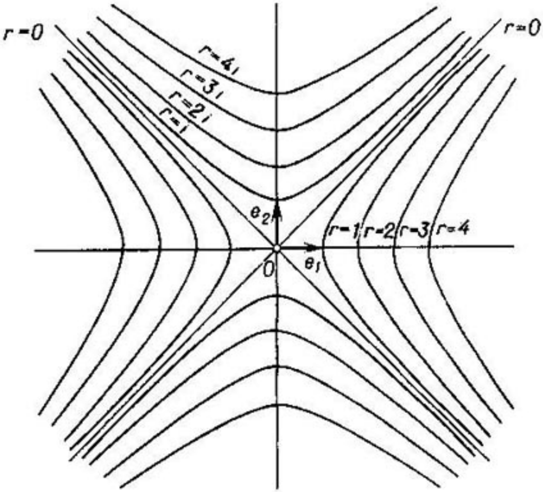


Fig. 21

(una hipérbola de eje real vertical). Estas dos familias de circunferencias están separadas por la circunferencia de radio *nulo*

$$x_1^2 - x_2^2 = 0$$

(dos rectas que son las asíntotas comunes de ambas familias de hipérbolas; véase la fig. 21).

Los vectores x e y son *ortogonales* (en este caso continuaremos escribiendo $x \perp y$) si el producto escalar de los mismos es igual a cero, es decir, si

$$(x,\ y) = x_1 y_1 - x_2 y_2 = 0.$$

En este caso tenemos

$$\frac{x_1}{x_2} = \frac{y_2}{y_1};$$

luego son inversos los coeficientes angulares de estos vectores considerados en la métrica euclídea; por consiguiente, si representamos en el plano euclídeo los vectores, ortogonales en la métrica

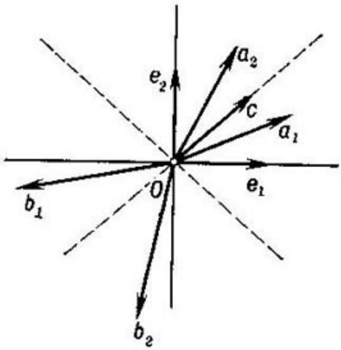


Fig. 22

seudoeuclídea, resulta que ellos son simétricos —en cuanto a la dirección— respecto a la bisectriz del primer y tercer cuadrantes (véase la fig. 22 en la que se tiene $e_1 \perp e_2$, $a_1 \perp a_2$ y $b_1 \perp b_2$).

Todo vector tal que $|x_1| = |x_2|$ es *ortogonal a sí mismo* y es *de longitud nula*. Para los vectores de longitud real se tiene $|x_1| > |x_2|$ y para los vectores de longitud imaginaria resulta $|x_1| < |x_2|$ (véase la misma fig. 22; los vectores e_1, a_1 y b_1 son de longitud real, los vectores e_2, a_2 y b_2 son de longitud imaginaria, mientras que el vector c es ortogonal a sí mismo y $|c| = 0$).

§ 4. Aplicaciones seudoortogonales

Una aplicación lineal $\mathcal{A}$ de un espacio seudoeuclídeo se llama *seudoortogonal* si conserva el producto escalar, es decir, si

$$(\mathcal{A}x, \mathcal{A}y) = (x, y)$$

para todos los $x, y \in R$.

Sea $\mathcal{A}$ una aplicación seudoortogonal de un plano seudoeuclídeo R y sea

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

su matriz en una base ortonormal e_1, e_2. Tenemos

$$\mathcal{A}e_1 = a_{11}e_1 + a_{21}e_2,$$
$$\mathcal{A}e_2 = a_{12}e_1 + a_{22}e_2.$$

Por definición,

$$(\mathcal{A}e_1, \mathcal{A}e_1) = (e_1, e_1) = 1,$$
$$(\mathcal{A}e_2, \mathcal{A}e_2) = (e_2, e_2) = -1$$

y

$$(\mathcal{A}e_1, \mathcal{A}e_2) = (e_1, e_2) = 0,$$

es decir,

$$\begin{aligned} a_{11}^2 - a_{21}^2 &= 1, \\ a_{12}^2 - a_{22}^2 &= -1, \end{aligned} \tag{4a}$$

y

$$a_{11} \cdot a_{12} - a_{21} \cdot a_{22} = 0. \tag{4b}$$

De la igualdad (4a) resulta que $a_{11} \neq 0$ y $a_{22} \neq 0$. De la igualdad (4b) se deduce que

$$\frac{a_{21}}{a_{11}} = \frac{a_{12}}{a_{22}}. \tag{5}$$

Indicando por β las razones iguales (5), obtenemos

$$\begin{aligned} a_{21} &= \beta a_{11}, \\ a_{12} &= \beta a_{22}. \end{aligned} \tag{6}$$

Introduciendo estos valores en las igualdades (4a), encontramos

$$a_{11}^2 - \beta^2 a_{11}^2 = 1, \text{ de donde } a_{11} = \frac{1}{\pm\sqrt{1-\beta^2}},$$

y

$$\beta^2 a_{22}^2 - a_{22}^2 = -1, \text{ de donde } a_{22} = \frac{1}{\pm\sqrt{1-\beta^2}};$$

por consiguiente, la matriz de la aplicación $\mathcal{A}$ es de la forma

$$A = \begin{bmatrix} \dfrac{1}{\pm\sqrt{1-\beta^2}} & \dfrac{\beta}{\pm\sqrt{1-\beta^2}} \\ \dfrac{\beta}{\pm\sqrt{1-\beta^2}} & \dfrac{1}{\pm\sqrt{1-\beta^2}} \end{bmatrix}, \tag{7}$$

con la particularidad de que ambos elementos de la primera columna y ambos elementos de la segunda columna se toman, como puede verse de las igualdades (6), con el mismo signo. Llamaremos *seudoortogonal* a toda matriz de este tipo.

Si indicamos por A_0 la matriz

$$A_0 = \begin{bmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} \\ \frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{bmatrix},$$

es fácil ver que

$$A_1 = \begin{bmatrix} -\frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} \\ -\frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{bmatrix} = A_0 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} \\ \frac{\beta}{\sqrt{1-\beta^2}} & -\frac{1}{\sqrt{1-\beta^2}} \end{bmatrix} = A_0 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} -\frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} \\ -\frac{\beta}{\sqrt{1-\beta^2}} & -\frac{1}{\sqrt{1-\beta^2}} \end{bmatrix} = A_0 \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

(las aplicaciones $\mathcal{A}_1$ y $\mathcal{A}_2$ difieren de $\mathcal{A}_0$ en una simetría axial, mientras que la aplicación $\mathcal{A}_3$ difiere en una simetría central).

Para los determinantes tenemos $|A_0| = |A_3| = 1$ y $|A_1| = |A_2| = -1$.

Observemos que, por ser $\frac{1}{\sqrt{1-\beta^2}} \geqslant 1$, existe un valor φ tal que $\frac{1}{\sqrt{1-\beta^2}} = \operatorname{ch} \varphi$ y $\frac{\beta}{\sqrt{1-\beta^2}} = \operatorname{sh} \varphi$; luego

$$A_0 = \begin{bmatrix} \operatorname{ch} \varphi & \operatorname{sh} \varphi \\ \operatorname{sh} \varphi & \operatorname{ch} \varphi \end{bmatrix}.$$

Si β varía continuamente desde 0 hasta 1 (o desde 0 hasta -1), los extremos de los vectores $\mathcal{A}_0 e_1$ y $\mathcal{A}_0 e_2$ se deslizan continuamente a lo largo de las circunferencias unidades $x_1^2-x_2^2 = \pm 1$, uno hacia el otro si $\beta \to 1$ y en direcciones contrarias si $\beta \to -1$.

Supongamos que en un plano seudoeuclídeo R se tienen dos bases *ortonormales* e_1, e_2 y e_1', e_2' y sea

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

la matriz del cambio de la primera base por la segunda. Consideremos la aplicación lineal $\mathcal{A}$ de matriz A en la base e_1, e_2 y demostremos que es *seudoortogonal*. Efectivamente, tenemos por hipótesis

$$\mathcal{A}e_1 = a_{11}e_1 + a_{21}e_2 = e_1'$$

y

$$\mathcal{A}e_2 = a_{12}e_1 + a_{22}e_2 = e_2'.$$

Si $x = x_1e_1 + x_2e_2$ e $y = y_1e_1 + y_2e_2$ son unos vectores cualesquiera de R, se tiene

$$\mathcal{A}x = x_1\mathcal{A}e_1 + x_2\mathcal{A}e_2 = x_1e_1' + x_2e_2'$$

y

$$\mathcal{A}y = y_1\mathcal{A}e_1 + y_2\mathcal{A}e_2 = y_1e_1' + y_2e_2'.$$

Puesto que ambas bases e_1, e_2 y e_1', e_2' son ortonormales, el producto escalar

$$(\mathcal{A}x, \mathcal{A}y) = x_1y_1 - x_2y_2 = (x, y).$$

Por consiguiente, la aplicación $\mathcal{A}$ es seudoortogonal y su matriz es de la forma (7).

§ 5. Espacio de sucesos. Principio de relatividad de Galileo

Supongamos que un punto M se desplaza por una línea recta l en la que se ha establecido *un sistema de referencia S*. Esto significa que dicha recta está provista de una escala con las divisiones respectivas y que en todos los puntos de la recta se tienen relojes sincronizados entre sí.

Supongamos que en el instante de tiempo t la coordenada del punto M es igual a x. Este hecho o, como diremos, «*suceso*» puede ser indicado en un plano P (de dos dimensiones) mediante el punto de coordenadas (x, t). El plano P se llama *espacio de sucesos*.

Las coordenadas del punto en el espacio de sucesos varían con el transcurso del tiempo, aun cuando el punto M no altera su posición sorbe la recta l (debido a la variación del tiempo t). Por consiguiente, la existencia del punto en el espacio y en el tiempo se manifestará por una *línea* en el plano P. Esta línea será recta si, y sólo si, el punto M se desplaza por la recta l con una velocidad constante u; en este caso su posición en el plano P se determina por la ecuación

$$x = ut+b,$$

donde $b = x(0)$ es la posición del punto en el instante $t = 0$. Si el punto M permanece inmóvil sobre la recta («se desplaza con velocidad nula»), la recta que le corresponde en el plano P será paralela al eje t.

Supongamos que a lo largo de la recta l se desplaza uniformemente con la velocidad v otro sistema de referencia S', con la particularidad de que los origenes de coordenadas de ambos sistemas coinciden en el instante inicial: $x = x' = 0$ para $t = 0$. En este caso la coordenada x del punto M en el sistema S y su coordenada x' en el sistema S' estarán ligadas por la relación

$$x = x'+vt.$$

Aquí se supone que el tiempo t en el sistema S y el tiempo t' en el sistema S' es el mismo, es decir, que $t = t'$ para un mismo suceso.

Las transformaciones

$$\begin{aligned} x &= x'+vt, \\ t &= t' \end{aligned} \tag{8}$$

o, que viene a ser lo mismo,

$$\begin{aligned} x' &= x-vt, \\ t' &= t \end{aligned}$$

se llaman *transformaciones de Galileo*. Derivándolas respecto a t, obtenemos de aquí

$$\frac{dx}{dt} = \frac{dx'}{dt} + v,$$

es decir,

$$u = u'+v, \tag{9}$$

donde u es la velocidad del punto en el sistema S y u' es su velocidad en el sistema S'. Esta es la ley de composición de velocidades en la mecánica clásica: la velocidad u del punto en el sistema antiguo de referencia es igual a su velocidad u' en el nuevo sistema sumada a la velocidad de «traslación» v (a la velocidad del nuevo sistema de referencia respecto al antiguo). Derivando una vez más respecto a t, obtenemos

$$\frac{d^2x}{dt^2} = \frac{d^2x'}{dt^2}.$$

Por consiguiente, *son idénticas las aceleraciones del punto M en el sistema S y en el sistema S'*, por donde llegamos a la conclusión de que fuerzas iguales provocan en ambos sistemas iguales efectos (que se describen por la segunda ley de Newton: la aceleración provocada por una fuerza F es proporcional a esta fuerza). Esto mismo se expresa en otras palabras diciendo que *las leyes de la mecánica son invariantes respecto a las transformaciones de Galileo* **(principio de relatividad de Galileo)**.

Volvamos a las fórmulas (8). Se ve de ellas que al pasar del sistema S al sistema S' las coordenadas de los puntos del espacio de sucesos se someten a una aplicación lineal de matriz

$$\begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix}. \tag{10}$$

Este hecho sugiere la idea de introducir en el espacio de sucesos *la métrica semieuclídea*. La distancia entre dos sucesos $A(x_1, t_1)$ y $B(x_2, t_2)$ obtiene entonces un sentido físico concreto: es igual a $|t_2 - t_1|$, es decir, *al intervalo temporal entre los sucesos A y B*.

Además, como el cambio de un sistema de referencia por otro se determina por la matriz (10), también resultará invariante el concepto de ángulo introducido en el § 2. Para ver su sentido físico, consideremos dos puntos M_1 y M_2 que se desplazan uniformemente a lo largo de la recta l. Indiquemos por u_1 y u_2 sus velocidades respectivas. En el plano P el movimiento de estos puntos se determina por unas rectas m_1 y m_2. Sea $A_0(x_0, t_0)$ el punto de intersección de estas rectas (esto significa que para $t = t_0$ ambos puntos M_1 y M_2 tenían la abscisa x_0 y se encontraban en un mismo lugar de la recta l). Supongamos que para $t = t_1$ la abscisa del punto M_1 es x_1 y que para $t = t_2$ la abscisa del punto M_2 es x_2. En estas condiciones el ángulo entre las rectas m_1 y m_2 (en la métrica semieuclídea) es igual al ángulo entre los vectores $\overline{A_0A_1}$ y $\overline{A_0A_2}$, donde $A_1(x_1, t_1)$ y $A_2(x_2, t_2)$ (fig. 23); luego es igual a

$$\left| \frac{x_2 - x_0}{t_2 - t_0} - \frac{x_1 - x_0}{t_1 - t_0} \right| = |u_2 - u_1|,$$

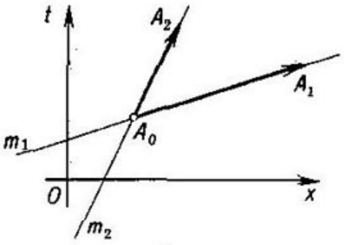


Fig. 23

es decir, es igual a *la velocidad relativa* del movimiento de estos puntos.

Si la distancia y el ángulo se interpretan de esta forma, los teoremas 1, 2 y 3 de la pág. 206 obtienen un sentido físico concreto que el lector podrá establecer.

§ 6. Principio de relatividad de Einstein

Es natural hacer la siguiente conclusión de la ley (9) de composición de velocidades: si el sistema de referencia S' se desplaza uniformemente con la velocidad v respecto a S y si la luz se propaga en el sistema S con la velocidad c, su velocidad en el sistema S' debe ser igual a $c-v$ en el sentido del movimiento del sistema S' y a $c+v$ en el sentido contrario. Sin embargo, Michelson, físico norteamericano, demostró experimentalmente en 1881 que *sobre la Tierra móvil la luz solar se propaga con la* ***misma*** *velocidad en todas las direcciones.*

Después de los intentos de varios científicos de concordar los resultados de los experimentos de Michelson con la teoría, fue publicado en 1905 un trabajo fundamental de Einstein en el que se exponía una nueva teoría del espacio y del tiempo, *la teoría de la relatividad especial*. Consideraremos aquí sólo los conceptos más importantes y elementales de esta teoría.

Einstein tomó como base de la teoría *la ley de constancia de la velocidad de la luz* en todos los sistemas inerciales de referencia[1].

[1] En Física se llama inercial a tal sistema de referencia en el que un cuerpo se desplaza uniforme y rectilíneamente cuando sobre él no actúan fuerzas exteriores.

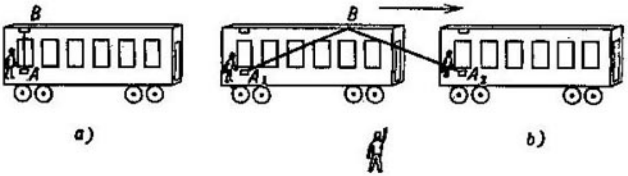


Fig. 24

Por consiguiente, *el principio de relatividad de Galileo consiste en la imposibilidad de determinar el movimiento uniforme de un sistema mecánico respecto a otro mediante cualesquiera experimentos mecánicos realizados dentro de este sistema. El principio de relatividad de Einstein afirma que es imposible lograr esto aun basándose en cualesquiera fenómenos ópticos* (relacionados, como se sabe, con el electromagnetismo) *y no sólo mecánicos.*

Pero al aceptar la ley de constancia de la velocidad de la luz, Einstein se vió obligado a *renunciar de la hipótesis de existencia del tiempo absoluto*, válido para medir los intervalos temporales simultáneamente en todos los sistemas de referencia.

Esta relatividad del tiempo necesariamente se deduce de la ley de constancia de la velocidad de la luz como puede verse del siguiente ejemplo sencillo[1]. Imaginémonos un tren, de dimensión lineal muy grande, cuya velocidad es comparable con la velocidad de la luz (*«el tren de Einstein»*). Supongamos que junto a la ventana de este tren se encuentra un observador que en un momento de tiempo determinado enciende una linterna enviando un rayo de luz hacia el techo. En el techo hay un espejo y el rayo después de reflejarse vuelve al observador. Desde el punto de vista de este observador la trayectoria del rayo de luz es el segmento *AB* recorrido dos veces (fig. 24, *a*). Para un observador, que se encuentra fuera del tren, la trayectoria del rayo de luz será la quebrada formada por los lados del triángulo isósceles cuya altura es igual a *AB* (fig. 24, *b*). Por consiguiente, desde el punto de vista del observador exterior la luz recorre una trayectoria mayor que desde el punto de vista del pasajero del tren. Puesto que la velocidad de la luz es constante, el tiempo que necesita la luz para recorrer esta trayectoria tomado

---

[1] Hemos tomado este y el siguiente ejemplos del folleto de L. D. Landau y Yu. B. Rumer "Qué es la teoría de la relatividad". Editorial MIR, 1972.

por el reloj del observador exterior será mayor que el tiempo tomado por el reloj del pasajero: el reloj dentro del tren *se retrasa* en comparación con el reloj de la estación.

También el concepto de simultaneidad se hace relativo al aceptar la ley de constancia de la velocidad de la luz lo que puede verse con claridad en el ejemplo siguiente. Supongamos que en el centro de un vagón de este mismo tren de Einstein se encuentra un observador que en un momento determinado del tiempo enciende una linterna. Las puertas del vagón están provistas de un mecanismo que permite abrirlas en cuanto la luz llegue a ellas. El observador que se encuentra en el centro del vagón verá como la puerta delantera y la trasera se abren simultáneamente. En cambio, desde el punto de vista de un observador exterior la puerta delantera se aleja del rayo de luz mientras que la puerta trasera se acerca a éste. Como la velocidad de la luz es constante, desde el punto de vista del observador exterior la luz llegará a la puerta delantera más tarde que a la trasera y ésta se abrirá antes.

Es más, incluso el orden en el que se realizan los sucesos puede resultar diferente para estos observadores. Si (debido a un defecto en el mecanismo de las puertas) la puerta trasera se abre transcurrido un tiempo después de que la luz llegue a ella y si esta diferencia en el tiempo es suficientemente pequeña, para el observador exterior la puerta trasera continuará abriéndose antes que la delantera mientras que para el observador que se encuentra en el centro del vagón el orden de estos sucesos será el inverso.

§ 7. Transformaciones de Lorentz

Nos vemos obligados, pues, a prescindir de la hipótesis de que el tiempo es el mismo en todos los sistemas de referencia que se desplazan uniformemente unos respecto a otros. No podemos aceptar ahora que $t' = t$ para un mismo suceso. ¿Cómo están relacionadas las coordenadas x, t de un punto en un sistema S con las coordenadas x', t' de este punto en un sistema S' que se desplaza uniformemente con la velocidad v respecto a S? En la mecánica clásica esta relación es lineal (transformaciones de Galileo). *Mantendremos esta hipótesis de que x', t' dependen linealmente de x, t.* El paso de S a S' corresponde entonces al paso a una base nueva en el espacio de sucesos. ¿Cúal es la métrica de este espacio?

Supongamos que los orígenes de los dos sistemas S y S' coinciden en un determinado momento del tiempo (inicial para *ambos* sistemas): $x = 0$ y $x' = 0$ para $t = 0$ y $t' = 0$. Supongamos, además, que en el momento $t = t' = 0$ se envía, desde el origen de coordenadas común de ambos sistemas, una señal luminosa que en el sistema S se recibe en el punto x y en el momento t y en el sistema S' se recibe en el punto x' y en el momento t'. Puesto que la velocidad de la luz c es constante, se tiene

$$\left|\frac{x}{t}\right| = \left|\frac{x'}{t'}\right| = c,$$

de donde resulta que $x^2 - c^2t^2 = 0$ y $x'^2 - c^2t'^2 = 0$. Luego, si la expresión

$$x^2 - c^2t^2 \tag{11}$$

es igual a cero en un sistema inercial de referencia, será igual a cero en todos los demás sistemas. Aceptaremos ahora adicionalmente que *la expresión* (11) *es, en general, un* ***invariante***, es decir, que *esta expresión es la misma en todos los sistemas inerciales de referencia.*

Tomando $x = x_1$ y $ct = x_2$ (y, respectivamente, $x' = x_1'$ y $ct' = x_2'$), podemos considerar nuestro espacio de sucesos como un *plano seudoeuclídeo* en el que la expresión (11), igual a

$$x_1^2 - x_2^2,$$

es el cuadrado de la distancia entre el punto (x_1, x_2) y el origen de coordenadas o, que es lo mismo, el cuadrado de la longitud del vector correspondiente. Pero la base en la que el cuadrado de la longitud de un vector toma esta forma es *ortonormal*. Por esta misma razón será ortonormal la base correspondiente del sistema S'; luego la matriz A del cambio de la base del sistema S por la base del sistema S' es seudoortogonal

$$A = \begin{bmatrix} \dfrac{1}{\pm\sqrt{1-\beta^2}} & \dfrac{\beta}{\pm\sqrt{1-\beta^2}} \\ \dfrac{\beta}{\pm\sqrt{1-\beta^2}} & \dfrac{1}{\pm\sqrt{1-\beta^2}} \end{bmatrix}$$

(y en cada una de las columnas se toma sólo un signo).

Por consiguiente,

$$e_1' = \frac{e_1 + \beta e_2}{\pm\sqrt{1-\beta^2}}, \qquad e_2' = \frac{\beta e_1 + e_2}{\pm\sqrt{1-\beta^2}}.$$

Consideremos primero el caso en el que ambos denominadores son positivos y la matriz A es

$$A_0 = \begin{bmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} \\ \frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{bmatrix}.$$

Las coordenadas x_1, x_2 y x_1', x_2' están ligadas entonces por las relaciones

$$x_1 = \frac{x_1'+\beta x_2'}{\sqrt{1-\beta^2}}, \qquad x_2 = \frac{\beta x_1'+x_2'}{\sqrt{1-\beta^2}}$$

o en las denotaciones antiguas

$$x = \frac{x'+\beta ct'}{\sqrt{1-\beta^2}}, \qquad t = \frac{\frac{\beta}{c}x'+t'}{\sqrt{1-\beta^2}}. \tag{12}$$

Expresando de aquí x' y t' mediante x y t, obtenemos las fórmulas

$$x' = \frac{x-\beta ct}{\sqrt{1-\beta^2}}, \qquad t' = \frac{-\frac{\beta}{c}x+t}{\sqrt{1-\beta^2}}. \tag{13}$$

¿Cuál es el sentido físico del parámetro β? Consideremos en el sistema S' un punto inmóvil M, por ejemplo, el origen de coordenadas $x' = 0$. Según la primera de las fórmulas (13), tenemos para este punto

$$x-\beta ct = 0$$

o

$$\frac{x}{t} = \beta c.$$

Pero $\frac{x}{t}$ es la velocidad del punto M en el sistema S y es igual, evidentemente, a la velocidad del sistema S' respecto a S. Por consiguiente,

$$v = \beta c \quad \text{y} \quad \beta = \frac{v}{c}.$$

Introduciendo este valor de β en las fórmulas (12) y (13), obtenemos

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad t = \frac{\frac{v}{c^2}x' + t'}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{14}$$

y

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad t' = \frac{-\frac{v}{c^2}x + t}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{15}$$

Las transformaciones (14) y (15) se llaman *transformaciones de Lorentz*. Observemos que las fórmulas (15) se obtienen de las fórmulas (14) cambiando simplemente el signo de v.

Hemos supuesto que en la matriz del cambio de la base e_1, e_2 por la base e_1', e_2' todos los denominadores son positivos. Demostremos que los demás casos pueden ser excluidos. Si en la segunda columna de la matriz del cambio aparecen los signos «$-$» (y en la primera, cualesquiera), obtenemos las fórmulas

$$x = \frac{\pm x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad t = \frac{\pm \frac{v^2}{c^2}x' - t'}{\sqrt{1 - \frac{v^2}{c^2}}};$$

aquí a un aumento de t' corresponde una disminución de t y esto es imposible, ya que en este caso el orden de *todos* los sucesos en el sistema S' será inverso al orden de estos mismos sucesos en el sistema S. Si los signos menos aparecen en la primera columna de la matriz del cambio (y en la segunda columna figuran los signos más), obtenemos las fórmulas

$$x = \frac{-x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad t = \frac{-\frac{v}{c^2}x' + t'}{\sqrt{1 - \frac{v^2}{c^2}}};$$

de estas fórmulas se puede pasar a (14) cambiando el signo de x', es decir, cambiando la dirección del eje Ox' por la contraria.

Las fórmulas de Lorentz tienen sentido sólo para $\left|\frac{v}{c}\right| < 1$, de donde resulta que $|v| < c$, es decir, que *es imposible el movimiento con la velocidad que sobrepasa la velocidad de la luz.*

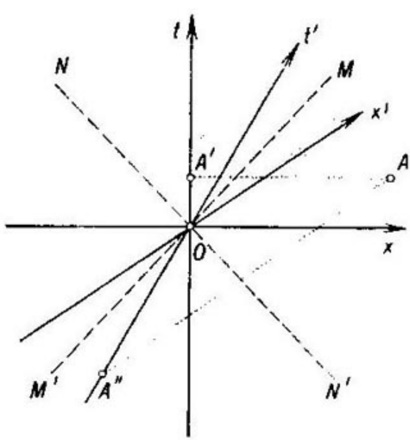


Fig. 25

Si v es pequeño en comparación con c, tenemos $\sqrt{1-\frac{v^2}{c^2}} \approx$ ≈ 1 y en este caso

$$x' \approx x - vt, \quad t' \approx t.$$

Por consiguiente, *siendo pequeño el valor de* v (en comparación con c) *las transformaciones de Lorentz se convierten en las transformaciones de Galileo de la mecánica clásica.*

Sean Ox y Ot los ejes coordenados del espacio de sucesos del sistema S y sean Ox' y Ot' los ejes del sistema S' (fig. 25). Como sabemos, los ejes Ox' y Ot' son simétricos respecto a las bisectrices MM' y NN' de los cuadrantes del primer sistema. El eje Ot' puede ser considerado como la gráfica del movimiento del origen de coordenadas de S' respecto a S: para todos sus puntos se tiene $x' = 0$. Recíprocamente, el eje Ot es la gráfica del movimiento del origen de coordenadas del sistema S respecto a S'. El valor absoluto de la tangente del ángulo que el eje Ot' forma con el eje Ox es igual a

$$\left|\frac{ct}{x}\right| = \frac{c}{|v|},$$

donde $\frac{x}{t} = v$ es la velocidad con la que el sistema S' se desplaza respecto a S. Puesto que $|v| < c$, resulta que el valor absoluto de esta tangente es mayor que la unidad, es decir, que *todos los ejes de tiempo Ot se encuentran dentro del ángulo MON y que, por consiguiente, todos los ejes espaciales Ox se encuentran dentro del ángulo MON'.*

Para las rectas MM' y NN' tenemos $\left|\frac{x}{ct}\right| = 1$, es decir, $|v| = c$; en todos los sistemas de referencia ésta es la gráfica del movimiento con la velocidad de la luz.

§ 8. Algunos resultados que se deducen de las fórmulas de Lorentz

1. Regla de composición de velocidades.

De la igualdad (15) obtenemos

$$\frac{dx'}{dt'} = \frac{dx'}{dt} : \frac{dt'}{dt} = \frac{\frac{dx}{dt} - v}{\sqrt{1-\frac{v^2}{c^2}}} : \frac{-\frac{v}{c^2}\frac{dx}{dt} + 1}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{\frac{dx}{dt} - v}{-\frac{v}{c^2}\frac{dx}{dt} + 1} = \frac{u-v}{-\frac{uv}{c^2}+1}$$

o

$$u' = \frac{u-v}{-\frac{uv}{c^2}+1}, \tag{16}$$

de donde resulta

$$u = \frac{u'+v}{\frac{u'v}{c^2}+1}.$$

Esta es la nueva *fórmula de composición de velocidades.*

Siendo $u = c$, de la fórmula (16) obtenemos

$$u' = \frac{c-v}{-\frac{v}{c}+1} = c$$

y, recíprocamente, siendo $u' = c$, también

$$u = \frac{c+v}{\frac{v}{c}+1} = c.$$

Si u y v son pequeños en comparación con c, se tiene $u' \approx u - v$.

2. *Relatividad de la simultaneidad.*

Supongamos que los sucesos A y B acaecen en el sistema S *para un mismo momento del tiempo* t en los puntos con abscisas diferentes x_1 y x_2. En el sistema S' estos sucesos acaecerán, según la segunda de las fórmulas (15), en los momentos

$$t_1' = \frac{-\frac{v}{c^2}x_1 + t}{\sqrt{1-\frac{v^2}{c^2}}} \quad \text{y} \quad t_2' = \frac{-\frac{v}{c^2}x^2 + t}{\sqrt{1-\frac{v^2}{c^2}}},$$

de donde resulta

$$t_2' - t_1' = \frac{\frac{v}{c^2}(x_1 - x_2)}{\sqrt{1-\frac{v^2}{c^2}}} \neq 0,$$

es decir, *dos sucesos simultáneos en un sistema de referencia no serán simultáneos en otro sistema.* Además, la diferencia $t_2' - t_1'$ puede ser positiva y negativa según el signo que tenga la diferencia $x_1 - x_2$. Esto queda muy claro gráficamente: si los sucesos A y B son simultáneos en el sistema S, el segmento AB debe ser paralelo al eje Ox y, si son simultáneos en el sistema S', este segmento debe ser paralelo al eje Ox'.

Es más, incluso el orden de los sucesos puede ser distinto en los sistemas S y S'. Por ejemplo, en la fig. 25 los sucesos A y A' son simultáneos en el sistema S ($AA' \| Ox$) con la particularidad de que A' acontece, obviamente, después de O y, por consiguiente, también A ocurre en S después de O. En el sistema S' los sucesos A y A'' son sumultáneos ($AA'' \| Ox'$) y, por consiguiente, el suceso A (igual que A'') antecede al suceso O.

Naturalmente, aquí surge el siguiente problema: ¿no podría ocurrir que, por ejemplo, un suceso O que en el sistema S es causa de un suceso A ocurra en otro sistema S' después de A lo que estaría en contradicción con el principio de casualidad? Demostremos que esto no tiene lugar.

Los puntos correspondientes a los sucesos que en el sistema S suceden después del suceso O son aquellos puntos, y sólo aquellos, que se hallan *por encima* del eje Ox; los puntos correspondientes a los sucesos posteriores al suceso O en el sistema S' son los puntos que se encuenran *por encima* del eje Ox'. Puesto que todos los ejes espaciales pasan por el interior del ángulo MON' (véase el final del § 7) y como es evidente que cualquier recta de este tipo es el eje espacial de un sistema de referencia, resulta que *la intersección de todos los semiplanos, superiores a algún eje espacial, es el ángulo MON formado por aquellos sucesos, y sólo aquellos, que son* ***posteriores*** *a O en todos los sistemas de referencia (este ángulo puede ser denominado «dominio del futuro»). Análogamente el ángulo M'ON' representa el conjunto de todos los sucesos que en todos los sistemas de referencia* ***anteceden*** *al suceso O (el «dominio del pasado»).*

En cambio, los puntos pertenecientes a los ángulos MON' y NOM' corresponden a los sucesos que en unos sistemas de referencia preceden a O y en otros sistemas siguen a O. Sin embargo, ninguno de estos sucesos no puede tener como causa el suceso O. Efectivamente, si el suceso O ha sido causa del suceso $A(x, t)$ (véase la misma fig. 25), debe existir una perturbación que se desplaza desde O hacia A. Pero esto es imposible ya que la longitud del vector $\overline{OA}$ es real y, por consiguiente, en todos los sistemas de referencia tenemos para este vector

$$x^2-c^2t^2 > 0, \text{ es decir, } x^2 > c^2t^2,$$

de donde resulta

$$\left|\frac{x}{t}\right| > c;$$

luego, la velocidad $u = \frac{x}{t}$ de esta perturbación debe ser superior a la velocidad de la luz lo cual es imposible.

Análogamente se demuestra que la ley de la causalidad es válida para cualesquiera dos sucesos A y B: si A puede ser causa de B, es decir, si existe una señal que se propaga (en el sistema de referencia dado S) desde A hacia B con una velocidad $v < c$, el suceso A antecede al suceso B *en todos* los sistemas inerciales de referencia.

3. *Contracción de las longitudes.*

Supongamos que una barra de longitud l reposa en un sistema S; indiquemos por x_1 y x_2, donde $x_1 < x_2$, las coordenedas de sus extremos; en estas condiciones

$$l = x_2-x_1.$$

Para determinar la longitud l' de la barra en un sistema S', debemos indicar las coordenadas de sus extremos en un momento cualquiera

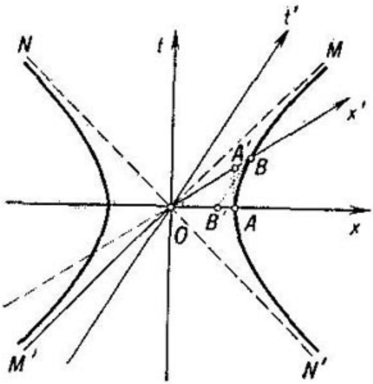


Fig. 26

t' (*¡el mismo para ambos extremos*!). Si estas coordenadas son x_1' y x_2', tenemos de acuerdo con la primera de las fórmulas (14)

$$x_1 = \frac{x_1' + vt'}{\sqrt{1-\frac{v^2}{c^2}}}, \quad x_2 = \frac{x_2' + vt'}{\sqrt{1-\frac{v^2}{c^2}}},$$

de donde resulta

$$l = x_2 - x_1 = \frac{x_2' - x_1'}{\sqrt{1-\frac{v^2}{c^2}}}$$

o, tomando en cuenta que la longitud l' de la barra en el sistema S' es igual a $x_2' - x_1'$,

$$l' = l\sqrt{1-\frac{v^2}{c^2}} < l.$$

Por consiguiente, *la longitud l' de una barra en un sistema de referencia respecto al cual la barra se desplaza es menor que su longitud en el sistema respecto al cual la barra reposa.*

Expliquemos este resultado gráficamente. Sea A el punto de intersección de la hipérbola $x^2 - c^2t^2 = l^2$ y del eje Ox (fig. 26); su distancia hasta el origen de coordenadas es igual en el sistema S a l. Si $AA' \parallel Ot$, los puntos A y A' se encuentran en el sistema S

a una misma distancia l del origen de coordenadas. (Es un mismo punto, que reposa en el sistema S, considerado en diferentes momentos del tiempo.) Pero en el sistema S' la distancia entre el punto A' y el origen de coordenadas es igual a OA'; luego es *menor* que OB igual a l.

Recíprocamente, el punto B, que es el cruce de la hipérbola $x^2-c^2t^2 = l^2$ y del eje Ox', se encuentra en el sistema S' a una distancia l del origen de coordenadas. Si $BB' \parallel Ot'$, el punto B' en el sistema S' se encuentra a la misma distancia l del origen de coordenadas; sin embargo, en el sistema S la distancia entre el punto B' y el punto O es igual a $OB' < OA = l$; luego, *la contracción* relativista, es decir, relacionada con la teoría de la relatividad, *de las longitudes es recíproca.*

Si v es pequeño en comparación con la velocidad de la luz, esta contracción de las longitudes en un sistema móvil es tan pequeña que en la práctica no puede ser observada. $\Bigg($Puesto que $l' = l\sqrt{1-\frac{v^2}{c^2}} \approx l\left(1-\frac{v^2}{2c}\right)$, resulta que la diferencia $l-l' \approx \frac{lv^2}{2c^2} = \frac{l}{2}\left(\frac{v}{c}\right)^2$ es del segundo orden respecto a $\frac{v}{c}$.$\Bigg)$ Por ejemplo, desde una nave cósmica (con la velocidad de 12 km/s) el diámetro de la Tierra (12 000 km) parecerá contraído solamente en 1 cm.

4. *Retraso del tiempo.*

Supongamos que en un reloj, inmóvil en un sistema S, ha transcurrido el tiempo T desde t_1 hasta t_2:

$$T = t_2 - t_1.$$

Determinemos el valor t_1' correspondiente a t_1, y el valor t_2', correspondiente a t_2, *en un mismo punto* con la abscisa x' en el sistema S'. De acuerdo a la segunda de las fórmulas (14), tenemos

$$t_1 = \frac{\frac{v}{c^2}x' + t_1'}{\sqrt{1-\frac{v^2}{c^2}}}, \quad t_2 = \frac{\frac{v}{c^2}x' + t_2'}{\sqrt{1-\frac{v^2}{c^2}}},$$

y de aquí resulta

$$T = t_2 - t_1 = \frac{t_2' - t_1'}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{T'}{\sqrt{1-\frac{v^2}{c^2}}}$$

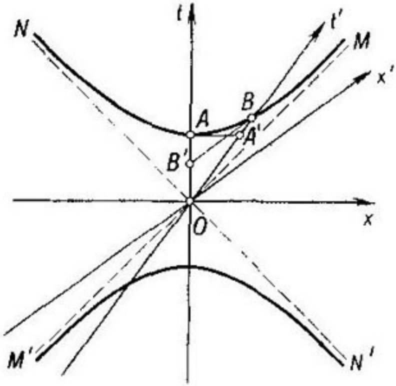


Fig. 27

donde $T' = t_2' - t_1'$ es el intervalo de tiempo que transcurre en el sistema S' mientras en el sistema S transcurre el tiempo T desde t_1 hasta t_2; por consiguiente,

$$T' = T\sqrt{1-\frac{v^2}{c^2}} < T.$$

Hemos encontrado que *en el sistema S' que se desplaza respecto al sistema S el tiempo transcurre más lentamente que en el propio sistema S*. Expliquemos este resultado gráficamente. Consideremos la hipérbola

$$x^2 - c^2t^2 = -c^2T^2$$

(fig. 27) y sea A el punto de intersección de esta hipérbola con el eje Ot; en estas condiciones, su distancia temporal hasta el punto O, es decir, el tiempo transcurrido desde el suceso O hasta el suceso A, es igual a T en el sistema S. Si $AA' \parallel Ox$, los sucesos A y A' son simultáneos en el sistema S. Pero en el sistema S' el tiempo transcurrido entre O y A' es igual a OA'; luego es *menor* que OB igual a T.

Recíprocamente, en el sistema S' el punto B se encuentra a una distancia temporal T del punto O. Si $BB' \parallel Ox'$, los sucesos B y B' son simultáneos en el sistema S'; sin embargo, en el sistema S la distancia temporal entre los puntos B' y O es igual a OB' y es menor que OA igual a T; luego, *el retraso temporal lorentziano es recíproco.*

Si la velocidad v es pequeña en comparación con la velocidad de la luz, este retraso del tiempo en un sistema de referencia móvil es nada más que $\frac{T}{2}\left(\frac{v}{c}\right)^2$, es decir, es de segundo orden respecto a $\frac{v}{c}$ y prácticamente no puede ser observado. Por ejemplo, para un cosmonauta las veinticuatro horas terrestres disminuyen en menos que $\frac{2}{10\,000}$ de segundo.

5. *Aumento de la masa de un cuerpo en movimiento.*

Sin exponer más resultados de la teoría de la relatividad, mencionaremos solamente el fenómeno del aumento de la masa de un cuerpo móvil.

Si sobre un cuerpo actúa una fuerza constante, la velocidad de su movimiento, en las condiciones corrientes, aumenta proporcionalmente al tiempo de acción de la fuerza. Sin embargo, debido a la existencia de la velocidad límite, esta proporcionalidad no puede mantenerse también para las velocidades grandes. Para las velocidades comparables con la velocidad de la luz el aumento ulterior de la velocidad se detiene: el cuerpo ofrece una mayor resistencia a la fuerza que sobre él actúa. Podemos decir que *aumenta la masa del cuerpo*. Resulta que

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}},$$

donde m es la masa del cuerpo en movimiento, v es su velocidad y m_0 es la masa de reposo, es decir, la masa que tiene el cuerpo respecto al sistema de referencia en el que este cuerpo reposa. Por ejemplo, si en un acelerador moderno un electrón alcanza una velocidad que sólo en 30 km/s difiere de la velocidad de la luz, la masa de este electrón aumenta casi 70 veces:

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{m_0}{\sqrt{\frac{c+v}{c}\cdot\frac{c-v}{c}}} \approx \frac{m_0}{\sqrt{2\cdot\frac{c-v}{c}}} =$$

$$= \frac{m_0}{\sqrt{2\cdot\frac{30}{300\,000}}} = \frac{100m_0}{\sqrt{2}} \approx 70\, m_0.$$

Hemos considerado el movimiento de un punto a lo largo de una línea recta. En el caso general, en el que un sistema espacial de referencia se desplaza uniforme y rectilíneamente respecto a otro sistema, la dirección de este desplazamiento puede ser tomada comó la del eje Ox y entonces en la mecánica clásica se tiene

$$x' = x - vt, \quad y' = y, \quad z' = z, \quad t' = t$$

mientras que en la teoría de la relatividad se tiene

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad y' = y, \quad z' = z, \quad t' = \frac{-\frac{v}{c^2}x + t}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

El espacio de sucesos es en este caso de cuatro dimensiones. La contracción de las longitudes (sólo en la dirección del desplazamiento) y el retraso del tiempo en un sistema de referencia móvil se manifiestan en la misma razón $\sqrt{1 - \frac{v^2}{c^2}}$.

CAPÍTULO IX

NOCIÓN DE TENSORES

§ 1. Ejemplos de tensores

Anticipemos a la definición general de un tensor algunos ejemplos.

1. *Función lineal.*

Sea $f(x)$ una función lineal (§ 1 del capítulo VI) en un espacio vectorial R de n dimensiones. Tomemos en R una base $e_1, e_2, \ldots, e_n$ y sea

$$x = x^1e_1 + x^2e_2 + \ldots + x^ne_n$$

un vector cualquiera de R. (Ahora convendremos en escribir el índice de la coordenada *arriba*; la oportunidad de este convenio quedará justificada más tarde.) En este caso

$$f(x) = a_1x^1 + a_2x^2 + \ldots + a_nx^n,$$

donde $a_i = f(e_i)$.

Consideremos una base nueva $e'_1, e'_2, \ldots, e'_n$ y supongamos que los nuevos vectores básicos se obtienen de los antiguos mediante las fórmulas

$$e'_i = c_i^1e_1 + c_i^2e_2 + \ldots + c_i^ne_n = \sum_{k=1}^{n} c_i^ke_k. \tag{1}$$

En la matriz del cambio

$$C = \begin{bmatrix} c_1^1 & c_2^1 & \ldots & c_n^1 \\ c_1^2 & c_2^2 & \ldots & c_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ c_1^n & c_2^n & \ldots & c_n^n \end{bmatrix} \tag{2}$$

convendremos ahora en emplear el índice superior para el número de la fila y el inferior para el número de la columna. Supongamos que en la base nueva se tiene $x = \sum_{i=1}^{n} x'^i e'_i$; en este caso

$$f(x) = \sum_{i=1}^{n} a'_i x'^i, \tag{3}$$

donde

$$a'_i = f(e'_i) = f\left(\sum_{k=1}^{n} c_i^k e_k\right) = \sum_{k=1}^{n} c_i^k f(e_k) = \sum_{k=1}^{n} c_i^k a_k. \tag{4}$$

Por consiguiente, *una función lineal $f(x)$ se determina en toda base por una fila de n números $a_1, a_2, \ldots, a_n$ y al pasar a una base nueva estos números se transforman según las fórmulas* (4), *es decir, se transforman igual que los vectores básicos* (1).

Convengamos ahora, para abreviar las denotaciones, en que (regla de Einstein) si en una expresión cualquiera se repite dos veces, una vez arriba y otra vez abajo, un mismo índice, digamos i, esto significa que se realiza la sumación respecto a este índice (en los límites $i = 1, 2, \ldots, n$) con la particularidad de que el signo de la sumación $\sum_{i=1}^{n}$ se omite. Por ejemplo, tenemos por definición

$$c_i^k e_k = \sum_{k=1}^{n} c_i^k e_k, \quad c_i^k x'^i = \sum_{i=1}^{n} c_i^k x'^i, \quad b_l^{pq} a_r^l = \sum_{l=1}^{n} b_l^{pq} a_r^l, \quad \text{etc.}$$

Empleando esta denotación podemos escribir la igualdad (1) en la forma

$$e'_i = c_i^k e_k,$$

la igualdad (3) en la forma

$$f(x) = a'_i x'^i,$$

y la igualdad (4) en la forma

$$a'_i = c_i^k a_k.$$

Análogamente, si en una expresión aparecen dos pares, o un número mayor de pares, de índices iguales (y cada uno de estos índices aparece una vez arriba y una vez abajo), aceptaremos que respecto a estos índices se realiza la sumación y que todos estos índices recorren independientemente los valores 1, 2, ..., n. Por ejemplo,

$$a_m^{ik}b_{lp}^m = \sum_{i,\, m=1}^{n} a_m^{ik}b_{lp}^m, \quad a_{(pj)}^{ijk} = \sum_{i,\, j=1}^{n} a_{ipj}^{ijk}, \text{ etc.}$$

2. *Vector*

Dada una base $e_1, e_2, \ldots, e_n$, todo vector x queda representado por la fila $(x^1, x^2, \ldots, x^n)$ de n números que son sus coordenadas. En una base nueva $e_1', e_2', \ldots, e_n'$ este mismo vector se representa por otra fila $(x'^1, x'^2, \ldots, x'^n)$ y si (2) es la matriz del cambio de la primera base por la segunda, tenemos, por lo visto en el § 4 del capítulo II,

$$x^k = c_j^k x'^j. \tag{5}$$

Esta es la expresión de las coordenadas antiguas mediante las nuevas. Busquemos de aquí las expresiones de las coordenadas nuevas x'^j mediante las antiguas x^j. Sea $C^{-1} = [b_k^i]$ la matriz inversa de la matriz C. De la igualdad

$$CC^{-1} = C^{-1}C = E$$

resulta que

$$c_k^i b_j^k = b_k^i c_j^k = \begin{cases} 1, & \text{si} \quad i = j, \\ 0, & \text{si} \quad i \neq j. \end{cases}$$

Tomando

$$\delta_j^i = \begin{cases} 1, & \text{si} \quad i = j, \\ 0, & \text{si} \quad i \neq j, \end{cases}$$

tenemos

$$c_k^i b_j^k = b_k^i c_j^k = \delta_j^i.$$

Multiplicando por b_k^i ambos miembros de la igualdad (5) (y, naturalmente, sumando respecto a k), obtenemos

$$b_k^i x^k = b_k^i c_j^k x'^j = \delta_j^i x'^j = x'^i$$

(puesto que $\delta_j^i = 0$ para $i \neq j$ y $\delta_i^i = 1$) o

$$x'^i = b_k^i x^k.$$

Por consiguiente, *las coordenadas nuevas x'^i de un vector x se obtienen de sus coordenadas antiguas x^i mediante la matriz C^{-1}, inversa de la matriz del cambio C, con la particularidad de que los coeficientes de los desarrollos de las x'^i por las x^i forman las filas de la matriz C^{-1}.*

Los dos ejemplos considerados (el de función lineal y el de vector) tienen algo común que permite incluirlos en los márgenes de una definición general. Tanto una función lineal como un vector se determinan en toda base por n números $-a_1, a_2, \ldots, a_n$ y $x^1, x^2, \ldots, x^n$, respectivamente— y *al pasar a una base nueva estos números se transforman linealmente:* mediante la matriz C, es decir, *igual* que los vectores básicos, en el caso de una función lineal y en el caso de un vector mediante la matriz C^{-1} que es la *inversa* de la matriz C. Los coeficientes de una forma lineal (así como las coordenadas de un vector) representan un ejemplo de un *tensor*, entendiendo por un tensor un sistema de números dados en cualquier base que se transforman linealmente al cambiar una base por otra. La definición exacta de este concepto será dada más tarde; mientras tanto sólo añadiremos que ambos tensores considerados son *univalentes* ya que se determinan por sistemas de números $a_1, a_2, \ldots, a_n$ y $x^1, x^2, \ldots, x^n$ que dependen de un índice. Los coeficientes de una forma lineal, que al pasar a una base nueva se transforman igual que los vectores básicos, forman un tensor *covariante*, es decir, un tensor que se transforma igual que los vectores básicos. En cambio, las coordenadas de un vector ofrecen un ejemplo de un tensor *contravariante*.

Consideremos otros dos ejemplos.

3. *Función bilineal.*

Sea $A(x, y)$ una función bilineal (§ 2 del capítulo VI) dada en un espacio vectorial R de n dimensiones. Siendo $x = x^ie_i$ e $y = y^ke_k$ dos vectores cualesquiera de R, se tiene

$$A(x, y) = A(x^ie_i, \quad y^ke_k) = x^iy^kA(e_i, e_k) = a_{ik}x^iy^k,$$

donde $a_{ik} = A(e_i, e_k)$, es decir, una función bilineal $A(x, y)$ se representa en una base dada $e_1, e_2, \ldots, e_n$ por una forma bilineal $a_{ik}x^iy^k$ (¡sumación respecto a i y k!) en las coordenadas de los vectores x e y con los coeficientes a_{ik}.

Pasemos a una base nueva $e'_1, e'_2, \ldots, e'_n$ y sea (2) la matriz del cambio. Si $x = x'^ie'_i$ e $y = y'^ke'_k$, tenemos

$$A(x, y) = A(x'^ie'_i, \quad y'^ke'_k) = x'^iy'^kA(e'_i, e'_k) = a'_{ik}x'^iy'^k, \tag{6}$$

donde

$$a'_{ik} = A(e'_i, e'_k) = A(c_i^j e_j, \quad c_k^h e_h) = c_i^j c_k^h A(e_j, e_h) = c_i^j c_k^h a_{jh}. \qquad (7)$$

Por consiguiente, una función bilineal $A(x, y)$ se determina en toda base por un sistema de n^2 números a_{ik} que dependen de dos índices y que, al pasar a una base nueva, se transforman según la ley (7), es decir, respecto a cada uno de los índices, igual que los vectores básicos. Este es un ejemplo de un tensor de *valencia dos* (por depender de dos índices) *covariante* respecto a ambos índices (o doblemente covariante).

4. *Aplicación lineal.*

Todo operador lineal (aplicación lineal) $\mathcal{A}$ de un espacio R de n dimensiones se representa en una base dada $e_1, e_2, \ldots, e_n$ por una *matriz* $A = [a_k^i]$ (también aquí el índice superior corresponde al número de la fila y el índice inferior, al número de la columna). Al pasar a una base nueva $e'_1, e'_2, \ldots, e'_n$ con la matriz del cambio C, la matriz A se transforma en

$$C^{-1}AC$$

(§ 3 del capítulo III). Veamos cómo se expresan los elementos $a_k'^i$ de la matriz $C^{-1}AC$ mediante los elementos a_k^i de la matriz A. El elemento de la p-ésima fila y de la k-ésima columna de la matriz AC es igual a $a_j^p c_k^j$. El elemento de la i-ésima fila y de la k-ésima columna de la matriz $C^{-1}AC$ es $b_p^i a_j^p c_k^j$, es decir,

$$a_k'^i = b_p^i a_j^p c_k^j = c_k^j b_p^i a_j^p. \qquad (8)$$

Por consiguiente, una aplicación lineal $\mathcal{A}$ se determina en toda base por un sistema formado de n^2 números a_k^i que dependen de dos índices, el inferior y el superior, y que al pasar a una base nueva se transforman según la fórmula (8), es decir, igual que los vectores básicos en cuanto al índice inferior y con la matriz inversa —«contravariantemente» a los vectores básicos— en cuanto al índice superior. Tenemos, pues, otro ejemplo de un tensor de *valencia dos* (porque depende de dos índices) que en este caso resulta una vez *covariante* y una vez *contravariante* (tensor bivalente mixto).

Consideremos el tensor bivalente mixto cuyas coordenadas en una base fija $e_1, e_2, \ldots, e_n$ se determinan por las igualdades

$$\delta_j^i = \begin{cases} 1, & \text{si} \quad i = j, \\ 0, & \text{si} \quad i \neq j. \end{cases}$$

En cualquier base nueva $e'_1, e'_2, \ldots, e'_n$ tenemos

$$\delta_j^{\prime l} = c_j^p b_q^l \delta_p^q = c_j^p b_p^l = \delta_j^l.$$

Por consiguiente, *las coordenadas del tensor δ_j^l son las mismas en todos los sistemas de coordenadas*. Esto se debe a que en la base inicial $e_1, e_2, \ldots, e_n$ los elementos δ_j^l forman la matriz unidad y, por lo tanto, el tensor correspondiente determina la aplicación idéntica cuya matriz es la misma en todas las bases.

§ 2. Definición y propiedades elementales de un tensor

Supongamos que en toda base de un espacio vectorial R de n dimensiones está dado un sistema de n^{p+q} números $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$ (denotados con q índices superiores y con p índices inferiores que toman independientemente los valores $1, 2, \ldots, n$) *y supongamos que al pasar a una base nueva con la matriz del cambio* (2) *estos números se transforman según la ley*

$$a_{k_1k_2\ldots k_p}^{\prime i_1i_2\ldots i_q} = c_{k_1}^{j_1} c_{k_2}^{j_2} \ldots c_{k_p}^{j_p} b_{h_1}^{i_1} b_{h_2}^{i_2} \ldots b_{h_q}^{i_q} a_{j_1j_2\ldots j_p}^{h_1h_2\ldots h_q}. \tag{9}$$

*Diremos entonces que se tiene un **tensor** $(p+q)$**-valente** p **veces covariante** y q **veces contravariante**. Los números $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$ se denominan **coordenadas** del tensor.*

Un *escalar*, es decir, una magnitud que tiene el mismo valor en todos los sistemas de coordenadas, puede ser considerado como un tensor de valencia *nula*.

Si las coordenadas de dos tensores de la misma estructura (es decir, que tienen igual número de índices covariantes y de índices contravariantes) *coinciden en una base, está claro que coinciden en todas las demás* (y, por consiguiente, estos tensores son iguales) ya que al pasar a una base nueva las coordenadas de ambos tensores se transforman igualmente. Por ello, *para definir un tensor de una estructura dada es suficiente dar sus coordenadas en un sistema de coordenadas*. Esto se puede hacer sin limitación alguna: para las coordenadas de un tensor en una base dada se pueden tomar números absolutamente arbitrarios. Efectivamente, supongamos que en una base $e_1, e_2, \ldots, e_n$ se han tomado arbitrariamente n^{p+q} números $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$. En otra base cualquiera las coordenadas del tensor correspondientes se determinan según la fórmula (9) y resta sólo demostrar que *al pasar de una base cualquiera $e'_1, e'_2, \ldots, e'_n$ a otra base* *cualquiera* $e_1'', e_2'', \ldots, e_n''$ *las coordenadas del tensor obtenido se transforman también según la fórmula* (9).

Demostremos esto en el caso de un tensor a_k^{ij} de valencia tres. Supongamos que al pasar de la base $e_1, e_2, \ldots, e_n$ a la base $e_1', e_2', \ldots, e_n'$ se tiene

$$e_j' = c_j^m e_m \tag{10}$$

y que al cambiar la base $e_1, e_2, \ldots, e_n$ por la base $e_1'', e_2'', \ldots, e_n''$ se tiene

$$e_i'' = \bar{c}_i^k e_k. \tag{11}$$

De la igualdad (10) resulta

$$b_k^j e_j' = b_k^j \bar{c}_j^m e_m = \delta_k^m e_m = e_k \tag{12}$$

y de la igualdad (11)

$$\bar{b}_m^i e_i'' = \bar{b}_m^i c_i^k e_k = \delta_m^k e_k = e_m. \tag{13}$$

Aquí $[b_k^j]$ es la matriz inversa de la matriz $[c_k^j]$ y $[\bar{b}_k^j]$ es la matriz inversa de la matriz $[\bar{c}_k^j]$. De las igualdades (11), (12), (10) y (13) se deduce que

$$e_i'' = \bar{c}_i^k e_k = \bar{c}_i^k b_k^j e_j' = d_i^j e_j'$$

y

$$e_j' = c_j^m e_m = c_j^m \bar{b}_m^i e_i'' = f_j^i e_i''.$$

Por consiguiente, la matriz del cambio de la base $e_1', e_2', \ldots, e_n'$ por la base $e_1'', e_2'', \ldots, e_n''$ es la matriz

$$[d_i^j] = [\bar{c}_i^k b_k^j]$$

y la inversa de ésta es la matriz

$$[f_j^i] = [c_j^m \bar{b}_m^i].$$

Tenemos

$$a_r'^{pq} = c_r^u b_s^p b_t^q a_u^{st} \tag{14}$$

y

$$a_k''^{ij} = \bar{c}_k^h \bar{b}_m^i \bar{b}_l^j a_h^{ml}. \tag{15}$$

De la igualdad (14) resulta

$$b_h^r c_p^m c_q^l a_r'^{pq} = b_h^r c_p^m c_q^l c_r^u b_s^p b_t^q a_u^{st} =$$
$$= (b_h^r c_r^u)(c_p^m b_s^p)(c_q^l b_t^q) a_u^{st} = \delta_h^u \delta_s^m \delta_t^l a_u^{st} = a_h^{ml}.$$

Introduciendo este valor de a_h^{ml} en la igualdad (15), obtenemos

$$a_k^{\prime\prime ij} = \bar{c}_k^h \bar{b}_m^i \bar{b}_l^j b_h^r c_p^m c_q^l a_r^{\prime pq} =$$

$$= (\bar{c}_k^h b_h^r)(\bar{b}_m^i c_p^m)(\bar{b}_l^j c_q^l) a_r^{\prime pq} = d_k^r f_p^i f_q^j a_r^{\prime pq};$$

luego, las fórmulas de transformación de las coordenadas del tensor al cambiar la base $e_1', e_2', \ldots, e_n'$ por la base $e_1'', e_2'', \ldots, e_n''$ tienen precisamente la estructura que se requiere. En el caso general la demostración es análoga. De aquí se deduce que, por ejemplo, un tensor x^i contravariante de una valencia puede ser considerado como el conjunto de las coordenadas de un *vector*. Efectivamente, si en una base cualquiera $e_1, e_2, \ldots, e_n$ tomamos el vector que tiene las coordenadas $x^1, x^2, \ldots, x^n$, las coordenadas de este vector y del tensor dado coincidirán en todas las demás bases. Igualmente, todo tensor doblemente covariante a_{ij} puede ser interpretado como el conjunto de los coeficientes de una *forma bilineal* y todo vector bivalente mixto a_j^i como el conjunto de los elementos de la matriz de una *aplicación lineal*.

Si las coordenadas de un tensor no varían al transponer dos cualesquiera de los índices del conjunto dado de los índices $i, j, \ldots, m$ (siendo todos estos índices solamente superiores o solamente inferiores), *se dice que el tensor es* ***simétrico*** *respecto a estos índices.* Como ejemplo de un tensor simétrico podemos señalar el conjunto de los coeficientes de una forma bilineal simétrica.

La propiedad de un tensor de ser simétrico no depende de cómo se escoja la base. Consideremos, por ejemplo, un tensor a_{ij}^k de valencia tres y supongamos que en una base $e_1, e_2, \ldots, e_n$ se tiene

$$a_{ji}^k = a_{ij}^k.$$

En la base $e_1', e_2', \ldots, e_n'$ tendremos

$$a_{ij}^{\prime k} = c_i^s c_j^t b_r^k a_{st}^r$$

y

$$a_{ji}^{\prime k} = c_j^s c_i^t b_r^k a_{st}^r.$$

Reemplazando aquí a_{st}^r por a_{ts}^r, encontraremos que

$$a_{ji}^{\prime k} = c_j^s c_i^t b_r^k a_{ts}^r.$$

Pero la suma no depende de cómo se denote el índice de sumación; por ello, reemplazando s por t y t por s, resulta

$$a_{ji}^{\prime k} = c_j^t c_i^s b_r^k a_{st}^r = a_{ij}^{\prime k}.$$

Un tensor se llama ***antisimétrico*** *respecto a unos índices dados* $i, j, \ldots, m$ (solamente superiores o solamente inferiores) *si sus coordenadas* $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$ *cambian de signo al realizar cualquier trasposición de los índices dados* (y conservan su valor absoluto). *La propiedad de un tensor de ser antisimétrico respecto a un grupo dado de índices no depende de cómo se escoja la base.*

Como ejemplo de un tensor antisimétrico podemos señalar una forma bilineal antisimétrica cualquiera.

§ 3. Operaciones con tensores

1. ***Adición.*** Supongamos que se tienen dos tensores de una *misma estructura:* $a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$ y $\bar{a}_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}$. La *suma* de estos tensores se define en cualquier sistema de coordenadas por la igualdad

$$d_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q} = a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q} + \bar{a}_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q}.$$

Es fácil ver que *la suma de dos tensores es un tensor de la misma estructura.*

2. ***Multiplicación.*** Sean dados dos tensores

$$a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q} \text{ y } \bar{a}_{j_1j_2\ldots j_s}^{m_1m_2\ldots m_r}$$

de *cualesquiera estructuras*. El *producto* de estos tensores en cualquier base se define como el conjunto de los $n^{p+q+s+r}$ números

$$f_{k_1k_2\ldots k_p j_1j_2\ldots j_s}^{i_1i_2\ldots i_q m_1m_2\ldots m_r} = a_{k_1k_2\ldots k_p}^{i_1i_2\ldots i_q} \bar{a}_{j_1j_2\ldots j_s}^{m_1m_2\ldots m_r}.$$

Demostremos que *el producto de dos tensores es también un tensor* (en nuestro caso, $(p+q+r+s)$-valente, $p+s$ veces covariante y $q+r$ veces contravariante). Efectivamente, en una base nueva tenemos

$$a'^{i_1i_2\ldots i_q}_{k_1k_2\ldots k_p} = c_{k_1}^{t_1}c_{k_2}^{t_2}\ldots c_{k_p}^{t_p}b_{l_1}^{i_1}b_{l_2}^{i_2}\ldots b_{l_q}^{i_q}a_{t_1t_2\ldots t_p}^{l_1l_2\ldots l_q}$$

y

$$\bar{a}'^{m_1m_2\ldots m_r}_{j_1j_2\ldots j_s} = c_{j_1}^{h_1}c_{j_2}^{h_2}\ldots c_{j_s}^{h_s}b_{g_1}^{m_1}b_{g_2}^{m_2}\ldots b_{g_r}^{m_r}\bar{a}_{h_1h_2\ldots h_s}^{g_1g_2\ldots g_r};$$

por esto

$$\begin{aligned} f'^{i_1i_2\ldots i_q m_1m_2\ldots m_r}_{k_1k_2\ldots k_p j_1j_2\ldots j_s} &= a'^{i_1i_2\ldots i_q}_{k_1k_2\ldots k_p}\bar{a}'^{m_1m_2\ldots m_r}_{j_1j_2\ldots j_s} = \\ &= c_{k_1}^{t_1}c_{k_2}^{t_2}\ldots c_{k_p}^{t_p}c_{j_1}^{h_1}c_{j_2}^{h_2}\ldots c_{j_s}^{h_s}b_{l_1}^{i_1}b_{l_2}^{i_2}\ldots b_{l_q}^{i_q}b_{g_1}^{m_1}b_{g_2}^{m_2}\ldots b_{g_r}^{m_r}\times \\ &\times a_{t_1t_2\ldots t_p}^{l_1l_2\ldots l_q}\bar{a}_{h_1h_2\ldots h_s}^{g_1g_2\ldots g_r} = c_{k_1}^{t_1}\ldots c_{k_p}^{t_p}c_{j_1}^{h_1}\ldots c_{j_s}^{h_s}b_{l_1}^{i_1}\ldots \\ &\ldots b_{l_q}^{i_q}b_{g_1}^{m_1}\ldots b_{g_r}^{m_r}f_{t_1t_2\ldots t_p h_1h_2\ldots h_s}^{l_1l_2\ldots l_q g_1g_2\ldots g_r}. \end{aligned}$$

La operación de multiplicación de tensores no es conmutativa. Consideremos, por ejemplo, el producto de dos tensores univalentes covariantes para $n = 2$. Sean a_1, a_2 las coordenadas de uno y b_1, b_2 las coordenadas del otro. El producto $c_{ik} = a_i b_k$ es un tensor doblemente covariante de coordenadas

$$c_{11} = a_1 b_1, \quad c_{12} = a_1 b_2, \quad c_{21} = a_2 b_1 \quad \text{y} \quad c_{22} = a_2 b_2.$$

El producto de estos mismos tensores tomados en el orden contrario es un tensor $d_{ik} = b_i a_k$ doblemente covariante de coordenadas

$$d_{11} = b_1 a_1, \quad d_{12} = b_1 a_2, \quad d_{21} = b_2 a_1 \quad \text{y} \quad d_{22} = b_2 a_2.$$

Este tensor difiere, en general, del primero.

Puesto que un *escalar*, es decir, una magnitud que tiene el mismo valor en todos los sistemas de coordenadas, es un tensor de valencia nula, resulta que al multiplicar un tensor por un escalar (es decir, al multiplicar por este escalar todas las coordenadas del tensor) se obtiene un tensor de la misma estructura.

La *sustracción* de tensores de una misma estructura se reduce a la multiplicación del sustraendo por -1 y a la adición (y se obtiene, obviamente, un tensor de esta misma estructura).

3. Contracción de un tensor.

Esta operacian es específica para los tensores y se define del modo siguiente. Sea dado, por ejemplo, el tensor a^{ij}_{mpq}. Tomemos en él dos índices cualesquiera, digamos j y p *(uno arriba y el otro abajo)*, escojamos entre todas las coordenadas del tensor aquellas en las cuales estos índices coinciden y sumemos todas estas coordenadas. Obtendremos

$$a^{ij}_{mjq} = a^{i1}_{m1q} + a^{i2}_{m2q} + \ldots + a^{in}_{mnq} = b^{i}_{mq}$$

que *es la contracción del tensor a^{ij}_{mpq} respecto a los índices j y p.*

Por ejemplo, el tensor a^{i}_{jk} tiene para $n = 2$ ocho coordenadas a^{1}_{11}, a^{1}_{12}, a^{1}_{21}, a^{1}_{22}, a^{2}_{11}, a^{2}_{12}, a^{2}_{21} y a^{2}_{22}. Contrayendo este tensor respecto a los índices i y j, obtenemos

$$a^{1}_{1k} + a^{2}_{2k} = b_k$$

o detalladamente: $b_1 = a^{1}_{11} + a^{2}_{21}$, $\quad b_2 = a^{1}_{12} + a^{2}_{22}$.

Contrayendo este mismo tensor respecto a los índices i y k, obtenemos $a^{i}_{ji} = c_j$, es decir,

$$c_1 = a^{1}_{11} + a^{2}_{12}, \quad c_2 = a^{1}_{21} + a^{2}_{22}.$$

Consideremos un *ejemplo* más. Contrayendo el tensor a_{klm}^{ij} (para $n = 2$) respecto a los índices j y l, obtenemos ocho valores $b_{km}^{i} = a_{kjm}^{ij}$ o más detalladamente

$$b_{11}^{1} = a_{111}^{11} + a_{121}^{12}, \quad b_{12}^{1} = a_{112}^{11} + a_{122}^{12},$$
$$b_{21}^{1} = a_{211}^{11} + a_{221}^{12}, \quad b_{22}^{1} = a_{212}^{11} + a_{222}^{12},$$
$$b_{11}^{2} = a_{111}^{21} + a_{121}^{22}, \quad b_{12}^{2} = a_{112}^{21} + a_{122}^{22},$$
$$b_{21}^{2} = a_{211}^{21} + a_{221}^{22}, \quad b_{22}^{2} = a_{212}^{21} + a_{222}^{22}.$$

Demostremos que *la contracción de un tensor conduce de nuevo a un tensor que tiene, en comparación con el tensor inicial, un índice inferior y un índice superior de menos.* Realicemos, por ejemplo, la contracción del tensor a_{mpq}^{ij} respecto a los índices j y p. Sea $a_{mjq}^{ij} = \bar{a}_{mq}^{i}$. En una base nueva las coordenadas del tensor inicial son

$$a_{mpq}^{\prime ij} = c_m^r c_p^s c_q^t b_g^i b_h^j a_{rst}^{gh}.$$

Escogiendo las coordenadas para las cuales $j = p$ y sumándolas respecto a $j = 1, 2, \ldots, n$, obtenemos

$$\bar{a}_{mq}^{\prime i} = a_{mjq}^{\prime ij} = c_m^r c_j^s c_q^t b_g^i b_h^j a_{rst}^{gh}.$$

Pero $c_j^s b_h^j = \delta_h^s$ y $\delta_h^s a_{rst}^{gh} = a_{rht}^{gh} = \bar{a}_{rt}^{g}$. Luego,

$$\bar{a}_{mq}^{\prime i} = c_m^r c_q^t b_g^i \bar{a}_{rt}^{g}.$$

En el caso general la demostración es análoga.

La operación de contracción puede ser aplicada a un tensor varias veces. Por ejemplo, contrayendo el tensor a_{pqr}^{ijk} (para $n = 2$) respecto a los índices i y q y respecto a los índices k y p, obtenemos el tensor $d_r^j = a_{kir}^{ijk}$ o más detalladamente:

$$d_1^1 = a_{111}^{111} + a_{121}^{211} + a_{211}^{112} + a_{221}^{212},$$
$$d_2^1 = a_{112}^{111} + a_{122}^{211} + a_{212}^{112} + a_{222}^{212},$$
$$d_1^2 = a_{111}^{121} + a_{121}^{221} + a_{211}^{122} + a_{221}^{222},$$
$$d_2^2 = a_{112}^{121} + a_{122}^{221} + a_{212}^{122} + a_{222}^{222}.$$

Contrayendo p veces un tensor p veces covariante y p veces contravariante, obtendremos, obviamente, un escalar (invariante), es decir, una magnitud que no depende de cómo se escoja la base. En esto consiste uno de los métodos para obtener invariantes numéricos. Por ejemplo, contrayendo el tensor a_j^i que determina una aplicación lineal $\mathcal{A}$, obtenemos el invariante a_i^i llamado *traza* de la aplicación (a_i^i es la suma de los elementos de la diagonal principal de la matriz A; en el § 7 del capítulo III hemos demostrado que esta suma es invariante: a_i^i es el coeficiente de λ^{n-1} en el polinomio característico da la aplicación $\mathcal{A}$).

Frecuentemente la operación de contacción se aplica al producto de dos tensores tomando los índices en *diferentes factores.* Si el producto de los tensores α_j^i y β_k^h se contrae respecto a los índices j y h, diremos simplemente que *los tensores* α_j^i *y* β_k^h *se contraen respecto a los índices* j y h. Por ejemplo, contrayendo un tensor a_i que define una función lineal $f(x)$) con un vector $x = (x^1, x^2, \ldots, x^n)$, se obtiene el escalar $a_i x^i$ igual, obviamente, a $f(x)$.

Contrayendo dos veces un tensor a_{ij} —que determina una función bilineal $A(x, y)$— con un par de vectores $x = (x^1, x^2, \ldots, x^n)$ e $y = (y^1, y^2, \ldots, y^n)$, obtenemos el escalar $a_{ij}x^i y^j$ igual al valor de la función $A(x, y)$ para los vectores dados x e y.

Contrayendo un tensor a_j^i —que determina una aplicación lineal $\mathcal{A}$— con un vector $x = (x^1, x^2, \ldots, x^n)$, obtenemos el tensor contravariante $y^i = a_j^i x^j$ que, según hemos visto en el § 1 del capítulo III, es el vector transformado $\mathcal{A}x$ (la imagen del vector x en la aplicación $\mathcal{A}$) dado por sus coordenadas y^i.

Sean α_j^i y β_k^h dos tensores que determinan las aplicaciones lineales $\mathcal{A}$ y $\mathcal{B}$ respectivamente. Contrayéndolos respecto a los índices j y h, obtenemos el tensor bivalente mixto $d_k^i = \alpha_j^i \beta_k^j$ que determina, por consiguiente, una aplicación lineal $\mathcal{D}$. Es fácil ver que la aplicación $\mathcal{D}$ es igual al producto $\mathcal{A}\mathcal{B}$ de las aplicaciones $\mathcal{A}$ y $\mathcal{B}$ (en el sentido del § 2 del capítulo III). La contracción $\alpha_j^i \beta_i^h$ de estos mismos tensores respecto a los índices i y k corresponde al producto $\mathcal{B}\mathcal{A}$ de estas mismas aplicaciones tomadas en el orden contrario.

4. Simetrización y alternación de un tensor.

Sea $A_{i_1 i_2 \ldots i_k}$ un tensor cualquiera en el que se han escogido unos k índices $i_1, i_2, \ldots, i_k$, todos ellos solamente superiores o solamente inferiores. El tensor

$$A_{(i_1 i_2 \ldots i_k)} = \sum_{(i_1, i_2, \ldots, i_k)} A_{i_1 i_2 \ldots i_k},$$

donde la suma se extiende a todas las permutaciones posibles $i_1, i_2, \ldots, i_k$ de los índices escogidos, es, obviamente, simétrico mientras que el tensor

$$A_{[i_1 i_2 \ldots i_k]} = \sum_{(i_1, i_2, \ldots, i_k)} (-1)^{[i_1, i_2, \ldots, i_k]} A_{i_1 i_2 \ldots i_k}$$

es antisimétrico. Las operaciones consistentes en pasar de un tensor dado $A_{(i_1 i_2 \ldots i_k)}$ a los tensores $A_{(i_1 i_2 \ldots i_k)}$ y $A_{[i_1 i_2 \ldots i_k]}$ se denominan *simetrización* y *alternación*, respectivamente.

Por ejemplo, en la pág. 169 las funciones bilineales $B(x, y)$ y $C(x, y)$ han sido obtenidas simetrizando y alternando, respectivamente, la función bilineal $A(x, y)$.

§ 4. Tensores en un espacio euclídeo

Supongamos ahora que R es un espacio n-dimensional *euclídeo*. Por supuesto, en este caso se conserva todo cuanto hemos dicho acerca de los tensores en un espacio lineal cualquiera. Pero en el espacio euclídeo los tensores poseen algunas propiedades nuevas.

Para cualesquiera dos vectores x e y de un espacio euclídeo R está definido el producto escalar (x, y) que es una función bilineal simétrica y que en una base dada $e_1, e_2, \ldots, e_n$ queda representado por la forma bilineal simétrica

$$(x, y) = g_{ik} x^i y^k,$$

donde $g_{ik} = (e_i, e_k)$. Las magnitudes g_{ik}, tomadas en todos los sistemas de coordenadas, forman como hemos visto en el § 1 un tensor doblemente covariante que se denomina *tensor métrico del espacio R*.

La contracción del tensor métrico g_{ik} con un vector $x = (x^1, x^2, \ldots, x^n)$

$$x_i = g_{ik} x^k \tag{16}$$

es un tensor univalente covariante. Los números x_i también caracterizan al vector x, es decir, en cierto sentido son también coordenadas de éste; se pueden llamar *coordenadas covariantes* del vector x a diferencia de sus coordenadas contravariantes x^i. Busquemos la interpretación geométrica de las coordenadas covariantes. Puesto que

$$x_i = g_{ik} x^k = (e_i, e_k)\, x^k = (e_i, e_k x^k) = (e_i, x),$$

resulta que *las coordenadas covariantes x_i son las proyecciones del vector x sobre los vectores básicos*. (Recordemos que las coordenadas contravariantes de un vector x son los coeficientes de su desarrollo

$$x = x^i e_i$$

respecto a la base $e_1, e_2, \ldots, e_n$).

En una base ortonormal se tiene

$$g_{ik} = (e_i, e_k) = \begin{cases} 1, & \text{si } i = k, \\ 0, & \text{si } i \neq k, \end{cases}$$

y, por consiguiente, $x_i = x^i$, es decir, las coordenadas covariantes y contravariantes de un vector coinciden.

La contracción doble $g_{ik}x^iy^k$ del tensor métrico g_{ik} con los vectores $x = (x^1, x^2, \ldots, x^n)$ e $y = (y^1, y^2, \ldots, y^n)$ es el *producto escalar* (x, y); la contracción doble $g_{ik}x^ix^k$ del tensor con el vector x es el *cuadrado escalar* (x, x) del vector x.

El determinante $|g_{ik}|$ de la matriz $[g_{ik}]$ es diferente de cero. Efectivamente, al pasar a una base nueva el rango de la matriz de una forma bilineal, en particular, el rango de la matriz $[g_{ik}]$, no varía. Pero en una base ortonormal $[g_{ik}]$ es la matriz unidad y su determinante es igual a la unidad; luego, el determinante de la matriz $[g_{ik}]$ es diferente de cero también en todas las demás bases.

Sea $[g^{ik}]$ la matriz *inversa de la matriz* $[g_{ik}]$ en una base determinada $e_1, e_2, \ldots, e_n$. En este caso tenemos $g^{ik}g_{kj} = \delta^i_j$ para todos los $i, j = 1, 2, \ldots, n$. Formemos el tensor *doblemente contravariante* cuyas coordenadas en la base $e_1, e_2, \ldots, e_n$ son iguales a g^{ik}; las coordenadas de este tensor en todas las demás bases se determinarán entonces por la fórmula (9). En toda base nueva $e'_1, e'_2, \ldots, e'_n$ tendremos (tomando en consideración el carácter tensorial de las operaciones de multiplicación y de contracción)

$$g'^{ik}\, g'_{kj} = \delta'^i_j = \delta^i_j,$$

es decir, *en todos los sistemas de coordenadas las coordenadas del tensor* g^{ik} *forman la matriz inversa de la matriz* $[g_{ik}]$. El tensor g^{ik} se denomina *tensor métrico contravariante*.

El paso de las coordenadas contravariantes de un vector a sus coordenadas covariantes según la fórmula (16) puede ser denominado *operación de descenso del índice*. Para *elevar el índice*, es decir, para pasar de las coordenadas covariantes de un vector a sus coordenadas contravariantes, multipliquemos por g^{ji} ambos miembros de la igualdad (16) (sumando, claro está, respecto a i); tendremos

$$g^{ji}x_i = g^{ji}g_{ik}x^k = \delta^j_k x^k = x^j.$$

En un espacio euclídeo la operación de descenso o de elevación del índice (que lleva el expresivo nombre de la *operación malabárica*) se puede aplicar a un tensor de cualquier estructura. Sea dado, por ejemplo, un tensor a_k^{ij} de valencia tres, una vez covariante y dos veces contravariante. Su contracción

$$g_{ih}a_k^{ij} = \alpha_{hk}^j \tag{17}$$

con el tensor métrico será un tensor dos veces covariante y una vez contravariante. La contracción

$$g_{pj}\alpha_{hk}^j = \beta_{phk} \tag{18}$$

será tres veces covariante y la contracción

$$g^{qk}a_k^{ij} = \gamma^{qij} \tag{19}$$

será, al contrario, un tensor tres veces contravariante. Si ambas bases $e_1, e_2, \ldots, e_n$ y $e_1', e_2', \ldots, e_n'$ son ortonormales, tenemos

$$g_{ik} = g^{ik} = \begin{cases} 1, & \text{si} \quad i = k, \\ 0, & \text{si} \quad i \neq k, \end{cases}$$

y las igualdades (17), (18) y (19) dan sucesivamente

$$a_k^{hj} = \alpha_{hk}^j, \tag{17'}$$

$$\alpha_{hk}^p = \beta_{phk}, \tag{18'}$$

$$a_q^{ij} = \gamma^{qij}, \tag{19'}$$

de donde cambiando la denotación del índice de la sumación obtenemos

$$a_k^{ij} = \alpha_{ik}^j = \beta_{jik} = \gamma_{kij}.$$

En este caso los índices covariantes y contravariantes se portan, al pasar a una base nueva, igualmente y *la ley de la transformación del tensor queda determinada exclusivamente por su valencia.*

Este último resultado puede explicarse también de la forma siguiente. Si ambas bases $e_1, e_2, \ldots, e_n$ y $e_1', e_2', \ldots, e_n'$ son ortonormales, la matriz C del cambio de la primera base por la segunda es ortogonal, es decir,

$$C^{-1} = C'.$$

Pero en este caso resulta $b_k^i = c_i^k$ y en la fórmula (9) para la transformación, por ejemplo, del tensor a_k^{ij}

$$a_k'^{ij} = c_k^p b_q^i b_r^j a_p^{qr},$$

podemos sustituir uno o ambos b por c o, al contrario, sustituir c por b. Por ejemplo, sustituyendo b^j_q por c^q_j, obtenemos

$$a'^{ij}_k = \sum_{q=1}^{n} c^p_k c^q_i b^j_r a^{qr}_p.$$

Aquí hemos tenido que emplear el signo de la suma $\sum_{q=1}^{n}$, porque el índice q, respecto al cual se realiza la sumación, aparece dos veces arriba. Poniendo $a^{qr}_p = \alpha^r_{qp}$, encontramos la ley de la transformación de α^j_{ik} en la forma

$$\alpha'^j_{ik} = c^p_k c^q_i b^j_r \alpha^r_{qp} = c^q_i c^p_k b^j_r \alpha^r_{qp},$$

es decir, el tensor α^j_{ik} es dos veces covariante y una vez contravariante. Pero sus coordenadas en ambas bases son iguales a las coordenadas respectivas del tensor a^{ij}_k una vez covariante y dos veces contravariante.

Por ejemplo, hemos visto (§ 6 del capítulo VI) que al cambiar una base ortonormal por otra la matriz de una forma bilineal (que es un tensor dos veces covariante) y la matriz de una aplicación lineal (que es un tensor una vez covariante y una vez contravariante) se transforman idénticamente.

CAPÍTULO X

CONCEPTOS PRINCIPALES DE LA TEORIA DE GRUPOS

§ 1. Ejemplos de grupos. Definición de un grupo

Consideremos *el conjunto de todos los números enteros*. Al sumar dos números enteros obtenemos de nuevo un número entero. Si uno de los sumandos es igual al número (entero) 0, la suma es igual al segundo sumando: $a+0 = a$; para todo número entero a existe el número (entero) $-a$, llamado opuesto de a, tal que su suma con a da 0. La operación de adición de números es, como se sabe, conmutativa ($a+b = b+a$ para dos números a y b cualesquiera) y asociativa ($(a+b)+c = a+(b+c)$ para tres números a, b y c cualesquiera).

En el conjunto de todos los números enteros, el subconjunto de todos los números *divisibles por un número dado* k posee estas mismas propiedades: es cerrado respecto a la operación de adición (la suma de dos números cualesquiera divisibles por k es divisible por k); contiene el 0 (el cero es divisible por un número cualquiera); si a es divisible por k, también $-a$ es divisible por k.

Propiedades análogas se cumplen en *el conjunto de todos los números racionales*, en el conjunto de *todos los números reales* y en el conjunto de *todos los números complejos*: cada uno de estos conjuntos es cerrado respecto a la operación de adición; el cero es a la vez un número racional, real y complejo; para todo número complejo a existe su opuesto $-a$ que es real, si a es real, y es racional, si a es racional.

En el conjunto de los números complejos (y, como consecuencia, también en el conjunto de los números reales y en el conjunto de los números racionales) la operación de adición es conmutativa y asociativa. Todos estos casos son ejemplos de «***grupos respecto a la adición***».

Consideremos ahora el conjunto de todos *los números reales*

diferentes de cero. El producto de dos números así es de nuevo un número diferente de cero; el producto de cualquier número a por el número 1 (real y diferente de cero) es igual a a; para todo número real a (¡diferente de cero!) existe el número inverso real a^{-1} (también diferente de cero) tal que su producto por a es igual a 1.

Propiedades análogas las posee el conjunto de todos *los números racionales diferentes de cero*, el conjunto de todos *los números reales positivos* o el conjunto de todos *los números racionales positivos*, así como el conjunto de todos *los números complejos diferentes de cero* o el conjunto de todos *los números complejos de módulo igual a* 1. Cada uno de estos conjuntos es cerrado respecto a la operación de multiplicación, todos ellos contienen la unidad y para todo elemento de estos conjuntos existe el elemento inverso que pertenece al mismo conjunto. La multiplicación de los números complejos (y, por consiguiente, también de los números reales y racionales) es conmutativa ($ab = ba$ para todos los a y b) y asociativa ($(ab)c = a(bc)$ para todos los a, b y c).

Todos estos casos son ejemplos de «***grupos respecto a la multiplicación***». Podemos indicar un ejemplo menos esperado: forman un grupo respecto a la multiplicación, por ejemplo, los números 1 y -1. Dicho sea de paso, el conjunto formado solamente por el número 1 (o por el número 0) también constituye un grupo respecto a la multiplicación (a la adición, respectivamente).

Es obvio que los números complejos 1, i, -1 y $-i$ también forman un grupo respecto a la multiplicación.

Pero, además de los números, se pueden sumar también los *vectores* de un espacio lineal R y esta operación de adición verifica las mismas leyes que la adición de los números: es conmutativa y asociativa, en R existe el elemento 0 tal que $x+0 = x$ para todo $x \in R$ y cualquiera que sea el vector $x \in R$ existe el vector opuesto $-x$ tal que $x+(-x) = 0$.

También se pueden sumar *las matrices de una misma estructura*, es decir, las $[m\times n]$-matrices. Esta operación de adición es asociativa y conmutativa, existe la matriz nula que es la matriz formada por ceros solamente y para toda matriz $[a_{ik}]$ existe la matriz opuesta $[-a_{ik}]$ tal que $[a_{ik}]+[-a_{ik}]$ es la matriz nula. Si se consideran solamente las matrices de *elementos enteros* a_{ik}, la suma será una matriz del mismo tipo, los elementos de la matriz nula serán números enteros y para toda matriz de elementos enteros la opuesta será también una matriz de elementos enteros. Todos estos casos son ejemplos de ***grupos respecto a la adición.***

Por otro lado, además de los números, se pueden multiplicar, por ejemplo, *las matrices cuadradas no degeneradas de un mismo orden n* con elementos reales (solamente racionales o, al contrario, solamente complejos). El producto de dos matrices así será también una matriz no degenerada (teorema 3 del capítulo III) de elementos reales (racionales o complejos, respectivamente); la matriz unidad es no degenerada y para toda matriz no degenerada existe la matriz inversa (también no degenerada y también de elementos reales, racionales o complejos, respectivamente). El producto de matrices es asociativo pero *no es conmutativo*. El conjunto de todas las matrices no degeneradas de orden n de elementos reales (racionales o complejos) ofrece un ejemplo de ***un grupo no conmutativo*** respecto a la multiplicación.

Demos ahora la definición general de un grupo.

Definición 1. *Se llama **grupo** a un conjunto G de elementos $a, b, \ldots$ para los que está definida una operación algebraica (denominada comúnmente **multiplicación** o **adición**) que a todo par ordenado de elementos a, b de G pone en correspondencia un tercer elemento $c = a \circ b$ con la particularidad de que se cumplen las condiciones siguientes:*

1. *Esta operación es asociativa: para cualesquiera tres elementos a, b y c de G se tiene*

$$(a \circ b) \circ c = a \circ (b \circ c);$$

2. *En G existe el elemento «**neutro**» e tal que*

$$a \circ e = e \circ a = a$$

para todo a de G;

3. *Para todo elemento a de G existe el elemento «**inverso**» a^{-1} tal que*

$$a \circ a^{-1} = a^{-1} \circ a = e.$$

*El grupo se llama **conmutativo** o **abeliano** si se cumple, además, la ley conmutativa:*

4. *Para cualesquiera dos elementos $a, b \in G$ se tiene*

$$a \circ b = b \circ a.$$

*El grupo se llama **finito** si consta de un número finito de elementos. El número de elementos de un grupo finito se denomina **orden** del mismo. Todo grupo que no sea finito se llama **infinito**.*

En el caso en el que la «operación de grupo» $a \circ b$ se denomina *adición* y se indica por el signo $+$, el grupo G se llama *grupo respecto a la adición* o *grupo aditivo*. En este caso el «elemento neutro» e se indica frecuentemente con el símbolo 0 y se llama *cero*, mientras que el elemento inverso de a se designa por $-a$ y se denomina *opuesto de a*. En el caso en el que la «operación de grupo» se denomine *multiplicación*, en lugar de $a \circ b$ se escribe ab, el grupo se llama *grupo respecto a la multiplicación* o *grupo multiplicativo* y el elemento neutro e se denota a veces simplemente por 1 y se denomina *unidad*.

Basándose en la ley asociativa, se puede definir el producto (la suma) de tres o de un número mayor de elementos del grupo. Puesto que $(ab)c = a(bc)$, tiene sentido hablar simplemente del producto abc de tres elementos tomándolo, por definición, igual a $(ab)c = a(bc)$. Igual que en el caso de los espacios lineales, se puede demostrar que *el elemento unidad (el elemento nulo) de un grupo es único y que el elemento inverso (opuesto) de un elemento dado es único.*

Es fácil ver que en todo grupo (digamos, multiplicativo) *tienen solución única las ecuaciones*

$$ax = b$$

(cuya solución es, obviamente, $x = a^{-1}b$) y

$$ya = b$$

(para la que se tiene $y = ba^{-1}$).

§ 2. Grupos de transformaciones

Una clase importante de grupos forman *los grupos de transformaciones*. Sea M un conjunto cualquiera. *Se llama transformación del conjunto M a toda aplicación biyectiva P de este conjunto sobre sí mismo*. Esto significa que para todo elemento x de M está definida unívocamente su *imagen* $Px = x' \in M$ y que, además, todo elemento x' de M es la imagen de un elemento único x denominado *imagen recíproca* de x'.

La multiplicación de las transformaciones consiste en la realización consecutiva de las mismas, es decir, se toma por definición que

$$(PQ)\,x = P(Qx).$$

La multiplicación de las transformaciones es una operación asociativa —esto se puede demostrar igual que en el caso de las aplicaciones lineales (§ 2 del capítulo III)— pero no es, en general, *una operación conmutativa* (la propia multiplicación de las aplicaciones lineales es ya no conmutativa). El papel de la unidad lo desempeña en este caso *la transformación idéntica* E que a todo elemento x de M le pone en correspondencia este mismo elemento. Para toda transformación P del conjunto M existe *la transformación inversa* P^{-1} que a todo elemento x' de M pone en correspondencia su imagen recíproca x (única por hipótesis); es obvio que

$$PP^{-1} = P^{-1}P = E.$$

Si el conjunto M es finito y consta de n elementos, todas las aplicaciones biyectivas posibles de este conjunto sobre sí mismo se denominan *sustituciones*, el grupo de transformaciones correspondiente se indica por S_n y se denomina *grupo de sustituciones* de n elementos o *grupo simétrico de orden* n.

Consideremos más detalladamente *el grupo simétrico de orden tres* S_3, es decir, el grupo formado por todas las aplicaciones biyectivas del conjunto compuesto de tres elementos a, b y c sobre sí mismo. Con tres elementos se puede formar un total de seis permutaciones diferentes

abc, *acb*, *bac*, *bca*, *cab* y *cba*;

luego, el número de distintas sustituciones de este conjunto es también igual a seis. Resulta cómodo indicarlas por

$$P_1 = \begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix}, \quad P_2 = \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix}, \quad P_3 = \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix},$$

$$P_4 = \begin{pmatrix} a & b & c \\ b & a & c \end{pmatrix}, \quad P_5 = \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}, \quad P_6 = \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix},$$

donde, por ejemplo, $\begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}$ es aquella aplicación del conjunto a, b y c sobre sí mismo para la que $a \to b$ (a se transforma en b), $b \to c$ y $c \to a$. No se consideran distintas las sustituciones que difieren sólo en el orden de secuencia de las columnas como son, por ejemplo,

$$\begin{pmatrix} b & a & c \\ c & b & a \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}.$$

La multiplicación de las sustituciones consiste en la realización sucesiva de las mismas; por ello se tiene, por ejemplo ,

$$P_6P_2 = \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix} \begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix} = \begin{pmatrix} a & b & c \\ c & b & a \end{pmatrix} = P_3$$

ya que en el primer factor $a \to a$, en el segundo $a \to c$ y, por consiguiente, en el producto $a \to c$, etc. En esta multiplicación la unidad es *la sustitución idéntica*

$$P_1 = \begin{pmatrix} a & b & c \\ a & b & c \end{pmatrix}$$

y para toda sustitución existe *la inversa*

$$P_2^{-1} = P_2, P_3^{-1} = P_3, P_4^{-1} = P_4, P_5^{-1} = P_6 \quad \text{y} \quad P_6^{-1} = P_5.$$

Para obtener la sustitución inversa de una dada basta cambiar entre sí sus filas

$$\begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}^{-1} = \begin{pmatrix} b & c & a \\ a & b & c \end{pmatrix} = \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix}.$$

El grupo S_3 puede ser representado por la siguiente tabla de multiplicación

	P_1	P_2	P_3	P_4	P_5	P_6
P_1	P_1	P_2	P_3	P_4	P_5	P_6
P_2	P_2	P_1	P_5	P_6	P_3	P_4
P_3	P_3	P_6	P_1	P_5	P_4	P_2
P_4	P_4	P_5	P_6	P_1	P_2	P_3
P_5	P_5	P_4	P_2	P_3	P_6	P_1
P_6	P_6	P_3	P_4	P_2	P_1	P_5.

donde en la columna de la izquierda aparecen los segundos factores P_i, en la fila superior aparecen los primeros factores P_k y en el cruce de las filas y de las columnas correspondientes figuran los productos P_iP_k de estos factores. Estas tablas se llaman *tablas de Cayley*. Es fácil ver que cada uno de los elementos aparece sólo una vez en toda fila y en toda columna de la tabla de Cayley (ya que multiplicando a la izquierda por P_i^{-1} la igualdad $P_iP_j = P_iP_k$ se obtiene $P_j = P_k$ y de la igualdad $P_jP_i = P_kP_i$ resulta que $P_j = P_k$).

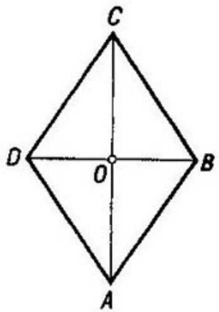


Fig. 28

El grupo S_3 *no es conmutativo*, ya que, por ejemplo,

$P_4P_5 = P_2$ y $P_5P_4 = P_3$

(la tabla de Cayley de este grupo no es simétrica respecto a la diagonal principal).

Otro ejemplo importante de un grupo de transformaciones es *el grupo de las rotaciones de un polígono regular de n lados*. Sea $A_1A_2\dots A_n$ un polígono regular y sea O su centro. Consideremos el conjunto de todas las rotaciones posibles del plano alrededor del punto O que hacen coincidir este polígono con sí mismo. Es obvio que existen n rotaciones de este tipo:

a_0, la rotación de ángulo 0 (la transformación idéntica),

a_1, la rotación de ángulo $\frac{2\pi}{n}$,

a_2, la rotación de ángulo $\frac{2\pi}{n}\cdot 2$,

. . . .

a_{n-1}, la rotación de ángulo $\frac{2\pi}{n}(n-1)$.

La multiplicación de las rotaciones consiste en la realización sucesiva de las mismas:

$a_k \cdot a_l = a_{k+l}$

aceptándose, claro está, que $a_n = a_0$ y que $a_{n+k} = a_k$. La rotación a_0 es el elemento unidad del grupo y $a_k^{-1} = a_{n-k}$.

Consideremos un ejemplo más: el grupo de superposiciones o *el grupo simétrico* de un rombo. Sea $ABCD$ un rombo (fig. 28).

Consideremos las siguientes transformaciones que sobreponen este rombo sobre sí mismo:

b_1, la transformación idéntica;
b_2, la simetría respecto a la diagonal AC;
b_3, la simetría respecto a la diagonal BD;
b_4, la simetría respecto al centro O.

El producto de dos cualesquiera de estas transformaciones es de nuevo una de ellas. Estas transformaciones forman un grupo con la siguiente tabla de Cayley

	b_1	b_2	b_3	b_4
b_1	b_1	b_2	b_3	b_4
b_2	b_2	b_1	b_4	b_3
b_3	b_3	b_4	b_1	b_2
b_4	b_4	b_3	b_2	b_1

§ 3. Subgrupo

Definición 2. *Se llama* ***subgrupo*** *de un grupo G a un conjunto G_1 de elementos del grupo G que constituye él mismo un grupo respecto a la operación definida en el grupo G.*

Así, el grupo aditivo de los números reales contiene el subgrupo de los números enteros y éste contiene los subgrupos de los números múltiplos de k cualquiera que sea k. El propio grupo de los números reales es un subgrupo del grupo de los números complejos.

El grupo multiplicativo de los números complejos diferentes de cero contiene el subgrupo de los números reales y éste contiene el subgrupo de los números racionales y el subgrupo de los números reales positivos.

Varios subgrupos interesantes contiene el grupo multiplicativo de las matrices no degeneradas de orden n (grupo lineal completo) formadas, por ejemplo, por elementos reales. Señalemos, en particular, el subgrupo de las matrices *ortogonales* y el subgrupo de las matrices *unimodulares* (es decir, de matrices de determinante igual a 1). También son subgrupos del grupo lineal completo el grupo de las matrices de *determinante igual* a ± 1, el grupo de las matrices de determinante *positivo*, el grupo de las matrices *diagonales*, el grupo de las matrices *escalares* (es decir, de matrices de tipo cE, donde c es un número cualquiera y E es la matriz unidad), el grupo de las matrices *triangulares* (es decir, de las matrices en las que son iguales a cero todos los elementos que aparecen por encima (por debajo) de la diagonal principal), etc.

Para comprobar que un subconjunto G_1 de un grupo G es un subgrupo de este último es necesario comprobar que el producto (la suma) de dos cualesquiera elementos de G_1 pertenece a G_1 y que si $a \in G_1$, también $a^{-1} \in G_1$. Pero esto resulta también suficiente, ya que la ley asociativa se cumplirá para los elementos de G_1 por ser válida en todo el grupo G y el elemento e pertenecerá a G_1 por ser igual al producto aa^{-1} (o a la suma $a+(-a)$). Está claro que *un subgrupo de un grupo conmutativo será siempre conmutativo.*

Es fácil concebir que *la intersección de dos subgrupos de un grupo G es un subgrupo de G.*

Así, en el grupo aditivo de los números enteros la intersección del subgrupo de los números múltiplos de tres y del subgrupo de los números múltiplos de dos es el subgrupo de los números múltiplos de seis.

Todo grupo tiene el subgrupo formado solamente por la unidad (el cero) y todo grupo es subgrupo de sí mismo. (Estos subgrupos se denominan *impropios.*)

§ 4. Isomorfismo de grupos

El grupo simétrico de tercer orden S_3 tiene tres subgrupos de segundo orden: $\{P_1, P_2\}$, $\{P_1, P_3\}$ y $\{P_1, P_4\}$ con las tablas de Cayley

	P_1	P_2
P_1	P_1	P_2
P_2	P_2	P_1

,

	P_1	P_3
P_1	P_1	P_3
P_3	P_3	P_1

,

	P_1	P_4
P_1	P_1	P_4
P_4	P_4	P_1

que considerados independientemente del grupo S_3 difieren uno del otro sólo en la designación de los elementos.

El grupo simétrico S_3 tiene un subgrupo A de tercer orden $\{P_1, P_5, P_6\}$ con la tabla de Cayley

	P_1	P_5	P_6
P_1	P_1	P_5	P_6
P_5	P_5	P_6	P_1
P_6	P_6	P_1	P_5

Comparémoslo con el grupo de rotaciones de un triángulo regular

	a_0	a_1	a_2
a_0	a_0	a_1	a_2
a_1	a_1	a_2	a_0
a_2	a_2	a_0	a_1

Estos grupos también difieren sólo en la denotación de los elementos. Se dice que estos grupos son *isomorfos*.

Definición 3. *Dos grupos G_1 y G_2 se llaman* ***isomorfos*** *si entre sus elementos se puede establecer una correspondencia biyectiva que conserve la operación de grupo, es decir, en la que para*

$$x_1, y_1 \in G_1 \quad \text{y} \quad x_2, y_2 \in G_2$$

y

$$x_1 \leftrightarrow y_1, \quad x_2 \leftrightarrow y_2,$$

se tenga

$$x_1 \circ y_1 \leftrightarrow x_2 \circ y_2.$$

Se puede demostrar que *todos los grupos de orden dos (así como todos los grupos de orden tres) son isomorfos entre sí*. Pero para el caso de orden cuatro existen ya *dos* grupos no isomorfos: *el grupo de rotaciones del cuadrado* y *el grupo simétrico del rombo*.

Todo grupo isomorfo al grupo de rotaciones de un polígono regular de n lados se llama ***grupo ciclio de orden n****;* este grupo está compuesto por las potencias de uno de sus elementos: $e, a, a^2, \ldots, a^{n-1}$. *Todo grupo isomorfo al grupo aditivo de los números enteros se llama* ***grupo cíclico infinito****;* este grupo también resulta generado por uno de sus elementos

$$2 = 1+1, \quad 3 = 1+1+1, \text{ etc.}$$

Observemos que las operaciones de grupos isomorfos pueden estar indicadas de distinta forma. Así, el grupo multiplicativo de los números positivos es isomorfo al grupo aditivo de los números reales. La correspondencia isomorfa entre estos grupos es la aplicación

$$a \leftrightarrow \ln a.$$

§ 5. Grupos de transformaciones de un plano

Consideremos ahora algunos grupos de transformaciones de un plano (euclídeo).

1. *Movimientos.*

Definición 4. *Se llama **movimiento** a toda transformación* (de un conjunto de puntos) *de un espacio euclídeo* (de un plano, en particular) *que conserva la distancia entre los puntos: si el punto A se transforma en A'* (en este caso escribiremos $A \to A'$) *y $B \to B'$, la longitud del segmento $A'B'$ es igual a la longitud del segmento AB.*

Estas transformaciones constituyen un *grupo:* si dos transformaciones no alteran las distancias entre los puntos, el producto de éstas posee la misma propiedad; la transformación idéntica que no varía la posición de ninguno de los puntos es un movimiento y la transformación inversa de un movimiento es, por supuesto, un movimiento.

En el grupo de los movimientos se destaca el subgrupo de *las traslaciones paralelas* que se define como sigue. Sea dado un vector a; se llama *traslación paralela de vector a* a la transformación T_a que a todo punto A del plano le pone en correspondencia el punto A' tal que el vector $\overline{AA'}$ es igual a a. Esta transformación es un movimiento: si $A \to A'$ y $B \to B'$, se tiene $\overline{AA'} = a$ y $\overline{BB'} = a$; luego, el cuadrilátero $AA'B'B$ es un paralelogramo y, por consiguiente, $A'B' = AB$ (fig. 29). El producto T_aT_b de dos traslaciones paralelas de vectores a y b es la traslación paralela T_{a+b} (de vector $a+b$) y la transformación inversa de la traslación paralela T_a es la traslación paralela T_{-a} (de vector $-a$).

Por consiguiente, *todas las traslaciones paralelas T de un plano forman un grupo* que es un subgrupo del grupo de los movimientos. De lo expuesto se deduce que este grupo es isomorfo al grupo (respecto a la adición) de los vectores de un plano. Si en la traslación paralela de vector $a = (a_1, a_2)$ el punto $A(x, y)$ se transforma en el punto $A'(x', y')$, resulta que el vector $\overline{AA'} = a$, es decir, que $(x'-x, y'-y) = (a_1, a_2)$ y, por consiguiente,

$$x'-x = a_1, \quad y'-y = a_2,$$

es decir,

$$x' = x+a_1,$$
$$y' = y+a_2.$$

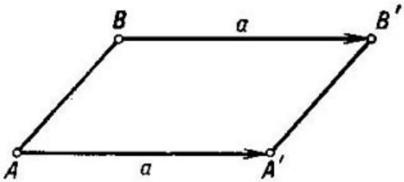


Fig. 29

Consideremos, por otro lado, *todos los movimientos posibles de un plano que conservan la posición de un punto determinado O;* es obvio que también constituyen un subgrupo del grupo de los movimientos del plano.

Sea H un movimiento cualquiera de punto fijo O. Indiquemos por $\mathscr{H}$ la correspondiente aplicación del espacio vectorial, es decir, la aplicación en la que el vector $\overline{OA}$ se transforma en el vector $\overline{OA'}$, donde A' es la imagen del punto A en la transformación H. Demostremos que $\mathscr{H}$ es una aplicación lineal.

1. Supongamos que $A \to A'$ y $B \to B'$; supongamos, además, que

$$\overline{OC} = \overline{OA} + \overline{OB}$$

y que $C \to C'$ (fig. 30, *a*). Tenemos $B'C' = BC$ (por ser H un movimiento), $BC = OA$ (porque $OBCA$ es un paralelogramo) y $OA = OA'$ (porque H es un movimiento). Luego, $B'C' = OA'$. Análogamente resulta

$$A'C' = AC = OB = OB'$$

y, por consiguiente, $A'C' = OB'$. De aquí se deduce que el cuadrilátero $OA'C'B'$ es un paralelogramo y, por ello,

$$\overline{OC'} = \overline{OA'} + \overline{OB'},$$

es decir, si

$$\overline{OA} \to \overline{OA'} \quad \text{y} \quad \overline{OB} \to \overline{OB'},$$

resulta que

$$\overline{OA} + \overline{OB} \to \overline{OA'} + \overline{OB'}.$$

2. Supongamos que $A \to A'$ y $B \to B'$, donde

$$\overline{OB} = \alpha \overline{OA}$$

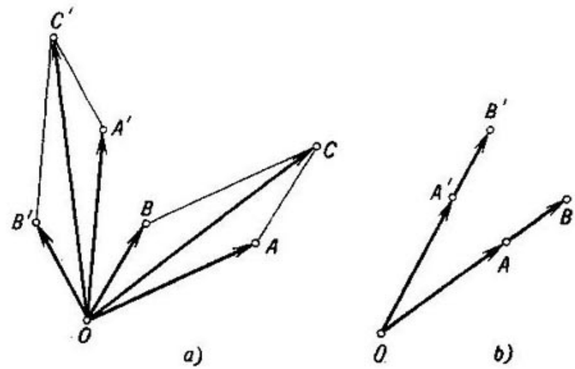


Fig. 30

(fig. 30, *b*). Puesto que los puntos O, A y B son colineales, de las igualdades

$$OA' = OA, \quad OB' = OB \quad \text{y} \quad A'B' = AB$$

resulta entonces que los puntos O, A' y B' son también colineales y están situados en el mismo orden que los puntos O, A y B; como, además,

$$OB' = OB = |\alpha|\cdot OA = |\alpha|\cdot OA',$$

tenemos

$$\overline{OB'} = \alpha\overline{OA}'.$$

Es decir,

$$\text{si } \overline{OA} \to \overline{OA}', \text{ resulta que } \alpha\overline{OA} \to \alpha\overline{OA}'.$$

La aplicación lineal $\mathscr{H}$ no altera las longitudes de los vectores y, por consiguiente, es *una aplicación ortogonal*, es decir, o bien *una rotación alrededor del punto O* o bien *una simetría respecto a un eje que pasa por el punto O* (§ 3 del capítulo V). La transformación (puntual) H tiene este mismo carácter y también puede ser denominada *ortogonal*.

Sea ahora *Q un movimiento cualquiera*. Tomemos en el plano un sistema cartesiano rectangular de coordenadas; sea O su origen y sea O' la imagen del punto O en la transformación Q. El producto $H = T_{O'O}\, Q$[1] del movimiento Q por la traslación paralela $T_{O'O}$ será un movimiento de punto fijo O. Por consiguiente, H es una transformación ortogonal, es decir, o bien una rotación alrededor del punto O o bien una simetría respecto a una recta que pasa por el punto O. Pero en este caso tenemos

$$Q = T_{OO'} H.$$

Hemos demostrado que *todo movimiento puede ser representado como el producto de una transformación ortogonal (es decir, de una rotación alrededor de un punto o de una simetría axial) y de una traslación paralela*[2].

2. *Transformaciones de semejanza.*

Definición 5. *Se llama* ***transformación de semejanza*** *de un coeficiente dado $k > 0$ a la transformación de un espacio euclídeo, en particular de un plano, en la que las longitudes de todos los segmentos se multiplican por un mismo número k:* si $A \to A'$ y $B \to B'$, *se tiene* $A'B' = k\,AB$.

Todas las transformaciones de semejanza de un plano constituyen, obviamente, *un grupo que contiene como subgrupo al grupo de los movimientos.*

Una transformación de semejanza que en un sistema determinado de coordenadas tiene la forma $A(x, y) \to A'(x', y')$, donde

$$x' = kx, \quad y' = ky \quad (k \neq 0), \tag{1}$$

se denomina ***homotecia*** *de centro en el origen de este sistema y de razón k.* Toda transformación de semejanza R con el punto fijo O determina en el espacio vectorial correspondiente una aplicación lineal (esto se comprueba igual que en el caso de los movimientos). Supongamos que en la transformación de semejanza R (de punto fijo O) las longitudes de todos los segmentos se multiplican por k. Tomemos un sistema de coordenadas de origen en el punto O e indiquemos por K la transformación (1). En estas condiciones la transformación $H = K^{-1}R$ no altera las longitudes de los segmentos

[1] Recordamos que el producto de estas transformaciones se lee de derecha a izquierda, es decir, primero se realiza la transformación Q y después la traslación paralela $T_{OO'}$ de vector $\overline{O'O}$.

[2] Se puede demostrar que el producto de una rotación y de una traslación paralela es una rotación alrededor de un punto nuevo.

y, por consiguiente, es una transformación ortogonal, es decir, una rotación o una simetría. Pero entonces $R = KH$; luego, una transformación de semejanza de punto fijo O (transformación de *semejanza con centro*) es el producto de una transformación ortogonal y de una homotecia de centro en el punto O.

Sea ahora U *una transformación de semejanza cualquiera* que transforma el origen de coordenadas O en el punto O'. El producto $T_{O'O}\,U$ será entonces una transformación de semejanza de punto fijo O, es decir, será el producto de una transformación ortogonal y de una homotecia:

$$T_{O'O}U = KH,$$

de donde resulta

$$U = T_{OO'}KH.$$

Hemos demostrado que *toda transformación de semejanza es o bien el producto de una rotación alrededor de un punto, de una homotecia y de una traslación paralela o bien el producto de una simetría axial, de una homotecia y de una traslación paralela*. En forma coordenada

$$x' = k\cos\varphi\cdot x - k\,\mathrm{sen}\,\varphi\cdot y + a_1,$$
$$y' = k\,\mathrm{sen}\,\varphi\cdot x + k\cos\varphi\cdot y + a_2$$

o, en un sistema nuevo de coordenadas,

$$x' = kx + a_1,$$
$$y' = -ky + a_2.$$

3. *Transformaciones afines.*

Supongamos que en un plano afín se ha escogido un sistema de coordenadas. Consideremos la transformación de este plano en la que el punto $A(x, y)$ se transforma en un punto $A'(x', y')$ tal que

$$x' = a_{11}x + a_{12}y + a_1,$$
$$y' = a_{21}x + a_{22}y + a_2, \tag{2}$$

donde

$$\delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

Esta transformación se denomina *afín*.

De la condición $\delta \neq 0$ se deduce que (2) es una transformación biyectiva y que, además, *la transformación inversa*

$$x = \frac{a_{22}x' - a_{12}y' - a_1a_{22} + a_2a_{12}}{\delta},$$

$$y = \frac{-a_{21}x' + a_{11}y' + a_1a_{21} - a_2a_{11}}{\delta}$$

de determinante

$$\begin{vmatrix} \frac{a_{22}}{\delta} & -\frac{a_{12}}{\delta} \\ -\frac{a_{21}}{\delta} & \frac{a_{11}}{\delta} \end{vmatrix} = \frac{1}{\delta^2}\begin{vmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{vmatrix} = \frac{1}{\delta} \neq 0$$

es también afín.

Si, además de la transformación (2), se tiene otra transformación afín

$$x'' = b_{11}x' + b_{12}y' + b_1,$$
$$y'' = b_{21}x' + b_{22}y' + b_2$$

de determinante

$$\delta' = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} \neq 0,$$

el *producto* de estas transformaciones

$$x'' = (b_{11}a_{11} + b_{12}a_{21})x + (b_{11}a_{12} + b_{12}a_{22})y + (b_{11}a_1 + b_{12}a_2 + b_1),$$
$$y'' = (b_{21}a_{11} + b_{22}a_{21})x + (b_{21}a_{12} + b_{22}a_{22})y + (b_{21}a_1 + b_{22}a_2 + b_2)$$

también será una transformación afín de determinante

$$\delta'' = \begin{vmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{vmatrix}$$

igual al producto $\delta\delta'$ (que, por consiguiente, también será diferente de cero).

Hemos obtenido de nuevo un grupo, *el grupo de las transformaciones afines de un plano*. Es fácil ver que la transformación afín (2) es el producto de la transformación *(afín de centro)*

$$\tilde{x} = a_{11}x + a_{12}y, \tag{3}$$
$$\tilde{y} = a_{21}x + a_{22}y$$

(a la que le corresponde en el espacio vectorial una aplicación lineal no degenerada) y de la traslación paralela T

$$x' = \tilde{x} + a_1,$$

$$y' = \tilde{y} + a_2.$$

Puesto que (§ 4 del capítulo V) toda aplicación lineal no degenerada es el producto de una aplicación autoconjugada $\mathcal{C}$ (es decir, de una dilatación a lo largo de dos direcciones recíprocamente ortogonales) y de una aplicación ortogonal $\mathcal{H}$ (es decir, de una rotación o de una simetría), podemos decir que *toda transformación afín A es de la forma*

$$A = THC = QC$$

(donde $Q = TH$ es un movimiento), es decir, *es el producto de una dilatación a lo largo de dos ejes recíprocamente ortogonales y de un movimiento*.

§ 6. Descomposición de un grupo por un subgrupo

Consideremos primero el **ejemplo** siguiente. Sea *G el grupo aditivo de los números enteros* y sea A su *subgrupo* formado por *los números múltiplos de k*. Descompongamos el grupo G en *clases* refiriendo a una misma clase los números que divididos por k dan el mismo resto. Luego, para que dos números x e y pertenezcan a una misma clase es necesario y suficiente que la diferencia de estos números sea divisible por k, es decir, que pertenezca a A:

$x - y = kn \in A$, de donde $x = y + a$ siendo $a \in A$.

Es obvio que de esta forma obtendremos k clases incluyendo el propio subgrupo A.

Esta descomposición del grupo de los números enteros por el subgrupo de los números múltiplos de k se puede representar para $k = 5$ mediante el siguiente esquema

A	..., −20, −15, −10, −5, 0, 5, 10, 15, 20, ...
$A+1$	..., −19, −14, −9, −4, 1, 6, 11, 16, 21, ...
$A+2$	..., −18, −13, −8, −3, 2, 7, 12, 17, 22, ...
$A+3$	..., −17, −12, −7, −2, 3, 8, 13, 18, 23, ...
$A+4$	..., −16, −11, −6, −1, 4, 9, 14, 19, 24, ...

Introduzcamos *la operación de adición en el conjunto de estas clases*. Sean dadas dos clases B y C. Tomemos en cada una de estas clases un elemento (un representante), digamos $b \in B$ y $c \in C$, sumémoslos y llamemos *suma* $B+C$ *de estas clases* a la clase que contiene la suma $b+c$. Es importante, por supuesto, que la clase a la que pertenece la suma $b+c$ no dependa de cómo se escojan los representantes b y c en las clases B y C. Comprobemos que esto es así. Si b' también pertenece a B y $c' \in C$, se tiene

$$b-b' = kn_1 \quad \text{y} \quad c-c' = kn_2$$

y por ello

$$(b+c)-(b'+c') = k(n_1+n_2)$$

es divisible por k, es decir, las sumas $b+c$ y $b'+c'$ pertenecen a una misma clase.

Por lo tanto, *la operación de* ***adición de clases*** *que hemos definido no depende de la elección de los representantes en las mismas*. La adición de clases es *asociativa* y *conmutativa* porque estas propiedades se verifican para la adición en el propio grupo G. La clase coincidente con el subgrupo A desempeña el papel del *cero*, ya que podemos tomar al cero como representante de la clase A y siempre se tiene $g+0 = g$, donde $g \in G$. Finalmente, para toda clase B existe la clase *opuesta:* si $b \in B$, la clase que contiene a $-b$ será la opuesta de B, ya que

$$b+(-b) = 0 \in A.$$

El grupo (de clases) que hemos obtenido se llama ***grupo cociente*** *del grupo de los números enteros respecto al subgrupo de los números múltiplos de k*. Es, obviamente, un grupo ***cíclico*** de orden k.

Apliquemos ahora una idea análoga en el caso general. Sea G un grupo arbitrario, esta vez multiplicativo, y sea A uno de sus subgrupos. Indiquemos por xA el conjunto de todos los elementos de tipo xa, donde $a \in A$; este conjunto xA se denomina ***clase contigua por la izquierda*** *del grupo G respecto al subgrupo A*.

Diremos que *todo elemento y perteneciente a la clase xA es* ***equivalente*** *a* x (escribiremos $x \sim y$). Señalemos las propiedades siguientes de este concepto:

1. *Todo elemento x es equivalente a sí mismo* ($x \sim x$), ya que $x = xe \in xA$.

2. *Si* $y \sim x$, *también* $x \sim y$. Efectivamente, si $y \sim x$, es decir, si $y \in xA$, se tiene $y = xa$, donde $a \in A$; luego $x = ya^{-1} \in yA$, es decir, $x \sim y$.

3. *Si $x \sim y$ e $y \sim z$, también $x \sim z$.* Por hipótesis, $x \sim y$, es decir, $x = ya_1$ e $y \sim z$, es decir, $y = za_2$, donde $a_1, a_2 \in A$. Pero en este caso se tiene $x = za_2a_1 \in zA$ y, por consiguiente, $x \sim z$.

4. *Si los elementos x e y son equivalentes, sus clases contiguas xA e yA coinciden.* Efectivamente, sea $x' \in xA$, es decir, $x' \sim x$. Puesto que $x \sim y$, resulta que $x' \sim y$, es decir, $x' \in yA$. Análogamente todo elemento y' de la clase yA pertenece también a xA y, por consiguiente, $xA = yA$.

5. *Si los elementos x e y no son equivalentes, las clases xA e yA no tienen elementos comunes.* Efectivamente, si $z \in xA \cap yA$, resulta que $z \sim x$ y $z \sim y$; luego, $x \sim y$.

Todo el grupo G resulta, pues, partido en clases contiguas por la izquierda (sin elementos comunes) respecto al subgrupo A siendo el propio subgrupo A una de estas clases. Si G es un grupo finito, todas las clases contiguas contienen un mismo número de elementos (los elementos de la clase contigua xA corresponden biyectivamente a los elementos del subgrupo A, porque si xa_1 es igual a xa_2 resulta que a_1 es igual a a_2). De aquí se deduce el siguiente teorema importante.

Teorema de Lagrange. *El orden de un subgrupo de un grupo finito es divisor del orden de este grupo.*

Demostración. Sea G un grupo finito de orden n y sea A su subgrupo de orden k. Realicemos la descomposición del grupo G en las clases contiguas por la izquierda respecto al subgrupo A. Sea j el número de las clases obtenidas; puesto que todas las clases constan de k elementos, resulta que el número total de elementos del grupo es

$$n = kj$$

y, por consiguiente, n es divisible por k. El número j se llama ***índice*** del subgrupo A en el grupo G.

Todo elemento g de un grupo G genera en éste un subgrupo cíclico $\{g\}$ formado por todas las potencias de este elemento. El orden del subgrupo $\{g\}$ se denomina ***orden del elemento*** g en el grupo G. Debido al teorema de Lagrange resulta que el orden de *todo elemento de un grupo finito es divisor del orden del grupo.*

Todo grupo finito cuyo orden es un número primo es cíclico, ya que el subgrupo cíclico generado por cualquiera de sus elementos (a excepción de e) debe coincidir con todo el grupo.

Análogamente a la descomposición por la izquierda puede ser construida la descomposición *por la derecha* del grupo G respecto al subgrupo A (en clases Ax). En el caso conmutativo ambas descomposiciones coinciden (es decir, constan de las mismas clases).

En un grupo no conmutativo las descomposiciones por la izquierda y por la derecha pueden resultar distintas. Consideremos, por ejemplo, la descomposición del grupo simétrico S_3 respecto a su subgrupo $B = \{P_1, P_2\}$. La descomposición por la izquierda está formada por las clases

$$B,\ P_5B = P_4B = \{P_4, P_5\},\ P_3B = P_6B = \{P_3, P_6\}$$

y la descomposición por la derecha consta de las clases

$$B,\ BP_6 = BP_4 = \{P_4, P_6\},\ BP_5 = BP_3 = \{P_3, P_5\}.$$

Al mismo tiempo, coinciden las descomposiciones por la izquierda y por la derecha de este grupo respecto a su subgrupo de tercer orden $A = \{P_1, P_5, P_6\}$: cada una de estas descomposiciones está formada por dos clases

$$A = \{P_1, P_5, P_6\} \quad \text{y} \quad B = \{P_2, P_3, P_4\}.$$

§ 7. Divisor normal

Generalicemos ahora la idea que nos ha conducido en el § 6 al concepto de grupo de clases (grupo cociente del grupo aditivo de los números enteros). Sea A un subgrupo de un grupo cualquiera G. Formemos todas las clases contiguas por la izquierda del grupo G respecto al subgrupo A e intentemos de definir la operación de multiplicación de estas clases de la forma siguiente: dadas dos clases B y C, tomemos en cada una de ellas un representante $b \in B$ y $c \in C$, multipliquemos estos representantes y definamos el producto BC como la clase a la que pertenece bc. Es preciso comprobar solamente *si esta definición del producto de clases no depende de cómo se escojan los representantes de las mismas*. Sea, pues, $b' \sim b$ y $c' \sim c$; ¿será equivalente el producto $b'c'$ a bc? Tenemos, por hipótesis, $b' = ba_1$ y $c' = ca_2$, de donde resulta que $b'c' = ba_1ca_2$. Si el grupo G es conmutativo, se tiene

$$a_1c = ca_1 \tag{4}$$

y $b'c' = bc(a_1a_2)$, es decir, $b'c' \sim bc$.

En un grupo no conmutativo la igualdad (4), en general, no tiene lugar. Sin embargo, para nuestros fines es suficiente que se cumpla la siguiente condición más débil que la de conmutatividad:

es suficiente que el producto a_1c pueda ser representado en la forma ca_3, donde a_3 es un elemento del subgrupo A diferente, en general, *de a_1*. Si esto es así, tenemos $b'c' = bc(a_3a_2)$ y $b'c'$ es equivalente a bc; luego el producto de clases *no depende de cómo se escojan los representantes de las mismas*.

Por esto, aceptaremos ahora que el subgrupo A satisface a la condición siguiente: para todo elemento $a \in A$ y un elemento cualquiera $g \in G$ existe un elemento $\tilde{a} \in A$ tal que $ag = g\tilde{a}$. Esto significa que para todo $a \in A$ y cualquier $g \in G$ el producto $g^{-1}ag = \tilde{a} \in A$. Indicando por $g^{-1}Ag$ el conjunto de todos los elementos de tipo $g^{-1}ag$, donde $a \in A$, introduciremos el concepto siguiente.

Definición 6. *Un subgrupo A de un grupo G se llama* ***subgrupo invariante*** *o* ***divisor normal*** *de éste si*

$$g^{-1}Ag \subseteq A$$

para todo elemento $g \in G$.

Teorema 1. *La intersección de dos divisores normales de un grupo G es también un divisor normal de G.*

Demostración. Sean A_1 y A_2 dos divisores normales del grupo G y sea $A = A_1 \cap A_2$. Sabemos que A es un subgrupo de G (§ 3). Ahora bien, puesto que $A \subseteq A_1$ y $A \subseteq A_2$, resulta que para cualquier elemento $g \in G$ se tiene

$$g^{-1}Ag \subseteq g^{-1}A_1g \subseteq A_1$$

y

$$g^{-1}Ag \subseteq g^{-1}A_2g \subseteq A_2$$

y, por consiguiente,

$$g^{-1}Ag \subseteq A_1 \cap A_2 = A,$$

es decir, A es un divisor normal.

Teorema 2. *Para que un subgrupo A de un grupo G sea un divisor normal de éste es necesario y suficiente que para todo elemento $g \in G$* se cumpla la igualdad

$$g^{-1}Ag = A. \tag{5}$$

Demostración. La suficiencia de la condición (5) resulta de la definición 6. Para demostrar la necesidad, supongamos que A es un divisor normal del grupo G; para todo elemento $g \in G$ se tiene entonces

$$g^{-1}Ag \subseteq A$$

y, por consiguiente,

$$gAg^{-1} = (g^{-1})^{-1} Ag^{-1} \subseteq A.$$

Por otro lado, tenemos

$$A = eAe = (g^{-1}g) A(g^{-1}g) = g^{-1}(gAg^{-1}) g \subseteq g^{-1} Ag,$$

es decir, $A \subseteq g^{-1}Ag$ y, por consiguiente, $g^{-1}Ag = A$.

Teorema 3. *Para que un subgrupo A sea un divisor normal de un grupo G es necesario y suficiente que coincidan las clases contiguas por la izquierda y por la derecha del grupo G respecto al subgrupo A.*

Demostración. De la igualdad

$$g^{-1}Ag = A$$

resulta que

$$Ag = gA,$$

es decir, para todo $g \in G$ coinciden las clases contiguas por la izquierda y por la derecha que contienen este elemento.

Recíprocamente, si para todo $g \in G$ es

$$Ag = gA,$$

se tiene

$$g^{-1}Ag = A$$

y A es un divisor normal.

Por ejemplo, en el grupo simétrico S_3 el subgrupo $A = \{P_1, P_2, P_6\}$ es un divisor normal, mientras que los subgrupos $\{P_1, P_2\}$, $\{P_1, P_3\}$ y $\{P_1, P_4\}$ no son divisores normales de S_3. En un grupo conmutativo cualquier subgrupo es un divisor normal.

§ 8. Grupo cociente

Sean G un grupo cualquiera, A un divisor normal del mismo y S el conjunto de todas las clases contiguas del grupo G respecto al subgrupo A (recordemos que en este caso las clases contiguas por la izquierda y por la derecha coinciden). *Introduciremos en el conjunto S una operación de multiplicación tomando*

$$xA\cdot yA = xyA.$$

Puesto que A es un divisor normal, *el producto $xA\cdot yA$ no depende de la selección de los representantes x e y de las clases que se multiplican.*

Demostremos que se ha obtenido un grupo.

1. **La asociatividad** de la multiplicación de las clases resulta de la asociatividad de la multiplicación en el grupo G:

$$(xA\cdot yA)\cdot zA = (xy)\,A\cdot zA = (xy)\,zA =$$
$$= x(yz)\,A = xA\cdot(yz)\,A = xA\cdot(yA\cdot zA).$$

2. **El elemento unidad** es el propio subgrupo A:

$$A\cdot xA = eA\cdot xA = exA = xA,$$
$$xA\cdot A = xA\cdot eA = xeA = xA.$$

3. La clase $x^{-1}A$ es **la inversa** de la clase xA, ya que

$$xA\cdot x^{-1}A = xx^{-1}A = eA = A,$$
$$x^{-1}A\cdot xA = x^{-1}xA = eA = A.$$

El grupo obtenido se indica por G/A y se denomina ***grupo cociente*** del grupo G relativo al divisor normal A.

El grupo cociente de un grupo conmutativo es conmutativo, ya que en este caso para dos clases cualesquiera se tiene

$$xA\cdot yA = (xy)\,A = (yx)\,A = yA\cdot xA.$$

El orden del grupo cociente de un grupo finito es igual al índice del divisor normal A en el grupo G (y, por consiguiente, es un divisor del orden n del grupo G).

El grupo cociente del grupo simétrico S_3 relativo a su subgrupo A consta de dos elementos $A = \{P_1, P_5, P_6\}$ y $B = \{P_2, P_3, P_4\}$ y es, por consiguiente, un grupo cíclico de orden dos (es fácil ver que $B^2 = A$).

Ejemplo. *Demuéstrese que en el grupo G de todas las matrices no degeneradas de orden n* (de elementos reales, por ejemplo) *es un divisor normal el subgrupo A de las matrices unimodulares* (es decir, de las matrices de determinante igual a 1) *y determínese el grupo cociente G/A.*

Solución. Las matrices unimodulares forman un subgrupo de G, ya que el producto de dos matrices unimodulares y la inversa de una matriz unimodular son matrices unimodulares (teorema 3 del capítulo III).

Además, el subgrupo A de las matrices unimodulares es un divisor normal de G: si la matriz $a \in A$ y, por consiguiente, $|a| = 1$, para toda matriz $g \in G$ se tiene

$$|g^{-1}ag| = |g^{-1}|\,|a|\,|g| = |g|^{-1}\,|a|\,|g| = |a| = 1,$$

es decir,

$$g^{-1}ag \in A.$$

Determinemos el grupo cociente G/A. Demostremos, ante todo, que una condición necesaria y suficiente para que dos matrices b y c pertenezcan a una misma clase contigua del grupo G respecto al divisor normal A es que sus determinantes sean iguales. Efectivamente, si $b \sim c$, es decir, si $b = ca$, donde $a \in A$ (y, por consiguiente, $|a| = 1$), tenemos $|b| = |c|\,|a| = |c|$.

Recíprocamente, si $|b| = |c|$, se tiene $b = c(c^{-1}b)$, donde

$$|c^{-1}b| = |c^{-1}|\,|b| = |c|^{-1}\,|b| = 1$$

y, por consiguiente, $c^{-1}b \in A$, es decir, $b \sim c$.

Luego, toda clase contigua de G respecto de A se caracteriza plenamente por el valor del determinante de las matrices que la componen. La multiplicación de clases corresponde a la multiplicación de cualesquiera representantes de las mismas, es decir, el producto de la clase B (formada por las matrices de determinante β) y de la clase C (formada por las matrices de determinante γ) es la clase BC de las matrices de determinante $\beta\gamma$. Hemos encontrado, pues, que *el grupo cociente G/A es isomorfo al grupo multiplicativo de los números reales diferentes de cero.*

§ 9. Divisores normales del grupo de transformaciones de un plano euclídeo y sus correspondientes grupos cocientes

A. Ejemplos de divisores normales

1. *El grupo de traslaciones paralelas es un divisor normal del grupo de movimientos de un plano.*

Demostración. Sea Q un movimiento cualquiera y sea T una traslación paralela. Es preciso demostrar que la transformación $Q^{-1}TQ$ es una traslación paralela. Sean A y B unos puntos arbitrarios del plano; supongamos que

$A \to A'$ y $B \to B'$ en la transformación Q;
$A' \to A''$ y $B' \to B''$ en la transformación T;
$A'' \to A'''$ y $B'' \to B'''$ en la transformación Q^{-1}.

En estas condiciones resulta que

$A \to A'''$ y $B \to B'''$ en la transformación $Q^{-1}TQ$

(fig. 31). Debemos demostrar que el segmento AA''' es igual y paralelo al segmento BB'''.

La transformación $Q^{-1}\,TQ$ es un movimiento; luego

$$A'''B''' = AB. \tag{6}$$

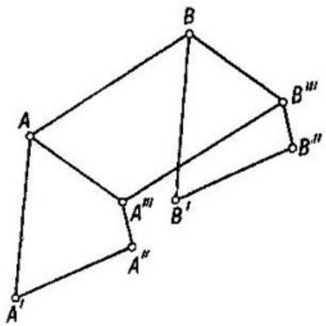


Fig. 31

Puesto que para el movimiento Q el punto $A \to A'$ y el punto $A''' \to A''$, tenemos

$$AA''' = A'A'' \tag{7}$$

y como, ademas, en este caso el punto $B \to B'$ y el punto $B''' \to B''$, tenemos también

$$BB''' = B'B''. \tag{8}$$

Por otro lado, T es una traslación paralela, luego

$$A'A'' = B'B''. \tag{9}$$

De las igualdades (7), (8) y (9) resulta que

$$AA''' = BB''' \tag{10}$$

y de las igualdades (6) y (10) se deduce que el cuadrilátero $AA'''\,B'''\,B$ es un paralelogramo y, por consiguiente, $AA''' \parallel BB'''$, es decir, la transformación $Q^{-1}TQ$ es una traslación paralela.

2. *El grupo de las traslaciones paralelas es también un divisor normal del grupo de las transformaciones afines.*

Demostración. Sea A una transformación afín cualquiera y sea T una traslación paralela. Debemos demostrar que la transformación $A^{-1}TA$ es una traslación paralela.

En el § 5 hemos demostrado que toda transformación afín A se puede representar en la forma $A = QC$, donde Q es un movimiento y C es una transformación autoconjugada. Pero entonces resulta que

$$A^{-1}TA = (QC)^{-1}T(QC) = C^{-1}(Q^{-1}TQ)C.$$

El producto $Q^{-1}TQ = T'$ es, según hemos demostrado en el p. 1, una traslación paralela. Consideremos la transformación $C^{-1}T'C$. Supongamos que el sistema de coordenadas se ha escogido de forma que la transformación C transforma el punto $A(x, y)$ en el punto $A'(\lambda_1 x, \lambda_2 y)$; la traslación paralela T' de vector (a_1, a_2) transformará el punto A' en el punto $A''(\lambda_1 x + a_1, \lambda_2 y + a_2)$. En estas condiciones, la transformación C^{-1} transformará el punto A'' en el punto $A'''\left(x+\frac{a_1}{\lambda_1}, y+\frac{a_2}{\lambda_2}\right)$ y, por consiguiente, la transformación $C^{-1}T'C$ es una traslación paralela de vector $\left(\frac{a_1}{\lambda_1}, \frac{a_2}{\lambda_2}\right)$.

3. *El grupo de las traslaciones paralelas también es un divisor normal del grupo de las transformaciones de semejanza, ya que este último es un subgrupo del grupo de las transformaciones afines.*

4. *El grupo de los movimientos es un divisor normal del grupo de las transformaciones de semejanza.* Efectivamente, sea U una transformación de semejanza cualquiera y sea Q un movimiento. En este caso $U^{-1}QU$ es una transformación de semejanza que no altera las longitudes de los segmentos (en la transformación U todas las longitudes se multiplican por k, en la transformación Q no varían y en la transformación U^{-1} se multiplican por k^{-1}) y es, por consiguiente, un movimiento.

B. Grupos cocientes

1. *El grupo cociente del grupo $\mathfrak{A}$ de transformaciones afines respecto al subgrupo $\mathfrak{C}$ de traslaciones paralelas.*

Dos transformaciones afines A_1 y A_2 son equivalentes (es decir, pertenecen a una misma clase respecto al subgrupo $\mathfrak{C}$), si, y sólo si, $A_1 = TA_2$, donde T es una traslación paralela. Si A es una transformación afín arbitraria que transforma el origen de coordenadas O en O', la transformación $T_{O'O}A$, que deja fijo el punto O, es equivalente a A: $T_{O'O}A \sim A$. Por consiguiente, toda transformación afín es equivalente a una transformación afín de punto fijo O. Por otro lado, si dos transformaciones afines A_1 y A_2 dd punto fijo O son equivalentes, es decir, $A_1 = TA_2$, resulta que el producto $A_1A_2^{-1} = T$ es una traslación paralela de punto fijo O es decir, es la transformación idéntica

$$A_1A_2^{-1} = E,$$

de donde se tiene que $A_1 = A_2$. Luego, todas las transformaciones afines de punto fijo O (las transformaciones afines de centro) constituyen un sistema completo de representantes de todas las clases contiguas del grupo $\mathfrak{A}$ respecto al subgrupo $\mathfrak{C}$. Por consiguiente, el grupo cociente $\mathfrak{A}/\mathfrak{C}$ es isomorfo *al grupo de las transformaciones afines de centro O* que en forma coordenada puede ser representado como

$$x' = a_{11}x+a_{12}y,$$
$$y' = a_{21}x+a_{22}y,$$

donde el determinante $\delta \neq 0$; luego, es isomorfo *al grupo de las matrices no degeneradas de segundo orden.*

2. *El grupo cociente del grupo $\mathfrak{B}$ de las transformaciones de semejanza relativo al subgrupo $\mathfrak{C}$ de las traslaciones paralelas.*

Toda transformación de semejanza es equivalente a una transformación de semejanza de punto fijo O (a una transformación de semejanza de centro) y dos transformaciones equivalentes de este tipo son idénticas. Por consiguiente, el grupo cociente $\mathfrak{B}/\mathfrak{C}$ es isomorfo al grupo de las transformaciones de semejanza de centro O. Pero toda transformación de este tipo puede ser representada como el producto KH de una transformación ortogonal H y de una homotecia K con un coeficiente determinado $k \neq 0$ (véase la pág. 262). Es verdad que esta representación no es unívoca. Si S es una simetría de centro O, se tiene $S^2 = E$ y

$$KH = KS^2H = (KS)(SH),$$

donde SH es también una transformación ortogonal y KS es una homotecia de coeficiente $-k$. Por ello, podemos limitarnos a considerar las descomposiciones de tipo KH, donde K es una homotecia de coeficiente *positivo*. Esta representación de una transformación de semejanza de centro será ya unívoca, porque de la igualdad

$$K_1H_1 = K_2H_2 \tag{11}$$

resulta

$$K_2^{-1}K_1 = H_2H_1^{-1};$$

luego la homotecia $K_2^{-1}K_1$ de coeficiente $k_1/k_2 > 0$ será también una transformación ortogonal (que no varía las longitudes de los segmentos), de modo que $k_1/k_2 = 1$ y $k_1 = k_2$, es decir, $K_1 = K_2$. De aquí y de la igualdad (11) resulta que también $H_1 = H_2$.

Por consiguiente, obtenemos un sistema completo de representantes de todas las clases de la descomposición del grupo $\mathfrak{B}$ respecto al divisor normal $\mathfrak{C}$ si tomamos todos los productos de tipo KH, donde H es una transformación ortogonal y K es una homotecia de coeficiente $k > 0$. Luego, el grupo cociente $\mathfrak{B}/\mathfrak{C}$ es isomorfo *al grupo de las matrices de tipo kh*, donde h es una matriz ortogonal de segundo orden y k es un número positivo.

3. *El grupo cociente del grupo de los movimientos relativo al subgrupo de las traslaciones paralelas* es isomorfo al grupo de los movimientos de punto fijo, es decir, es isomorfo *al grupo de las matrices ortogonales de orden dos*.

4. *El grupo cociente del grupo de las transformaciones de semejanza* $\mathfrak{B}$ *relativo al subgrupo* $\mathfrak{R}$ *de los movimientos*.

Para toda transformación de semejanza existe una equivalente transformación de semejanza de centro O. Cualquier transformación de este tipo se puede representar como al producto KH, donde H es una transformación ortogonal y K es una homotecia de coeficiente $k > 0$ (p. 2). Si las transformaciones K_1H_1 y K_2H_2 son equivalentes, se tiene $K_1H_1 = K_2H_2Q$, donde Q es un movimiento; como la homotecia $K_2^{-1}K_1 = H_2QH_1^{-1}$ de coeficiente $k_1/k_2 > 0$ no varía las longitudes de los segmentos (por ser un movimiento), resulta que $k_1/k_2 = 1$, de donde $K_1 = K_2$. Recíprocamente, si $K_1 = K_2 = K$, las transformaciones KH_1 y KH_2 son equivalentes, porque $KH_1 = KH_2(H_2^{-1}H_1)$, donde $H_2^{-1}H_1$ es un movimiento. Por consiguiente, obtenemos un sistema completo de representantes de la descomposición del grupo $\mathfrak{B}$ respecto al subgrupo $\mathfrak{R}$ tomando, por ejemplo, todas las homotecias de centro O y de coeficiente positivo. Por ello, el grupo cociente $\mathfrak{B}/\mathfrak{R}$ es isomorfo *al grupo multiplicativo de los números positivos*.

INDICE ALFABÉTICO

A NUESTROS LECTORES:

"Mir" edita libros soviéticos traducidos al español inglés, francés y árabe. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica: manuales para los centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia ficción. Dirijan sus opiniones a Editorial "Mir", Rizhski per., 2, 129820, Moscú, I-110, GSP, URSS.

Gmurman V.

OBRA DE REFERENCIA PARA RESOLVER LOS PROBLEMAS DE LA TEORIA DE LAS PROBABILIDADES Y DE LA ESTADISTICA MATEMATICA

Por su contenido, esta obra corresponde al curso de la teoría de las probabilidades y de estadística matemática y contiene más de 500 problemas para todas las partes del curso. Al principio de cada párrafo el autor expone los conocimientos teóricos necesarios, las fórmulas requeridas y muestra los ejemplos cómo resolver problemas típicos. Los problemas destinados para resolverse independientemente se dan según sus difficultades crecientes. Además de los problemas clásicos acerca de lanzamientos de monedas o dados, este libro contiene un gran número de problemas que ilustran la aplicación de la teoría de las probabilidades en la técnica (fiabilidad, control estadístico de la calidad de producción y tratamiento de los datos experimentales). El mérito fundamental de este libro consiste en que el autor, en todos los problemas, pone en primer plano su esencia probabilística, reduciendo al mínimo el cálculo numérico.

Esta obra puede servir para enseñar a resolver los problemas de la teoría de las probabilidades, de estadística matemática y para los estudiantes de las especialidades ingeniero-técnicas y económicas.

ACABA DE APARECER

Gmurman V.

LA TEORIA DE LAS PROBABILIDADES Y LA ESTADISTICA MATEMATICA

Este tratado ha sido escrito de conformidad con el nuevo curso de la teoría de las probabilidades y de la estadística matemática. El material de esta obra se presenta en tres partes fundamentales: las dos primeras se dedican a la teoría de las probabilidades y la última, a la estadística matemática. En este manual se estudian los temas siguientes: probabilidad de las hipótesis; fórmulas de Bayes; distribución de Poisson; nociones de la estadística matemática y noción de correlación.

La cuarta edición de este libro contiene algunos capítulos nuevos: distribución exponencial; verificación de las hipótesis estadísticas y análisis dispersional uniforme.

En el libro se presta gran atención a los métodos estadísticos de elaboración de los datos experimentales; las tablas que se exponen para los cálculos son muy cómodas. Cada capítulo contiene problemas y sus respuestas que han sido elegidos adecuadamente. Además, todo capítulo va acompañado del análisis de las soluciones de los problemas del material correspondiente.

Esta obra contiene diecisiete capítulos, varias tablas de números, veintidós figuras y un gran número de ejemplos teóricos y técnicos; se recomienda para los estudiantes de las facultades ingeniero-técnicas y económicas.